POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Identification of Misbehavior Detection Solutions and Risk Scenarios in Advanced Connected and Automated Driving Scenarios

Tesi di Laurea Magistrale in Telecommunication Engineering - Ingegneria delle Telecomunicazioni

Author: **Cristina Insalaco**

Student ID: 10568530
Advisor: Prof. Monica Barbara Nicoli
Co-advisors: Luis Jofre Roca, Luca Montero Bayo
Academic Year: 2021-2022

Abstract

The inclusion of 5G cellular communication system into vehicles, combined with other connected-vehicle technology, such as sensors and cameras, makes connected and advanced vehicles a promising application in the Cooperative Intelligent Transport Systems.
One of the most challenging task is to provide resilience against misbehavior i.e., against vehicles that intentionally disseminate false information to deceive receivers and induce them to manoeuvre incorrectly or even dangerously. This calls for misbehaviour detection mechanisms, whose purpose is to analyze information semantics to detect and filter attacks. As a result, data correctness and integrity are ensured.
Misbehaviour and its detection are rather new concepts in the literature; there is a lack of methods that leverage the available information to prove its trustworthiness. This is mainly because misbehaviour techniques come with several flavours and have different unpredictable purposes, therefore providing precise guidelines is rather ambitious. Moreover, dataset to test detection schemes are rare to find and inconvenient to customize and adapt according to needs.
This work presents a misbehaviour detection scheme that exploits information shared between vehicles and received signal properties to investigate the behaviour of transmitters. Differently from most available solutions, this is based on the data of the on-board own resources of the vehicle. Computational effort and resources required are minor concerns, and concurrently time efficiency is gained. Also, the project addresses three different types of attack to show that detecting misbehaviour methods are more vulnerable to some profile of attacker than others.
Moreover, a rich dataset was set up to test the scheme. The dataset was created according to the latest standardised evaluation methodologies and provides a valuable starting point for any further development and research.

Keywords: V2X, 5G, cyber security, misbehaviour detection, connected vehicles, advanced driving systems.

Contents

1 | Introduction

Over the past twenty years, driving systems have undergone an unprecedented evolution. Vehicles communicate to each other and concepts such as connected vehicles are increasingly present in everyday life. Future vehicles will be connected and exchange information, and within the transportation field connectivity is a driving enabler for the provision of value-added services.

Recently, these advanced vehicles are exponentially gaining momentum thanks to the outstanding technological advancements which need to meet demand within the automotive industry for more advanced and automated driving applications. For instance, 5G is expected to further extend vehicle automation, supporting a wide range of societal benefits, the Internet of Things (IoT) shall contribute to collect additional data, that in turn will be essential to develop new services for vehicle users.

Such promising technologies are evolving rapidly and are expected to transform the whole driving experiences, provide safer cars, and improve travel efficiency. Developing vehicle automation would also create new employment opportunities and open business opportunities. The very basis for enabling this breakthrough is to rely on appropriate communication technologies, referred to as *vehicular communications*.

Vehicular communications are a type of communication system specifically designed to support a fast-moving environment with very stringent requirements in terms of latency and reliability, and where data traffic and users are extremely dense. Current mobile network technologies do not meet these needs, while 5G new radio (5G-NR), the latest mobile network generation, is expected to increase network capacity, improve reliability and availability, and lower latencies.

Leveraging such communication systems, vehicles in proximity exchange information such as their position, their speed, and their role – car, ambulance, track, etc, and also data collected from on-board sensors. As a result, users' awareness increases and self-decision

taking is enabled. Many more advantages are achieved: mobility facilitation, more efficient traffic management, greener transports, better road safety, etc.

On the other hand, many challenges come into view. Apart from the technological support perspective, there are many open discussions concerning security and data truthfulness: it is essential to provide vehicles with reliable data to ensure safe driving experiences. Let set an example to deeply get the importance of this part. There are two vehicles, vehicle A receives a message from vehicle B informing that vehicle B is behind, but in reality, vehicle B is in front of A and driving in the opposite direction. Vehicle A, thinks the road is free and decides to do an overtaking, but then it crushes against vehicle B.

This example represents the worst-case scenario that could happen and is a real issue. Besides, less harmful behaviour can occur. For instance, a vehicle wants to benefit a clear road to get home as quickly as possible after work. It sends many messages to make the road appear congested when it is not, and other vehicles may choose a different route. In the end the malicious vehicle has the road for himself.

These two examples above give an idea of the meaning of **misbehaviour detection**. Misbehaviour implies that there is someone, an attacker, who is not behaving as expected and whose intents are malicious. Falsifying the information transmitted to surrounding vehicles or infrastructure is a real threat as it can lead to several damaging inconveniences.

Misbehaviour detection falls under the wide umbrella of cybersecurity, and it goes beyond privacy matters or ensuring safe and secure communications. Guaranteeing access to the network only to legitimate users is undeniably important, but this does not obstacle misuse and injection of false data. The very first purpose in misbehaviour detection mechanisms is to analyse the received data and determine whether these data stick to reality or not. In other words, vehicle trustworthiness is investigated by means of behaviour analysis. Any useful information is included in these analysis; data from on-board sensors, other surrounding vehicles, past and future movements, communication activity such as message frequency etc.

The misbehaviour detection world is extremely broad and diverse, it is still going through a very immature stage, many aspects still need to be studied. Moreover, it comes with many complexities, responsibilities and legal issues, and thrilling challenges. The research is becoming increasingly rich in works along this direction and certainly, in the coming years, these topics will bounce back again and again as it is fundamental to provide

guarantees in this regard and to ensure reliability.

1.1. Motivation

Misbehaviour detection is a rather new concept, and this makes it difficult to approach it in a unified direction. Implementing a scheme to analyse the behaviour of vehicles may involve data from one vehicle only as well as data from several vehicles. To evaluate the information one can use data from sensors, information got from network entities, pedestrian devices e.g., smartphones or smartwatches, and also data included in messages received by other vehicles.

In addition, misbehaviour may come into many different flavours, and, as explained in the introduction, many different goals are possible. As a result, it is even more a complex and unpredictable reality that calls for studies, research and investigation. To be precise, the literature suggests many deep learning-based methods [36], while methods leveraging characteristics at the physical layer are quite unexplored. However, in general, there is a lack of guidelines to implement a scheme assessing the trustworthiness of the information a vehicle receives. And, moreover, there are no directives on how to certify the performances of a suggested method.

Beyond matters related to misbehaviour detection schemes, another significant limitation is that there is very little availability of dataset to be used as tests. Since these methods should all stick to common standards, having an available dataset to use would simplify and significantly speed up research.

In the end, two main motivations moved and led this work:

1. Developing a misbehaviour detection scheme that leverages physical properties of signals received.
2. Creating a dataset to easily customize as needed to support future works.

1.2. Purpose

The main purpose of the conducted work is to implement a misbehaviour detection mechanism. In particular, the idea is to leverage characteristics from the physical layer to prove the correctness of the received data. The primary goal is to suggest a method able to discriminate malicious data but also to ensure that this method works well in case of

non-compromised data. To this end, the method developed for this work is subjected to a double test:

- **Reliability test**. Trustworthy data are used to examine the capabilities and performances of the method.
- **Misbehaviour detection test**. Falsified data are considered to assess the success of the method for its intended purpose.

Moreover, a reusable dataset for later studies was made. As mentioned in the previous section regarding the motivation, there is a lack of data for testing these methods. With this motivation, one of the goals of this work was to implement a sufficiently considerable dataset compliant to the communication standards to be used in future works.

1.3. Structure of the thesis

This first chapter gave a glimpse on the main topics handled in this thesis and illustrated its purposes. In the subsequent chapter, theoretical topics about vehicular communications are deeply discussed. Chapter 2 is organized into two main sections, first the emerging technologies and applications are explained and then topics related to misbehaviour detection are covered.

Chapter 3 presents the methods and tools used to implement the misbehaviour detection scheme. It also includes an overview of the simulation scenarios reproduced and the standard documentation taken as reference.

Chapter 4 reports the obtained results. First, a part is presented showing the performance of the method in case of faithful data, then the ability to detect malicious data is assessed.

Chapter 5 closes the work by reviewing the main achievements, pointing out weaknesses and shortcomings and suggesting possible future works.

2 | Background

This first chapter is devoted to an overview of some theoretical concepts which are fundamentals to fully understanding the discussed topics. First of all, some basics related to the vehicular communications are provided, showing the most advanced technologies involved and mentioning the most recent developments. A good deal of room is dedicated to a comprehensive explanation of the standardised messages these communication systems adopt. Then, misbehaviour detection takes up the second part of the chapter; after explaining its meaning, an insight of existing misbehaviour detection identification techniques is given.

2.1. Vehicular Communications

The vehicular communications are a communication technology that enables vehicles to exchange information with other vehicles and with nearby roadside infrastructure. The type of information transmitted may concern either vehicles' information such as speed and positioning, or data collected by vehicle's sensors. With this information, vehicles awareness increases, and autonomous decisions can be taken.

Vehicular communications made possible the development of many applications related to the world of the cooperative intelligent transport systems (C-ITS).

2.1.1. Cooperative Intelligent Transport Systems

C-ITS are systems in which vehicles and road infrastructure cooperate to improve the traffic safety and support traffic efficiency applications. C-ITS can be envisioned as a more advanced subset of the Intelligent Transport Systems (ITS) [33]. This is a wider term that refers to integrated systems combining several IT engineering branches with transport engineering to plan, design, operate and manage transport systems.

ITS applications are expected to significantly improve efficiency, safety and security of road transport, minimize environmental impact, whilst opening up a wide range of new business and market opportunities. For years, ITS have been part of our daily lives, some

good examples are navigation systems like the GPS, parking management systems that inform drivers about parking occupancy level, speed detection tools and red light cameras detecting vehicles crossing the street on a red light.

Figure 2.1: Examples of Intelligent Transport Systems applications.

For its side, C-ITS involve real-time information exchange among road users, which can be cars as well as portable devices, and infrastructure enabling advanced applications and services. Broadly speaking, there are two main application fields: traffic safety or traffic efficiency [20]. The former aims to reduce the number and severity of road accidents, the latter to make mobility and traffic management more efficient, reduce fuel consumption, provide infotainment services, etc. In the case of traffic safety, drivers can obtain useful information and assistance from vehicles in proximity but also pedestrians or road infrastructure. Instead, efficiency applications, also referred as non-safety applications, mainly rely on a communication between vehicles and external networks.

Development opportunities are not missing, and benefits are clear: better traffic-flow solutions, safer driving conditions, more comfortable driving experience. As a result, the transport system is populated with new applications enabled by vehicular communications. In short, connectivity among vehicles and their surroundings is a promising key enabler for the provision of C-ITS applications.

2.1.2. Standardisation bodies

Further on, an overview of these applications with their requirements will be given and, it will be observed that they have stringent and varied requirements. Thus, it is fundamental to set rules and to employ appropriate communication technologies and protocols. Besides, the services provided have to be safe: vehicles, devices and network infrastructures are to be put in a position of communicating coherently and smoothly. With these motivations, the V2X communications, like any type of communication system, must be based on standards.

In telecommunications, standards are the result of agreements involving many diverse bodies, called Standards Developing Organizations (SDO), and they define specification models. These represent a guideline that implementation entities have to be compliant with to ensure efficient, safe and high-quality solutions. Therefore, standards are paramount for the deployment of communication systems to guarantee proper interconnection as well as implementations' interoperability.

The V2X standards consist of a large number of specifications from diverse domains ranging from radio and protocols to security and applications and, as a result, multiple standardization and regulation bodies are involved. Going into details of each organization is beyond the scope of this work, however it is necessary to mention some of them to better understand the upcoming discussions.

- **3rd Generation Partnership Project** (3GPP) develops standards for mobile telecommunications systems based on cellular technologies. It is a landmark for telecommunication SDOs around the world, and provides them with reports and specifications on radio access, core transport network and service capabilities. 3GPP documentation is structured in releases, which consists of a set of features and specifications. Typically, a new 3GPP release is delivered every 1.5 years and describes solutions and features referring to a specific mobile communication generation i.e., 2G, 3G, 4G or 5G. Since Release 14 (2016), 3GPP has been actively performing normative work on the applicability of vehicular communications to mobile networks, known as Cellular V2X (C-V2X) technologies.
- **Institute of Electrical and Electronics Engineers** (IEEE) is an association for electronic engineering, electrical engineering, and associated disciplines. The IEEE Standards Association (IEEE-SA) body develops global standards in a broad range of industries, including telecommunication. Its main contribution to V2X communications is the IEEE 802.11p standard, a WiFi-based protocol that supports dedicated short-range communications (DSRC), a radio access technology that enables V2X communications.
- **European Telecommunications Standards Institute** (ETSI) is a standardization organization for ICT standards in Europe. The ETSI Technical Committee ITS (CT-ITS) aims to achieve global standards for C-ITS. Its regulations deal with a wide range of areas: communication architecture, security access layer-agnostic protocols, spectrum requirements, communication management etc. Also, CT-ITS develops conformance test specifications, which are essential for the commercial deployment of the technology.

- **5G Infrastructure Public Private Partnership** (5GPPP) is a joint initiative between the European Commission and the European ICT industry, which aims to deliver solutions, architectures, technologies, and standards for the ubiquitous next-generation communication infrastructures. 5GPP has shown how 5G can enable the next generation of connected and automated driving and related critical services, that cannot be implemented using today's communication technologies [8].
- **5G Automotive Association** (5GAA) links the mobile communications business with car manufacturers. Its direction is to harmonize the automotive and telecommunications industries to embrace and accelerate the global deployment of ITS solutions. Among all, there is a close cooperation with 3GPP. 5GAA operates on several parallel activities: system architectures, use cases and technical requirements, security, testbeds and performance evaluation, business models.
- **Society of Automotive Engineers** (SAE) International is a global association of engineers and related technical experts in the aerospace, automotive, and commercial-vehicle industries with tasks including voluntary consensus standards development Since April 2014, SAE has been developing DSRC-based V2X application standards in collaboration with ETSI and IEEE.

2.2. V2X Applications

3GPP started shaping V2X use cases and their requirements from Release 14 [3]. At that time, Long-Term Evolution (LTE) was the available mobile technology employed, and the very first V2X applications (e.g., collision warning, emergency stop warning, automated parking assistance, traffic route information support) addressed context with more relaxed requiremennts than more recent applications [32].

As V2X began to catch the eye of industries, many new and more demanding applications appeared. And, despite intensive work to make LTE supporting V2X applications, there were still too many limitations. Besides the stringent requirements in terms of latency, reliability and throughput, such a dense and highly mobile environment was the real issue for LTE.

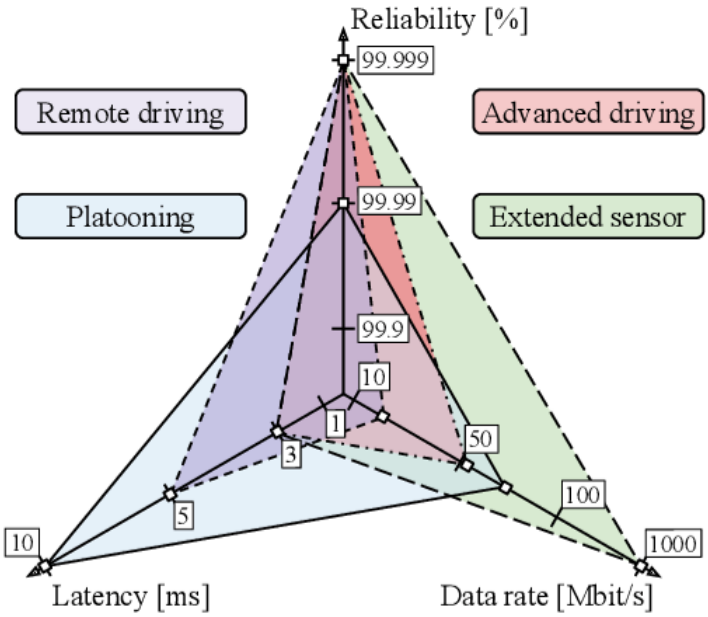


Figure 2.2: Requirements of advanced V2X use cases.

The next round of standardizations, Release 15, is 5G-based, and this technology is meant be a good candidate to support this type of services. 3GPP TR 22.186 [4] and TS 22.886 [6] are the two reference documents that introduce advanced V2X applications with related details and requirements. Four main application categories are described: vehicles platooning, remote driving, advanced driving, and extended sensors.

Vehicles platooning use cases support the formation and management of a platoon i.e., a group of vehicles travelling together. The vehicles of the platoon exchange periodically information, for example the leading vehicle of the platoon every x seconds sends information on the gap between two adjacent vehicles or about the speed to be maintained.

Remote driving applications enable the vehicle to be driven by a remote entity. This would make it possible for passengers who cannot drive themselves or use vehicles in dangerous areas e.g., on a construction site.

Extended sensors category is based on exchange of sensor data, either raw or processed, collected through local sensors. Its purpose is to increase the perception of the surroundings beyond the capabilities of the vehicle alone.

Advanced driving use cases lead to semi-autonomous and fully autonomous systems. Vehicles share data obtained from their local sensors with vehicles and entities in proximity. In addition, vehicles share their driving intention in order to coordinate their trajectories or manoeuvres.

The advanced driving group leads to the concept of level of automation, which is also related to the difference between semi-autonomous and fully autonomous systems. Indeed, any use cases come with requirements depending on the automation level to be supported [12].

Automation Levels

The Society of Automotive Engineers (SAE) has defined six automation levels [30], from zero to five, and the higher the level, the higher the level of driving automation.

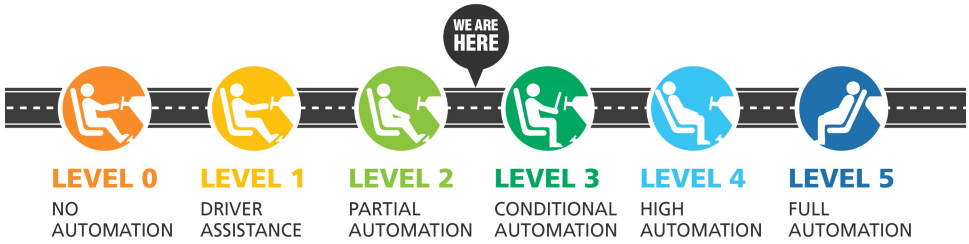


Figure 2.3: SAE Automation Levels

In **level 0** there is no automation, and human performs all driving tasks and it's up to him to monitor the environment. It is also known as *no automation* level. In **level 1**, or *driver assistance* level, the human driver can enjoy vehicle assistance features e.g., adaptive cruise control. Level 0 and level 1 do not require connectivity, any automated function relies on information from on-board sensors and data processing is performed in the vehicle itself.

Partial automation level - **level 2** - the vehicle is able to autonomously steer and accelerate/decelerate depending on the circumstances, while the human is in charge of performing all remaining tasks and must be able to take control at all times. In **level 3**, the vehicle can carry out most of the driving tasks and has environmental detection capabilities, from its side, the human shall monitor the environment and be ready to take control if the vehicle cannot handle the situation. This level is also referred to as *conditional automation* level.

In **level 4**, or *high automation* level, the driving system can intervene in all kind of dynamic driving tasks still accordingly to the current environment, whether or not drivers respond appropriately to a request to intervene. However, in certain conditions, the vehicle may not be able to operate in automated mode. This is the case of severe weather

conditions and snowy or slippery roads. In that case, the driver has to take control. In **level 5** the vehicle performs all the driving by itself under all roadway and environmental conditions. This level is also known as *full automation* level, and no human interaction is required.

Achieving level 3 is the first step to switch the approach from manual to automated. The responsibility would leave the human side, and move closer to an autonomous system. To support such scenario, connectivity becomes a mandatory component for providing the required functionality.
Today, the most promising countries are Europe, China and the US, where it is provided level 2 and is enabled through ADAS based on on-board sensors like cameras, lidar and radar. Moving upwards in the level involves numerous legislative, regulatory, and technical challenges, and the impact of 5G will likely be significant [20]. Likely, higher automation levels will first be available on highways and later in urban, suburban, and rural areas.

2.2.1. V2X Communication Types

According to the previous description of the V2X applications and considering their strict requirements shown in figure 2.2, the services that can be developed, and their requirement are widely diverse. Meeting the requirements of any use case is challenging and suitable communication technologies shall be employed. Due to this diversity, V2X communications is categorised into four communication groups that differ in the type of entity involved in the communication.

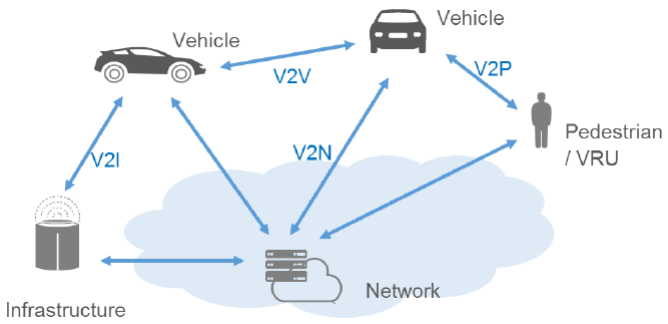


Figure 2.4: Vehicular Communication Types

- **Vehicle-to-Vehicle** (V2V). Vehicles in proximity exchange information one to each in a direct mode i.e., without passing through a network node. Communication is

mainly broadcast-based, and the information exchanged is about location, dynamics (e.g. speed and acceleration), and vehicle attributes.

- **Vehicle-to-Infrastructure** (V2I). Communication involves a vehicle and a so-called Road Side Unit (RSU). RSUs are communication units supporting V2X communication and they can be incorporated in either a Base Station (BS) node or a stand-alone road infrastructure placed in the street.
- **Vehicle-to-Pedestrian** (V2P). Communication occurs between a vehicle and a Vulnerable Road User (VRU) which can be a pedestrian or a cyclist. Unlike V2V applications, the device of the non-vehicular entity has a lower battery capacity, therefore, in V2P scenarios, messages from VRUs are less frequent.
- **Vehicle-to-Network** (V2N). Vehicles communicate with BSs and remote servers a communication network. For example, it is useful to send alerts to vehicles regarding an accident happened a few kilometres ahead. In this way, the communication range is extended, as road conditions are spread in larger areas. V2N can be also employed to collect data on road occupancy and status (e.g., alert for slippery road surfaces, presence of dangerous objects, etc..) or to deliver multimedia content to the vehicle.

All these communication services are commonly referred to as Vehicle-to-Anything (V2X).

2.3. V2X Communication Technologies

Today, solutions for the V2X communications are either WiFi-based or rely on cellular networks. More precisely, the former uses the IEEE 802.11p technology combined to other national protocols like IEEE 1609.x standards in the U.S. and GeoNetworking in Europe. The latter – also called Cellular V2X (C-V2X) - was supported by LTE first and then by 5G.

Both the technologies matured in the past few years and have evolved, and each comes with advantages and limitations [33]. WiFi ones, for instance, do not call for pre-installed infrastructure but cannot apply in any coverage situations: in tunnels or harsh area where there is no coverage, communication cannot be established. Moreover, new emerging applications in vehicular communications demand for higher requirements in terms of data rate, reliability, and latency.

According to the research, upcoming cellular networks, like 5G and beyond, are expected to accommodate these features. Moreover, cellular-based solutions would provide connec-

tivity even in out-of-coverage situations, enable long-range communications and support multi-cast communications.

2.3.1. IEEE 802.11p

WiFi-based V2X communications have gone through a longer period of studies and investigations, as a result they reached a more mature stage than cellular-based solutions. Now, the two leading technologies are Dedicated Short-Range Communications (DSRC), developed in the U.S., and C-ITS, developed in Europe. All over the globe, many solutions are adopted, each with its own characteristics and operating band, DSRC and C-ITS are the major ones.

Dedicated Short-Range Communications

DSRC was introduced in 2004 and the two most involved SDOs are IEEE and SAE. This solution relies on a protocol stack that is the result of the combination of IEEE 802.11 and IEEE 1609 standards, commonly known as Wireless Access for Vehicular Environment (WAVE). More precisely, the 802.11p amendment is adopted, which is specifically designed to support V2X communications.

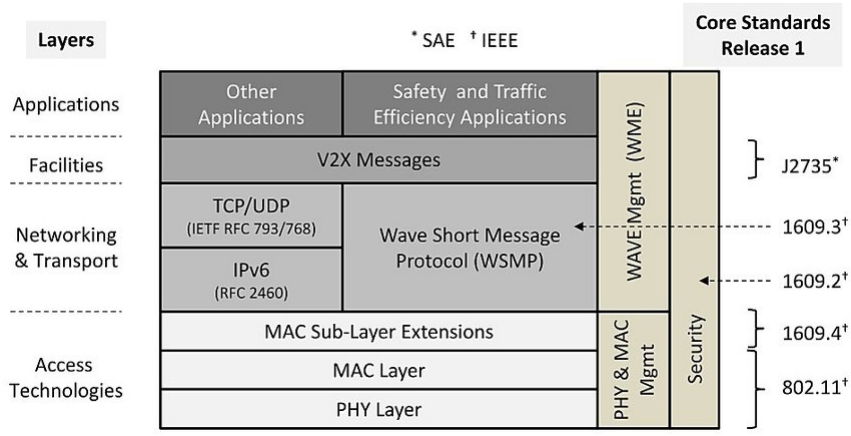


Figure 2.5: DSRC protocol stack and related core standards [21]

The most relevant innovation is to allow for communication outside the context of a basic service set (BSS). A BSS is a group of stations that all belong to network and employ specific standard terminology. Devices need to be a member of a BSS in order to exchange messages, and joining a BSS implies to go through control procedures such as authentica-

tion and association. Allowing the outside the context of a BSS mode (OCB), means that devices in communication range are able to exchange data immediately, without prior exchange of control information to join the network. As a result, use of timing resources is optimized.

Besides, IEEE shapes the physical (PHY) and the medium access control (MAC) layers of the IEEE 802.11a protocol with suitably modified features [24]. The transmission scheme stays the Orthogonal Frequency Division Multiplexing (OFDM), but 10 MHz channel spacing is used instead of 20 MHz. This helps in coping with longer delay spreads that can occur in outdoor environments. At the MAC layer, instead, stations use an enhanced distributed channel access (EDCA) that applies the Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) which effectively allows for data traffic prioritization [21].

Then, moving up in the stack, to accommodate V2X applications demanding for direct communication, IEEE 1609 standards are used instead of the traditional ones like TCP, UDP and IP. The Wave Short Message Protocol (WSMP) is the core of this specification block. It is a single hop network protocol with minimum header of few bytes, and it also provides the multiplexing of messages to upper layer protocol entities based on service IDs. Other standards then define security and authentication procedures (IEEE 1609.2), multi-channel operation management (IEEE 1609.4) and so on so forth.

On top of that, SAE standards define syntax and semantics of V2X messages. More details on the messages exchanged in V2X communication will be given in section 2.4, here some basic concepts are only introduced [1]. The most relevant message in DSRC is the Basic Safety Message (BSM); it conveys core state information about the sending vehicles, including position, dynamics, status, and size. It is sent periodically at a rate of 10 Hz maximum i.e., ten messages per second, and has to contain some mandatory information.

ITS-G5

ITS-G5 is the equivalent DSRC version applied in Europe. This second WiFi-based technology comes with many similarities to the American one, with some differences though.

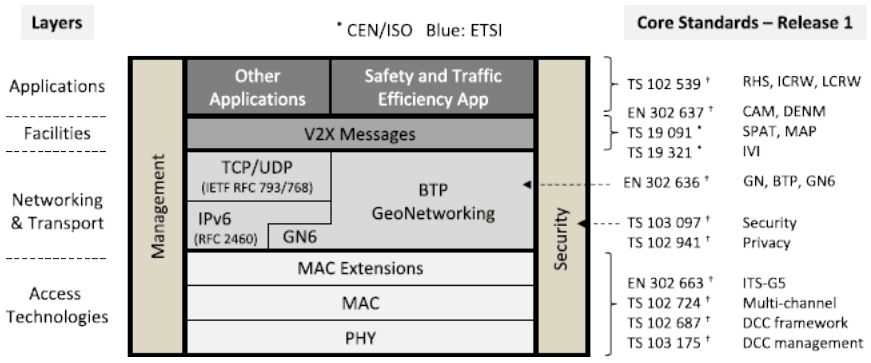


Figure 2.6: ITS-G5 protocol stack and related core standards [21].

Both the PHY and MAC layer are the same as DSRC, but instead of using WAVE standards, ITS-G5 specifies an ad hoc routing protocol for multi-hop communications, termed GeoNetworking which uses geographical coordinates for addressing and forwarding[13].

At the facilities layer, ETSI is still in charge of pinpointing V2X messages. Foremost, the *Cooperative Awareness Message* (CAM) periodically collects critical vehicle state information in support of safety and traffic efficiency applications. It can be considered as an equivalent to the BSM in the DSRC protocol stack. Then, the *Distributed Environmental Notification Message* (DENM) disseminates safety information in a geographical region, and it is triggered by an event; its role is to alert surrounding vehicles and network users and entities of something that has happened. The *Collective Perception Message* (CPM) allows for sharing of information about detected objects between V2X-enabled vehicles, thus surrounding vehicles become aware of non-visible obstacles to them, for example.

2.3.2. Cellular V2X

Considering the ambitious upcoming V2X applications, it is an open question to what extend the WiFi-based systems mentioned above are capable to meet such requirements. Those technologies are considered appropriate solutions for the basic applications of 3GPP Rel-14, however, network performances need to be empowered, and here the potential of cellular V2X comes into play.

3GPP started shaping C-V2X in 2014 to extend the support for V2X connectivity in mobile networks, thus first works were based on the LTE standard specifications. At that time, LTE has been introduced primarily to build up higher data rate systems, as a result its usage for V2X communication had several limitations [33]. Out-of-coverage circumstances were unsatisfactory, transmission scheduling was inefficient and signalling procedures led to many overheads.

To overcome these issues, the 3GPP introduced in later releases a feature known as Proximity Service (ProSe). ProSe allows devices in communication range to exchange data directly without passing through a BS. These services define sidelink communication, in contrast to the conventional up- and downlink ones, and data are transmitted over a subset of the uplink resources [10]. The most important benefit is that out-of-coverage communication is fulfilled as devices are allowed to select resources autonomously without involving orchestration of a base station. Nonetheless, ProSe is unspecialised for vehicle speeds and needs to be enhanced in terms of functionality and performance to be applied to the latest V2X scenarios [22].

The thing is that LTE technology was not specifically designed for V2X applications, and despite the remarkable technological developments, those standards hardly met the highly demanding latency and reliability requirements. With these motivations, the next generation of cellular communication systems, 5G, was developed to meet these demands as well, and indeed the newest 3GPP specifications show much progress towards the integration of V2X communications into cellular networks [32].

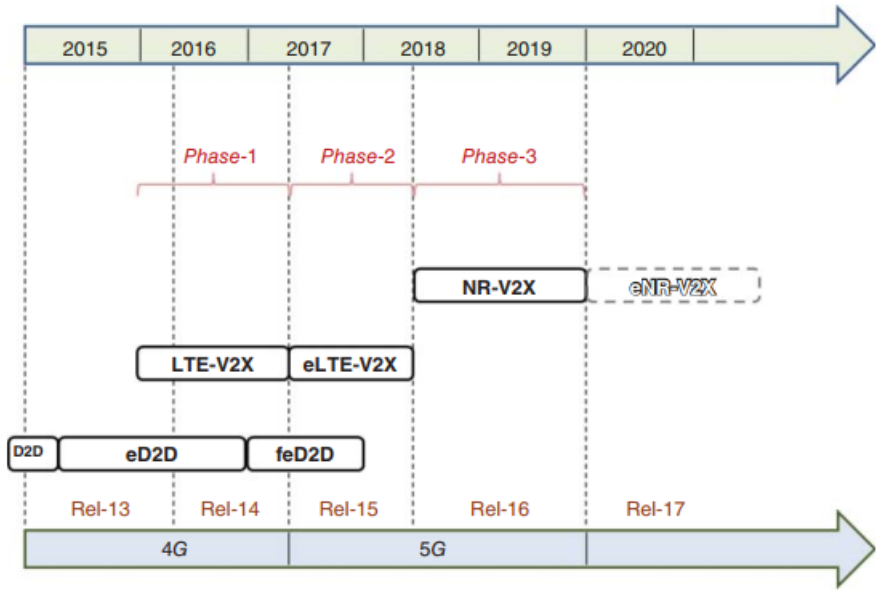


Figure 2.7: C-V2X standardization evolution.

The very game changer in the evolution of cellular V2X technologies is New Radio (NR) V2X, a 5G-based solution that will leverage many benefits of 5G in terms of performance, efficiency and security, retaining backward compatibility [33]. LTE-based solutions supported basic safety applications, while now NR V2X promises to serve also the advanced one mentioned in 2.2. In particular, it is meant to fit good highly dynamic environments,

and dense traffic situations, to increase service availability and scale efficiently to the diversity of requirements.
Some of the new features in NR V2X: more efficient establishment and management of direct communications (e.g., widely used in V2V, V2P and V2I contexts), enhancement of groupcast and unicast communications (important for platooning applications), support for transmitting smaller packets, employment of beamforming antennas [20].

By far, one of the most relevant research topic in NR V2X is the improvement in direct communication mode. Its advancement brings numerous benefit and overcomes many limitations.

Device-to-Device Communications

Device-to-Device (D2D) communication enable users that are in proximity to communicate in a direct mode i.e., their traffic does not go through network entities, such as Base Stations (BS). That means transmission can take place without involving network infrastructure.

Benefits are that communication is possible also in out-of-coverage environments, energy consumption is reduced, and transmissions are more efficient in terms of spectral usage, data rate and latencies. Moreover, BSs workload would be relaxed in terms of amount of traffic to process, that is a significant advantage considering the ever-growing number of vehicles, devices and infrastructure involved in a vehicular communication scenario. All these features are compliant with V2X application requirements, thus it is a suitable and promising technology for V2X communications.

3GPP introduced D2D communications for the first time in Release 12 to support the so-called Proximity Services (ProSe), and devices were enabled to discovery and communicate with each other directly. D2D communications are possible by establishing a direct radio link between user equipments (UEs) which takes the name sidelink (SL). Sidelink communications complement uplink and downlink transmission modes and occur over a subset of the uplink time-frequency resources.

3GPP has defined two resource allocation modes in the context of a direct communication: scheduled and autonomous. In the scheduled mode, the BS manages the allocation of resources to users, alternatively users select the resources from a pre-configured re-

source pool. This second option, in particular, makes the transmission possible regardless of network coverage.

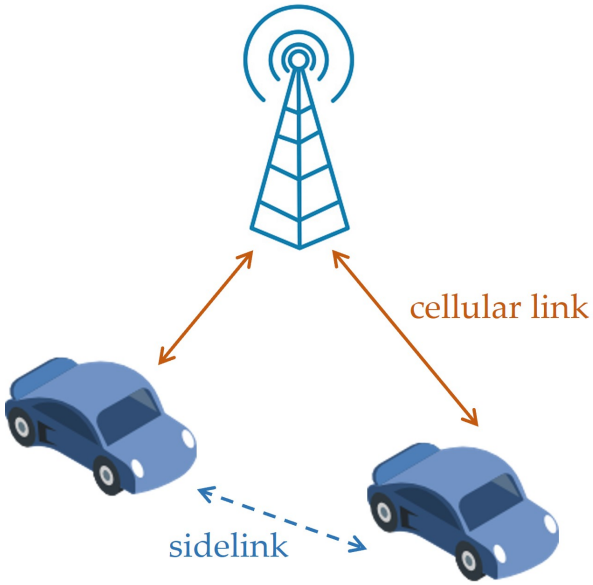


Figure 2.8: Device-to-Device communication.

In its earliest stages, ProSe addressed to public safety applications, while it was not specifically targeted at the vehicular world, whose applications and requirements were beyond imagination [9]. As a matter of fact, this technology has evolved over the years to meet the operating conditions of V2X.

Release 16 was the first specification to include sidelink in the 5G framework with a focus on V2X, enhancing reliability, latency, capacity, and flexibility [2]. Its development now addresses the advanced V2X use cases (sect 2.2), unicast and multicast transmissions are now supported, and also new resource allocation modes and algorithm are defined for the sidelink [31]. For instance, enabling groupcast communication facilitates platooning applications, where data need to be addressed only to the vehicles involved in the platoon.

2.4. V2X Messages

As already mentioned in the previous sections, to support V2X applications, vehicles send and receive messages. These messages shall be normalized, again to implement a robust,

coherent and reliable communication service.
Depending on the country, there are different standards regarding V2X messages. The focus here is on European standards, where the association in charge of defining these messages is ETSI, and, so far, three messages have been defined: Cooperative Awareness Message (CAM) [16], Decentralized Environmental Notification Message (DENM) [15] and Collective Perception Message (CPM) [17].

Cooperative Awareness Message

CAM provides a basic awareness service that periodically transmits information about positions, movement, basic attributes, and basic sensor information related to the originating ITS station. These messages are disseminated to neighbouring ITS stations that are located within a single hop distance from the transmitter. By receiving CAM messages, the surrounding ITS stations are aware of other stations in range as well as their positions, movement and relevant characteristics.

The general structure of a CAM consists of a header and a body. The header includes information about the message type, the ID of the originating device and the generation time. On the other hand, the body is a collection of information, some of which shall always be sent, while others are optional. Bearing this in mind, the frequency of message generation is variable and ranges from 1 Hz to 10 Hz. The generation frequency is affected also by the type of transmitting ITS station, and by specific phenomena.
For example: a new CAM is triggered if the absolute difference between the current and the last CAM speed is more than 1 m/s, or also if the current and the last CAM position differ by more than 5 m. Still, an emergency vehicle may generate CAMs more frequently.

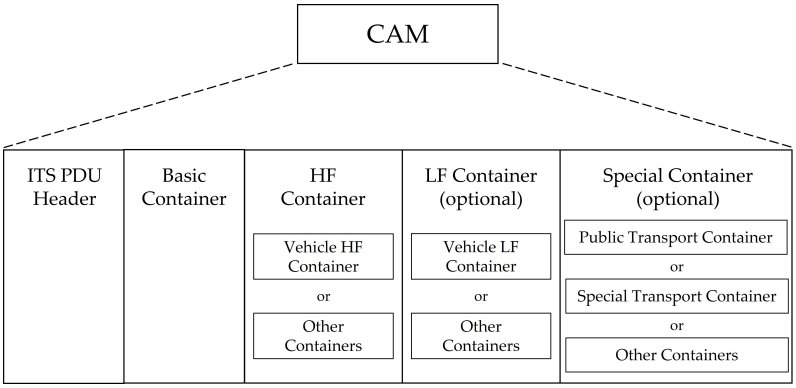


Figure 2.9: CAM structure [16]

The body of a CAM counts a number of so-called containers that are distinguished according to the type of information they carry. The type of originating ITS station (cars, trucks, RSUs etc...) also determines a different CAM content. ETSI standard mentions basic container, high frequency containers, low frequency containers and special containers.

The **basic container** provides the type of originating ITS and the latest geographic position of the transmitter at the CAM generation instant. **High frequency containers** collect highly dynamic i.e., fast-changing, information of the originating ITS like heading or speed. **Low frequency containers** include static and not highly dynamic information of the originating ITS station like the status of the exterior lights of the vehicle. **Special containers** originate from vehicles which carry out specific roles such as public transport or ambulances or vehicles transporting dangerous goods.

Decentralized Environmental Notification Message

DENM constitutes another type of application support facility that provides a notification service about detected events. Differently to CAMs, these messages are not always on, an event shall happen to trigger them, and they are transmitted in a multi-hop way to cover a concrete geographic dissemination area.

An event is characterized by an event type e.g., traffic jam, accident, its position, a detection time, a destination area indicating the geographical area over which the DENM is meant to be propagated, and a transmission frequency. Conceptually, when a new event is detected, a new DENM is generated that reaches all ITS stations in range. In turns, these stations retransmit the message to their stations in range and so on until the dissemination area constraint is satisfied.

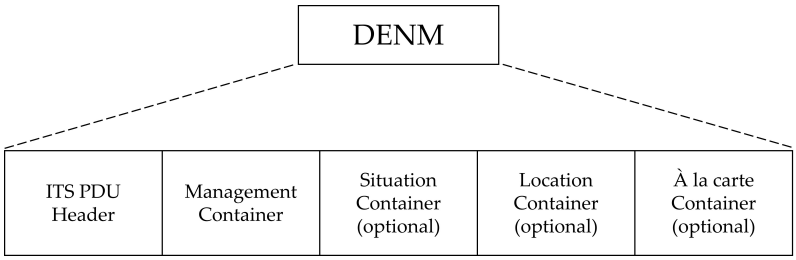


Figure 2.10: DENM structure [15].

The DENM structure consists of a header, like the one used in CAMs, and a body containing event information. The body is composed of four containers: management container, situation container, location container and à la carte container. Similarly to CAMs, not all the containers are intended to be included in a DENM.
The **management container** includes information about the event and its management, such as its version, expiration time, sending frequency etc. The **situation container** provides specific details about the event informed in the message According to some standards, it is specified the class of the event and its severity. The **location container** gathers the event location data i.e., the coordinates of the danger. The **à la carte container** includes information specific to the use case which requires the transmission of additional information that is not included in the three other containers.

Collective Perception Message

CPM is the most recent message ETSI has defined. Its goal is to support the so-called collective perception (CP) service i.e., sharing on-board sensors information (detected items or people) among nearby vehicles. Collective perception provides a view of the surrounding of other vehicles, extending the individual perception of a vehicle, eliminating blinding spots, and improving the quality and reliability of individual measurements. The identified objects are either static i.e., do not move and are located on the driving lanes, or dynamic i.e., move or have the ability to move.

Structurally, CPMs are coherent to its siblings, CAMs and DENMs. It consists of the usual header, and a body comprising a number of containers, some of them mandatory while others not: management container, station data container, sensor information containers, perceived objects container and free space addendum containers.

The **management container** is the only mandatory one and contains basic information about the transmitting station (vehicle or RSU) like its type and position. The **station data container** includes additional information about the transmitter, such as its speed, heading, and acceleration. The **sensor information container** provides details about the on-board sensors in the sending station. The **perceived objects containers** include information about the detected objects like their distance to the vehicle or their speed. A CPM adds a perceived object container per each detected objects, up to 128 objects in one message. The **free space addendum container** describes the free space areas within the sensor detection areas i.e., different areas of the detection area of a sensor are associated to a confidence level.

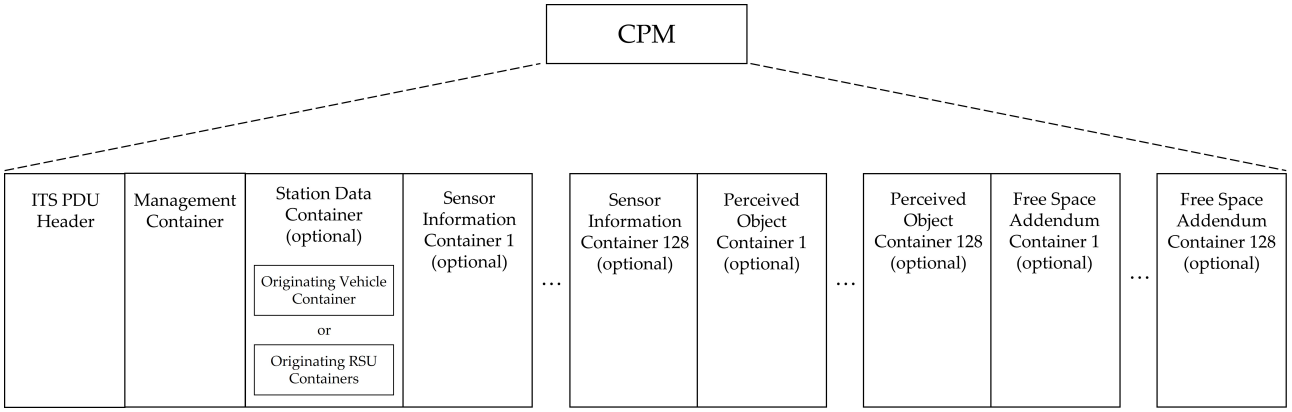


Figure 2.11: CPM structure [17].

About periodicity, the collective perception service follows the cooperative awareness approach. Messages are periodically sent at a frequency ranging from 1 Hz to 10 Hz, and it depends on the dynamic state of the detected object, similarly to what is done for CAMs. Highly dynamic objects are more often included in transmitted CPMs than slow or static ones.

On top of that, ETSI points out that CPMs are to be sent for relevant objects only that are detected with sufficient confidence level. The specifications also mention the importance of coordinating CPMs transmission with CAMs transmission, optimize the channel utilization is a priority. To this purpose, generation rules have been defined to rule how often a vehicle shall generate a CPM and the information it should include [35].

2.5. Misbehaviour Detection

It is now presented the second part of the background chapter which is dedicated to covering topics about misbehaviour detection.

In any communication system, securing transmissions against malicious attackers is a fundamental requirement for secure and reliable operation of applications [36]. V2X communications are not an exception and are exposed to this issue as well. Verifying the integrity and correctness of the information received is commonly known as *misbehaviour detection*, and it is another story with respect to security and privacy.

Security is about authentication and authorization processes to provide permissions in the form of digital certificates to any entity willing to participate in a network. Only once the identity of devices has been validated, are such permissions issued and the device can transmit packets within the network. IEEE [25] and ETSI [18] [19] extensively produced regulatory material targeting this matter which define a signature- and certificate-based approach.

On the other hand, privacy addresses protection of personal data, including the person's name, phone number, driving style, and vehicle data like the license plate number. In this direction, V2X protocols and messages support the adoption of pseudonyms so that identifiers, addresses and certificates cannot be linked to a person or vehicle. Furthermore, standards require V2X devices to regularly change their pseudonyms as a defense measure against traceability attacks [14].

Misbehavior detection is considered an essential second layer of security for networks, and their impact, in particular in vehicular contexts, might be very damaging as accidents and even loss of life can occur. For example, in a platoon an attacker might compromise a message indicating that one of the vehicles is breaking. The other vehicles would then react by braking in turn, probably causing a collision. Misbehaviour detection methods are supposed to prevent these undesirable situations and intended to enable V2X applications to safely operate.

According to the literature [36], misbehaviour is performed by nodes belonging to networks and not by external entities intruding into the communication. These nodes are called *misbehaving nodes*, and transmit erroneous data on purpose i.e., with malicious intents, while simultaneously making the network to behave as expected [28]. As a result, the reliability of the data and the trustworthiness of nodes are compromised, and the goal of misbehaviour detection is to determine whether a certain message or signal constitutes unexpected behaviour.

Security and Privacy in V2X systems

The current European standardisation for securing V2X communications adopts asymmetric cryptography using a Public Key Infrastructures (PKI) that manage security credentials. A PKI sends out digital certificates to On Board Units (OBUs) and Road Side Units (RSUs), and entities which have received such a certificate are then referred to as End Entities (EEs). Users' privacy is protected by a pseudonym scheme i.e., by frequently

changing the pseudonym used to authenticate V2X messages. This mechanism avoids vehicle tracking, or at least, makes it more difficult.

Every EE is assigned two certificates, and their format, within the European area, is ruled by the ETSI TS 103 097 standard [18].

- A **long-term certificate** that is received during the enrolment phase and identifies and accounts of a transmitting station. These certificates are named *Enrolment Certificates* (EC) and are used to sign requests sent to PKI authorities, for instance to get pseudonym certificates.
- **Multiple short-term certificates**, also referred to as *pseudonym certificates* or *Authorization Tickets* (AT), that are used to sign messages. For privacy purposes, EEs regularly change their identity, which means changing the pseudonym. Usually, twenty multiple short-term certificates are issued per week.

Along with long-term certificates, EEs are provided with a verification public key, that is used by the receiver to verify the signature of messages, and an encryption public key that is used to encrypt data solely intended for the owner of that certificate.

PKI authorities are organized according to a hierarchical structure. *Root Certification Authorities* (RCA) act as trust anchors[1] and control several subordinate *Certification Authorities* (CA) and end-entities such as vehicles or RSUs. RCAs are managed by various stakeholders (car manufacturer, telecommunication providers, European/national governments etc.), thus cooperation and cross-certification between RCAs is made possible.

CAs are Trusted Third Parties (TTP), and their role is to sign and deliver digital certificates. Depending on the type of certificate issued, one can distinguish between Enrolment Authorities (EA) and Authorization Authorities (AA). The former approve the trust of stations and validate their identity, and are therefore responsible for releasing long-term certificates. The latter, instead, are in charge of providing the multiple ATs to be used in V2X communication.

[1] In cryptographic systems, a trust anchor is an authoritative entity for which trust is assumed and not derived.

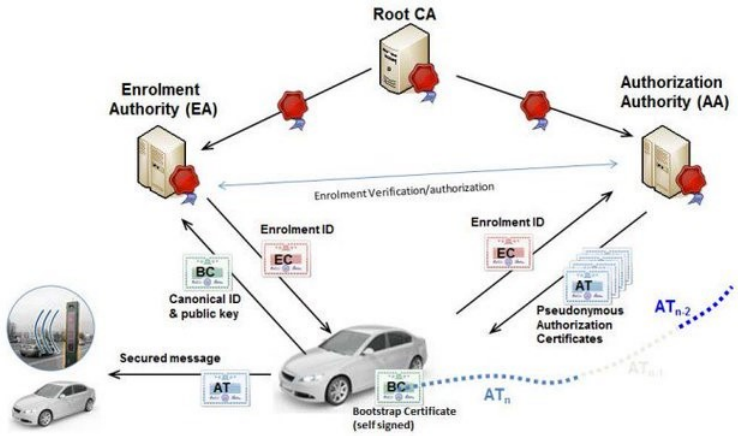


Figure 2.12: PKI structure in ETSI ITS trust model [27].

Obtaining a pseudonym certificate is a three-stage procedure. First, the EE needs to be registered with the EA. In case of OBUs, this step is expected to be carried out by the car manufacturer that provides the vehicle's canonical identifier and the associated public key. Then, it follows the enrolment certificate request to the EA, which is signed with the public key associated with the registered canonical identifier. Once these two steps have been executed, the EE can request its twenty weekly pseudonym certificates, and this request is signed using the enrolment certificate.

The security method mentioned above provides sender authenticity and message integrity, and can effectively protect against attackers that transmit messages using commodity hardware without key material [11]. However, messages correctness is not guaranteed; it is not implausible for an attacker to own a vehicle or to extract key material from old communication units to obtain the proper certificates and then to transmit malicious data. Or, also, attackers might be able to take over the software and create arbitrary fake messages. This is the purpose of misbehaviour detection: understanding whether and when any semantically correct information has been compromised and react accordingly. Hence, misbehaviour detection in V2X applications is a key driver for secure and safe operation.

2.6. Types of attack

To implement any misbehaviour detection methods, it is fundamental to understand how the information itself might be compromised. This means understanding what the type of information flowing in a V2X communication looks like. Moreover, it is necessary to investigate what the profiles of possible attackers might be. Concerning the messages,

this is something already explored in section 2.4, while now this section gives a glimpse on types of misbehaviour that can be performed i.e., what types of attacks might occur.

Possible attacks on the C-ITS are divided into two main groups: cyber and physical [36]. **Cyber attacks** aims for compromising computer systems, some examples are: denial of service (DoS) attacks, data injection attacks and Sybil attacks. On the other side, **physical attacks** aims for physical processes surrounding the vehicle itself. This second family group is more about vandalism, for instance attackers might blind cameras by pointing a laser or damage vehicle electronics (e.g., sensors and processors) by transmitting electromagnetic pulses [28]. The impact of cyber attacks is potentially more dangerous and widespread than physical attacks. Furthermore, physical attacks are unpredictable; there is no way of implementing a prevention system able to detect such attacks and protect the vehicle and the environment. For these reasons, the work focuses on the cyber category.

Cyber-based attacks are well-known issues in information systems. Their intent can be to steal data as much to destroy a system in order to disable it. The literature records several types of attacks and attackers' behaviour [23, 36].

An adversary performing a **DoS attack** sends tons of requests to the network causing an overload and a sudden stop in communication between network nodes. A quite known type of DoS attack is the *jamming attack* where the attacker disrupts the communication channel and can filter and limit incoming messages. Upon these attacks, communication performance dramatically drops: latencies increase and reliability of the network is reduced. Another common DoS attack is the *JellyFish attack* that exploits vulnerabilities in congestion control protocols and delays or even drops packets. In a V2X scenario such a situation is catastrophic: messages do not propagate and information sharing is undermined. *Flooding attacks*, instead, crowd the network with bursts of data packets making the network resources (e.g., bandwidth, power, etc.) unavailable to legitimate nodes.

Sybil attacks pretend a vehicle has multiple identities either at the same time or in succession. Consequences are not that far from the ones of a DoS attack: resources unavailability and network destabilization. In addition, misleading information is disseminated and false representations of reality are given to surroundings vehicles and RSUs. For instance, with a Sybil attack, a road would appear to be bottle necked, while in reality it is not, and so the attacker can enjoy the road for himself.

The real threat to V2X applications are **false data injection attacks**. Attackers send

messages whose real-world information is twisted e.g., false messages (like wrong positioning data) or incorrect information (about traffic conditions for instance) are broadcast to the network. Such attacks have widely different intentions: generating spurious braking, provoking irregular manoeuvres, disrupting road traffic etc [23]. As a result, all these actions would trigger inappropriate reactions that can be life-threatening for the driver himself as well as for surrounding vehicles, and could endanger passengers' lives.

2.7. Misbehaviour detection methods

A small taxonomy of existing misbehaviour detection techniques is now given to facilitate understanding of the method implemented for this work.

A first broad distinction in security matter is drawn between reactive and proactive approach. *Proactive* refers to mechanisms including a security policy such as digital signature, PKI, certificates, etc. Such solutions are good at defending the system against potential external attackers, but cannot prevent insider attackers from generating legitimate false information. On the other side, *reactive* solutions assume that malicious activity can be present within the system and consist of a detection and reaction step. Misbehaviour detection is positioned in this second category.

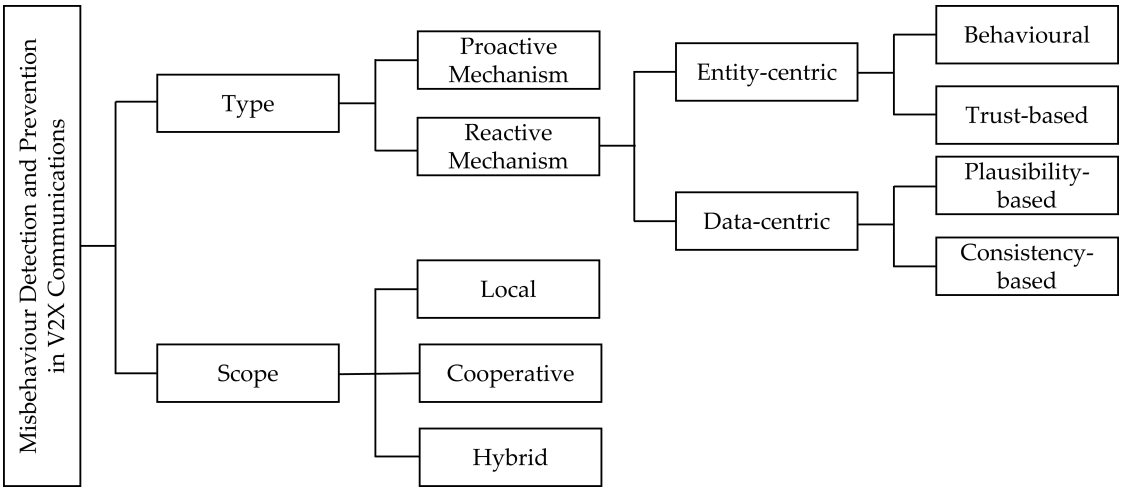


Figure 2.13: Taxonomy of V2X misbehavior detection/prevention approaches [23]

Reactive mechanisms can be then distinguished in node-centric and data-centric. *Node-centric* approaches investigate the behaviour of the nodes participating in the network. This category can be further subdivided into behavioural and trust-based. *Behavioural*

solutions observe patterns in the behaviour of nodes at the protocol level e.g., by checking packets frequency and messages format. *Trust-based* detection is based on reputation systems or voting scheme that assess the trust of nodes, assign trust score and elect legitimate nodes. In contrast to node-centric solutions, many nodes are involved and, often, infrastructural support may be included to support decision-making and remove malicious nodes.

Data-centric techniques, instead, use the content of the message to determine the validity of the received data, no matter who the sender is, and these approaches can be either plausibility-based or consistency-based. *Plausibility-based* mechanisms leverage an underlying model of data to verify if the transmitted information is consistent with the model. For example, to verify position information, two successive messages are considered and the distance travelled is compared with the speed. *Consistency-based* solutions compare the content of a new packets with packet from many nodes previously received, and determine the truthfulness of the new data. For instance, to check the correctness of the information speed of a packet, it is compared with the average speed of previously collected packets. Similarly to node-centric mechanisms, these two subcategories differ for the number of nodes involved. Plausibility-based schemes consider packets from individual senders, while consistency-based demand data from several participants.

In the end, depending on the scope, detection mechanisms can operate locally, in a cooperative way or mix these two modi operandi - referred to as hybrid schemes. *Local detection* checks consistency internally i.e., according to vehicle OBUs and optionally sensors without involving other vehicles. *Cooperative detection* implies collaboration between vehicles and RSUs if possible. *Hybrid* solutions partially perform detection in back-end systems. Behavioural and plausibility schemes generally operate locally, while consistency and trust-based rely on cooperation among vehicles/RSUs to detect inconsistencies.

2.8. Summary

This chapter exhaustively introduced the vehicular communications explaining how it works, the latest supporting technologies, application areas and security aspects and issues.

Vehicles exchange messages that contain useful information both for safety and traffic efficiency purposes, and also to improve the driving experience. There are standards governing this type of communication, and the format and syntax of these messages.

Establishing such standards is crucial to ensure homogeneity and compliance, so that communication can take place successfully and smoothly.

In the second part, misbehaviour concepts are addressed. Misbehaviour means that someone with malicious intent, an attacker, manipulates the data. This comes with many purposes, and may cause different unpleasant situation. Due to its variety and unpredictability, misbehaviour is considered one of the most challenging issues to handle in these applications; relying on reliable information is a cornerstone for such applications. Research is strenuously working on these topics: the literature is rich in techniques and approaches. However, misbehaviour detection methods are still matter of open discussion.

3 | Method

In this section, the reader finds a detailed explanation of the conducted work; all the employed tools, and the methodologies are given here. The project aims to develop and test a misbehaviour detection method; as a first step, an attack type and a simulation scenario had been defined. Setting up this simulation environment also served to build our data set i.e., to have a consistent collection of data to use in the testing phase. Next the method for detecting misbehaviour had been outlined, and, in the end, the testing part. This latter part involved evaluating the reliability of the implemented function using truthful data, and then assessing the ability of the method to discriminate malicious behaviour, and here compromised data were included.

3.1. Type of attack

As discussed in the background chapter, in section 2.6, the literature shows many possible type of attacks with several purposes. Some aim to disrupt communication between nodes, others to represent non-compliant traffic conditions with reality and still others to inject false data. Consequences of such attacks range from selfish purposes, such as benefiting from less jammed roads, to more malicious and dangerous aims, like causing car crashes.

False data injection are broadly the most harmful attack [36] since vehicular communications effectively need to deal with reliable and trustworthy information, otherwise autonomous decision making, for instance, cannot be supported. With these motivations, the simulated scenarios implement a false data injection attack where the malicious vehicle transmits a false position. Compromising the position is justified on the grounds that those data are one of the most available and frequent information transmitted as it shall always be included in every CAM message. Indeed, CAMs are essential drivers for cooperative driving applications, therefore it is meaningful to provide protection at this level.

Figure 3.4 shows an instance of a simulated case. More details on the implementation of

such scenarios are given in the following section, here some basic notions are just given. There are two vehicles travelling along the same direction, and one of them falsify its position. This vehicle is called *malicious vehicle*, while the honest part is referred to as *trustworthy vehicle*. In the figure vehicle 2 behaves maliciously, and instead of sending the green trajectory, that is the real one, sends the red one. Hence, these two vehicles have two different goals. The trustworthy vehicle has to detect whether CAMs received from the malicious vehicle are reliable or not. The malicious vehicles has to mislead the trustworthy vehicle by injection false data in its CAMs.

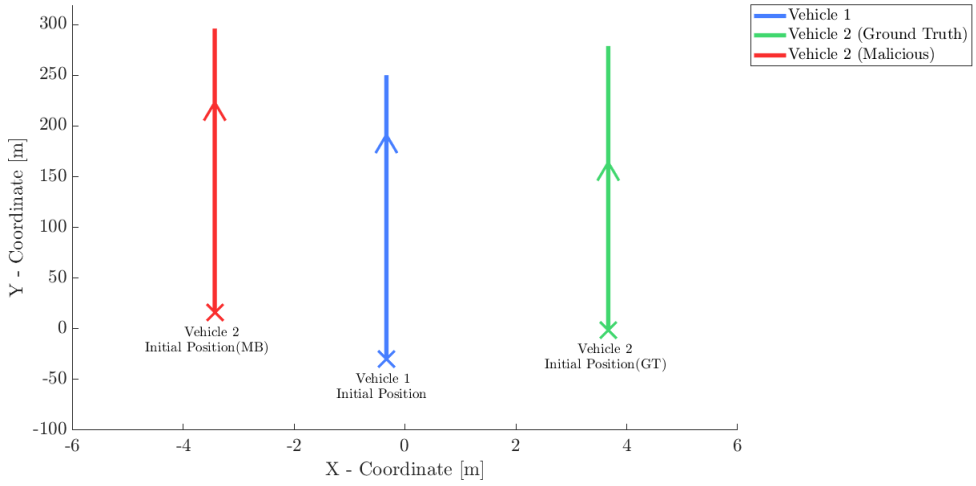


Figure 3.1: Vehicles trajectories

Three different ways of falsifying the position are considered, hence three different attack behaviours have been investigated. This choice is inline with approaches applied in the literature [34] which show how the type of attack behaviour can affect the misbehaviour detection itself. That is, depending on the detection method, some attacks may be detected more easily than others.

- **Constant position**. A fixed position is enclosed in any CAM. For the entire length of the simulation, the misbehaving vehicle sends a position that does not change over time.
- **Random offset**. A random offset is added to the real position. In particular, a maximum distance in the Y direction of up to 50 meters is allowed, while in the X direction the upper limit is 8 meters, that corresponds to two lanes of distance.
- **Random position**. The starting point of the trajectory is randomly drawn within the simulation area, and the false trajectory evolves like a real one but starting from

a different point.

At the beginning, a fourth type of attack was also evaluated in which a constant offset was added to the real positions. This attack behaviour led to inconsistent results and redundant observations, so it was discarded in the discussion.

Ultimately, in any implemented case, the false trajectories result being shifted in the simulation area. The intention was to limit significant differences with the real trajectories and to ensure that the falsified trajectory was always falling within the simulation area: a trajectory 200 meters away from the real one is fairly easy to detect. In other words, the generated attacks aim to produce competitive trajectories, also to limit cases to be discarded due to obvious incompatibilities.

3.2. Scenario

This second section discusses the simulation setup and the tools used. Most of the work has been carried out in Matlab, and external libraries have been included. The envisioned simulation environment is in accordance with the details of the 3GPP TR 37.885 standard [5]. This document specifies all the evaluation methodologies to be used in analysing 5G V2X use cases compliant with the ones outlined in the 3GPP TR 22.886 [4]. Road configurations, users type, and all the design considerations in terms of traffic model, channel model, deployment, etc. are described.

The 3GPP TR 37.885 document addresses two key scenarios: urban grid and highway, and considers both below and above 6 GHz frequency bands. Accordingly, any detail is described: BS and RSU deployment, road configuration, antenna models, mobility models, vehicles types and dimensions, traffic models, channel models. For instance, the standard specifies the antenna height of vehicles - which varies whether it is a truck or a car -, UEs transmitting power - which differs depending on the operating frequency band, the bandwidth dedicated to uplink and downlink transmissions, and so on.

Summarising, this standard is a reference and provides guidelines that both academia and industry have to stick to it so as to have a coherent methodology.

3.2.1. Scenario setup

Simulations for this work are about a highway context and V2V communications only is involved i.e., RSUs, devices of pedestrians and any network entity are not considered and

direct communication only applies.
The road configuration in figure 3.2 counts three 4-metre-wide lanes [5] in each direction where two vehicles are placed at a maximum distance of 100 meters. The two vehicles drive in the same direction and their speed takes a value in the range [50, 75 100, 125, 150] km/h. The carrier frequency is 5.9 GHz, the transmitting power applied is 23 dBm and the antenna gain, for both the nodes, is equal to 6 dBd. Both the vehicles are considered able to generate CAM messages at a frequency equal to 10 Hz (10 messages per second).

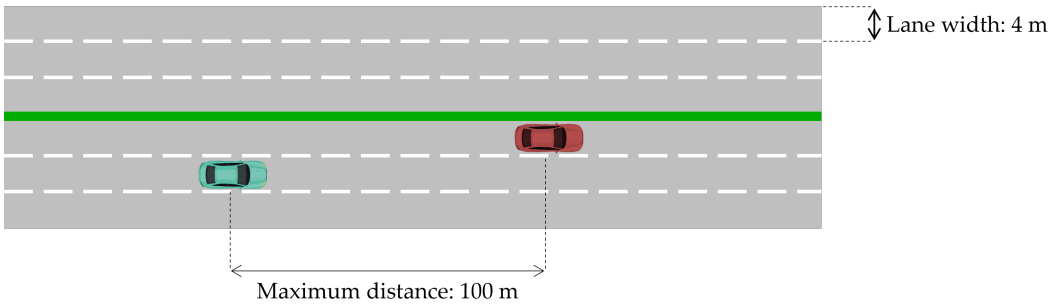


Figure 3.2: Road configuration compliant to 3GPP TR 37.885 [4]

In the 3GPP report [5], three possible transmission conditions are discussed: line of sight (LOS), line of sight blocked by buildings (NLOS) and line of sight blocked by vehicles (NLOSv). Since the developed scenario involves two vehicles only and the maximum distance is relatively small, a LOS condition is assumed. This assumption simplifies the work, but the choice is still valid since the path loss model defined in case of LOS and NLOSv does not change [4]. In highways, considering a LOS path blocked by buildings is rather unlikely.

The path loss model formula from 3GPP TR 37.885 is shown in equation (3.1):

$$PL = 32.4 + 20log_{10}(d_{3d}) + 20log_{10}(f_c) \tag{3.1}$$

d_{3d} denotes the Euclidean distance between transmitter and receiver in 3D space in meters[1].
f_c denotes the center frequency in GHz.

Matlab was used to develop the scenarios, and the approach goes through four steps:

1. **Trajectories generation**. Vehicles trajectories are generated by randomly posi-

[1]The two vehicles height is assumed to be the same

tioning the two vehicles within the simulation area. Once, the initial positions are defined, tracks are recorded over a 10-second time span.

2. **Position falsification**. Taking the two generated trajectories, one of them is falsified depending on the attack type in analysis.

3. **Channel simulation**. Considering all the parameters above mentioned and the data from the trajectory generation step[2], the channel between the two vehicles is simulated in order to record received power values.

Trajectories Generation

Firstly, the two vehicles are placed in the simulation area that involves three 200-metre highway lanes: the first vehicle is randomly positioned, then the second one is placed such that it is within the highway scene and at a maximum distance of 100 meters. A speed value is chosen for both vehicles that is to be kept constant throughout the simulation and, accordingly, the trajectories of vehicles along the route are recorded. That is, the position of the two vehicles is tracked every time a CAM message is sent. By setting CAM frequency at 100 ms and simulation time at one minute, 600 messages per simulation are collected.

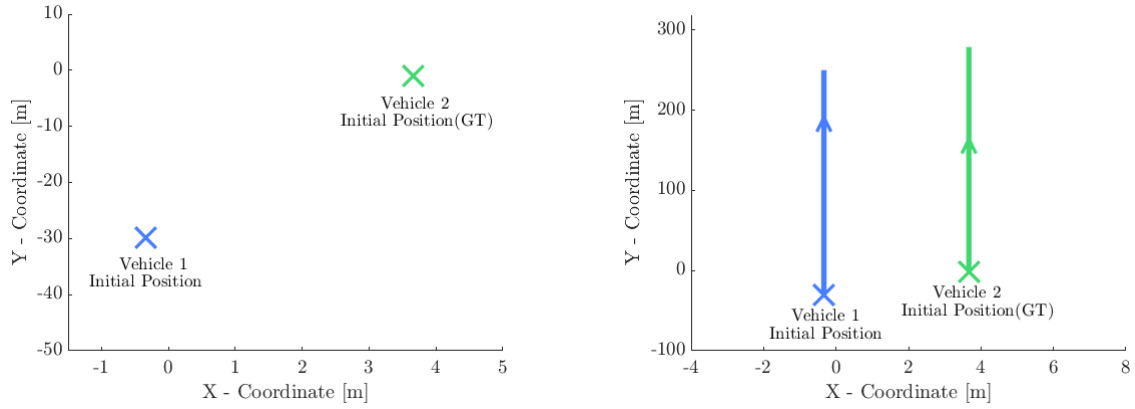
Figure 3.3 provides a representation of this first step. On the left 4.1a, the two crosses represent the initial position of the vehicles, those are going to be the starting point for the two trajectories. The frame displays a case where the two vehicles are one lane apart, as their distance in the X-direction is four meters. On the right, the route of the vehicles during the simulation is depicted. Both cars travel in the same direction, describing a linear trajectory, and the total length of trajectories depends on the speed value.

Position falsification

Once true paths are settled, one trajectory is then falsified i.e., the malicious vehicle replaces its actual location with a different one and, in this way, the trustworthy vehicle receives data on a position that is not compliant with the reality.

The way the position is distorted depends on the type of attack, and as discussed in the section about the attack types (Section 3.1), the position of the malicious vehicle is modified so that it is still within the simulation area.

[2]To simulate the channel, the true trajectories are employed



(a) Initial position of the vehicles

(b) Trajectories evolution

Figure 3.3: Actual trajectories generation.

For the constant position attack, the position of the malicious vehicle is at the center of the simulation area. For the random position case, the distance is randomly assigned such that the maximum distance to real positions is maximum 50 meters. For the random offset case, maximum 50 meters in the Y direction and 8 meters - 2 lanes, in the X direction are added. These choices allow to narrow down unrealistic situations in which the received position is clearly misleading.

Figure 3.4 depicts the final configuration. *Vehicle 1* sends reliable messages, while *Vehicle 2* falsifies its position. The real trajectory that vehicle 2 follows is the green one, and the red one is the sent trajectory. As a result, vehicle 1, relying on the received CAMs, gets that vehicle 2 is travelling on the red line instead of the green line.

Channel simulation

The data obtained at the trajectory generation step are then used to simulate the channel between the two vehicles. This second part still has been developed in Matlab and the simulation environment relies on the **QuaDRiGa** libraries that are used for modeling realistic radio channel for system-level simulations of mobile radio networks compatible with 3GPP specifications [29]. This tool is well appreciated for the simulation of vehicular scenarios as it solidly suits to their fast-changing and mobile characteristics.

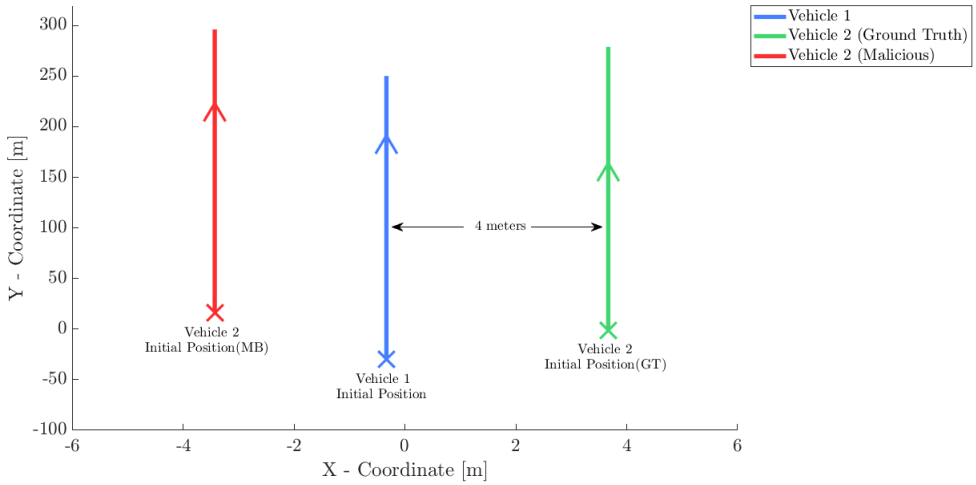


Figure 3.4: Vehicles trajectories

QuaDRiGa supports a fully-fledged three dimensional geometry-based stochastic channel model and contains a collection of features inherited by previously existing channel models, WINNER [26] and SCM [7]. These two are channel models simulator introduced by 3GPP. Modeling novelties introduced by QuaDRiGa provide features to enable quasi-deterministic multi-link tracking of users movements in changing environments.

QuaDRiGa follows a geometry-based stochastic channel modeling which creates arbitrary double directional radio channels. Channel parameters are determined stochastically, based on statistical distributions extracted from channel measurements. Its approach consists of taking a trajectory evolving over a period of time and splitting this trajectory into a number of segments; for each channel segment the channel parameters are calculated from the distributions.

QuaDRiGa is implemented in Matlab using an object oriented framework, and its installation does not require any changes to the system settings. This tool consists of a number of model files that can be added to the Matlab work environment and the user interface is built upon classes which can be customized by the user. Users provide trajectories and configure the network layout inline with the propagation environment to be simulated. In particular, the user:

- Sets the transmitter position
- Defines antenna properties for the transmitter and the receiver
- Defines the user trajectory

- Defines states (or segments) along the user trajectory
- Assigns a propagation environment to each state

The required information derives from the previous part, and trajectories are divided into segments according to CAM frequency i.e., the segment length coincide with the space travelled between the receiving time of one message and the next one.

Then, according to the scenario instantiated by the user, QuaDRiGa provides the corresponding channel characteristics for each segment such as delay spread, K factor, shadow fading. For example, the delay spread for a given configuration is defined as log-normal distributed with a range from 40 to 400 ns. Each segment of the trajectory is assigned a value within this range; e.g., 307 ns for segment 1, 233 ns for segment 2, 152 ns for segment 3 and so on.

Once these parameters are all set, for each segment of the trajectory, the antenna amplitudes and phases are calculated and returned to the user.

The scenarios supported by QuaDRiGa include standardised models, including the one outlined in the 3GPP TR 37.885 standard [5] - the reference standard used in this work. At the end of this part, in the output, users get the received power by mobile terminals involved in the simulation.

3.3. Misbehaviour detection method

This section gets to the heart of the work, and provides an exhaustive discussion of the implemented misbehaviour detection method. The deployed scheme is inline with a signal-based plausibility approach and exploits the received signal strength indication (RSSI) to compare it with the position transmitted in CAM messages.

Plausibility methods involve comparing experimental data with an underlying reference model and assigning it a numeric value, often called **plausibility value**. The analysed data are supposed to be as closely as possible to this model, and the plausibility value strictly depends on this compliance: the more similar the data to the model, the higher the value.

The rationale for implementing the reference model is the received power. Indeed, assuming the propagation conditions are known, given the received signal power, it is pos-

sible to derive the distance between the transmitting and receiving node. Leveraging the signal strength to get an accurate positioning of terminals generally leads to imprecise estimations, hence localisation is not the primary goal. Rather, the purpose is to keep track of the received power over time, recreate a trajectory and then compare it with the positions received in CAMs.

To put this explanation into context, consider this example: a vehicle continuously receives CAM messages from another vehicle and these messages provide a position that places the transmitter 40 meters away. The receiver derives the distance from the received signal strength, and obtains that the transmitter should be 10 meters away. This distance is inconsistent with the data in the CAM, and thus the transmitting vehicle is compromising the communication.

First, considering the received power, the experimental path loss is extracted:

$$PL_{exp} = P_t + 2G - P_r \tag{3.2}$$

P_t denotes the transmitted power in dBm.
P_r denotes the received power in dBm.
G denotes the antenna gain [3].
The transmitted power and the antenna gain are known, and respectively equal to 23 dBm and 6 dBd. The received power is a data that is extracted from the signal itself [4].

Then, the path loss model used to perform the simulations is applied and the term regarding the distance between transmitter and receiver is computed. The reference path loss model is the one from the 3GPP TR 37.885 standard that has already been introduced at the beginning of this chapter in section 3.2, the formula is here shown for the sake of simplicity.

$$PL_{th} = 32.4 + 20log_{10}(d_{3d}) + 20log_{10}(f_c) \tag{3.3}$$

Experimental distance terms result from:

$$d_{exp} = PL_{exp} - 32.4 - 20log_{10}(f_c) \tag{3.4}$$

[3]Antenna gains of the receiver and transmitter are the same
[4]The received power is the output of the channel simulations

These values are the ones compared with the theoretical ones i.e., with the received distance values.
Experimental distance values are then grouped into bunches of twenty items, normally distributed and mean value and standard deviation are stored. This approach implies that the algorithm needs to collect twenty messages to perform, since CAM frequency is fixed to 100 ms, it takes two seconds to catch on.

Now, these experimental distributions are leveraged to assess the likelihood of the theoretical distances and to provide a plausibility value for the received data. To this end, an heatmap is constructed around the receiver position based on the received power values. Hypothetical trajectories are evaluated, compared with the experimental distribution calculated in the previous step, and a likelihood value is assigned to the trajectory under analysis. Once the entire space has been evaluated, the likelihood values are divided by the maximum likelihood. This is the final value assigned to the hypothetical trajectory, and it represents its plausibility value. Since likelihoods are divided by the global maximum, plausibility values are expressed in terms of **relative likelihood**.

In other words: the receiver node, for various distance values, computes theoretical path loss values and distances, and compare them with the received data. The further these values from the theoretical model, the lower the likelihood. Then, once all hypothetical trajectories have been considered, it is assigned a plausibility value corresponding to the relative likelihood of values with respect to the most plausible trajectory. Essentially, the purpose is to evaluate n trajectories, identify the most plausible one based on the received power and assign each trajectory to a plausibility value.
Figure 3.5 helps in visualising this part.

The plausibility values result from these formulas:

$$p_i = \frac{q_i}{\max\limits_{j=1,..,n} q_j} \qquad \text{i=1,2,...,\ n} \tag{3.5}$$

$$q_i = \text{normpdf}(d_i, \mu, \sigma) \tag{3.6}$$

d_i denotes the distance between the i-th hypothetical trajectory and the receiver.
μ is the experimental mean value of the model distribution.
σ is the experimental standard deviation of the model distribution.

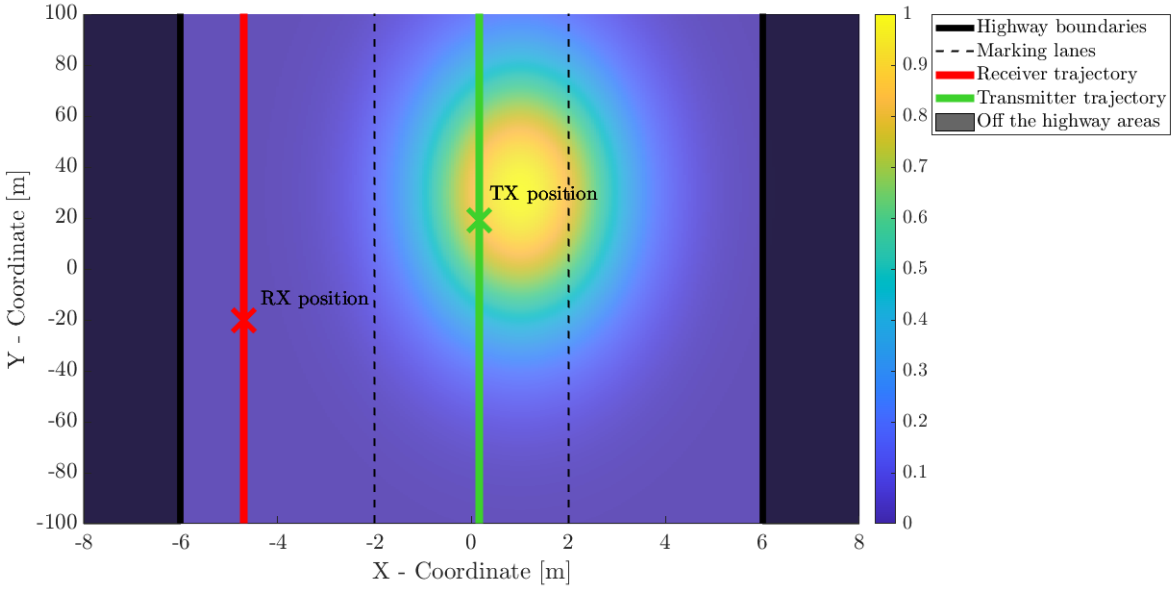


Figure 3.5: Heatmap around the receiver node for the transmitter node position

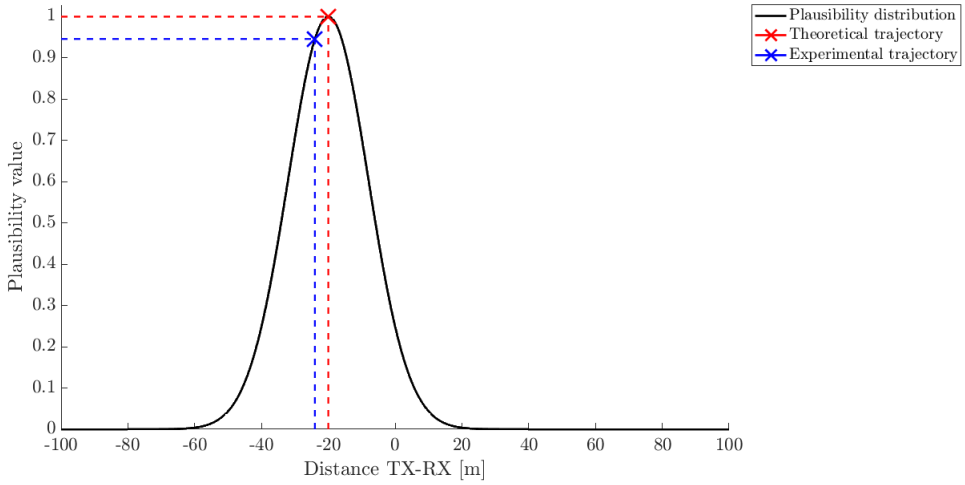
q_i denotes the likelihood of the i-th hypothetical trajectory path given current μ and σ.
q_i denotes the final plausibility value to i-th trajectory.
normpdf is a function that takes as input the distance d_i, the mean value and the standard deviation, and gives as output the probability q_i.

Equation 3.6 expresses the likelihood of a trajectory placed at a distance d_i, while equation 3.5 computes the relative likelihood i.e., the plausibility value.

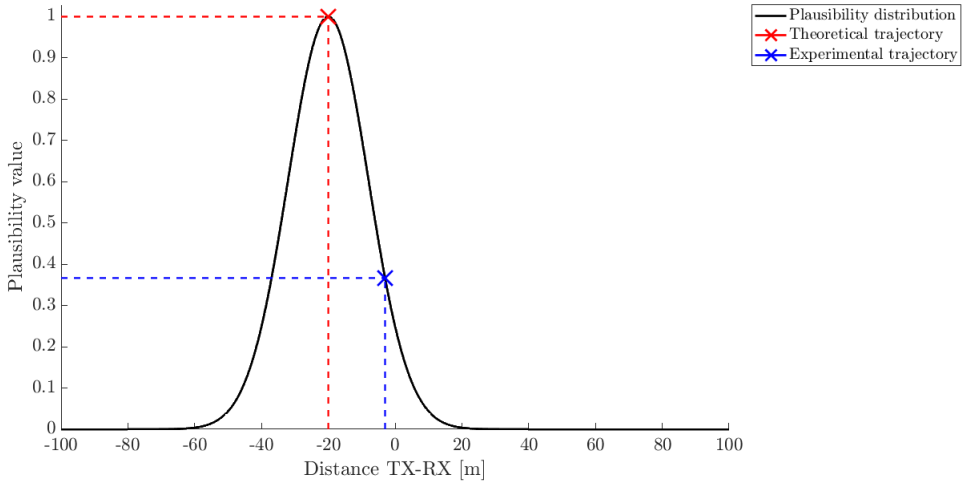
As time goes by, the heatmap and the distribution parameters are constantly updated according to the reference model deduced from the theoretical data.

Besides assigning a plausibility value to the received trajectory, the distance value corresponding to the maximum likelihood value i.e., the value corresponding to a plausibility value equal to one, is also given in the output. This information is useful to compare the received experimental data to the theoretical values one expects i.e., to evaluate the reliability of the received data. Here the idea is: depending on the power values received, the most plausible trajectory is found and compared to the received one. If the two are close in distance, the transmitting vehicle is honest, otherwise it is malicious.
Providing the trajectory corresponding to the maximum plausibility value is also useful to evaluate in a second moment the performances of the method itself.

Figure 3.6 provides an example by comparing an honest case with a malicious one. According to the plot, the most plausible trajectory, the one in red, is positioned 20 meters away approximately. Figure 4.6a shows the received trajectory, the one in blue, is assigned to a plausibility value between 0.9 and 1, and the distance difference with the red trajectory is approximately 5 meters. How to evaluate such numbers is explained in the next chapter that is about results, for the time being this case is assumed reliable.
Figure 4.6c, on the contrary, assigns to the experimental trajectory a lower plausibility value, lower than 0.4, and its distance to the ground truth trajectory is almost 20 meters. This case is considered malicious.



(a) Trustworthy scenario



(b) Malicious scenario

Figure 3.6: Plausibility distribution and comparison between the experimental trajectory and the most plausible trajectory.

3.4. Evaluation methodology

This last part of the chapter is devoted to describing how the performance of the developed misbehaviour detection method is evaluated.
The analysis advances in two parallel directions, carrying out two different tests, respectively referred to as *reliability test* and *misbehaviour detection test*. The first one aims to quantify the effectiveness of the method, i.e. to see whether it is something actually reasonable. The second effectively aims to assess the primary purpose of this work, namely the capability of the method to spot compromised data by leveraging the received data and the transmitted signals properties.

3.4.1. Reliability test

The goal of the assessment is to evaluate whether the methods actually works in case of non-compromised data only i.e., analysis are performed assuming no misbehaviour is taking place. Intuitively, high plausibility values are expected as the received data should be compliant with the theoretical expectations.

As already stated when the method was introduced in Section 3.3, using the receiver power level to derive distances is not really reliable in terms of accuracy due to the unpredictability of the propagation conditions; considerable differences between predicted and experimental distance values are inevitable. However, it is emphasised that accurate localisation is out of the scopes of this work.

Plausibility values are here compared to distances between the two vehicles and speed. The purpose is to investigate how plausibility changes with inter-vehicular distance, whether greater distances between receiver and transmitter impact on plausibility, if differences between experimental and theoretical trajectories are affected by the distance between the two vehicles and so on. Then, different speed values are considered, and it is observed whether greater speed impact on the results.

3.4.2. Misbehaviour detection test

The second analysis involves falsified data, and the goal is to figure it out whether the method is capable of identifying misbehavior.

Similarly to the reliability tests, plausibility is related to speed, and differently, distances between true trajectories and falsified ones is taken into account. Besides, analysis ex-

tend to different type of attacks. This allows investigating whether the method is more vulnerable against some specific attacks i.e., whether some types of attack are easier to detect than others.

3.5. Summary

This chapter discussed the methods used to develop the work. The simulation environment entirely relies on 3GPP specifications described in the technical report TR 37.885 and the employed tool is Matlab, with the addition of ad hoc external libraries to support the channel simulation part.

Scenarios are designed according to this 3GPP document and type of attacks are outlined taking inspiration from the most recent researches in the literature. In particular, it is assumed that malicious entities falsify their position and, trustworthy nodes aim to detect such misbehaviour by leveraging the received power of signals.

The receiver performs a method to analyse the received information and assigns a so-called plausibility value that is a measure of the reliability of the received information.

4 | Results and Discussion

In this chapter the results of the tests and their discussion are given. The analysis consists of two distinct parts: in the *reliability test* only real data are involved while in the *misbehaviour detection test* compromised data are included.

In the reliability test case, the purpose is to evaluate the effectiveness and the making of the implemented method in terms of reliability. No compromised data is included, thus all the received trajectories are theoretically plausible. On the other hand, misbehaviour detection method tests handle falsified data, consequently here the primary goal of the method – i.e., the capability of identifying misbehaviour, is tested.

Data involved to produce the results are the plausibility values that the misbehaviour detection method gives in the output, the distance between transmitter and receiver, their speed and the type of attack. Both the true distance, referred to as *ground truth distance*, and the false one, *compromised distance*, are included.

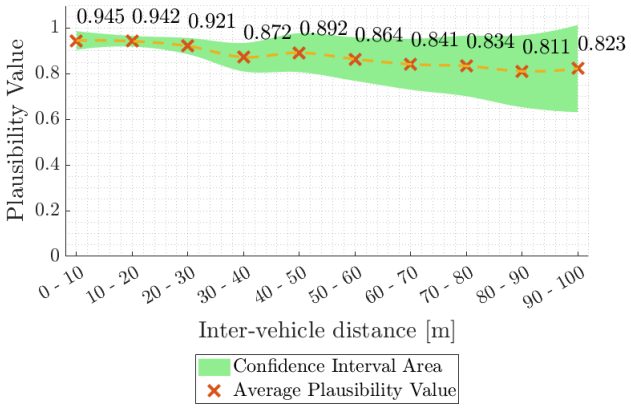
Broadly, both tests compare the plausibility values with the speed of the vehicles, then for the reliability case the plausibility is related to the distance between the two vehicles. Differently, misbehaviour detection tests compare the plausibility with the distance difference between the ground truth trajectories and the falsified ones. Besides, different type of attacks are involved too.

4.1. Reliability test

Reliability tests rely on real data. The main objective is to evaluate the impact of inter-vehicle distances and different speed values on plausibility values. The term inter-vehicle distance refer to the distance between transmitter and receiver. First, the relation between inter-vehicle distances and plausibility is investigated and the velocity is held constant. In a second stage, different speed values are discussed.

4.1.1. Plausibility VS Inter-vehicle distance

The results here presented refer to a speed value equal to 100 km/h and, according to the simulation setup, the maximum distance between the two vehicles is equal to 100 meters. The simulations are then collected in groups, referred to as *distance groups*, based on the inter-vehicle distances. For each distance group the average plausibility value is computed along with its confidence interval area.



(a) Plausibility VS Inter-vehicle distance plot

Distance [m]	Plausibility
0 - 10	0.945
10 - 20	0.921
20 - 30	0.921
30 - 40	0.872
40 - 50	0.892
50 - 60	0.864
60 - 70	0.841
70 - 80	0.834
80 - 90	0.811
90 - 100	0.793

(b) Average plausibility values per distance group

Figure 4.1: Plausibility VS Inter-vehicle distance.

Figure 4.1 shows the average plausibility trend for increasing inter-vehicle distance values. More distant vehicles decrease the plausibility value and increase its uncertainty. The plausibility value achieved is competitive and the downward trend makes sense: greater distances worsen propagation conditions and consequently the distance estimate is less accurate and more fluctuating.

Figure 4.2 adds interesting observations by comparing the plausibility value with the absolute difference of the most plausible trajectory returned by the function and the experimental one. Remember that the most plausible trajectory is one of the outputs of the misbehaviour detection scheme.

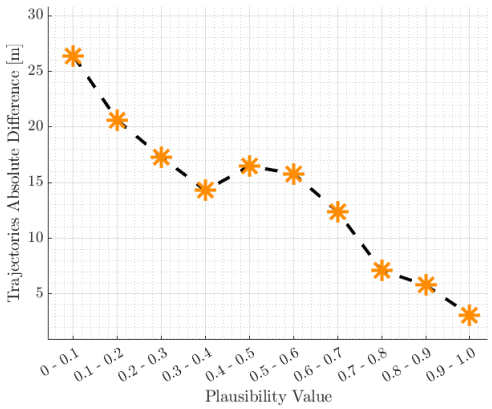


Figure 4.2: Plausibility VS Trajectories absolute difference.

The two plots in figure 4.1 and 4.2 suggest that the method is more reliable when the two vehicles are closer together, and higher plausibility values result in smaller errors in terms of absolute distance difference between the real and the estimated trajectory. Furthermore, plausibility values higher than 80% achieve a difference equal to 7.1 meters.

By combining the information from these two plots discussed, it is possible to tune a threshold for the plausibility value. That is, cases achieving values above the threshold are considered reliable. This choice is a matter of trade-off between the acceptable difference in distance one accepts and the plausibility percentage.

4.1.2. Plausibility VS Speed

Here different speed values are considered, and the impact on the plausibility is evaluated. The speed values taken into account are: 50, 75, 100, 125 and 150 km/h.

The trend of the plausibility is still observed (Figure 4.3) with different velocities: closer vehicles obtain higher plausibility values. In the figure, the confidence intervals are not reported as they follow the pattern analysed in figure 4.1a. Besides, it can be noticed that higher speed values generally achieve higher plausibility values.

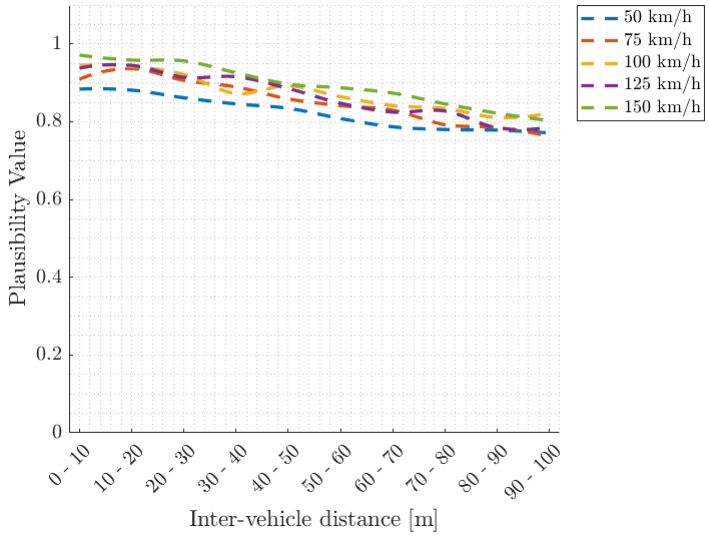


Figure 4.3: Plausibility VS Inter-vehicle distance at different speed values.

A possible explanation now follows. Driving at higher speed and maintaining the same CAM frequency, means that between one CAM and the next, cars travel longer distances and this implies that plausibility distributions are characterised by higher standard deviation values. Nonetheless, the difference between the experimental result and the ground truth increases at higher velocities. Figure 4.4 supports this explanation. Only data referring to speed values equal to 50 km/h, 100 km/h and 150 km/h are reported to make it more visible.

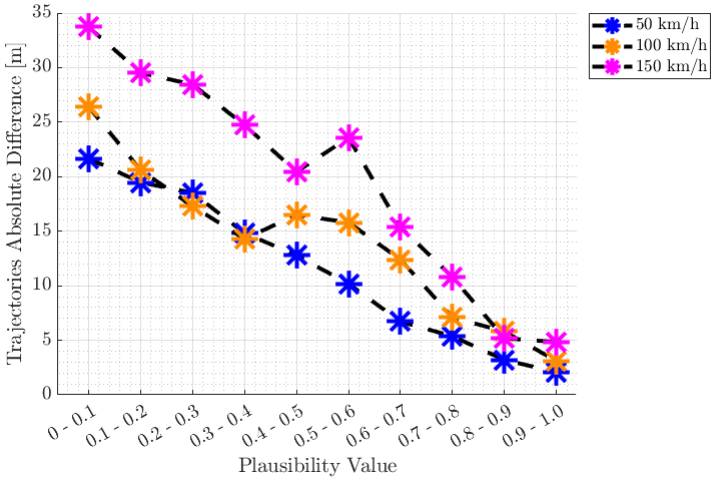


Figure 4.4: Plausibility VS Distance difference evaluated at different speed values.

This analysis on the speed extend the conclusions drawn in previous section where plau-

sibility was related to inter-vehicle distance only (Section 4.1.1). The applied speed value impacts on the threshold one chose to evaluate the effectiveness of the scheme, and according to the results shown above, higher velocities should force higher plausibility threshold values.

4.2. Misbehaviour detection test

This second part of the results discusses the ability of the implemented misbehavior detection method to actually individuate misleading data. Here, the type of attack is also included in any analysis performed. Therefore, considerations are made on cases of the same type of attack with different velocity values, as well as on cases with the same velocity but different attack type.

For convenience, the three attack types implemented are reported:

- **Constant position**: the true position is changed to one that remains unchanged for the duration of the simulation.
- **Random offset**: the true position is subject to the addition of a random offset.
- **Random position**: the true position is replaced with another random position.

First, data are selected depending on their plausibility values. Then, per each attack type, plausibility values are compared to the difference between the true and falsified trajectories. In the end, type of attacks are compared to the speed value.

4.2.1. Data over threshold

Data skimming is ruled by a threshold that is set to 50%, and values below these threshold are rejected for their small plausibility value. Each type of attack is combined with each speed value to evaluate the percentage of data which are over this threshold i.e., data which can be considered as reliable.

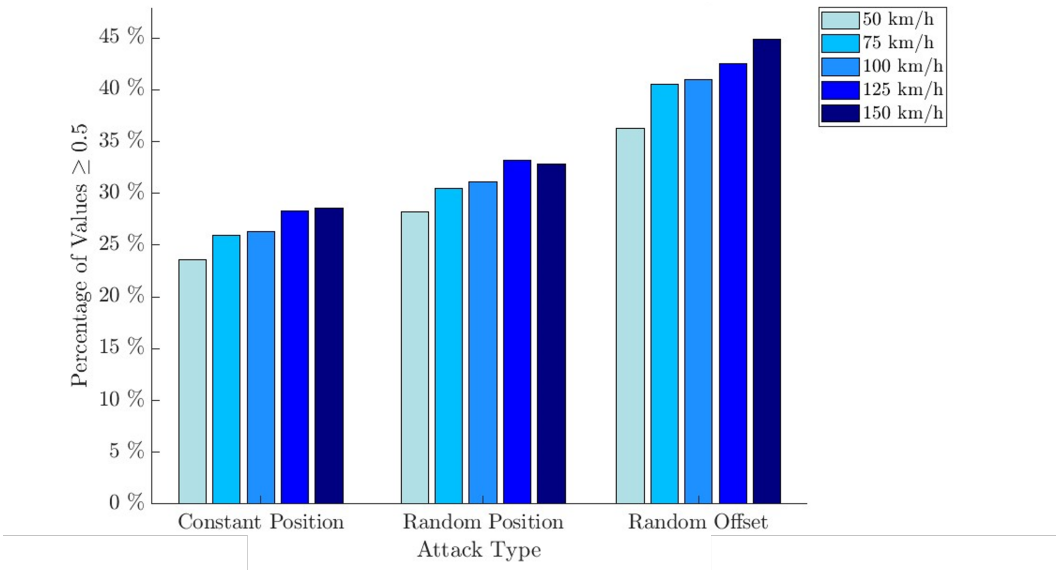


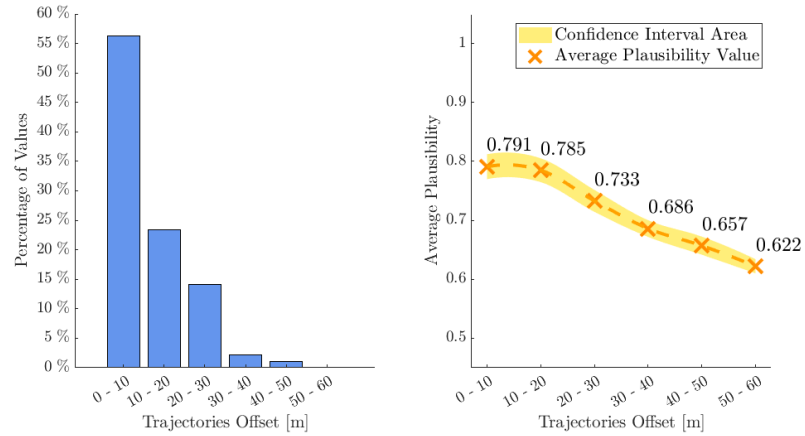
Figure 4.5: Percentage of data over the plausibility threshold per each type of attack and speed value.

Referring to the graphs in figure 4.5, regardless of the type of attack, the percentage of over threshold values increases with speed. This is inline with the considerations stated in the previous section. Indeed, it was shown that for greater velocities the average plausibility is higher, hence it makes sense that in this case the higher the velocities, the more the values over the plausibility threshold.

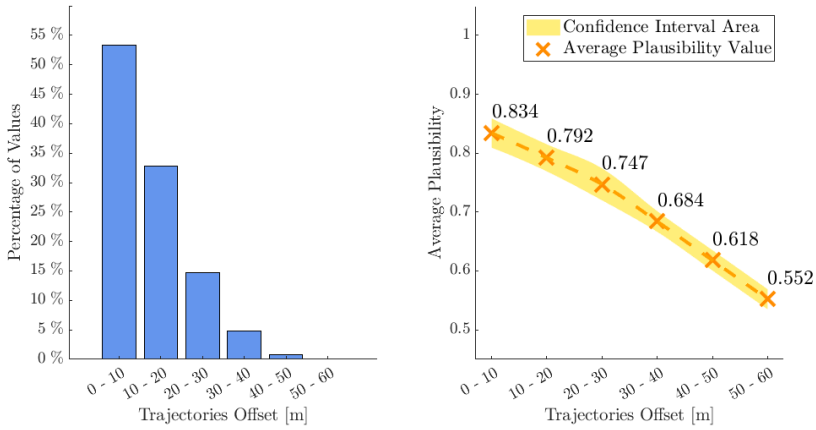
Then, considering a fixed value for the speed and comparing the three types of attack, the random offset case is the one achieving the greatest number of plausible values. This means that random offset attacks are more difficult to detect with respect to the others, and this is a direct consequence of the nature of the attack itself. Constant position attacks are expected to reproduce trajectories that are far removed from the real ones, so they are incompatible with the estimation of the misbehaviour detection function, i.e. they are associated with a low plausibility value and are rejected. Similarly, the random position quite easily reproduces positions that are far removed from the real ones. The random offset turns out to be the most deviant.

4.2.2. Plausibility VS Trajectories difference

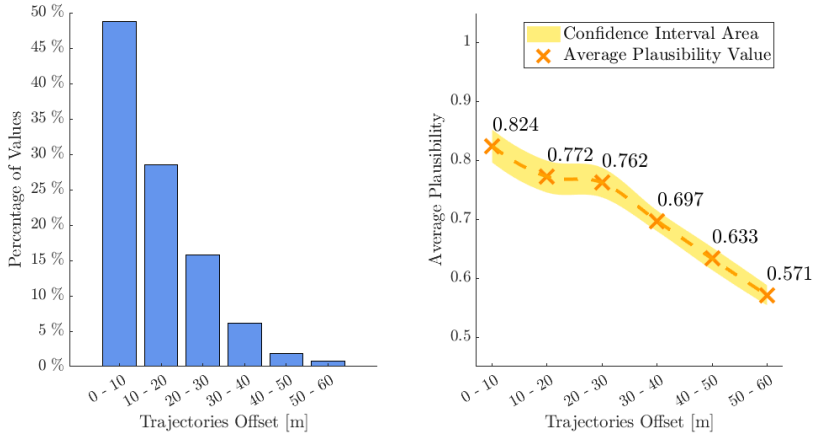
This part is dedicated to look into the relation between the plausibility value assigned to trajectories and the difference between true and falsified trajectories. This analysis is extended to different type of attacks and different velocities as well.



(a) Constant position



(b) Random position



(c) Random offset

Figure 4.6: Plausibility values VS Difference between the true and falsified trajectory at 100 km/h. On the left, data points are grouped depending on the difference of trajectories (trajectories offset). On the right, the value of average plausibility is related to trajectories offset.

A fixed speed value of 100 km/h is considered and the behaviour in the three types of attacks is compared. Data are grouped according to the difference between the true trajectory and the falsified one, that is referred to as *trajectories offset*, and the percentage of data per trajectory offset is given. This first relation is shown in the left-hand side of Figure 4.6. Most of values relate to cases where the true and falsified trajectory are closer together, then the greater the offset the smaller the amount of data. There is no significant difference between the three types of attack with regard to the distribution of data in these distance groups i.e., this descending trend where more than 50% of data belongs to cases with trajectories offset in the range [0-10] meters is observed.

Then, again considering data grouped as described above, the average plausibility is analysed. Graphs on the right hand-side of Figure 4.6 show that the smaller the difference between the true and falsified trajectory, the higher the average plausibility value. This plausibility trend is consistent with any type of attack. Moreover, as the trajectory offset increases, the fluctuations of data around the mean value decrease.

The graphs shown in Figure 4.6 suggest that in any attack type the great majority of data is about cases where the falsified trajectory is quite close to the real one, and, furthermore, these data are associated to higher plausibility values. In other words: the closer the compromised trajectory, the more difficult it is to detect misbehaviour. This result is confirmed also considering the variations of data which show higher uncertainties for smaller trajectory offsets. When the falsified trajectory is considerably distant from the true one, this is associated with lower plausibility and minor variance.
This result is compliant to expectations: if malicious positions are comparable to reality one receives, it is more difficult to detect misbehaviour.

Some considerations on differences between the three types of attacks now follows. The graphs feature different numerical values, but all have a common trend: majority of data and higher average plausibility values when the falsified trajectory is relatively similar to the real one, and minor variance around the mean value for consistent trajectory offset values. However, the constant position attack presents lower values for the average plausibility value, emphasizing the idea that it is easier to detect than the other two attacks. Indeed, showing lower plausibility values implies that the received data are more suspicious. Moreover, since the case of random offset attack, since plausibility values are slightly higher than random position attack, the idea that it is the most misleading attack is supported.

4.2.3. Plausibility VS Velocity

In conclusion, average plausibility values are related to different speed values and results involve the type of attack too. Here average plausibility values are expressed as weighted mean values, where weights are given by the percentage of data per distance group.

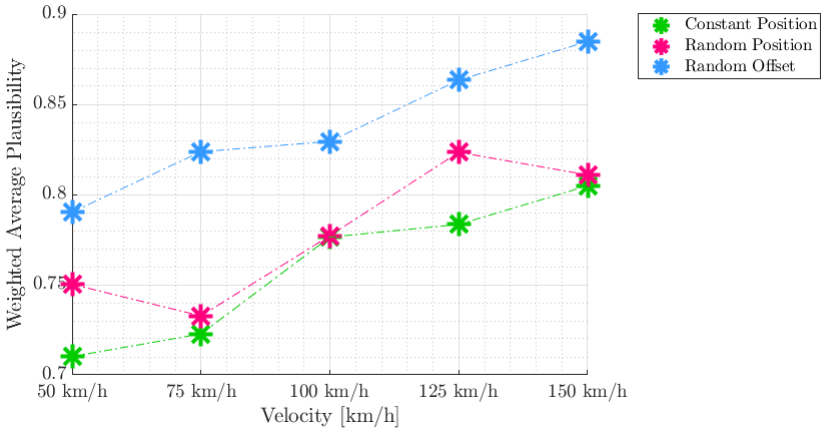


Figure 4.7: Weighted average plausibility VS Velocity.

The graph in Figure 4.7 strengthens some of the considerations made up to now. First of all, as the speed value increases, the plausibility assumes higher values, and this supports what written in Section 4.1: trajectories are more likely to be considered as plausible, with the drawback of causing greater differences between prediction and ground truth.

Then, the attack type with the highest mean plausibility values is the random offset one. This observation confirms what stated in previous paragraphs (Section 4.2.1 and 4.2.2) regarding the ease of detecting one attack rather than others. Since data from a random offset attack show higher plausibility values, trajectories implemented with this type of attack are less likely to be considered malicious. On the contrary, constant position and random offset result in lower plausibility values i.e., higher probability of being discarded.

5 | Conclusions

The following lines summarize the main aspects discussed in the work and provide also some prospects for future work.

This work discussed topics related to misbehaviour detection in V2X communication. Misbehaviour detection concepts are rather new and unexplored issues that need to progress in the following years. The target of a researcher is to exploit available resources such as data from sensors, information contained in messages exchanged by vehicles, prediction schemes based on vehicles' behaviour, and physical property of transmitted signals, to verify the received information is trustworthy.

Previous chapters largely presented the misbehaviour detection scheme implemented for this work. Its goal is to investigate whether the position received by a vehicle has been compromised or not, and to do so power signal strength is leveraged. Essentially, the received information is compared to a theoretical model and then a so-called *plausibility value* is assigned. Ideally, the higher the value, the more reliable the data. And this means that if compromised data are associated to large plausibility values, the attacker is fooling the receiver and achieving its purpose.

Many variables were considered in the analysis: several types of attacks and velocity values, and different differences between the falsified and real trajectories. The scheme works discretely when the difference between real and falsified data is rather small, approximately 10 15 meters, and specifically a constant position attack is rather easy to detect. On the other way around, false trajectories that are considerably far - approximately 30 meters - from the real one are associated to very low plausibility values i.e., misbehaviour is easily detected.

The most challenging aspect is to tune a threshold that discriminates malicious and legitimate data, as it was proved plausibility assignation depends on the speed of the vehicles and on their distance. Also the type of attack affects the method itself: some attacks are

easier to detect than others. Therefore, fixing a unique threshold for any application case would not make any sense at all: any case shall be properly contextualised.

These results bring to light the importance and the need of developing these topics; developing a misbehaviour detection scheme is arduous as it deals with many variables. The suggested scheme has great potentials and the results are encouraging, also its computational simplicity is an interesting plus. However, there are some limitations and some refinement is needed: sometimes predictions are too far apart to expectations and data are too sensible to variations. These two limitations are a direct consequence of the fast changing conditions in the propagation environment.

As future developments, it is necessary to refine the implemented method so that it works with less uncertainty and can provide more stable results. Furthermore, it would be interesting to think of an approach that can work correctly and uniformly in the presence of any attack, so as to weaken this dependency between the type of attack and the results. This would considerably reduce the vulnerabilities and the limitations. Lastly, it is crucial to identify a way to establish thresholds for plausibility values in order to properly process the received information. This last aspect, however, is a matter for standardisation bodies that shall provide further guidelines.

Bibliography

[1] SAE J2735: Dedicated Short Range Communications (DSRC) Message Set Dictionary (TM) . , SAE, September 2015.

[2] TS 23.287, Architecture enhancements for 5G System (5GS) to support Vehicle-to-Everything (V2X) services. , 3GPP, September 2020.

[3] 3GPP. TR 22.885 v14.0.0, Study on LTE Support for V2X Services, . URL `http://www.3gpp.org/DynaReport/22885.html`.

[4] 3GPP. TR 22.886 v16.2.0, Study on enhancement of 3GPP Support for 5G V2X Services. , 3GPP, .

[5] 3GPP. TR 37.885 v15.3.0, 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Study on evaluation methodology of new Vehicle-toEverything (V2X) use cases for LTE and NR; (Release 15). , 3GPP, .

[6] 3GPP. TS 22.186 v16.2.0, Service requirements for enhanced V2X scenarios. , 3GPP, .

[7] 3GPP. Spatial channel model for Multiple Input Multiple Output (MIMO) simulations. , 3GPP TR25996, .

[8] 5GPPP. 5G Automotive Vision, 2015-Oct-20. URL `https://5g-ppp.eu/wp-content/uploads/2014/02/5G-PPP-White-Paper-on-Automotive-Vertical-Sectors.pdf`.

[9] M. H. Adnan and Z. Ahmad Zukarnain. Device-to-device communication in 5g environment: Issues, solutions, and challenges. *Symmetry*, 12(11), 2020. URL `https://www.mdpi.com/2073-8994/12/11/1762`.

[10] A. Asadi, Q. Wang, and V. Mancuso. A survey on device-to-device communication in cellular networks. *CoRR*, 2013. URL `http://arxiv.org/abs/1310.0720`.

[11] E. Barker. Digital signature standard (dss), 2013-07-19 2013.

[12] A. Bazzi, A. Berthet, C. Campolo, B. Masini, A. Molinaro, and A. Zanella. On the

Design of Sidelink for Cellular V2X: A Literature Review and Outlook for Future. *IEEE Access*, PP:1–1, 07 2021.

[13] ETSI (2014). Intelligent Transport Systems (ITS); Vehicular Communications; GeoNetworking; Part 1: Requirements, EN 302 636-1. , ETSI, April 2014.

[14] European Commission, Directorate-General for Mobility and Transport (2019). Commission Delegated Regulation (EU) of 13.3.2019 supplementing Directive 2010/40/EU of the European Parliament and of the Council with regard to the deployment and operational use of cooperative intelligent transport systems.

[15] European Telecommunications Standards Institute. Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part 3: Specifications of Decentralized Environmental Notification Basic Service. , ETSI TS 102 637-3, 201.

[16] European Telecommunications Standards Institute. Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part 2: Specification of Cooperative Awareness Basic Service. , ETSI TS 102 637-2, 2010.

[17] European Telecommunications Standards Institute. Intelligent Transport System (ITS); Vehicular Communications; Basic Set of Applications; Analysis of the Collective - Perception Service (CPS). , ETSI TR 103 562 V2.1.1, December 2019.

[18] European Telecommunications Standards Institute (2017). Intelligent transport systems (ITS); security; security header and certificate formats. , TS 103 097 v1.3.1.

[19] European Telecommunications Standards Institute (2018). Intelligent transport systems (ITS); security; trust and privacy management. , TS 102 941 v1.2.1.

[20] M. Fallgren, M. Dillinger, Mahmoodi, and T. Svensson. *Cellular V2x for Connected Automated Driving*. John Wiley Sons Inc, 1 edition, May 2021.

[21] A. Festag. Standards for vehicular communication—from ieee 802.11p to 5g. *e i Elektrotechnik und Informationstechnik*, 132, 09 2015.

[22] M. Harounabadi, D. M. Soleymani, S. Bhadauria, M. Leyh, and E. Roth-Mandutz. V2x in 3gpp standardization: Nr sidelink in release-16 and beyond. *IEEE Communications Standards Magazine*, 5(1):12–21, 2021.

[23] M. Hasan, S. Mohan, T. Shimizu, and H. Lu. Securing vehicle-to-everything (V2X) communication platforms. *CoRR*, abs/2003.07191, 2020. URL `https://arxiv.org/abs/2003.07191`.

[24] Institute of Electrical Engineers (2011). IEEE standard for information technol-

ogy— telecommunications and information exchange between systems—local and metropolitan area networks—specific requirements, part 11: wireless LAN medium access control (MAC) and physical layer (PHY) specifications. , IEEE, March 2012.

[25] Institute of Electrical Engineers (2017). IEEE standard for wireless access in vehicular environments— security services for applications and management messages amendment 1. Std 1609.2a. , IEEE.

[26] P. Kyösti, J. Meinilä, L. Hentila, X. Zhao, T. Jämsä, C. Schneider, M. Narandzi'c, M. Milojevi'c, A. Hong, J. Ylitalo, V.-M. Holappa, M. Alatossava, R. Bultitude, Y. Jong, and T. Rautiainen. Ist-4-027756 winner ii d1.1.2 v1.2 winner ii channel models. *Inf. Soc. Technol*, 11, 02 2008.

[27] B. Lonc and P. Cincilla. Cooperative its security framework: Standards and implementations progress in europe, 2016.

[28] J. Petit and S. E. Shladover. Potential Cyberattacks on Automated Vehicles. *IEEE Transactions on Intelligent Transportation Systems*, 16(2):546–556, 2015.

[29] S. Jaeckel, L. Raschkowski, K. Borner, L. Thiele, F.Burkhardt and E.Eberlein. "QuaDRiGa - Quasi Deterministic Radio Channel Generator, User Manual and Documentation", 2019.

[30] SAE International. SAE J3016 (2014): Taxonomy and definitions for terms related to onroad motor vehicle automated driving systems, Januray 2014. URL `http://standards.sae.org`.

[31] K. Sehla, T. M. T. Nguyen, G. Pujolle, and P. B. Velloso. Resource Allocation Modes in C-V2X: From LTE-V2X to 5G-V2X. *IEEE Internet of Things Journal*, 9 (11):8291–8314, 2022.

[32] H. Seo, K.-D. Lee, S. Yasukawa, Y. Peng, and P. Sartori. LTE Evolution for Vehicle-to-Everything Services. *IEEE Communications Magazine*, pages 22–28, 2016.

[33] R. Shrestha, S. Y. Nam, R. Bajracharya, and S. Kim. Evolution of v2x communication and integration of blockchain for security enhancements. *Electronics*, 9, 2020. URL `https://www.mdpi.com/2079-9292/9/9/1338`.

[34] S. So, J. Petit, and D. Starobinski. Physical layer plausibility checks for misbehavior detection in v2x networks. pages 84–93, 05 2019.

[35] G. Thandavarayan, M. Sepulcre, and J. Gozálvez. Generation of Cooperative Per-

ception Messages for Connected and Automated Vehicles. *CoRR*, abs/1908.11151, 2019. URL `http://arxiv.org/abs/1908.11151`.

[36] R. W. van der Heijden, S. Dietzel, T. Leinmuller, and F. Kargl. Survey on Misbehavior Detection in Cooperative Intelligent Transportation Systems. *IEEE Communications Surveys Tutorials*, 21(1):779–811, 2019.

List of Figures

List of Symbols

3GPP 3rd Generation Partnership Project
4G Fourth-Generation
5G Fifth-Generation
5GAA 5G Automotive Association
5G-NR 5G New Radio
5GPPP 5G Infrastructure Public Private Partnership
AA Authorisation Authorities
ADAS Advanced Driver Assistance Systems
AT Authorization Ticket
BS Base Station
BSM Basic Safety Message
BSS Basic Service Set
C2C Car 2 Car
CA Certification Authorities
CAM Cooperative Awareness Message
CP Collective Perception
CPM Collective Perception Message
C-ITS Cooperative Intelligent Transportation Systems
C-V2X Cellular V2X
DENM Decentralized Environmental Notification Message
DoS Denial of Service
DSRC Dedicated Short Range Communications
EA Enrolment Authorities
EC Enrolment Certificates
EE End Entities
ETSI European Telecommunications Standards Institute
HARQ Hybrid Automatic Repeat reQuest
ICT Information and Communications Technology
IEEE Institute of Electrical and Electronics Engineers

IEEE-SA Institute of Electrical and Electronics Engineers Standard Association
IP Internet Protocol
ITS Intelligent Transport System
LOS Line of Sight
LTE Long Term Evolution
NLOS Non-Line of Sight
NLOSv Non-Line of Sight Vehicle
NR New Radio
OBU On Board Unit
OCB Outside Context of a Basic Service Set
OFDM Orthogonal Frequency Division Multiplexing
PKI Public Key Infrastructure
ProSe Proximity-based Services
QoS Quality of Service
RCA Root Certification Authorities
RSSI Received Signal Strength Indicator
RSU Road Side Unit
SAE Society of Automotive Engineers
SCM Spatial Channel Model
SDO Standards Developing Organizations
TCP Transmission Control Protocol
TTP Trusted Third Party
UDP User Datagram Protocol
UE User Equipment
V2I Vehicle-to-Infrastructure
V2N Vehicle-to-Network
V2P Vehicle-to-Pedestrian
V2V Vehicle-to-Vehicle
V2X Vehicle-to-Everything
VRU Vulnerable Road User
WAVE Wireless Access for Vehicular Environment
WINNER Wireless World Initiative New Radio