Security for autonomous vehicles as seen from a societal and systems perspective – Final report

Mikael Asplund (LiU), Valentina Ivanova (RISE), Felipe Boeira (LiU)

1 Introduction

This is the final project report for the Vinnova-funded project Security for autonomous vehicles as seen from a societal and systems perspective. The project is funded through the Drive Sweden program as a strategic initiative. The report is written in English to facilitate international collaboration and dissemination.

1.1 Project aims

The original formulation of the project aims was (translated from Swedish):

1. To contribute to a continued societal development and enable the implementation of intelligent transportation system by broadly meeting the challenge of increased cyberthreats. The purpose of the project is therefore to increase the engagement and knowledge regarding security for intelligent transportation systems.
2. To gather stakeholders from both academia and industry, as well as from governmental agencies and municipalities with an interest and need to understand how system safety for transport systems are affected by digitalization and new threat models
3. To be the foundation for a larger cooperation project with a larger set of actors and longer time horizon. The need analysis, comparison study and data collection are important components to form the basis of a larger application in the area.

1.2 Project partners

The partners in the project have been:

- Linköping University (LiU)
- RISE Research Institutes of Sweden
- Scania
- Ericsson
- Transdev
- Combitech
- Nationellt Forensiskt Center (NFC)

The project has been led by LiU. RISE has been responsible for data collection activities, and all partners have been participating in discussions and workshops.

1.3 Planned project activities

The following activities were planned:

- Performing **workshops** with the purpose of gathering stakeholders, collecting needs and conveying new results.
- **Analysis** of the ELIN (Ride the future) platform, in particular with the Control Tower functionality that is planned in the SHOW project, with regard to cybersecurity.
- **Comparison study**, external analysis and identification of challenges, as well as inspiration from other sectors.
- **Data collection** from the vehicles in the ELIN platform to support further investigations.
- Preparation for a **project application**.

1.4 Ride the future platform

The Ride the future platform (aka ELIN) is a collaboration project with eight partners across the Östergötland region. The current physical manifestation of the platform are two electric autonomous buses that serve the Linköping University campus Valla area (one of them is shown in Figure 1). The buses have been supplied by two different manufacturers, Navya and Easymile. The buses started operating during 2020 and were actively transporting passengers until the Covid-19 situation caused a stop of passenger service. We have still been running the buses to perform experiments and studies relating to the different research project connected to the platform.

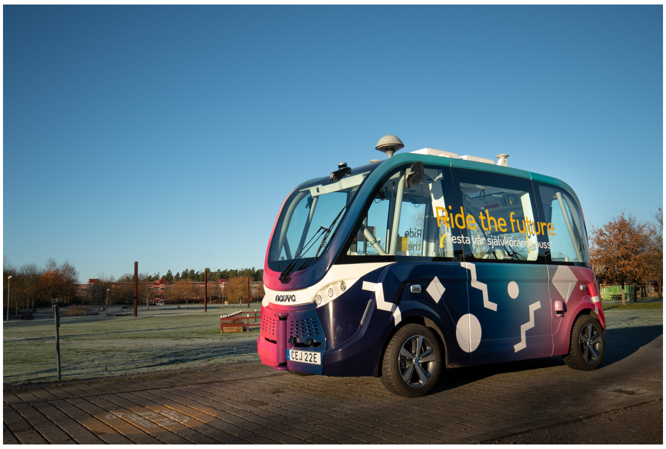


Figure 1. One of the buses in Ride the future

Work is now ongoing to expand the operating domain of the buses to a residential area, which includes a school and an elderly home.

1.5 Document outline

The rest of this document is organized as follows. Section 2 lists the activities that have been performed in connection to this project. Section 3 presents the system analysis and literature review that are part of the project deliverables. Section 4 presents results from the technical activities that have been performed mainly at LiU and RISE, and Section 5 discuss some of the lessons learnt during the project. Finally, we present some general recommendations in Section 6.

2 Performed activities

The project has been significantly affected by COVID-19. The original plan was to hold a number of physical workshops to discuss different aspects of security of autonomous vehicles. This was made difficult due to travelling restrictions, furloughs in participating organizations as well as an increased workload due to distance mode teaching at LiU. Because of this, the project was granted an extension to end of March which allowed some more activities to be performed as planned.

Another fact that forced some replanning was the somewhat limited access to the buses in the Elin platform. While we have been able to access data through the APIs provided by the manufacturers, this data has been restricted by non-disclosure agreements and only physically accessible by RISE. Moreover, the permission to use the buses in traffic as given by Transportstyrelsen also restricts the ability to in any way modify the buses (e.g., to explore remote operation). Therefore, the analysis activity originally planned has been repurposed to investigate other aspects of security for autonomous vehicles.

We briefly describe the activities that have been performed in the project.

Workshops

The project has had three internal project-wide workshops (2019-12-12, 2020-06-16, 2021-03-30). These workshops have been used to exchange information, discuss the direction of the work, present results and get feedback. Moreover, an external workshop was arranged on November 5, 2020. The workshop was attended by more than 20 participants and contained both presentations from external actors and results from the project.

Comparison study

This mainly theoretical activity has been centred around understanding the implications of cybersecurity on autonomous vehicles. The inputs to this study have been presentations and discussions in workshops, scientific literature, and other relevant material such as industry standards. The results from this work are presented in Section 3.

Data collection

The data collection activities began with an integration period (during spring 2020) when the data flows in Apache Nifi (see below for a description) were developed and tested to confirm that all the data is received correctly. As a result, the data collection process for one of the shuttles was established during late spring/early summer 2020. Data retrieval from the other shuttle has been complicated by problems with certificates, which were mostly resolved during the summer months of 2020. The data collection process for the second shuttle was established during the autumn 2020. Both shuttles provide data with similar content but through a different API endpoints and structure. The data mainly consists of telematics data and data describing the site and routes. Due to the agreements concerning the co-suppliers of the vehicles, and requirements for road safety, it has proved difficult to obtain more detailed data that would be of value from a cybersecurity perspective (both in terms of analysis of possible vulnerabilities and from a forensic perspective). The results of this activity are presented in Section 4.1.

Secure localization

Since we were not able to penetrate the technical side of the Ride the future platform from cybersecurity perspective, we have performed two technical activities that look at security problems for connected and autonomous vehicles. The first of these is a physical implementation of a protocol for secure localization called Vouch/Vouch+ [1]. The purpose of this activity is to investigate if such a scheme can have a realistic use case in the Ride the future platform. The outcome of this work is presented in Section 4.2.

Formal verification of security protocols

The other technical strand that we have been working on as part of this project emerged as a use case from the domain of vehicular platooning for heavy-duty vehicles. In this work we have been analysing a communication protocol for platoons from a security perspective. The outcome of this work is presented in Section 4.3.

Project applications

The final activity that we outlined at the beginning of the project was to form a larger consortium for applying for a larger project proposal. LiU and Scania both participated in a call that focus on safety and security for autonomous vehicles to the European Marie Curie ITN call. At the moment there are no plans for making an extended project application in this particular constellation, but work is on-going towards calls in the Horizon Europe program for cybersecurity for connected and autonomous vehicles, and the contacts established as part of this project will serve as the base for future applications in the area.

3 System analysis and literature review

3.1 Security and safety issues related to autonomous buses

The first question to address when thinking about security for autonomous vehicles is whether there are any specific cybersecurity concerns related to autonomous vehicles. To make the matter very concrete, it is natural to consider the question of whether there is a higher risk of cyberattacks associated with autonomous vehicles compared to regular vehicles. Risk here should be understood as the combination of likelihood of an attack and the consequence that such an attack can have.

3.1.1 Does automation increase the likelihood of cyberattacks?

Considering first the likelihood, the question of whether an AV is more likely to be the subject of a cyberattack than a regular vehicle does not have an obvious answer, since it depends on what we mean by a regular vehicle. A vehicle that provides advanced driver assistance and a high degree of connectivity will exhibit most of the technologies that are present in AVs, and therefore potentially have a comparable exposure to threats. However, it is to be noted that AVs require a certain base set of technologies to work, and therefore one needs to consider their impact of cybersecurity. We go through these very briefly.

- **Sensing technologies.** An AV needs to know its environment and surroundings. Today the use of satellite navigation, cameras, LIDARs, and radars (often in some combination) constitute the main means to acquire this knowledge. Each of these have been shown to be vulnerable to attacks [2, 3], but usually they are very specific and require physical proximity.
- **Automated decision making.** The automated decisions that must be present in an AV can broadly be divided in two classes, rule-based and learning-based. Rule-based decision making is the less problematic of these from a security standpoint, since it is no different from other software present in a modern vehicle (with regards to likelihood, we will return to the problem of consequence). Learning-based decision making on the other hand causes new problems that are still not solved to a satisfactory degree. First, there is a lack of verification procedures. Essentially, it is very difficult to verify the correct behaviour of the system since it is not subject to inspection and analysis (contributing to the need and interest in explainable AI). Second, an attacker with access to a trained model can analyse its behaviour and create malicious input that it knows will be erroneously treated [4]. Finally, there is the possibility of attackers somehow impacting the training data to explicitly create such backdoors that make it easier to deploy this kind of attacks. However, taken together the introduction of automated decision making is less of an issue when it comes to increases

likelihood of cyberattacks, and more an issue of verification of the safety of the intended functionality (SOTIF).
- **Interconnected subsystems.** An often-neglected requirement of AVs is the need for subsystems to be interconnected. In a regular vehicle, it is possible to completely separate or at least have very strict interfaces between subsystems. In an AV, they must be tightly integrated through the control of the vehicle. This significantly increases the exposure to attacks since gaining access to one subsystem can potentially result in gaining access to other subsystems as well.
- **Connectivity.** The extent to which AVs at large need to be connected is still an open question. However, for autonomous buses, which is the study object here, there are some indications that connectivity will be an essential part of future automated public transport systems. There are three reasons why public transport AVs require more connectivity compared to private AVs. (1) Unless a safety driver is present in all vehicles the AV cannot resort to a human operator in the vehicle itself in case of problems and will therefore likely require remote assistance. (2) Fleet management (e.g., instructing the vehicle on which route to follow) will be an essential part of any automated public transport system. (3) Requirements on public safety will mean that the AVs must be possible to supervise and potentially stop in case of some emergency event (e.g., fire, crimes, etc). The extent to which this added connectivity is also tied to the control of the vehicle is an open question, but the mere fact that communication is required does impact the cybersecurity assessment and in particular the likelihood of attack.

In summary, the two main factors that potentially increase the likelihood of a cyberattack against public transport AVs are increased connectivity and a higher degree of interconnectivity between subsystems.

3.1.2 Does automation increase the consequence of cyberattacks?

The question of whether automation increases the consequences of an attack is deceptively easy to answer with a clear "yes of course!". However, there are nuances to this question that deserve to be further investigated. Below we briefly touch upon three of the most important factors to consider in this regard.

- **Actual effects.** The concrete effects of a cyber-attack cannot be characterized at a general level as it is dependent on the specifics of the vehicle design and in particular what safeguards and protection mechanisms that are in place. However, it is important to realize that there are many potential effects that can be used by cybercriminals other than the most spectacular ones where an attacker takes direct control over a vehicle. Examples of possible attacks include preventing a vehicle from starting, theft of personal information, playing extremely loud music to distract the driver, turning of all lights in the dark, and so on. Many of these have a significant impact on safety, thus motivating the integration of safety and security analyses [5, 6]. Also, see He et al. [7] for an analysis where such consequences are mapped to individual assets (e.g., subsystems).

- **Scale.** It is important to differentiate between attacks against individual vehicles (which can be extremely serious since a misbehaving heavy can kill many people), and large-scale attacks. Unfortunately, in many of the situations where an individual vehicle is vulnerable to an attack, then it is likely that others are as well. This potentially opens up for economically incentivized extortion attacks through for example a denial-of-service attack where parts of the traffic system are incapacitated. It is clear that in both small-scale and large-scale attacks the consequences of cyber-attacks are increased for AVs compared to regular vehicles.
- **Societal effects.** Public perception and acceptance of new technologies depend on a large number of factors. Already today there is considerable concern that AV technologies are poorly protected from hackers and terrorists (e.g., Ahmed et al. [8] report 68% of participants have this concern in a relatively large study). If there are actual cases of severe incidents caused by cyberattacks this could have very large consequences on the way society thinks and relates to AVs.

3.1.3 Security engineering at what level?

In the two previous subsections we have seen that automated driving has an impact on both the likelihood and consequences of cyberattacks, and therefore cybersecurity risk assessment becomes central. Moreover, there is clearly a connection between security and safety since (1) a lack in security can have significant consequences on safety, (2) measures to increase safety can have negative security implications, (3) measures to increase security can have negative safety implications, and (4) they have many commonalities in how they should be analysed and treated with regards to risk assessment, documentation, processes and incident reporting.

There is much to say about these issues, and we relate some of the academic discussions on the topic in Section 3.3. In this report we focus on risk assessment from a larger perspective with the question of who owns cybersecurity issues. We have identified the following levels of security analyses:

- **Component level.** Performing best-practice security engineering at the component level is conceptually not that complicated. A component usually has a well-defined function specification and a limited set of interfaces which makes it feasible to perform the security analysis. Moreover, this is covered by the ISO/SAE 21434 standard that states "This document specifies requirements for cybersecurity risk management regarding engineering for concept, development, production, operation, maintenance, and decommissioning for road vehicle electrical and electronic (E/E) systems, including their components and interfaces." The beneficiary of a security analysis at component level would typically be the original equipment manufacturers (OEMs).

- **OEM level.** The risk assessment at vehicle-level is more complicated, but again the ISO/SAE 21434 standard provides much needed support to perform such analyses not only at the production phase, but also during operation, maintenance and decommission. While the standard is process-oriented with requirements on procedures and mechanisms that should be in place, it sets a much higher standard level than what has previously been required.
- **Transportation system level.** Autonomous and semiautonomous vehicles will be part of a much larger transportation system. This will encompass vehicles from a large number of manufacturers as well as different operators, service providers and governmental stakeholders. At the moment there is a lack of risk management relating to cybersecurity at the transportation system level. This is due to several reasons. First, it is not clear who has the responsibility to perform such risk management. Second, the road-based transportation system is still often considered by many as mainly physical infrastructure (roads, lamps, drainage, etc) and not so much a digital infrastructure. To some extend this is still true, but as discussed in this report will change as more and more vehicles become reliant on cloud connectivity of some sort.
- **Societal level.** Finally, at the societal level, the issue of cybersecurity for autonomous vehicles is really just in its infancy. In a report from December 2019 [9] the Swedish Defence Research Agency states that (our translation) "Cybersecurity regulation for the transport sector in its entirety is composed of a fragmented and complex legislation", as well as "In order to further streamline the national application of these regulations, increased collaboration between competent authorities, industry and the research community will be needed. The work of designing, disseminating, and training on new regulations and guidelines must continue to concretize organizational and technical requirements as well as overarching legal principles." Our assessment after this pre-study project is that this conclusion still holds. The UN Regulation on Cybersecurity and Cyber Security Management Systems is a first step in the right direction but is still at a very high level and requires more efforts to make concrete and applicable in a practical setting.

3.1.4 Liability and insurance

In the external workshop Dr. Sara Landini from University of Florence presented results from studies they have made on liability and insurance and how these have been dealt with at a European level. In her 2020 paper on this topic [10] she proposes that strict legal liability might not be sufficient to prevent incidents and calls for mechanisms where insurance can play a larger role. She concludes that

"From the analysis of the interventions at the level of EU legislation, there is a certain lack of communication between the various regulatory areas. Protection of road victims, data protection in cyberspace and producer responsibility are strongly correlated in reality. These are regulations that must be constructed as 'communicating vessels' and not as closed and self-referential areas. The national legislator and the community legislator must recover a vision that is close to the problem and at the same time coherent with the general framework of value, principles and social instances."

3.2 Scientific literature

We briefly survey some of the scientific literature in the area, starting with connected vehicles at large, narrowing down to works on formal analysis of security in this context, accounting for false information, and works that try to handle this issue.

Much of the research on cybersecurity in relation to connected vehicles have been conducted in the framing of vehicular networks. Given the legacy from research on MANETs many of the earlier works in this area focused on security issues with regards to routing [11]. As the technology of inter-vehicle communication matured and became more standardized, the focus shifted to the challenges of authentication, privacy [12], and how to prevent spreading of false information [13].

Security in the context of vehicular platoons have been investigated by Studer et al. [14] who employ a combination of ensuring validity of data over time to verify that a vehicle is travelling in the same convoy, and distance-based verification using time-of-flight of messages and MAC-layer timestamps. Ucar et al. [15] show that visible light communication can reduce but not remove the risks associated with attackers outside a platoon.

Vehicular network security standardization and its formal analysis is rather recent. Whitefield et al.[16] analyze V2X certificate revocation of malicious or misbehaving vehicles with the Rewire scheme using Tamarin. In their analysis, they are able to identify an authentication weakness and propose an extension to mitigate it. Li et al. propose a lightweight privacy-preserving authentication protocol that is verified with BAN logic and Proverif [17]. Dahl et al. [18] also use the ProVerif tool to analyze location privacy of cars in a vehicular network. Assuming that cars continuously broadcast their identity and location to their surroundings to facilitate safety applications, this information can also be used to track car movements.

Chen et al. [19] presented an example of what can occur if the design of connected vehicular applications has not sufficiently accounted for the possibility of false information. The authors analyze the Intelligent Traffic Signal System (I-SIG), which is currently being launched as a pilot application by the US department of transportation. The authors show that by spoofing messages, a single attacker is able to disturb the signaling control algorithm, resulting in massive traffic jams. Moreover, our previous work [20] demonstrates that the ability to create fake identities coupled with false message contents can cause severe collisions for vehicular platooning scenarios.

Common for these works is that they consider attacks that use incorrect information regarding speed and location of one or multiple vehicles in the area. Counteracting such attacks require some form of location verification. There is a wide range of works that try to tackle this long-standing problem using various assumptions. For example, the work by Yan et al. [21] use antenna arrays to verify that claimed locations are consistent with the angle at which signals arrive to a base station. However, given the inherent uncertainty associated with radio propagation and the potentially complex attack scenarios, collaborative algorithms are needed to differentiate false alarms from real attacks. Existing work for collaborative location verification such as that by Zhu and Cao [22] are not adapted for the high-speed scenarios associated with vehicular applications.

4 Results from technical activities

4.1 Data collection

4.1.1 Data interfaces available from buses

Data from two autonomous shuttles operating in campus Valla of Linköping University have been collected in connection to this project. They are from two different suppliers – EasyMile and Navya. Due to contract agreements, we are not allowed to reveal information about the structure of the vehicle APIs. Both companies have developed public API interfaces using HTTPS and WebSocket Secure (WSS) communication protocols to deliver data from their vehicles. Data from the vehicles is first delivered to the respective supplier cloud-based platform and then to the shuttles' operators (see Figure 2). Only authenticated users can receive data from the vehicles. In both cases authentication is done by sending an access token with each API call in case of REST API or a heartbeat message (sent with a certain frequency) in the case of WSS. The access token is renewed once per day by supplying authentication credentials (received from the company) to an external identity provider.

The different data interfaces in the SHOW project are shown in Figure 2. The description here considers the interface between the suppliers' cloud and Apache NiFi deployment at RISE.

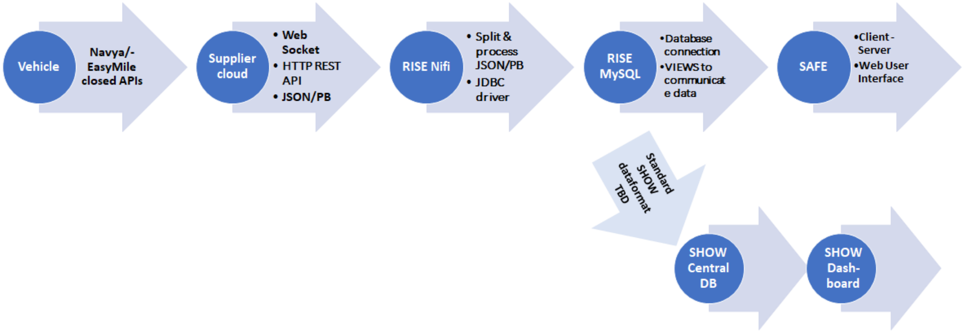


Figure 2. Data interfaces in the SHOW project.

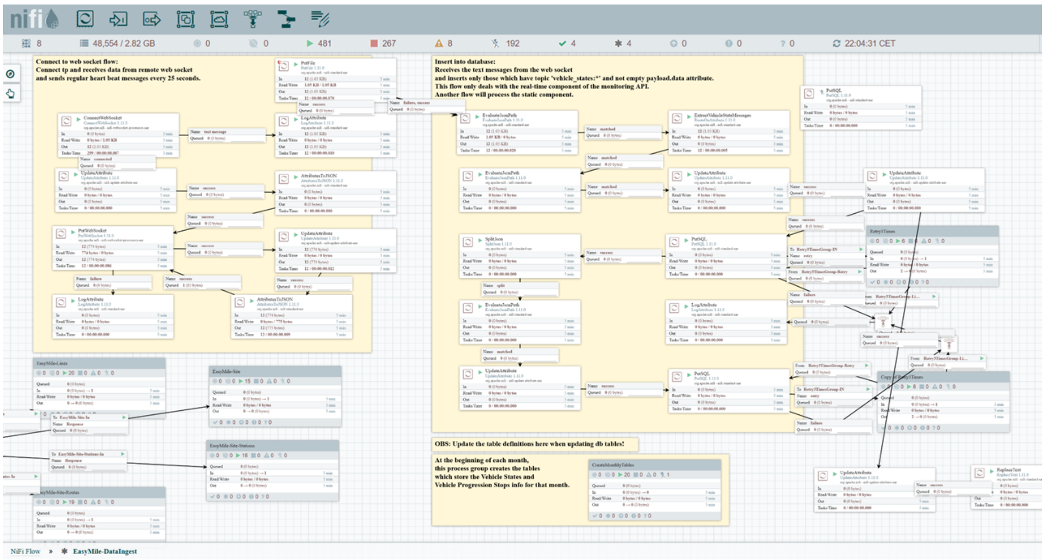
Two categories of data are distinguished in both vehicles. (the data categories listed below are derived from the data available in the APIs but values are not necessary supplied for all attributes. Additionally, these data might have been received by each of the vehicles or both). The data from both vehicles is similar with some differences in the structure of the endpoints and attribute names:

- Static data retrieved once per day using HTTP REST API endpoints. The static data focuses on the description of the site with the location of the stations and their coordinates as well as preconfigured routes and lines. Additionally, some information about the vehicle equipment is also provided.
- Dynamic (near real-time) data pushed once or twice per second via WebSocket Secure. The dynamic data can also be separated into several categories:
    - Vehicle identifier and a timestamp of the message.
    - Current vehicle location and heading – including GPS coordinates and GNSS correction if available. Correspondence between the pre-recorded reference 3D map and current LiDAR readings.
    - Current state of the vehicle's sensors and actuators –
        - Vehicle control - speed, acceleration, mileage, and steering data as well as connection status (connected/disconnected).
        - Signal lights and wipers.
        - Passenger access and accessibility equipment (door and ramp) as well as number of passenger and payload.
        - Battery status and state.
        - Internal, external, engine and battery temperatures.
    - Mode of operation – autonomous or manual navigation, operating on a predefined route (with or without stopping at each stop) or on-demand service, in-use status - transporting passengers or in stand-by before/after the current trip.

  - Next stops and waiting times as well as progression to the next stop.
  - Events – describing unexpected situations and the location of the vehicle at that time.

4.1.2 Data collection framework

Data collection was implemented using Apache NiFi which we have used for data collection in another project too. Apache NiFi is an open-source project designed to automate data flows (including data collection, pre-processing, transformation, and routing) between software systems. Figure 3 shows an example data collection flow. It adopts flow-based programming paradigm with a web-based user interface to design and implement data flows between various systems. Data flows are developed by connecting NiFi modules (called processors) with various functionalities to each other. NiFi provides a large number of processors covering various functionalities, for instance receive/send HTTP requests, evaluate JSONPath expressions, insert into database, and it is also easily extensible – allowing for development of custom processors. NiFi is very scalable and capable to operate within clusters, thought for this project, we didn't need to use these capabilities. After the data have been received, either from the REST APIs or WebSocket, it has been pre-processed to extract relevant values and then inserted into a MySQL database. Both Apache NiFi and MySQL deployments are on RISE premises – see Figure 2 above.



Connect to web socket flow:
Connect tp and receives data from remote web socket
and sends regular heart beat messages every 25 seconds.

Insert into database:
Receives the text messages from the web socket
and inserts only those which have topic 'vehicle_states:*' and not empty payload.data attribute.
This flow only deals with the real-time component of the monitoring API.
Another flow will process the static component.

OBS: Update the table definitions here when updating db tables!

At the beginning of each month,
this process group creates the tables
which store the Vehicle States and
Vehicle Progression Stops info for that month.

Figure 3. Example data collection flow in Apache NiFi

4.1.3 Collected data

The data collection activities begin with an integration period when the data flows in Nifi were developed and tested to confirm that all the data is received correctly. Then the data flows can be used in production. We have started collecting the first

data at the beginning of June 2020 and then continue again in September as the vehicles were not in operation during the summer.
Static data is collected once per day and dynamic data is collected only when the vehicles are in operation – once or twice per second. The operating schedule has been reduced due to the Covid-19 pandemic and additionally technical issues with the shuttles (for instance, waiting to replace LIDARs) and often only one of the shuttles was in operation during a working day. Data collected spans most of the forementioned categories and attributes with the noticeable exception of the lack of the vehicle events data where no such data have been transmitted from the APIs even though the endpoints were working.

4.2 Secure localization

In this section we present the work we have done to implement a prototype for secure localization. The problem that is addressed in this work can be described as follows. Consider two autonomous vehicles A and B that are approaching an intersection from two different directions. Both vehicles would benefit from knowing the other's position and speed to be able to plan ahead. However, if either provides false information to the other, it can result in negative consequences or even hazard. Therefore, it would be beneficial if the location information sent (for example) from A to B could be independently verified by B to be plausible, then it could proceed with a higher confidence in the information than if it just had to trust A's information.

In work that preceded this project participants at Linköping University in collaboration with researchers from Brazil developed a secure localization scheme called Vouch+ to achieve exactly this goal. During this project we have been implementing this protocol as a proof-of-concept (PoC) prototype to deploy on the physical bus platform provided by Ride the future. In the remainder of this section we first very briefly describe the secure localization scheme, give an overview of the implementation efforts and finally describe some of the lessons learnt and the next steps in this direction.

4.2.1 Secure localization scheme

The Vouch+ decentralized proof-of-location scheme is composed of four main components, (i) a proof acquisition and (ii) a proof dissemination protocol, (iii) a plausibility verification module, and (iv) a reaction strategy. The first module ensures that location proofs are created and provided to nodes that wish to prove their location. The seconds determines how these proofs are disseminated to neighbor nodes that want to verify the location of the sender. The third module resides in the verifier nodes and applies a plausibility check to decide whether the beacons it receives can be trusted or not. Finally, the reaction strategy determines how to act if the location messages cannot be trusted. More details on how the scheme works can be found in [1].

4.2.2 Implementation of Vouch+ on Ride the future buses

Vouch+ has been extensively studied in simulation settings, but so far not in a more realistic environment. Some of the required technologies to do a full-scale implementation of Vouch+ are currently not attainable in a short pre-study project with a restricted budget as this one. However, to do a PoC implementation one can use other alternative technologies just to test the ideas and make initial experiments on timing and accuracy and see how this affects the usefulness of the scheme in a setting of autonomous buses.

The realization of the proof-of-concept requires can be roughly divided into design, implementation, and integration/testing. In the design phase, we had to solve two main problems. First, what hardware to use that could meet the requirements on fast prototyping while still providing sufficient capabilities to make the protocol implementation feasible. Second, how to structure the software components of the implementation, what framework to use etc.

For the hardware we did a survey of the market and found that the best option was to purchase several small handheld devices on which it was possible to install a regular Linux distribution. In combination with communication interfaces such as Bluetooth, Wifi, 4G and GPS, these battery-driven units fulfilled most of the desired capabilities. Figure 4 shows one of these devices.

Figure 4. Gemini PDA used for PoC, photo by Gareth Halfacree

The software design involved both deciding on the basic framework to use as well as structuring the code for the PoC. We decided that an implementation in Python would be appropriate for this work as it provides rapid prototyping, plenty of libraries and support for hardware interfaces. The software structure was developed using UML and can be seen in Figure 5.

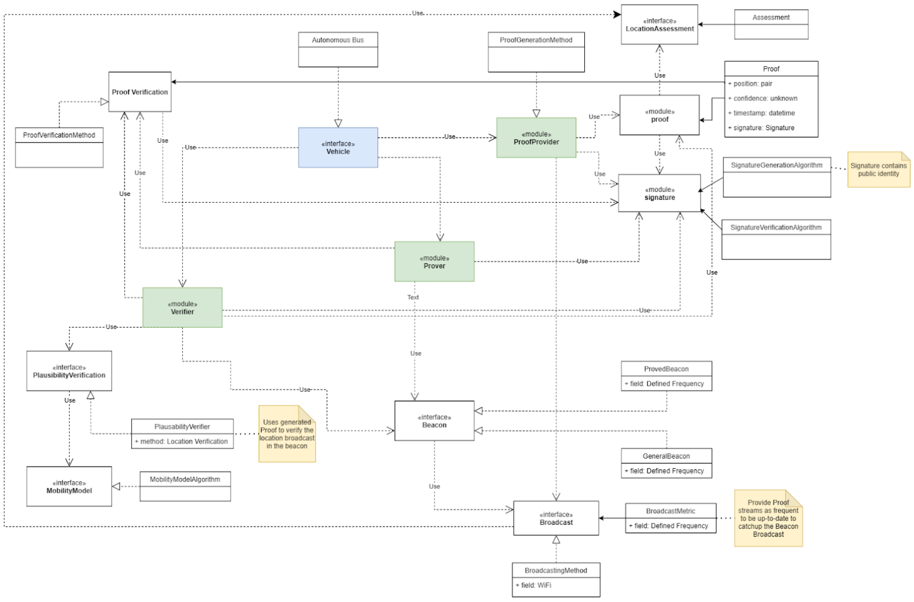

Use
«interface» LocationAssessment
Assessment
Autonomous Bus
ProofGenerationMethod
Proof Verification
Proof
+ position: pair
+ confidence: unknown
+ timestamp: datetime
+ signature: Signature
ProofVerificationMethod
«interface» Vehicle
«module» ProofProvider
«module» proof
«module» signature
SignatureGenerationAlgorithm
Signature contains public identity
SignatureVerificationAlgorithm
«module» Prover
Text
«module» Verifier
«interface» PlausibilityVerification
PlausabilityVerifier
+ method: Location Verification
Uses generated Proof to verify the location broadcast in the beacon
«interface» Beacon
ProvedBeacon
+ field: Defined Frequency
GeneralBeacon
+ field: Defined Frequency
«interface» MobilityModel
MobilityModelAlgorithm
«interface» Broadcast
BroadcastMetric
+ field: Defined Frequency
Provide Proof streams as frequent to be up-to-date to catchup the Beacon Broadcast
BroadcastingMethod
+ field: WiFi

Figure 5. Design model for the Vouch+ implementation

As for the actual implementation of the PoC we were able to implement most of the key components, but not all. The current implementation allows sending and receiving cryptographically signed messages according to the CAM message standard (part of ETSI ITS), over a network link. However, we are still missing some parts related to localization of other nearby devices as there has been issues with both the GPS interfaces and Bluetooth scanning.

4.2.3 Next steps

While the implementation of the PoC cannot be considered fully complete, we think that it has come a long way and that work should be continued. Currently, a bachelor student is working with finalizing the implementation with the ambition to perform field experiments before summer. The purpose of these test is to assess the usefulness of the PoC in the context of autonomous buses in an urban setting with a low to moderate speed.

4.3 Formal verification of security protocols for vehicular control

An important topic in this project is that of remote control of autonomous buses to sort out situations that the vehicle itself cannot. As discussed in Section 3.1.1, this is likely to be necessary in the case of autonomous public transport vehicles, but at the same time raises the risk level considerably.

Therefore, it was deemed relevant from the perspective of this project to investigate to what extent it is possible to increase the assurance level of the remote control link so that it can be deemed secure to the level that it can be trusted with safety-critical decisions. We now proceed to describe some outcomes of a study to formally prove security properties of one such example of a remote control situation. The analysis focus on a high-level cyber-physical protocol that is currently in the pre-standardisation phase and described in the European Ensemble project, and which also builds on the existing ETSI ITS-G5 and IEEE vehicular networking standards (including security). Together, these form an interesting study object since (i) they will have a real and significant impact on the way future commercial vehicles are operated and controlled, (ii) they represent a typical standardisation product composed of multiple cross-references documents (in our case 8 documents and 617 pages), and (iii) the protocol and the associated security specification describe a complex system with dynamically joining and leaving nodes and a non-trivial cryptographic key structure.

4.3.1 Background

In recent years organisations such as the Institute of Electrical and Electronics Engineers (IEEE) and European Telecommunications Standards Institute (ETSI) have been actively working towards standardisation of vehicular network protocols and applications. The Ensemble protocol is built on top of these existing standards and makes use of their services. The Ensemble protocol itself works as a group formation protocol with key establishment and distribution. To formally prove security properties of a protocol such as Ensemble, there are at least three things to consider, how to model the protocol, how to model the attacker, and what security properties to verify.

4.3.2 Methodology

There are two protocol models generally considered for creating cryptographic protocol representations: computational and symbolic. The symbolic model – which we use in this work - allows the reasoning to be automated, although complex protocols usually require the solver to be guided with some proof strategy. Given proper heuristics, the Tamarin protocol verified tool has been shown to work with protocols that exhibit complex state machines that may include loops and agent memory [23, 24,25,26].

Some of the complexities to model protocols lie in collecting information that is scattered across different documents and connecting information that is defined sparsely, as well as interpreting possibly ambiguous specifications with regards to, for example, whether to include certain optional fields in a message. We improved this process by introducing an intermediate step based on compiling the formal descriptions of message structures present in the standards into final message specifications, which were much easier to use as a basis to create formal models from.

The security verification of Ensemble was performed through two Tamarin model variants which we call static and dynamic. The static model contains a fixed set of vehicles, whereas the dynamic model allows an unbounded amount of vehicles to join and leave the system. Our models consist of rules that represent the public key infrastructure, initialisation of the vehicles and platoons, and sending and receiving messages. The rules and properties take approximately 900 lines of text to be defined.

The combination of keys defined in Ensemble, the public key infrastructure, and ephemeral keys used in message profiles from the security standards considerably increases the complexity of our analysis. Our strategy towards making the analysis tractable is to define the relations between the keys and break the complexity into smaller parts that can be combined to prove the security properties. By defining an order relation over the set of key we were able to analyse the security properties of the system despite these problems.

4.3.3 Results

We prove three kinds of security properties, model liveness, secrecy and authenticity. Intuitively, liveness means that the model is sound and that there are no inherent modelling issues that stops is from "running". Secrecy means that confidential data such as keys cannot be learnt by the attacking node, and authenticity means that when node A believes it has been interacting with node B, this is also the case and there has not been a so-called man-in-the-middle that intercepted the interaction and modified packets. The formal specification of these properties is are expressed as lemmas, with a total of 20 lemmas for all properties.

We verified these properties using four different configurations in the modelling framework. The configurations are "Bare Tamarin", which is the default configuration used in the protocol verifier tool that we used. The "Lemma reuse" configuration allows reusing some of the previously used lemmas, which can be seen as a way to allow the manual construction of lemmas with increasing complexity to guide the proof process. The final two configurations rely on the contributions we made in the proof constriction process and which takes as input the ordering of keys that we identified. The difference between these two variants is the extent to which this ordering is allowed to guide the prover.

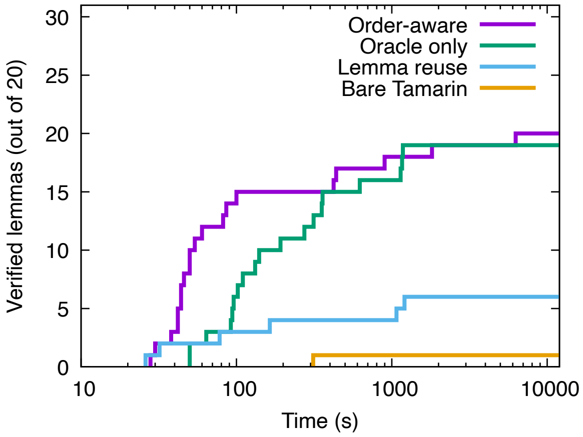


Figure 6. Number of lemmas verified over time.

Figure 1 shows the results of using these configurations on the static model. The x-axis shows the time (logarithmic scale) and the y-axis shows the number of lemmas that could be verified in this time. As can be seen, with our contributions to the formal verification approach, all 20 security lemmas could be proven, which is much better than the 6 lemmas that could be proven without this intervention.

4.3.4 Conclusions

Our analysis of this vehicular communication protocols has revealed a number of interesting facts. First, this pre-standard has a much higher base level of security when compared to many other domains that have been digitalized in the last few years. There is a defence-in-depth approach to how messages are secured, and it follows recent best-practices for security engineering. Second, just as has previously been noted by security researchers in relation to the 5G standard [23] the fact that protocol behaviour and structure is defined through a set of inter-related standard documents create both complexity and risk of design faults and make it difficult to formally analyse their security properties. We have found a clear need for such documents to clearly specify not only message structure, but also dynamic behaviour in a well-defined unambiguous formal language. Finally, our work covers only a small part of the full set of protocols that will be in play in future connected and autonomous transportation systems, and much more work is needed to analyse and verify security properties in these related areas.

5 Lessons learnt and future recommendations

This project had a very wide approach from the onset toward security for autonomous vehicles in the sense that it includes both a systems perspective (which includes technical as well as non-technical issues) and a societal perspective which considers how security of autonomous vehicles impacts society and how society impacts security considerations. We relate some of the lessons we have learnt in this study and provide recommendations for future research and innovation efforts.

5.1 Comparison study

From the literature study, workshop discussions and other meetings and interactions that have taken place during this project, we have found that there is a discrepancy between the visions of future connected autonomous transportation systems and the societal support and regulation that would ensure that these systems are safe and secure. On the other hand, it is an extremely active area with many national and international efforts. The UN Regulation on Cybersecurity and Cyber Security Management Systems is a major step in the right direction as it mandates that there are cybersecurity management systems in place to manage these risks. It also requires that there is incident reporting system in place, but unfortunately, does not put in place a comprehensive incident management system that would feedback this information to other manufacturers or suppliers.

5.2 Data collection

Lessons learnt can be roughly divided into two parts – regarding the data collection framework and data itself.

Nifi proved to be suitable for the data collection task and after the initial learning phase flows were developed and maintained with an ease. One aspect worth mentioning is that Nifi does not support undo or save operations of the flows (so changes are lost). In order to maintain version control of the flows an additional tool (Apache Nifi Registry) has been installed. This tool resembles version control approaches where flows can be committed, versioned and retrieved accordingly.

Both shuttles supply data with custom-structured APIs although they are similar in content. This demanded development of a custom data collection process for each of the two shuttles. This has opened the question about the standardization of the data and API structure to facilitate the addition of other vehicles from a different manufacturer/supplier and integrating the data before analysis. In the framework of the SHOW project, we reviewed existing public transportation standards and drafted (together with partners) a common vehicle model to be used within the project. While three widely used public transportation standards exist (NeTEx/Siri, GTFS and NOPTIS in the Nordics) no standard for structuring the data from autonomous/intelligent vehicle exists. While existing public transportation standards can be used to supply data relevant to passengers' transportation, further standardization efforts are needed in connection to the autonomous operation of the vehicles.

5.3 Secure localization and formal verification

Our other two technical tracks in this project have been valuable to provide both a practical and theoretical perspective of what cybersecurity challenges remain in the area of connected and autonomous vehicles. These studies have shown that there is much still to be done at the technical level and that while the process requirements set out through the UN regulations as well as the ISO/SAE 21434 standard needs to be complemented with more technical method-oriented standard to strengthen cybersecurity engineering practices in the industry.

5.4 Future recommendations

Based on the literature studies, practical experiments, workshop discussions and other inputs gathered, and the lessons learnt, we make the following recommendations to decision makers and stakeholders in the area of autonomous transport.

- Authorities need to start creating regulation and clear requirements on cybersecurity for the transportation sector. This should include risk assessment and risk management, incident management (not just reporting), as well as requirements on mechanisms and methods and not only process.
- The interplay between connectivity and autonomy of vehicles needs further research as it is a non-trivial issue where safety and security are at odds.
- Test sites such as the Ride the future research platform are very valuable for practical studies on autonomous systems, but as they are based on commercial products, it is very difficult to perform independent academic studies on topics relating to cybersecurity. Therefore, it is important to require openness and transparency from providers of systems that build future transportation systems. History has shown that security by obscurity does not work, therefore open standards and protocols are imperative.

6 References

[1] F. Boeira, M. Asplund, and M. Barcellos, Decentralized Proof of Location in Vehicular Ad Hoc Networks, Computer Communications, 2019. doi:10.1016/j.comcom.2019.07.024.

[2] Yulong Cao, Chaowei Xiao, Benjamin Cyr, Yimeng Zhou, Won Park, Sara Rampazzi, Qi Alfred Chen, Kevin Fu, and Z. Morley Mao. 2019. Adversarial Sensor Attack on LiDAR-based Perception in Autonomous Driving. In Proceedings of the

2019 ACM SIGSAC Conference on Computer and Communications Security (CCS '19). Association for Computing Machinery, New York, NY, USA, 2267–2281. DOI:https://doi.org/10.1145/3319535.3339815

[3] P. Kapoor, A. Vora and K. Kang, "Detecting and Mitigating Spoofing Attack Against an Automotive Radar," 2018 IEEE 88th Vehicular Technology Conference (VTC-Fall), 2018, pp. 1-6, doi: 10.1109/VTCFall.2018.8690734.

[4] Nicolas Papernot, Patrick McDaniel, Ian Goodfellow, Somesh Jha, Z. Berkay Celik, and Ananthram Swami. 2017. Practical Black-Box Attacks against Machine Learning. In Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security (ASIA CCS '17). Association for Computing Machinery, New York, NY, USA, 506–519. DOI:https://doi.org/10.1145/3052973.3053009

[5] A. Lautenbach and M. Islam, 'Security models', Deliverable HEAVENS D2.0, v2.0, Mar. 2016.

[6] G. Macher, H. Sporer, R. Berlach, E. Armengaud, and C. Kreiner, 'SAHARA: A Security-Aware Hazard and Risk Analysis Method', in DATE 2015, doi: 10.7873/DATE.2015.0622.

[7] Qiyi He, Xiaolin Meng, Rong Qu, "Towards a Severity Assessment Method for Potential Cyber Attacks to Connected and Autonomous Vehicles", Journal of Advanced Transportation, vol. 2020, Article ID 6873273, 15 pages, 2020. https://doi.org/10.1155/2020/6873273

[8] Sheikh Shahriar Ahmed, Sarvani Sonduru Pantangi, Ugur Eker, Grigorios Fountas, Stephen E. Still, Panagiotis Ch. Anastasopoulos, Analysis of safety benefits and security concerns from the use of autonomous vehicles: A grouped random parameters bivariate probit approach with heterogeneity in means, Analytic Methods in Accident Research, Volume 28, 2020, 100134, ISSN 2213-6657, https://doi.org/10.1016/j.amar.2020.100134.

[9] Erik Zouave, Sabrine Wennberg, Margarita Jaitner, Lag och cybersäkerhet i smart vägtrafik, FOI-R--4811—SE, 2019

[10] Sara Landini, The Insurance Perspective on Prevention and Compensation Issues Relating to Damage Caused by Machines, The Italian Law Journal, 2020, http://hdl.handle.net/2158/1187890

[11] T. Leinmuller, E. Schoch, and F. Kargl. Position verification approaches for vehicular ad hoc networks. Wireless Communications, IEEE, 13(5), 2006. doi: 10.1109/WC-M.2006.250353.

[12] P. Papadimitratos, L. Buttyan, T. Holczer, E. Schoch, J. Freudiger, M. Raya, Z. Ma, F. Kargl, A. Kung, and J.-P. Hubaux. Secure vehicular communication systems: design and architecture. Communications Magazine, IEEE, 46(11), 2008. doi: 10.1109/MCOM.2008.4689252.

[13] N. Bissmeyer, K. Schroder, J. Petit, S. Mauthofer, and K. Bayarou. Short paper: Experimental analysis of misbehavior detection and prevention in vanets. In Fifth IEEE Vehicular Networking Conference, VNC 2013, pages 198–201. IEEE Communications Society, 2013.

[14] A. Studer, M. Luk, and A. Perrig. Efficient mechanisms to provide convoy member and vehicle sequence authentication in vanets. In Third International Conference on Security and Privacy in Communications Networks (SecureComm), 2007. doi: 10.1109/SECCOM.2007.4550363.

[15] S.Ucar, S.C.Ergen, and O.Ozkasap. Security vulnerabilities of ieee 802.11p and visible light communication based pla- toon. In 2016 IEEE Vehicular Networking Conference (VNC), 2016. doi: 10.1109/VNC.2016.7835972.

[16] J. Whitefield, L. Chen, F. Kargl, A. Paverd, S. Schneider, H. Treharne, and S. Wesemeyer. Formal analysis of v2x re- vocation protocols. In G. Livraga and C. Mitchell, editors, Security and Trust Management. Springer, 2017.

[17] X. Li, T. Liu, M. S. Obaidat, F. Wu, P. Vijayakumar, and N. Kumar, "A lightweight privacy-preserving authentication protocol for vanets," IEEE Systems Journal, vol. 14, no. 3, pp. 3547–3557, 2020.

[18] M. Dahl, S. Delaune, and G. Steel. Formal analysis of privacy for vehicular mix-zones. In D. Gritzalis, B. Preneel, and M. Theoharidou, editors, Computer Security – ESORICS 2010, volume 6345 of LNCS. Springer, 2010. doi: 10.1007/978-3-642-15497-3_4.

[19] Chen, Q. A., Yin, Y., Feng, Y., Mao, Z. M., & Liu, H. X. (2018, February). Exposing Congestion Attack on Emerging Connected Vehicle based Traffic Signal Control. In Network and Distributed Systems Security (NDSS) Symposium 2018.

[20] Felipe Boeira, Marinho P. Barcellos, Edison Pignaton de Freitas, Mikael Asplund and Alexey Vinel, On the Impact of Sybil Attacks in Cooperative Driving Scenarios, in proceedings of IFIP Networking 2017 Conference and Workshops, 2017. doi:10.23919/IFIPNetworking.2017.8264890

[21] S. Yan, R. Malaney, I. Nevat, and G. W. Peters. 2016. Location Verification Systems for VANETs in Rician Fading Channels. IEEE Transactions on Vehicular Technology 65, 7 (July 2016). https://doi.org/10.1109/TVT.2015.2453160

[22] Z. Zhu and G. Cao, APPLAUS: A Privacy-Preserving Location Proof Updating System for location-based services, 2011 Proceedings IEEE INFOCOM, Shanghai, 2011, pp. 1889-1897. doi: 10.1109/INFCOM.2011.5934991

[23] David Basin, Jannik Dreier, Lucca Hirschi, Saša Radomirovic, Ralf Sasse, and Vincent Stettler. 2018. A Formal Analysis of 5G Authentication. In Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security (CCS '18). Association for Computing Machinery, New York, NY, USA, 1383–1396. DOI:https://doi.org/10.1145/3243734.3243846

[24] C. Cremers, M. Horvat, S. Scott, and T. van der Merwe, “Automated analysis and verification of tls 1.3: 0-rtt, resumption and delayed authentication,” in 2016 IEEE Symposium on Security and Privacy (SP), 2016, pp. 470–485.

[25] R. Ku ¨nnemann, “Automated backward analysis of pkcs#11 v2.20,” in Principles of Security and Trust, R. Focardi and A. Myers, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2015, pp. 219–238.

[26] C. Cremers, M. Dehnel-Wild, and K. Milner, “Secure authentication in the grid: A formal analysis of dnp3: Sav5,” in Computer Security – ESORICS 2017, S. N. Foley, D. Gollmann, and E. Snekkenes, Eds. Cham: Springer International Publishing, 2017, pp. 389–407.