H. Roßnagel, C. H. Schunck, S. Mödersheim (Hrsg.): Open Identity Summit 2021,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2021 37

Towards the COSCA framework for "COnseptualing Secure CArs"

Giampaolo Bella[1] Pietro Biondi[2] Gianpiero Costantino[3] Ilaria Matteucci[4] Mirco Marchetti[5]

Abstract: Cyber risks associated with modern cars are often referred to safety. However, modern cars expose a variety of digital services and process a variety of personal data, at least of the driver's. This paper unfolds the argument that car (cyber-)security and drivers' privacy are worthy of additional consideration, and does so by advancing "COSCA", a framework for "COnceptualising Secure CArs" as interconnected nodes of the Next Generation Internet. COSCA adopts an innovative socio-technical approach. It crowdsources drivers' perceptions on core privacy topics and it classifies the data collected by cars and processed by manufacturers pursuant the General Data Protection Regulation. These steps inform a risk assessment which highlights the more relevant mitigation strategies and cyber security technologies. Finally, COSCA aims at designing novel interfaces to enable drivers to exercise their rights about personal data collection and processing.

Keywords: automotive; cybersecurity; framework; privacy

1 Introduction

The digital side of innovations in the automotive field has been thriving in the last decade, not just towards the traditional goal of vehicle and passengers' safety. The inside of cars increasingly integrates infotainment, passengers' physical preferences, such as on mirror positions, seating configuration and air conditioning setup, as well as non-physical preferences, such as on music to play, preferred numbers to call and on-line payment details, and driver generated data such as driving style and location. The outside perhaps is even more varied, with a number of vehicle-to-vehicle and vehicle-to-infrastructure new application scenarios, and many technologies to support them at various layers, such as Dedicated Short-Range Communications (DSRC) and Wireless Access in Vehicular Environment (WAVE).

While autonomous driving is a compelling ultimate aim, it is not the only one, as it is clear that cars are taking the shape of a unique hub offering drivers and, potentially also

[1] Dipartimento di Matematica e Informatica, Università degli Studi di Catania, Catania, Italy. giamp@dmi.unict.it
[2] Dipartimento di Matematica e Informatica, Università degli Studi di Catania, Catania, Italy. pietro.biondi@phd.unict.it
[3] Istituto di Informatica e Telematica, Consiglio Nazionale delle Ricerche,Pisa, Italy. gianpiero.costantino@iit.cnr.it
[4] Istituto di Informatica e Telematica, Consiglio Nazionale delle Ricerche,Pisa, Italy. ilaria.matteucci@iit.cnr.it
[5] Dipartimento di Ingegneria "Enzo Ferrari", Università degli Studi di Modena e Reggio Emilia, Modena, Italy. mirco.marchetti@unimore.it

passengers, a novel or reshaped range of digital services through an integrated user interface and a *sui generis* user experience. For example, cars come with Internet connectivity (via an embedded chipset or a dedicated SIM card) and can download over-the-air updates from the manufacturer. They can also provide services through an app that the driver runs on their smartphone to remotely operate car functions like power doors, sunroof, air conditioning, headlight, horn and even engine start/stop. Additionally, the app enables the driver to remotely locate the car (through the car's onboard GPS) and remotely operate geo-fencing, namely to set an area on a map and be alerted should the car ever exceed that area [AC20].

It becomes fair to claim that cars are progressively resembling computers, hence offering digital services while treating a variety of passengers' personal data and, consequently, attracting various malicious aims.

In this context, we strongly argue that the goal of increasing vehicle safety must be paired up with a similar increase in car security, drivers' privacy and, ultimately, overall passengers' trust in the technologies operating their cars. COSCA is an NGI TRUST H2020 project [CO20] founded on the European forefront of safety and privacy standards for the "driver-vehicle" system in the landscape of connected vehicles [Un20]. COSCA leverages the relevant contributions of the ETSI Cyber [ET20], Human Factors and Intelligent Transport Systems technical committees, in addition to Enisa's recommendations on IoT and Smart Infrastructures [EN20]. The project is also based on major international standards such as the ISO/IEC 27000 series [In18].

The main aim of the COSCA project is to define the COSCA framework for "COnceptualising Secure CArs", which revolves around four main activities and combines them innovatively. The first is an understanding of how drivers and passengers feel about their privacy and what level of trust they pose in their vehicle. The second is a classification of the types of data collected by car manufacturers according to Regulation (EU) 2016/679, the "General Data Protection Regulation" (GDPR) [Eu16]. The third is a risk assessment exercise based on ISO/IEC 27000 series geared towards risks for car safety and drivers' privacy. The fourth and final element is the set of measures for car security such as security protocols and their threat models, as well as the human-computer interfaces to enable drivers to consciously use the systems.

This paper details the four activities and explains the current state of their developments, thus providing a snapshot of where the project stands when it is *half way through its lifetime*. The organisation of the prose is simple. Section 2 outlines the COSCA framework, Section 3 discusses how drivers' privacy concerns and trust perceptions can be leveraged and Section 4 explains what personal data is collected and treated by relevant car brands. This knowledge collectively informs the risk assessment exercise on car security and drivers' privacy conducted in Section 5, while the resulting measures are discussed in Section 6. Finally, Section 7 presents some related work and Section 8 concludes.

2 The COSCA framework

The four activities of the COSCA framework are outlined in this Section and expanded in the sequel of the paper.

Activity 1. *Crowd-sourcing drivers' privacy concerns and trust perceptions.* Our framework adopts a socio-technical approach because it bases its contribution on people. This is achieved by leveraging crowd-sourcing to study how real people understand and perceive car security and their own privacy while they drive. We argue that this is innovative and not yet available; for example, the mentioned works about securing communication within vehicles often fail to fully ground the technical details upon drivers' privacy concerns and trust perceptions.

Activity 2. *Classifying car collected data.* While modern infotainment systems come with the usual cumbersome privacy policies that should explain what sort of data will be collected about the driver, this information is not always crystal clear to every driver. Each car may collect specific categories of personal data, some brands may collect more or less sensitive categories and treat them somehow adequately, hence we must dig out such information and evaluate how data is treated. In consequence, the COSCA framework prescribes the analysis of what data relevant car brands collect and treat.

Activity 3. *Assessing car security and drivers' privacy risks per car brand.* The previous elements inform a risk assessment exercise inspired to an ISO/IEC methodology suitably adapted to the risks of car security and drivers' privacy. The mitigation strategies stemming from the risk assessment suggest and motivate the development of techniques such as security protocols, secure cryptographic key distribution and storage, freshness methods and related threat models. Such techniques find their precise and well grounded motivation as a mitigation of assessed risks.

Activity 4. *Devising measures for car security and drivers' privacy.* There is a stringent need to devise a common set of measures for car security at all levels, from the lowest up to the human-computer interface. Weaknesses may derive not only from the known limitations of old technologies such as the CAN Bus but also from the system's interaction with the driver. Policies should explain in simple terms the types of data that a car collects from the driver, data on which the car manufacturer may act as data Controller or Processor, perhaps also through appropriate cues.

Figure 1 shows the COSCA framework with its activities and and their interrelations. It can be noted that the first two activities fuel up the third activity. Finally, the third activity determines the fourth activity, and the framework is completed.

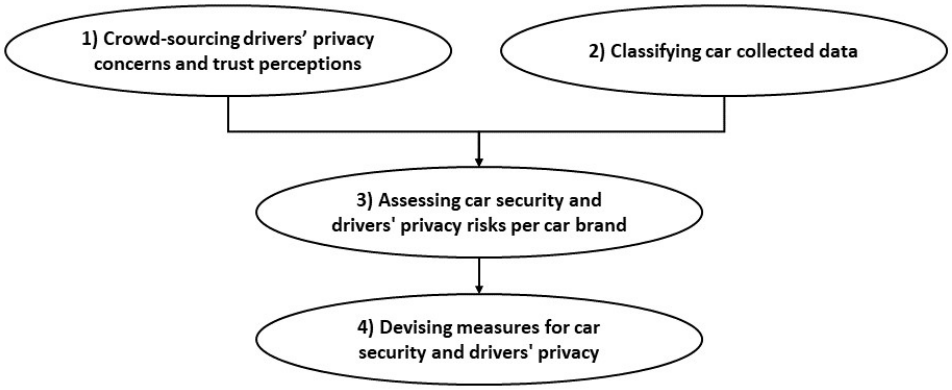

Fig. 1: The four activities in the COSCA framework

3 Crowd-sourcing drivers' privacy concerns and trust perceptions

Nowadays, there would be limited use in addressing a problem that drivers did not feel. Despite a few recent headlines on attacks to real cars [VM15], there is limited literature demonstrating how drivers feel about their privacy in their cars and what level of trust they pose e.g. in the interconnected infotainment.

This Section explains the details of activity 1, which prescribes the extraction of people's practical perceptions on (the potential endangerment to their) privacy while they are driving their cars and, at the same time, the trust they place on their vehicles. Such extraction produces a model of privacy and trust, nuanced through the lens of car drivers.

Several steps must be taken. Initially, a questionnaire must be designed with appropriate questions to distil out drivers' privacy concerns and, correspondingly, their trust perceptions. Then, the questionnaire must be submitted to the drivers through crowdsourcing. This is a fundamental step because it allows us to get answers from all over the world, and a valid platform that can be used is Prolific [Pr20]. Finally, the answers of the questionnaires must be studies through statistical analysis in order to derive meaningful correlations [Wi20a, Wi20b].

We completed this work collecting answers from a sample of 1101 respondents, which would be reliable because setting a margin of error of 4% and confidence level of 99% over a large population produces an ideal sample size of 1037 respondents for any population that exceeds 1 million [Qu20]; however, a limitation is that uniformity of distribution of respondents could not be ensured on Prolific. Relevant correlations have arisen through the answers. In particular, most of the participants agree that the systems and technologies implemented in modern cars are increasingly similar to modern computers. Another correlation shows that most drivers disagree that it is necessary for their car to collect their personal data due to the fact that they deem that collection unnecessary to the full functioning of the car. Another correlation shows that a large part of the sample thinks that their data is analysed

and studied by the vehicle systems to evaluate some personal aspects. These correlations highlight privacy concerns.

4 Classifying car collected data

History shows that a poorly designed, interconnected fridge could expose people's emails [Ne15]. A similar attack could leverage a hacked car to harvest the personal data that the car handles. We must understand and gather what types of data categories cars are collecting — and their manufacturers are treating. Notably, these may even include special categories of data, according to the GDPR, such as religious beliefs and health data.

First, we must build a knowledge base on the current categories of data collected by the major car manufacturers. According to the GDPR, each manufacturer is a data Controller and possibly a data Processor. Manufacturers are therefore responsible for the protection measures they implement on the personal data they collect. To collect this information we can use various ways, such as exercise the right of access with the manufacturer, inspecting and comparing privacy policies or empirically penetrating available cars. This work has been completed about the top ten best-selling car manufacturers of the first quarter of 2019 [Be19] extended with KIA and Tesla. Once data is obtained, we can build metrics regarding privacy policies, such as calculating the level of readability, defining taxonomies based on keywords or noting correlations between car manufacturers. In particular, if special categories under the GDPR are present, then they would require special protection measures as encryption while at rest and in transit.

5 Assessing car security and drivers' privacy risks per car brand

In all areas of cybersecurity, risks must not be presupposed, they must be assessed. This is especially true for risks affecting car security and drivers' privacy. A structured, standard risk assessment methodology ought to be used to qualitatively assess such risks, then prioritise them and ultimately mitigate them.

The third activity aims to conduct a cybersecurity risk assessment exercise geared at assessing the risks of car attacks and breaches of drivers' personal data. Leveraging the outcomes of activity 1 and activity 2, a standard methodology must be tailored to the aims of this activity, yielding the Driver Risk Assessment Methodology (DRAM). This methodology must be then used to conduct the Driver Risk Assessment Exercise (DRAE). In particular, the likelihood parameter will be set considering the outcomes of activity 1, and the impact parameter will be determined by the outcomes of activity 2. Finally, appropriate driver risk mitigation strategies (DRMSs) must be defined, which may include technical measures as well as socio-technical measures to increase people's understanding and ultimate compliance with the technical measures, such as gamification-enhanced user experiences and specific human training activities.

This work is half-way through. The DRAM is ready, based on ISO/IEC 27005:2018 [In18], also intertwined with other approaches:

- the STRIDE approach [Mi20], to tailor the general security threats to the specific identified assets;
- the threats outlined in premise no.75 of GDPR, such as involvement of "vulnerable natural persons", to strengthen the account for data protection threats;
- the Guidelines on Data Protection Impact Assessment (DPIA) by the European Article 29 Data Protection Working Party [Eu19], to capture threats specifically inspired to a DPIA exercise.

We have started the DRAE exercise but it is yet to be completed over the twelve car brands, so the DRMSs are yet to be devised.

6 Devising measures for car security and drivers' privacy

The fourth activity concludes the conceptual development of the COSCA framework for securing modern cars. Inspired by the mitigation strategies produced by the risk assessment exercise from activity 3, this activity covers the technical measures needed for securing in-vehicle communications and all personal data, at rest or in transit, also extending to the human-computer interface and overall user experience.

The first step to take here leads to the relevant technical measures through a systematic classification of all necessary technical elements, such as security protocols, secure cryptographic key distribution and storage, freshness and threat models, followed by the abstract design of their possible combinations. The second step produces the expected socio-technical measures by conducting a systematic classification of interface design methodologies followed by the abstract definition of an appropriate car interface that leverages hypermedia, machine readable privacy terms and GDPR compliance.

This work is still in its early phases. While there is vast literature related to the techical aspect of cyber security and privacy, the design of user-empowering human-computer interfaces appears to be generally neglected. We will tackle this task by conducting an abstract analysis of the user experience by means of tools such as Cognitive Walkthrough. This will guide the use of audiovisual cues to enable drivers to consciously express their informed consent to data processing and thus exercise their rights.

7 Related Work

The COSCA framework embraces the key aspects of cybersecurity and privacy in the automotive field, hence it is convenient to organise the related work along the four activities that the framework combines.

7.1 Crowd-sourcing drivers' privacy concerns and trust perceptions

In 2014, Schoettle and Sivak [SS14] surveyed public opinions in Australia, the United States and the United Kingdom regarding connected vehicles. Their research noted that people (drivers as well as non-drivers) expressed a high level of concern about the safety of connected cars, which does not seem surprising on the basis of the novelty of the concept at the time. However, participants demonstrated an overall positive attitude towards connected car technology, with particular interest in device integration and in-vehicle internet connectivity.

Moreover, in 2016, Derikx et al. [De16] investigated whether drivers' privacy concerns can be compensated by offering monetary benefits. They analysed the case of usage-based auto insurance services where the rate is tailored to driving behaviour and measured mileage and found out that drivers were willing to give up their privacy when offered a small financial compensation.

These works are only loosely related to the crowd-sourcing activity conducted within the COSCA framework because they do not significantly contribute to understanding how drivers "feel" about the security of the cars they drive as well as about how their personal data are treated onboard. It is worth remarking that these matters apply to all modern, interconnected cars. COSCA does not assess the possible inter-relation with incentives such as monetary rewards.

7.2 Classifying car collected data

As mentioned above, modern vehicles generate, collect and share a variety of data. Such data can be often associated with a physical person and, therefore, qualify as personal data and must be treated in full compliance with the GDPR.

We have noticed that there is scant literature focusing on protecting drivers' data. The "CANDY" attack reconfirms how data can be stolen following security weaknesses, which can derive from optimistic network isolation assumptions made at application layer [G.18].

A few works emphasise the overarching problem of how to effectively transmit the contents of a lengthy policy to people. Those wishing to use a service routinely accept the terms of

the service provider without fully understanding them [PLM19]. However, privacy policies in this context are no exception and remain difficult to understand [Wi16].

The literature recalled above reinforces the motivation for the data classification activity conducted within the COSCA framework because no comparative analysis of car manufacturer's privacy policies is available.

7.3 Assessing car security and drivers' privacy risks per car brand

The increasing complexity and connectivity of cars raise the necessity of conducting a security and privacy risk assessment exercise. It is clear that car manufacturers are best positioned to do that exercise because they are fully knowledgeable on the technologies embedded in their cars.

While best practices and recent standards mandate risk assessment processes (e.g. ISO 26262), their outcomes are not publicly available, hence the motivation to do the exercise at end-user level. This is the core of the risk assessment activity within the COSCA framework, which intends to rely solely upon information that the layman can obtain. Therefore, this COSCA activity may only rely on analysing available privacy policies and on conducting empirical tests on how cars operate.

In consequence, an account on the security risks in combination with those on the freedom of people deriving from how their personal data is treated is an original contribution of COSCA.

7.4 Devising measures for car security and drivers' privacy

Recent literature comprises many examples of attacks against modern vehicles, as well as possible defences. Rather than proposing novel defensive techniques, COSCA aims at generalising the main technical countermeasures that are required to protect user's data collected, transmitted and processed by modern vehicles. We remark that this activity includes the design of user-empowering interfaces specifically tailored to the automotive domain. To the best of our knowledge, this aspect has not been analysed before.

8 Conclusion

Modern cars are increasingly interconnected systems, both internally and to an external environment formed by other cars and dedicated infrastructures. Cars also interact more and more tightly with its users, mainly with drivers but sometimes also with passengers. However, users are usually given cumbersome security and privacy policies that should

explain how a car is secured and what types of data it collects about those who use it. We recognise such policies as the essential means that car manufacturers leverage to inform people. While this currently has a variety of limitations in conveying information to people, we conjecture that clarity and effectiveness of privacy policies may become a valuable asset to influence people's choices of a particular car brand in the future.

In this paper we have presented the COSCA framework that conceptualises secure and privacy preserving cars through the execution of four main activities.

These involve studying whether drivers have privacy concerns and whether they trust the treatment of their personal data by the manufacturers of the cars they drive. Such data must be comparatively understood across a number of relevant manufacturers (ideally all manufacturers) in light of the GDPR. The activities continue with a risk assessment exercise aimed at assessing the cybersecurity of cars from a traditional IT point of view as well as the privacy of drivers from a GDPR perspective. The final activity pertain to socio-technical measures that are motivated by the risk assessment and therefore concerns the design of driver empowering interfaces.

The project is currently underway and we expect the project to contribute to improving the user perception of the drivers by making them aware of the data that is collected, of how they are protected. Furthermore, by defining the Human-Machine Interface we want to allow drivers to exercise their rights over the data processed by the car.

Acknowledgement

This work has been supported by the COSCA research project [CO20] (NGI TRUST 2nd Open Call (ref: NGI TRUST 2019002)).

Bibliography

[AC20] ACKO: , Connected Cars: What is it? Features and Benefits. https://www.acko.com/car-guide/connected-cars-features-benefits/, 2020.

[Be19] Bekker, Henk: , Q1/2019 Europe: Best-Selling Car Manufacturers and Brands. https://www.best-selling-cars.com/europe/q1-2019-europe-best-selling-car-manufacturers-and-brands/, 2019.

[CO20] COSCA Team: , COnceptualising Secure CArs (COSCA) Website. https://cosca-project.dmi.unict.it/, 2020.

[De16] Derikx, Sebastian et al.: Can privacy concerns for insurance of connected cars be compensated? Electronic Markets, 2016.

[EN20] ENISA: , ENISA Good practices for IoT and Smart Infrastructures. https://www.enisa.europa.eu/topics/iot-and-smart-infrastructures/iot/good-practices-for-iot-and-smart-infrastructures-tool, 2020.

[ET20] ETSI: , European Telecommunications Standards Institute. https://www.etsi.org/, 2020.

[Eu16] European Union: , General Data Protection Regulation (EU Regulation 2016/679). https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L:2016:119:FULL, 2016.

[Eu19] European Union: , Guidelines on Data Protection Impact Assessment (DPIA). https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611236, 2019.

[G.18] G. Costantino et al.: CANDY: A Social Engineering Attack to Leak Information from Infotainment System. In: IEEE 87th VTC. 2018.

[In18] International Organization for Standardization: , Information technology — Security techniques — Information security risk management. https://www.iso.org/standard/75281.html, 2018.

[Mi20] Microsoft: , The STRIDE Threat Model. https://docs.microsoft.com/en-us/previous-versions/commerce-server/ee823878(v=cs.20), 2020.

[Ne15] Neagle, Colin: , Smart refrigerator hack exposes Gmail login credentials. https://www.networkworld.com/article/2976270/smart-refrigerator-hack-exposes-gmail-login-credentials.html, 2015.

[PLM19] Pardo, Raúl; Le Métayer, Daniel: Analysis of Privacy Policies to Enhance Informed Consent. In: Data and Applications Security and Privacy XXXIII. Springer, 2019.

[Pr20] Prolific: , Prolific platform. https://www.prolific.co/, 2020.

[Qu20] Qualtrics: , Determining sample size: how to make sure you get the correct sample size. https://www.qualtrics.com/experience-management/research/determine-sample-size/, 2020.

[SS14] Schoettle, B.; Sivak, M.: A survey of public opinion about connected vehicles in the U.S., the U.K., and Australia. In: ICCVE. 2014.

[Un20] Union, European: , Guidelines 1/2020 on processing personal data in the context of connected vehicles and mobility related applications. https://edpb.europa.eu/sites/edpb/files/consultation/edpb_guidelines_202001_connectedvehicles.pdf, 2020.

[VM15] Valasek, Chris; Miller, Charlie: , Remote Exploitation of an Unaltered Passenger Vehicle. http://illmatics.com/Remote%20Car%20Hacking.pdf, 2015.

[Wi16] Wilson, Shomir and Schaub, Florian et al.: The Creation and Analysis of a Website Privacy Policy Corpus. ACL, Berlin, Germany, 2016.

[Wi20a] Wikipedia: , Pearson Product-Moment Correlation. https://en.wikipedia.org/wiki/Pearson_correlation_coefficient, 2020.

[Wi20b] Wikipedia: , Spearman's rank correlation coefficient. https://en.wikipedia.org/wiki/Spearman%27s_rank_correlation_coefficient, 2020.