Privacy in Connected Vehicles: Perspectives of Drivers and Car Manufacturers[★]

Andrea Fieschi[2], Yunxuan Li[1], Pascal Hirmer[1], Christoph Stach[1], and Bernhard Mitschang[1]

[1]IPVS, University of Stuttgart, Stuttgart, Germany
{firstname.lastname}@ipvs.uni-stuttgart.de
[2]Mercedes-Benz AG, Stuttgart, Germany
{firstname.lastname}@mercedes-benz.com

Abstract. The digital revolution has led to significant technological advancements in the automotive industry, enabling vehicles to process and share information with other vehicles and the cloud. However, as data sharing becomes more prevalent, privacy protection has become an essential issue. In this paper, we explore various privacy challenges regarding different perspectives of drivers and car manufacturers. We also propose general approaches to overcome these challenges with respect to their individual needs. Finally, we highlight the importance of collaboration between drivers and car manufacturers to establish trust and achieve better privacy protection.

Keywords: Connected Vehicles · Privacy · Anonymization.

1 Introduction

Connected Vehicles (CVs) are a revolutionary advancement in the field of transportation that combines traditional vehicles with modern technology to enhance their capabilities. CVs are vehicles that are equipped with modern applications (apps) and are capable of accessing the internet, collecting and processing real-time data from multiple sources, and interacting with their external environments [4]. With these capabilities, CVs have become a significant source of data extraction, providing insights into driving behavior, vehicle performance, and other valuable data points. While these vehicle data can be useful for achieving autonomous driving or providing personalized services to drivers, they also contain sensitive information that could potentially identify the driver. Hence, privacy protection has become an emerging concern in the automotive industry.

In domains such as IoT and smartphones, privacy protection solutions are available. Nonetheless, Connected Vehicle Environments (CVEs) possess specific characteristics that need to be taken into account [16]. The solution proposed in other domains can be used as inspiration but not directly translated to the CV domain. While CVs can communicate with various entities in CVEs, such as roadside units, this paper focuses on privacy protection issues regarding data exchange between vehicles and the cloud. For

[★] Supported by SofDCar (19S21002), which is funded by the German Federal Ministry for Economic Affairs and Climate Action, Mercedes-Benz AG, GSaME.

instance, car manufacturers collect data in order to provide services. This data exchange and its privacy implications are at the center of attention in our discussion.

In our previous work Fieschi et al. [8], we explored the significance of privacy in CVEs, listing examples of data collection use cases, e.g., battery improvement, live traffic monitoring, and "pay how you drive" car insurance. However, stakeholders, such as individual drivers and car manufacturers, hold varying interests in privacy protection in CVEs. While drivers have a vested interest in protecting their personal information, such as location and driving habits, car manufacturers seek to improve their products through the analysis of privacy-protected data.

In this paper, we analyze the current situation, outline a first approach to overcome the discussed privacy challenges while accommodating the individual needs of both parties, and define the ground for future research. This paper explores the key privacy challenges in the CVE from the perspectives of drivers and car manufacturers, which are discussed in Sect. 2 and Sect. 3, respectively. Thus, in Sect. 4, we outline the general requirements for cooperation and building trust between drivers and car manufacturers. Finally, we summarize the paper and give an outlook on future work in Sect. 5.

2 Privacy from the Driver's Perspective

Based on domain expert discussions, we have derived a privacy attack model for CVEs from the driver's perspective. As depicted in Fig. 1, this model considers the underlying CV as *trusted-and-secure*. This implies that any personal data stored in the CV cannot be accessed or shared without the driver's consent, and all computations performed within the CV are secure and resilient to attempts to compromise them. However, remote services, such as applications whose computation is executed external to a CV, are considered as *honest-but-curious*. That is, these services comply with legal and driver-consent policies regarding the processing, storage, and sharing of personal data. Nevertheless, as drivers lost control of their personal data when sharing them with remote services, they still have concerns that remote services would derive sensitive information from the collected data.

2.1 Privacy Challenges for Drivers

Despite the desire of drivers to protect their personal data, their general demand is to continue utilizing as many user-dependent applications enabled by the CVs as possible, such as using navigation or fatigue detection services. To ensure the functionality of these services, certain vehicle data must be shared, such as location data for navigation services. Furthermore, sharing a greater quantity and higher quality of data allows the service provider to conduct more detailed analyses and, therefore, offer better customized services. However, the increased volume and precision of the data shared by drivers also pose greater privacy risks, as they may reveal sensitive information about their driving behavior and activities. Thus, a challenge of preserving privacy in CVEs is to balance the trade-off between privacy protection and service quality.

Another challenge of preserving privacy in CVEs is to achieve Situation-Awareness [12]. As the sensitivity of a data point is related to when and where as well as for what purpose the data is being collected [9], drivers' privacy requirements can also change when the

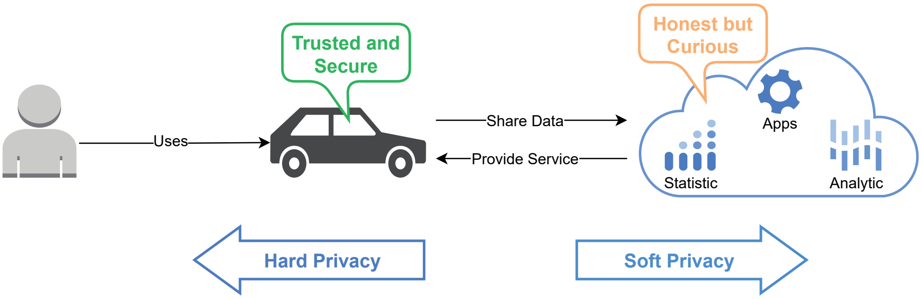

Fig. 1. Privacy Attack Model from Driver's Perspective for Car to Cloud Environment.

situation changes. For instance, drivers may agree to share their unmodified location and speed data with a data collection company for analysis purposes when they are driving adhere to traffic regulations. However, in the occurrence of an accidental speeding violation, drivers would revise their privacy requirements to hide their speeding behavior. Hence, approaches to privacy protection in CVEs must consider the dynamic and context-dependent nature of drivers' privacy needs.

Although privacy is a highly concerning issue in many domains, users often struggle to manage their privacy settings effectively. For instance, Ramokapane et al. [14] found that many smartphone users find it difficult to customize privacy features provided by their smartphone manufacturers, as they lack knowledge on how to configure them. From our research project, we noticed that the aforementioned challenge is magnified in the automobile domain since CVs typically have significantly more data sources and potential data consumers than smartphones. Consequently, managing the fine-graind and situation-aware privacy policy for a CV can easily create information and choice overhead for drivers. As a result, the difficulties in managing privacy settings would contribute to the so-called "privacy paradox" [13], where people claim to be concerned about their privacy but still share a lot of private information.

2.2 Privacy Protection Approaches for Drivers

In 2008, Danezis [6] proposed two concepts of privacy: hard privacy and soft privacy. The concept of hard privacy aims to minimize the amount of personal data shared, thereby decreasing the level of trust required between the data subject and the data collector. On the other hand, the concept of soft privacy assumes that the data subject does not have full control of their personal data and, therefore, has to trust the honesty and competence of data controllers. Under this assumption, soft privacy aims to ensure consent-based data processing through policies, access control, and audit.

As depicted in Fig. 1, we argue that both concepts are essential in preserving privacy in CVE. The general concept is to achieve hard privacy before vehicle data leave the CV while ensuring soft privacy for data that is shared with different remote services. To achieve data minimization of hard privacy, services must provide drivers with essential metadata, such as what vehicle data are collected and for what purpose. They should also support drivers in managing their privacy policies in a fine-granular manner. Based on the assumption that services are honest-but-curious, the service's metadata is considered reliable and will be used to conduct data minimization.

In accordance with the concept of hard privacy, drivers are advised to block any unnecessary data sharing for the desired service functionality based on the information provided in the services' metadata. This would provide a basic level of privacy protection against the curious nature of different services. For the data that are necessary for the computation of the desired service, data minimization can still be achieved through different approaches, such as reducing the accuracy of the vehicle data. To balance the trade-off between privacy protection and service quality, different Privacy Enhancing Technologies (PETs) [3], such as obfuscation and pseudonymization, can be utilized to distort or anonymize vehicle data so that the sensitive information is removed and the perturbed data are still precise enough to ensure service functionality. Furthermore, there is the challenge of handling scenarios where the privacy requirements of drivers may change depending on the situation. As a result, different PETs used in CVs should be developed in a modular manner, and the data processing in CVs should also support live adaptation, allowing for the dynamic integration, replacement, or removal of PETs in the vehicle's data pipeline.

To utilize service functionalities, it is inevitable that drivers have to share certain vehicle data with the corresponding service providers. As drivers no longer have control over the shared data, we can only ensure soft privacy for them. To mitigate privacy leakage risks, a Service Level Agreement (SLA) can be established between the driver and the service provider before the driver uses the service for the first time. Through the privacy section of the SLA, the service provider should enable drivers to explicitly express how their shared data can be further processed, stored, or published. However, as drivers usually do not have insight into data processing, it is important for them to receive transparent information regarding how their data is being processed by the service provider.

Additionally, to assist drivers with a basic understanding of privacy in customizing their privacy policies for CVs and managing their privacy preferences in SLAs, user-friendly privacy management mechanisms, such as the privacy context model dedicated to CVs [11], have to be developed. Overall, by adopting the concepts of hard and soft privacy, we can strike a balance between protecting drivers' privacy while still ensuring various service functionalities.

3 Privacy from the Car Manufacturer's Perspective

From the point of view of a data-collecting company, privacy protection is important for multiple reasons. Firstly, companies have an ethical obligation to ensure privacy protection for their users, thereby adhering to ethical guidelines and minimizing the risk of privacy violations. Secondly, legal compliance is crucial, as the General Data Protection Regulation (GDPR), enforced by the European Union, mandates strict restrictions and limitations on data collection to safeguard user privacy. Lastly, the implementation of robust data protection measures can be particularly appealing to customers. Prioritizing privacy and making it a core value of a company will help gain further trust with the general public and add value to its products.

3.1 Privacy Challenges for Car Manufacturers

From a data science perspective, CVs represent an immensely valuable source of data, as they allow manufacturing companies to monitor how their products perform in real-world scenarios, gain insights into usage patterns and preferences, and identify opportunities for improvement or redevelopment in the next iteration. However, it is important to note that the data collected from these vehicles can be closely linked to the behavior of the drivers. As a result, the improper use of CVs can lead to the risk of leaking personal information, such as the position of the car, their general behavior behind the wheel, and other habits that are kept inside the vehicle. It is imperative that this information remains secure and inaccessible to unauthorized parties, and if possible not linkable to a specific person, i.e., anonymized. Drivers must have the assurance that any data they choose to share will only be used to enhance their service and experience and that none of the collected information will be used against them. Therefore, manufacturers must ensure that adequate privacy measures are in place.

With regard to privacy, data collection use cases can be mainly divided in user dependent and user independent use cases [8]. These come with different and specific privacy challenges. User dependent use cases need to collect data and send information back to the same specific user, so the data need to be protected but connected to an identifiable source. User independent use cases collect data to then provide a service to entire fleets, anonymization becomes an option with the extra challenge of guaranteeing a high level of anonymity. In Fig. 2 we have a graphical representation of these two kinds of data collection.

It should be noted that the collection of data from cars raises privacy concerns, not only for the driver but also for individuals who are merely in the vicinity of the vehicle. Camera data, for instance, may include images of pedestrians that could potentially identify them and disclose their whereabouts at a specific time. Additionally, companies must ensure that they do not collect data that goes beyond what is necessary. If a driver declines to share their identity in connection with the collected data, the data collector must ensure that the driver's identity cannot be inferred by analyzing patterns in the data.

Offering strong privacy protection policies presents several challenges and difficulties for companies. Firstly, drivers must provide consent for data collection, and the data can only be used for agreed purposes. Any further analysis of the data is prohibited. Preprocessing may be necessary on board the vehicle before data transfer, and communication channels must be secure. Another challenge is the inability to use data retrospectively for unforeseen purposes, requiring new data collection campaigns and new user agreements.

3.2 Privacy Protection Approaches for Car Manufacturers

If a privacy-conscious company wishes to provide even stronger protection, additional computational and design efforts are required. Changing data collection campaigns, such as incorporating new data types to be gathered, may necessitate redesigning and new user agreements have to be stipulated with the driver, which can prolong the time needed to provide data to analysts. A significant challenge is information loss, as data privacy often comes at the cost of sacrificing some information that raw data would convey. The

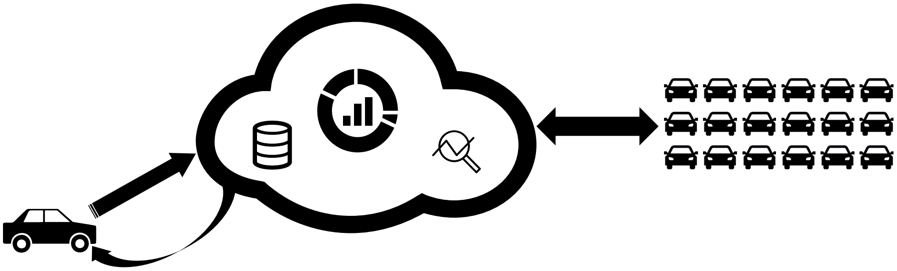

Fig. 2. Manufacturer's Perspective: Single customer-related services (left) and fleet-focused services (right).

trade-off between privacy protection and information content requires consideration, as stronger protection may necessitate relinquishing additional data.

This shows that privacy is not an element that can simply be added in hindsight as a plug-in element to the data flow chain. Privacy needs to be taken into account during the development of the data collection use case, every privacy methodology should fit in a frame of privacy by the design. From the early stage of design the developer should take into account the importance of privacy and it's implications. Important elements to consider are: the data type needed for the service and how they could be used to violate the users privacy; which amount of information loss will occur with different PETs; how the data collection could be structured differently in order to have a better ratio of quality of service and privacy protection.

A privacy-conscious car manufacturer models its data acquisition scenarios after various privacy-preserving methodologies to protect individual privacy. These methodologies include differential privacy [7], which involves perturbing the data to achieve privacy protection; federated learning [10], which processes the data on-board and collaboratively trains networks; homomorphic encryption [2], which protects the data during processing without decrypting the information; and k-anonymization [15], which groups data-points into equivalence classes of size k in order to protect the individual's identity. Each of these privacy-preserving approaches requires adaptation of the data collection scenario to meet their respective paradigms. Companies can gain enhanced user trust and competitive advantage by guaranteeing strong privacy policies. The trusted status of the company can encourage users to share more data, resulting in added value for the company, subsequently resulting in new and improved products for the customers. In summary, companies offering strong privacy protection policies must overcome numerous challenges and obstacles. However, the advantages of gaining user trust and enhanced value outweigh the challenges. Companies can achieve this by implementing robust security measures, ensuring transparency, obtaining user consent, and complying with relevant privacy regulations.

4 Discussion

Privacy protection in CVEs is a complex task, as the perspectives of drivers and car manufacturers may differ. From a driver's point of view, preserving the privacy of their personal data refers to balancing privacy protection and service quality in their privacy policies. Additionally, the dynamic and context-dependent nature of drivers' privacy needs must also be recognized. To overcome these challenges, we proposed a first approach in this paper, which ensures drivers' privacy by utilizing various PETs to achieve hard privacy before any vehicle data leave the CV and to ensure soft privacy through SLAs for data that is shared with different remote services.

However, achieving this goal requires the collaboration of remote service providers and car manufacturers. Primarily, remote services must be transparent about the data they collect and the purpose for which it will be used. Additionally, both service providers and car manufacturers should give drivers the freedom to customize their privacy preferences in a precise manner, which may include refusing requests for unnecessary vehicle data or reducing the quality of data that is necessary for the desired service functionality. In return, drivers must understand that these actions may result in a reduction in service quality.

From the car manufacturer's perspective, there is a strong interest in making privacy protection a priority and a key value of their company. That comes at their advantage since it also fulfills an ethical obligation and a legal compliance requirement to protect users' privacy. Collecting vehicles' data comes with a variegated constellation of challenges: providing high-performing services without collecting more data than necessary, implementing a privacy-preserving structure that allows guaranteeing strong privacy protection, gaining the trust of the drivers, and having them agree to share informative data about their cars.

A transparent data handling from the manufacturer needs to be matched with users willing to express their privacy requirements and understand the risks of agreeing to share data. Understanding that a very low amount of information will not allow the service to be top-notch but guarantee a stronger level of privacy is also a concept that the driver needs to understand fully; this should by any means come with the implication that top-notch services cannot guarantee privacy protection though, that always needs to be a priority. Drivers that communicate privacy preferences and well-thought-out boundaries are of highly valuable worth to a privacy-conscious car manufacturer.

Privacy protection in CVEs is not a one-sided issue. While implementing excessive PETs within CVs would compromise data quality of shared data, the scarcity of privacy protections within CVs also shifts greater responsibility to car manufacturers to meet the driver's privacy requirements. Thus, PETs used in CVs must be chosen carefully to enable privacy protection while ensuring sufficient data quality. However, there are still limited PETs available that are designed specifically for the privacy protection of CVs. To tackle this challenge, privacy mechanisms from other domains could be adapted in CVEs. For instance, the PRIVACY-AWARE concept proposed by Alpers et al. [1] for mobile devices, or the state-of-the-art PETs summarized by Curzon et al. [5] for smart cities. Nevertheless, there is still room for developing new PETs dedicated to CVs that can guarantee privacy without compromising service quality.

Car manufacturers and drivers have various challenges to overcome, various sets of requirements they need to evaluate, and the common goal of safeguarding people's privacy. Cooperation between the two parties and efficient as well as open communication about this topic is the way to be taken to improve privacy while still allowing services to become more sophisticated. In the meantime, laws and regulations governing the collection and processing of personal data should be enhanced and improved regularly to keep pace with technological advancements. With an infrastructure that allows drivers to fully express their privacy preferences without burdening them with a cumbersome task, and with transparent data handling from the data collectors' side, the potential for enhanced privacy protection and improved service performance can be greatly increased.

5 Summary and Future Work

In conclusion, privacy is a crucial factor to consider for both car manufacturers and drivers. While car manufacturers need to implement robust privacy measures to protect sensitive data collected from vehicles, drivers need to be aware of their privacy rights and take steps to safeguard their personal information. Failure to prioritize privacy can lead to severe consequences such as data breaches or loss of trust between manufacturers and customers. Therefore, it is imperative for all stakeholders to recognize the importance of privacy in the automotive industry and take appropriate measures to ensure that privacy is protected.

To better understand how car manufacturers can cooperate with drivers regarding privacy protection, we plan to conduct a user study to comprehend drivers' privacy awareness and requirements in CVEs as well as interviews with domain experts to gain insights into manufacturers' strategies and legal constraints. Furthermore, we also plan to research existing PETs specifically designed for CVEs as well as PETs utilized in other relevant domains to assess the feasibility and potential applicability of these technologies in the CVEs. This would help us identify suitable PETs for CVs that can guarantee privacy protection without compromising service quality. Overall, our research will further explore the effective approaches and mechanisms that facilitate collaboration in privacy protection between car manufacturers and drivers in CVEs.

References

[1] Alpers, S., Oberweis, A., Pieper, M., Betz, S., Fritsch, A., Schiefer, G., Wagner, M.: PRIVACY-AVARE: An approach to manage and distribute privacy settings. In: Proceedings of the 2017 3rd IEEE International Conference on Computer and Communications (ICCC), Chengdu, China, December 13–16, 2017, pp. 1460–1468, IEEE, Piscataway, NJ, USA (2017), ISBN 978-1-5090-6353-6, https://doi.org/10.1109/CompComm.2017.8322784

[2] Armknecht, F., Boyd, C., Carr, C., Gjøsteen, K., Jäschke, A., Reuter, C.A., Strand, M.: A Guide to Fully Homomorphic Encryption. IACR Cryptology ePrint Archive pp. 1192:1–1192:35 (Dec 2015)

[3] van Blarkom, G.W., Borking, J.J., Olk, J.G.E. (eds.): Handbook of Privacy and Privacy-Enhancing Technologies: The case of Intelligent Software Agents. PISA Consortium, The Hague, The Netherlands (2003), ISBN 978-90-74087-33-9

[4] Coppola, R., Morisio, M.: Connected Car: Technologies, Issues, Future Trends. ACM Computing Surveys **49**(3), 46:1–46:36 (Oct 2016), ISSN 0360-0300, https://doi.org/10.1145/2971482

[5] Curzon, J., Almehmadi, A., El-Khatib, K.: A survey of privacy enhancing technologies for smart cities. Pervasive and Mobile Computing **55**, 76–95 (Apr 2019), ISSN 1574-1192, https://doi.org/10.1016/j.pmcj.2019.03.001

[6] Danezis, G.: Introduction to Privacy Technology. Talk, COSIC / ESAT, KU Leuven (Jul 2007), URL http://www0.cs.ucl.ac.uk/staff/G.Danezis/talks/Privacy_Technology_cosic.pdf

[7] Dwork, C., Roth, A.: The Algorithmic Foundations of Differential Privacy. Foundations and Trends in Theoretical Computer Science **9**(3–4), 211–407 (Aug 2014), ISSN 1551-305X, https://doi.org/10.1561/0400000042

[8] Fieschi, A., Hirmer, P., Sturm, R., Eisele, M., Mitschang, B.: Anonymization Use Cases for Data Transfer in the Automotive Domain. In: Proceedings of the 2023 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops), Atlanta, GA, USA, March 13–17, 2023, pp. 98–103, IEEE, Piscataway, NJ, USA (2023), ISBN 978-1-6654-5382-0, https://doi.org/10.1109/PerComWorkshops56833.2023.10150357

[9] Gharib, M., Giorgini, P., Mylopoulos, J.: An Ontology for Privacy Requirements via a Systematic Literature Review. Journal on Data Semantics **9**(4), 123–149 (Dec 2020), ISSN 1861-2040, https://doi.org/10.1007/s13740-020-00116-5

[10] Konečný, J., McMahan, B., Ramage, D.: Federated Optimization: Distributed Optimization Beyond the Datacenter. CoRR: A Computing Research Repository **arXiv:1511.03575**, 1–5 (Nov 2015), https://doi.org/10.48550/arXiv.1511.03575

[11] Li, Y., Hirmer, P., Stach, C.: CV-Priv: Towards a Context Model for Privacy Policy Creation for CVs. In: Proceedings of the 2023 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops), Atlanta, GA, USA, March 13–17, 2023, pp. 583–588, IEEE, Piscataway, NJ, USA (2023), ISBN 978-1-6654-5382-0, https://doi.org/10.1109/PerComWorkshops56833.2023.10150231

[12] Li, Y., Hirmer, P., Stach, C., Mitschang, B.: Ensuring Situation-Aware Privacy for Connected Vehicles. In: Proceedings of the 12th International Conference on the Internet of Things (IoT), Delft, Netherlands, November 7–10, 2022, pp. 135–138, ACM, New York, NY, USA (2023), ISBN 978-1-4503-9665-3, https://doi.org/10.1145/3567445.3569163

[13] Norberg, P., Horne, D.R., Horne, D.A.: The Privacy Paradox: Personal Information Disclosure Intentions versus Behaviors. Journal of Consumer Affairs **41**(1), 100–126 (Mar 2007), ISSN 0022-0078, https://doi.org/10.1111/j.1745-6606.2006.00070.x

[14] Ramokapane, K.M., Mazeli, A.C., Rashid, A.: Skip, Skip, Skip, Accept!!!: A Study on the Usability of Smartphone Manufacturer Provided Default Features and User Privacy. Proceedings on Privacy Enhancing Technologies **2019**(2), 209–227 (Dec 2018), ISSN 2299-0984, https://doi.org/10.2478/popets-2019-0027

[15] Samarati, P., Sweeney, L.: Generalizing Data to Provide Anonymity When Disclosing Information (Abstract). In: Proceedings of the Seventeenth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems (PODS), Seattle, Washington, USA, June 1–4, 1998, p. 188, ACM, New York, NY, USA (1998), ISBN 978-0-89791-996-8, https://doi.org/10.1145/275487.275508

[16] Wang, H., Liu, T., Kim, B., Lin, C.W., Shiraishi, S., Xie, J., Han, Z.: Architectural Design Alternatives Based on Cloud/Edge/Fog Computing for Connected Vehicles. IEEE Communications Surveys & Tutorials **22**(4), 2349–2377 (Sep 2020), ISSN 2373-745X, https://doi.org/10.1109/COMST.2020.3020854

All links were last followed on July 27, 2023.