Driving for Big Data? Privacy Concerns in Vehicular Networking

David Eckhoff *Member, IEEE* and Christoph Sommer *Member, IEEE*

1 Introduction

For the first time, the 2013 LA Auto Show hosted a dedicated Connected Car Expo where manufacturers showed off their plans for upcoming networked cars. This underlines that vehicular networks, created from communicating smart vehicles, have ceased to be a vision and are now rapidly becoming a reality. The ability of these vehicles to communicate with each other and/or the infrastructure using short range wireless technology is envisioned to improve road safety and decrease the number of traffic fatalities, and also offer comfort services such as parking spot assistance, wireless payment, or traffic light assistance systems.

However, the ever present dark side of vehicular networks is often ignored: They accumulate enormous amounts of private user data such as detailed location information, and due to their decentralized nature, a large portion needs to be broadcast wirelessly – for everybody to hear. This can be exploited by operators, other drivers, or arbitrary people, be it for profit, surveillance, or overly restrictive law enforcement. The situation becomes worse when, mandated by governmental institutions, providers are unable to install privacy protection measures and are forced to disclose user data.

The difference of vehicular networks to many of today's offered services and systems is that turning off the device to preserve your privacy to a certain degree will no longer be possible as they might become legally mandated. Moreover, one of the benefits of these networks is safety, something that most users will probably value higher than even their privacy.

The standardization process for vehicular network technology is advancing and we observe an alarming tendency that effective privacy protection is not an integral part and is also often neglected in field trials. In a worst case scenario, this might lay the groundwork for the creation of the *transparent driver* and also for building the infrastructure for a surveillance society. With current versions of European (ETSI ITS G5) and U.S. (IEEE WAVE) standards in mind we want to pessimistically discuss possible privacy and traffic surveillance issues in future vehicular networks, outline research directions that could address these problems, and identify open challenges in these efforts.

2 Privacy in Vehicular Networks

For a vehicular network to be successful, it needs to be secure, enforcing authentication, authorization, and accounting. Naturally, this requires identifiers and conflicts with the drivers' desire to enjoy privacy. The necessity for privacy protection in vehicular networks has certainly been understood from the very beginnings – and position papers have been published in this same magazine as far back as in 2004 [1]. Here Hubaux et al. proposed the use of Electronic License Plates, that is, pseudonymous identifiers which could only be resolved to a driver's real identity by law enforcement.

The basic need for privacy protection is also recognized by the two most prominent families of standards for short range radio vehicular networks, namely the European ETSI ITS G5 and the U.S. IEEE WAVE. In general it can be said that the ETSI standards cover privacy aspects more detailed. Still, while ETSI 102731-v1.1.1 notes that "No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence", concrete protection measures are either absent, insufficiently precise, or not effective.

To meet the specific security and privacy challenges of vehicular networks, both IEEE 1609.2-2013 and ETSI 102941-v1.1.1 describe the deployment of the following Public Key Infrastructure (PKI) variant: Vehicles have one pre-installed certificate, the *base identity*, which is only used to request pseudonyms from a (possibly governmental) Certificate Authority (CA). A pseudonym is a certificate itself and only valid when (directly or through a chain) signed by the CA. Each vehicle uses a pseudonym to sign and send messages over the wireless channel. A receiving vehicle will only accept a message if it has been signed with a valid pseudonym.

- *D. Eckhoff is with the Department of Computer Science, University of Erlangen, Germany;*
  C. Sommer is with the Institute of Computer Science, University of Innsbruck, Austria.
  E-mail: eckhoff@cs.fau.de, sommer@ccs-labs.org

- *You are reading the authors' pre-print version. The published version has undergone extensive copy-editing to fit a smaller format. It can be found on IEEE Xplore, under DOI 10.1109/MSP.2014.2*

Yet, the use of a pseudonym instead of a real identity alone does not solve one of the central threats to drivers' privacy: Even if drivers cannot immediately be identified, they can be re-identified. By picking up their signed wireless broadcasts at different locations their movement can easily be tracked. Even without additional knowledge, this allows their true identity to be revealed [2]. Current research therefore proposes that vehicles keep a pool of pseudonyms.

While it would be beneficial for the anonymity of a driver to use a different pseudonym for each message, this would very likely confuse safety applications of other vehicles. Therefore, pseudonyms are only changed according to certain pseudonym changing strategies. Unfortunately, the IEEE and ETSI standards do not list or suggest any pseudonym changing strategy, but only mention the need to "use a pseudonym that cannot be linked to [...] the user's true identity" (ETSI 102893-v1.1.1) and suggest to change it frequently "[...] to avoid simple correlation between the pseudonym and the vehicle" (ETSI 102940-v1.1.1). A common approach in field trials is to change the pseudonym from time to time to prevent the linking of messages with different pseudonyms. Unfortunately this was already shown to be ineffective [3] if the change (that is, two consecutive messages with a different pseudonym) was overheard by an attacker. Countermeasures thus include silent times after each pseudonym change [4] or only changing the pseudonym when this is believed to cause confusion for an attacker [5]. However, these approaches can possibly interfere with traffic safety applications and are therefore unlikely to be deployed.

Independent of pseudonymous identifiers, both ITS G5 and WAVE mandate the periodic (1 Hz–10 Hz) broadcasting of unencrypted awareness messages, which contain identifying information. ETSI 302637-20-v1.3.0 Cooperative Awareness Messages (CAMs) and SAE J2945.1-2.2 Basic Safety Messages (BSMs) include the vehicle's current direction, position, speed, and acceleration. Furthermore, messages to inform other vehicles of road hazards include fields or even sequence numbers that allow the re-identification of vehicles and "[...] may be problematic with respect to privacy protection" (ETSI 102893-v1.1.1). Although the standards mark some message fields as *optional*, the decision whether to include them will not be made by the driver but by the on-board unit. It is an open challenge to identify how often and what additional data must be transmitted to still allow proper operation of safety applications without making vehicles unique.

ETSI 102893-v1.1.1 states that tracking can be prevented by either the use of pseudonyms or by sending encrypted messages. We deem these statements a dangerous simplification: First, they do not hold when pseudonym changes can be tracked, or the path traveled using a single pseudonym can be related to home or work addresses [2]. Second, the use of encryption can alleviate but never solve tracking, as it only blinds identifying information to non-participating entities. Yet, periodic awareness messages, the basis of many envisioned applications, need to be readable by everyone and are thus sent unencrypted.

Aside from live tracking, there is a second dimension to this problem: It is frequently necessary to revoke all pseudonyms of a vehicle, e.g., when it is (deliberately or unintentionally) sending false messages or when it is sold. Publishing a list of all or many pseudonyms belonging to one vehicle will retrospectively reveal location information of the driver. A possible solution for this is to only disclose current and future pseudonyms of a vehicle [6]. Although the responsible IEEE 1609.2-2013 standard acknowledges the need for privacy protection, this issue is not addressed.

3 Automated Surveillance

Pseudonyms are the most important privacy measure in future Intelligent Transportation Systems (ITS), but even if they cannot be linked to each other, the problem remains that each pseudonym can still be resolved to a base identity by the authority that signed it, meaning that complete privacy cannot be achieved. This ensures accountability: it allows the identification of vehicles that send false messages, the recovery of stolen vehicles, or the detection of hit-and-run offenses – but it could also change traffic supervision as we know it.

A vehicle that continuously broadcasts its current velocity will also do so when the driver is speeding. These messages could be received by provider operated Roadside Units (RSUs) for automated ticketing or (if no RSU is in direct transmission range) be collected by other vehicles and forwarded later on. With similar ease, most other traffic offenses (red light running, improper passing, no turn signaling, etc.) can be detected by examining the *exteriorLights*, *pathHistory*, or *steeringWheelAngle* field of one or few CAMs or BSMs.

Although there are approaches to prevent the resolving of pseudonyms (pseudonym swapping, blind signatures), they will not be deployed as they are not "[...] supporting law enforcement access under appropriate circumstances" (IEEE 1609.2-2013). It is therefore of utmost importance to lawfully control *when* pseudonyms can be resolved, for example by requiring the cooperation of multiple institutions. Although ETSI 102941-v1.1.1 recommends separating base identifier assignment from signing of pseudonyms, this would not offer additional privacy protection when a third party is able to access all data arbitrarily.

We are well aware that from today's view such a scenario certainly seems far-fetched; however, currently envisioned ITS give the operator or the government the ability to deploy these or similar "features" at will. Once on-board units are widely deployed or even legally mandated, the penetration rate of equipped vehicles will increase, making this kind of traffic supervision far more interesting for certain stake-holders.

4 Outlook

Communicating vehicles will change road traffic as we know it and help create Intelligent Transportation Systems (ITS). The fact that this technology is mostly beneficial is beyond controversy; however, certain implications of such a system raise concerns.

Current proposals for ETSI ITS G5 and IEEE WAVE can severely compromise drivers' privacy. In a worst case scenario, the envisioned systems can be instrumented to deploy a fully automated traffic surveillance system. On the plus side, both families of standards are currently under development. Thus, there is still time for both academia and industry to work toward an integration of applicable privacy measures before the roll-out phase.

To achieve this, three things are needed. First, we need to fully understand how privacy provisions affect other applications such as safety or comfort, allowing us to draw a reasonable line at the amount and accuracy of information included in periodic safety messages. For this, we need to be able to measure the effectivity of privacy protection in such simulations that we already use for performance evaluations. Finally, to convince decision makers to employ certain privacy measures we require easy-to-understand and meaningful privacy metrics – unlike the ones currently used in vehicular network privacy research.

This is the time to put a stronger emphasis on privacy in on-going standardization efforts, putting in place measures for the technical protection of users' location information. Retrofitting privacy is bound to fail; therefore field trials all over the world should understand privacy as an integral part to serve as an actual example for future implementations.

References

[1] J.-P. Hubaux, S. Čapkun, and J. Luo, "The Security and Privacy of Smart Vehicles," *IEEE Security and Privacy*, vol. 2, no. 3, pp. 49–55, May 2004.

[2] P. Golle and K. Partridge, "On the Anonymity of Home/Work Location Pairs," in *7th International Conference on Pervasive Computing*, vol. LNCS 5538. Nara, Japan: Springer, May 2009, pp. 390–397.

[3] B. Wiedersheim, Z. Ma, F. Kargl, and P. Papadimitratos, "Privacy in Inter-Vehicular Networks: Why simple pseudonym change is not enough," in *7th IEEE/IFIP Conference on Wireless On demand Network Systems and Services (WONS 2010)*, Kranjska Gora, Slovenia, February 2010.

[4] L. Huang, K. Matsuura, H. Yamane, and K. Sezaki, "Enhancing Wireless Location Privacy Using Silent Period," in *IEEE Wireless Communications and Networking Conference (WCNC 2005)*, New Orleans, LA, March 2005.

[5] M. Gerlach and F. Güttler, "Privacy in VANETs Using Changing Pseudonyms - Ideal and Real," in *65th IEEE Vehicular Technology Conference (VTC2007-Spring)*, Dublin, Ireland, April 2007, pp. 2521–2525.

[6] J. J. Haas, Y.-C. Hu, and K. P. Laberteaux, "Efficient Certificate Revocation List Organization and Distribution," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 3, pp. 595–604, March 2011.