Before the
Federal Trade Commission and
National Highway Traffic Safety Administration

Request for Comments re: Connected Cars Workshop, Project No. P175403

Comments of Electronic Frontier Foundation
May 1, 2017

Submitted by:

Jeremy Gillula, Ph.D.
Seth Schoen
Lee Tien
Jamie Williams
Electronic Frontier Foundation
815 Eddy Street
San Francisco, CA 94109
Telephone: (415) 436-9333
jeremy@eff.org

RE: Connected Cars Workshop and P175403

The Electronic Frontier Foundation (EFF) submits the following comments in response to the Federal Trade Commission and National Highway Traffic Safety Administration's call for comments in advance of its Connected Cars Workshop to be held on June 28, 2017. EFF is a member-supported, nonprofit, public interest organization dedicated to protecting privacy, civil liberties, and innovation in the digital age. Founded in 1990, EFF represents tens of thousands of dues-paying members, including consumers, hobbyists, computer programmers, entrepreneurs, students, teachers, and researchers. EFF and its members are united in their commitment to ensuring that news technologies—including connected vehicle technologies—do not undermine privacy and security.

Internet-connected vehicles implicate serious privacy and security concerns—concerns which in turn threaten physical safety. For the purposes of these comments, EFF restricts its attention to the Dedicated Short Range Communication (DSRC) Vehicle-to-Vehicle (V2V) system NHTSA proposed in its Notice of Proposed Rulemaking (NPRM), Docket No. NHTSA-2016-0126—technology that is referred to in the FTC's detailed notice regarding the forthcoming workshop. EFF also restricts its attention to three specific questions about the privacy and security of connected vehicles that we are best positioned to answer—Questions 2, 6, and 8.

<u>Question 2</u>: How do these vehicles integrate data into their functionality? How do consumers benefit from the collection and use of their information?

In its NPRM, NHTSA proposes a standard that would require vehicles to transmit non-identifying information such as their position, acceleration, and the status of various vehicle functions (brakes, traction control, stability control, etc.). 82 FR 3898. The standard would also require vehicles to transmit information that could be used to uniquely identify and track them, including vehicle dimensions and path history for the past 300m. Finally, these Basic Safety Messages (BSMs) would be signed with a cryptographic key, attested as belonging to a vehicle by a certificate. Per the NPRM, NHTSA proposes a system of rotating certificate credentials, whereby each V2V-enabled vehicle would have 20 valid certificates per week, which would change at least once every 5 minutes.

Theoretically, this information is intended only to be used for applications related to vehicle safety—*e.g.*, so that a vehicle can warn its driver when a collision might occur, or so that infrastructure like a stoplight could optimally control its timing. Thus, the *theoretical* consumer benefit is increased safety.

In practice, however, this information could—*and will*—be used by a variety of malicious actors to violate the privacy, security, and safety of drivers. Indeed, the privacy and security concerns implicated by V2V technology, as outlined herein and in EFF's comments on NHTSA's NPRM, Docket No. NHTSA-2016-0126[1]—threaten to undermine any safety benefit the technology promises. It is imperative that these concerns be addressed.

<u>Question 6</u>: What privacy and security issues might arise from consumer operation of connected vehicles, including use of third-party aftermarket products that can plug into vehicle diagnostic systems, geolocation systems, or other data-generating aspects of connected vehicles?

As we explained in EFF's comments on NHTSA's NPRM, NHTSA's "proposal, while well intentioned, will not actually protect privacy. Vehicles transmitting V2V communications will still be trackable. Specifically, NHTSA has failed to adequately account for the need to protect privacy against *systematic attempts* that will undoubtedly be made to monitor and record BSMs for the purpose of tracking vehicles.

"Even assuming rotating certificates were an appropriate approach…rotating through a mere 20 different identities, every 100 minutes, over the course of one week will not protect a vehicle's privacy. While a human being might find it confusing to remember 20 different identities for the same vehicle, it would be straightforward for a computer to analyze data

[1] *See* EFF Comments to NHTSA re V2V Notice of Proposed Rulemaking (Apr. 12, 2017), https://www.eff.org/document/eff-comments-nhtsa-re-v2v-notice-proposed-rulemaking.

collected via a sensor network and identify a vehicle over the course of one day—including associating the full set of certificates assigned to the vehicle.

"After a sensor network has determined the identity of a vehicle over the course of a day—via its 20 rotating certificates—it would be able to immediately identify the vehicle for the remainder of the week. The vehicle would be completely deanonymized for the course of that week, and for the corresponding week in any subsequent year. The sensor network would merely have to complete the same process every week, but this would be feasible given the straightforward nature of the process. And because 'human mobility traces are highly unique,'[2] it would be easy, in the case of a vehicle used in its typical way, to recognize and track a vehicle from week to week, even as the vehicle's list of 20 assigned certificates changed. Indeed, a 2009 study by the Palo Alto Research Center (PARC) showed that 5% of Americans—*i.e.*, more than 15 million people—could be uniquely identified by simply pairing data regarding their home and work areas.[3] And this does not even take into account other locations that people routinely or habitually visit—such as schools, daycare facilities, yoga studios, or grocery stores—or patterns regarding the times at which they visit these locations. [By] combining vehicle location history aggregated over time with other information or data sources (such as databases tracking employment and home addresses), it will likely even be possible to identify who exactly is behind the wheel.

"Recent computer science research has achieved remarkable success at classifying and recognizing entities from noisy data, readily finding the nearest matches even when data sets are not precisely identical. This has often enabled practical deanonymization of pseudo-identifiers—an entire field of research in its own right. In an age of machine learning and artificial intelligence, NHTSA should be aware that computer algorithms will be able to quickly and easily analyze V2V data to track vehicles over time—notwithstanding attempts to prevent this with key rotation. While this form of deanonymization [may not be 100% accurate], early attempts at deanonymizing location-related data sets have shown striking success."

The proposed standard also fails to account for how trivial it will be for a systematic tracking system to link together messages transmitted by the same vehicle just before and just after it switches to a new certificate. This is because BSMs will include vehicle dimensions as well as the GPS trace of each vehicle for at least the last 300m. Given the wide variety of vehicles on the road and their differing dimensions, these two data points alone would be

---

[2] *See* Yves-Alexandre de Montjoye et al., Unique in the Crowd: The privacy bounds of human mobility, Scientific Reports 3, Article Number 1376 (Mar. 23, 2013), http://www.nature.com/articles/srep01376 (finding that "four spatio-temporal points are enough to uniquely identify 95% of the individuals" in a study of 15 months of human mobility data for 1.5 million individuals, with a dataset where the location of an individual is specified hourly).

[3] Philippe Golle & Kurt Partridge, Palo Alto Research Center, On the Anonymity of Home/Work Location Pairs (2009), https://crypto.stanford.edu/~pgolle/papers/commute.pdf.

sufficient for a tracking system to conclude that the different certificates belonged to the same vehicle.

Additionally, the metadata that can be derived from the chain of trust built between the leaf certificate and the root certificate pursuant to the proposed Public Key Infrastructure (PKI) can also be used to associate BSMs with the vehicles that transmitted them. For instance, in a design (like the one proposed) involving multiple root certificate authorities (CAs), intermediate CAs, and pseudonym CAs—all controlled by different manufacturers and departments within these manufacturers—it would be trivial to significantly narrow down the number of vehicles from which any given BSM could have possibly transmitted by simply analyzing the path between the leaf certificate and the root certificate.

Combining BSM content and metadata with the relative simplicity of tracking a single vehicle that is transmitting numerous messages per second over a short period, it is evident that simply observing a vehicle for the relatively short rotation period (*i.e.*, 100 minutes) would likely be sufficient to fully deanonymize it. There are many parties who would be interested in vehicle tracking, and NHTSA's NPRM provides such parties with a straightforward way of doing so.

In particular, vehicle tracking is an already established industry. As we explained in our comments to NHTSA, data on vehicle locations is in high demand from companies like "banks, insurance companies, credit reporting agencies, and 'auto recovery' (*i.e.*, [repossession]) companies [who] assert that [this] data can help these companies find fraud and identity theft.[4] This data may also be of interest for marketing and advertising purposes, *e.g.*, a grocery store sending coupons or directing ads to individuals or households [whose] vehicle was recorded as parking in the lot of a competitor. It may [also] be of interest to divorce attorneys." And of course, vehicle location data is frequently by law enforcement.[5]

As we further stated in our comments to NHTSA, "There are currently two main private companies—DRN and MVTrac—that hire contractors to collect license plate data from cars across the United States.[6]" At the moment, these two companies are limited to using line-of-sight Automatic License Plate Reader (ALPR) technology, and even with that limitation, "DRN's database contains over 2 billion records, and MVTrac said in 2012 that it has data on a 'large

---

[4] *See, e.g.*, DRN: Vehicle Location Data for Auto Lenders, Insurance Carriers and Recovery Professionals, http://drndata.com/ ("Our data helps lenders make right party contact to reduce charge-offs, insurers improve pricing at underwriting and claims investigations, and gives recovery agents the technology they need to recover more vehicles.").

[5] Cyrus Farivar, NYPD to conduct "virtual stakeouts," get alerts on wanted cars nationwide, *Ars Technica* (Mar. 2, 2015), https://arstechnica.com/tech-policy/2015/03/nypd-to-conduct-virtual-stakeouts-get-alerts-on-wanted-cars-nationwide/.

[6] Julia Angwin & Jennifer Valentino-Devries, New Tracking Frontier: Your License Plates, *Wall Street Journal* (Sept. 29, 2012), https://www.wsj.com/articles/SB10000872396390443995604578004723603576296.

majority' of the vehicles in the United States." There is no doubt that companies like DRN and MVTrac will augment their existing ALPR systems with the cheaper, more effective tracking capabilities that mandated V2V systems would provide them—giving companies and law enforcement an almost dystopian ability to track the movements of the 95% of American households that have a vehicle.[7]

EFF's NPRM comments also highlight serious security concerns with NHTSA's V2V proposal—concerns also raised by several other commenters. As one commenter, Alex Kreilein, a former Department of Homeland Security lead cybersecurity strategist and the cofounder and managing partner of SecureSet, noted in a report submitted in response to NHTSA's NPRM, "the addition of DSRC exposes a *new, additional* attack surface to vehicles which may already be vulnerable through different means."[8] EFF agrees with Kreilein's concerns regarding V2V security and the shortfalls of the NPRM's proposal. If V2V were to be deployed, it is imperative that it be deployed—and used—with extreme caution. The current proposal fails to lay out either a security framework or a compliance regime, putting the safety and lives of individuals at risk. Indeed, while some attackers may aim to merely block traffic, frighten motorists, or damage a competing manufacturers' business or reputation, other may seek to cause catastrophic car crashes.

Given the very serious and complicated privacy and security concerns raised by V2V technology, combined with the very high cost of deploying the technology—estimates that do not even take into account all of the research and development needed to solve the varied privacy and security issues—NHTSA's V2V proposal does not make sense from a cost-benefit perspective. Indeed, as outlined by NPRM commenter Brad Templeton,[9] a developer of and commentator on self-driving cars, software architect, and Internet entrepreneur, it will take a great deal of time and resources before there is any "payoff" in terms of increased safety. As As NHTSA's NPRM itself acknowledges, it would take decades before a significant percentage of vehicles on the road were equipped with V2V, not even taking into account that some of these devices will go unrepaired or un-updated. *See* RT 82 FR 3989-3990.

Meanwhile, as we noted in our NPRM comments, "the cost of implementing V2V would be great. NHTSA estimates that the total annual cost to comply with its proposed V2V mandate

---

[7] *See* Robin Chase, Car-sharing Offers Convenience, Saves Money and Helps the Environment, You Asked: Does Everyone in American Own a Car, U.S. Department of State, Bureau of International Information Programs, https://photos.state.gov/libraries/cambodia/30486/Publications/everyone_in_america_own_a_car.pdf.

[8] Alex Kreilein, Dedicated Short Range Communications (DSRC) Expose Critical Gaps in Security and Privacy, SecureSet, 2 (Mar. 29, 2017) (emphasis in original), http://glenechogroup.isebox.net/securesetaccelerator/dedicated-short-range-communications-dsrc-expose-critical-gaps-in-security-and-privacy.

[9] See Brad Templeton, V2V and the challenge of cooperating technology, http://www.templetons.com/brad/robocars/v2vdata.html.

would range from $2.2 billion on the low end to $5.0 billion on the high end, corresponding to a cost per new vehicle of roughly $135 to $301. 82 FR 398X. Not taking into account opportunity costs, NHTSA estimates that under its proposed rule—assuming that a final rule was issued in 2019, the phase-in period began in 2021, and compliance was required by 2023—the breakeven year would be between 2029, at its most liberal estimate, and 2036, at its most conservative estimate—*i.e.*, over 15 years after the final rule was issued. 82 FR 3859. The investment NHTSA is proposing is no small matter: at an annual cost of between $2.2 billion and $5.0 billion, the proposal would cost $33 to $75 billion over the 15-year-period during which the monetary investment in the technology surpassed any gains. In other words, we would see no benefit from this technology until after $33 to $75 billion was already spent. And these estimates do not even take into account the costs associated with the privacy and security risks introduced by the technology, which will include accounting for security breaches and identify theft."

Give the exponential rate of technological development in mobile data networks alone, it's likely that V2V technology will become obsolete long before consumers would as a practical matter receive any benefit from the technology. Thus, pursuing other technologies—including 5G cellular networks, either through vehicle-to-cloud or phone-to-phone communications, and potentially third-party aftermarket products—is the smarter strategy.

<u>Question 8</u>: What are the roles of the FTC, NHTSA, and other federal government agencies with regard to the privacy and security issues concerning connected vehicles?

As we explained in our comments to NHTSA, "Data regarding an individual's physical location—including details regarding the particular route taken or the physical start and end points of a trip—is extraordinarily sensitive. It can paint an intimate portrait of a person's daily life and reveal private information, such as confidential personal and professional relationships, medical information, religious affiliation, participation in stigmatized activities, and more. *See United States v. Jones*, 132 S.Ct. 945, 955 (2012) (Sotomayor, J., concurring) (noting that location information, such as GPS data, can "generate[] a precise, comprehensive record of a person's public movements that reflects a wealth of detail about her familial, political, professional, religious, and sexual associations"). Disclosing a person's physical location also facilitates stalking and a wide range of crimes against people or their property (such as burglary when a person is known to be out of town)."

As federal agencies tasked with preserving the safety and privacy of Americans, it is unquestionably the responsibility of the FTC and NHTSA to ensure that any government-mandated V2V technologies not only increase drivers' safety, but also protect their privacy and security.

Unfortunately, thus far NHTSA has not demonstrated a full understanding of the privacy and security dangers associated with its V2V proposal. In its NPRM, NHTSA suggests that the proposed standard would "make it difficult to track through space and time specific vehicles,

owners or drivers on a persistent basis." 82 FR 3869. This is flatly incorrect, and is based on the *false* assumption that there is not significant economic incentive to create a widespread tracking system. Indeed, NHTSA's proposed technology would in fact allow tracking of vehicles at much farther distances than possible with ALPR (much farther than the standard's suggested 300m[10]), while simultaneously not requiring line-of-sight as ALPR does. Further, NHTSA seems to be completely unaware of the extensive literature on practical deanonymization of pseudo-identifiers, as it does not contain any analysis of the difficulty of re-associating observed identities or the likelihood of a large tracking network observing the same identity twice.[11]

As such, it is the FTC's role as one of the primary agencies protecting Americans' privacy to educate NHTSA about the dangers of its proposed V2V standard. While the scheduled workshop is an excellent first start, EFF calls on the FTC to further highlight and explain to NHTSA all of the ramifications of its proposed V2V standard, including by conducting its own analysis of the associated privacy and security harms.

In the meantime, unless and until the privacy and security concerns with NHTSA's proposed V2V standard are resolved, it would be irresponsible for NHTSA to recommend—let alone mandate—the technology.

Respectfully submitted,

Jeremy Gillula, Ph.D.
Seth Schoen
Lee Tien
Jamie Williams
Electronic Frontier Foundation
815 Eddy Street
San Francisco, CA 94109
Phone: (415) 436-9333
jeremy@eff.org

[10] There are many examples o f radio signals being successfully received and interpreted at distances far greater than expected. *See, e.g.*, Peter Shipley, Open WLANS: The early results of WarDriving (2001), https://blyx.com/public/wireless/open_wireless_lans.pdf (researchers were able to make a connection to a network with an intended range of around 150 feet from around 25 miles away).

[11] The likelihood of observing the same identifier twice is very high and typically requires far fewer observations than intuition would suggest. Take the well-known "birthday paradox": In a room of just 23 people there's a 50% chance of two people having the same birthday; in a room of 70, that chance increases to 99.9%. *See* Wikipedia, Birthday Problem (last updated Apr. 8, 2017), https://en.wikipedia.org/wiki/Birthday_problem.