MICHIGAN ACADEMICIAN XLIV (2016), 000–000

Properties of Vehicle Network Privacy

GEORGE CORSER
Saginaw Valley State University

HUIRONG FU
Oakland University

MATHIAS MASASABI
Oakland University

AND

LARS KIVARI
Oakland University

ABSTRACT

Contemporary automobiles utilize a range of safety devices, such as seatbelts and airbags, to reduce injuries in accidents. Future vehicles, especially autonomous cars, will also utilize computer networks to avoid accidents. These vehicular ad hoc networks, *VANETs, may one day save thousands of lives and billions of dollars, reduce fuel consumption and pollution, and expand ubiquitous connectivity and mobile application functionality to the world's roadways. One problem: privacy. VANETs may expose motorists to surveillance by eavesdroppers, from casual stalkers to Big Brother. The problem has perplexed researchers for decades, perhaps partly because the desired properties of vehicle network privacy have not been sufficiently defined. The purpose of this paper is to provide a taxonomy to classify privacy properties in vehicular contexts.*

Safety belts and airbags save lives, but automobiles would be even safer if crashes were prevented before they ever happened. Jumbotron-style billboards inform drivers to avoid bottlenecks of traffic congestion, but traffic flow would be even smoother if individual drivers received traffic notifications from their own dashboards based on optimal routes to each driver's destination. Better yet, a central traffic management system could inform self-driving cars, which could re-route vehicles on the fly without any driver interaction. Not just fewer accidents, but also completely avoided accidents. Not just faster commutes ~~not~~ only for one driver, but also less traffic congestion for an entire transportation system. These are major goals of *vehicular ad hoc networks*, VANETs, or more simply, *vehicle networks*.

The phrase "connected cars" often implies infotainment and other telematics functionality, rather than safety and traffic management functionality. The term VANET often implies vehicle-to-vehicle (V2V) safety communications, though VANETs could perhaps include other functionality, including telematics. Some authors use the phrase "vehicle network" to refer to connectivity between components within a vehicle—but in this paper the terms vehicle network and VANET, are used interchangeably to indicate inter-vehicle connectivity.

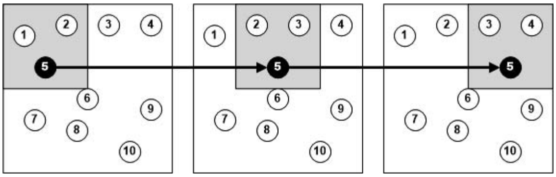
Autonomous cars are designed to operate using data from several sources, including the Internet, line-of-sight on-board sensors, and V2V data from VANETs. These self-driving vehicles will make decisions based on highly localized data. If connected to a wider network, however, self-driving cars could automate actions received from a central traffic authority. That's one reason why vehicle networks are important for autonomous cars.

One problem with vehicle connectivity is vulnerability to unwanted surveillance. With volumes of network communication data being transmitted rapidly, how can drivers keep from being spied on by the very systems designed to save their lives? Without privacy protections in place, wireless eavesdroppers or malicious network administrators could track specific vehicles, or cross-reference vehicles' precise origin and termination geographical coordinates with home and work addresses, using Google Maps or some similar map database, perhaps revealing (deanonymizing) a vehicle at a given location at a given time. Because motor vehicles tend to move at high speeds in predictable patterns along prescribed routes, their mobility patterns may make vehicles more vulnerable to location privacy attacks than, say, pedestrian mobile phone users. Vehicle surveillance can be appropriate, as when trucking fleet operators keep tabs on the locations of their big rigs. Some insurance companies exchange surveillance for lower insurance premiums, reducing premiums in exchange for electronic confirmation of driving behavior.

But there are times when such transparency is undesirable. The vehicular location privacy problem is important because driver location data can be misused. Employers might monitor an employee's car parked at a competitor's office (revealing an employee's job interview) or at specialized medical facilities (revealing an employee's health condition). It is not difficult to construct further privacy breaches arising from vehicle surveillance by spouses and ex-spouses, or paparazzi and other stalkers. Proposed national level legislation in the United States to address digital location privacy threats includes the Location Privacy Act and the Geolocation Privacy and Surveillance Act.

The location privacy challenge from a technical standpoint is large-scale and complicated in VANETs. Equipment supporting wireless/wifi networks is already being installed in new vehicles. Industry representatives estimate that 90% of vehicles will be wifi-connected within the decade (Bush 2013). Location based service (LBS) usage continues to grow rapidly (Johnson 2012) and is expected to expand to VANET platforms (Koslowski 2012). Standards governing VANETs provide data fields for future privacy protocols, but the protocols themselves remain open ~~area~~ of research.

Spatial cloaking has been a standard solution to the LBS location tracking problem. The idea is, if *k* LBS users are operating in a spatial area, *s*, then *k*,*s*-privacy (a derivative form of *k*-anonymity (Sweeney 2002)) is achieved (Lu 2008). The problem is, if LBS requests are repeated frequently over time, and only one of the *k* LBS users is consistent throughout the set of cloaked requests, then that user is exposed. Researchers have modified spatial cloaking to preserve *k*-anonymity even when LBSs receive frequent requests. However, no research has been performed which addresses the following problems. First, cloaking requires a trusted third party (TTP) or cloaking proxy, which may be unnecessary additional overhead. Second, cloaking is ineffective in low vehicle densities, especially if only one user is using LBS in the given vicinity~~.~~



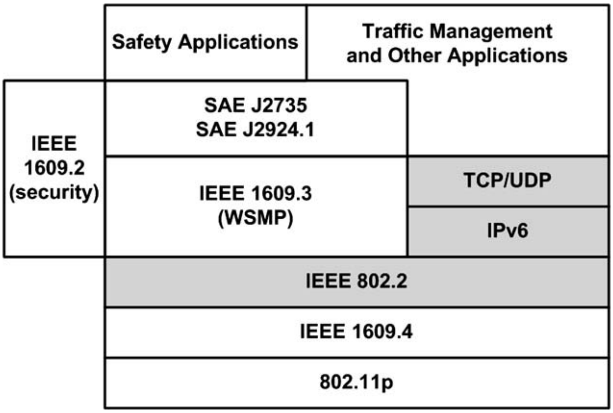
Because of the success of cloaking, other privacy methods remain relatively under-researched. In vehicular settings, *dummy event* and *active decoy* methods may prove especially effective. A dummy event is a message containing false data, sent in order to help conceal a genuine message. Dummy events and genuine messages are sent by the same genuine entity. Dummy events function analogously to aircraft countermeasures, such as flares. An active decoy, on the other hand, is a dummy event sent by an entity other than the genuine one. Naturally, any privacy protocol must accommodate safety considerations. This paper examines the desired properties of vehicle networks, including safety properties, as they relate to protecting driver location privacy.

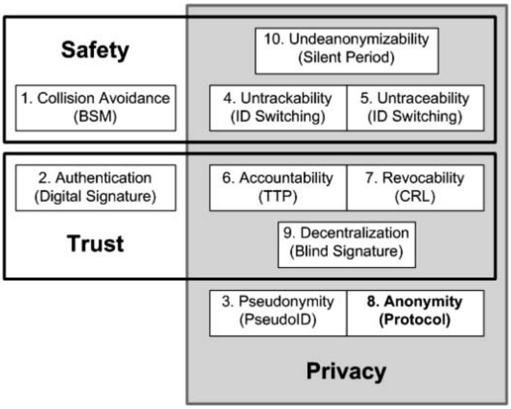
VEHICLE NETWORK PRIVACY PROPERTIES: INTRODUCTION

The literature refers to several forms of privacy. *Identity privacy* is sometimes referred to as anonymity, pseudonymity or unlinkability with personally identifiable information. Identity privacy can be achieved by the use of pseudonyms. *Location privacy* refers to the unlinkability of personal information with a geographical position, and further, the unlinkability of one pseudonym with another by using location data. *Query privacy* would make unlinkable to the user's personal information, not only the location of the user, but also the particular query made or query service used. Privacy must be constrainable, as in the cases of *conditional privacy* and *revocability*. This paper focuses on identity privacy and location privacy, ever cognizant that privacy concerns remain subordinate to safety considerations.

Vehicle safety has historically been a matter of *crash mitigation*. Safety belts, air bags, collapsible steering wheels and other technologies have been designed to reduce the severity of the consequences of a crash. A new direction in automotive safety has arisen in *crash prevention*. Vehicle networks have been proposed which would allow every car to compute its trajectory and the trajectories of other vehicles to alert drivers regarding potential crashes faster than human response alone would achieve. Such technology could eliminate 80% of crashes of unimpaired motorists. Besides saving lives, crash prevention technologies such as those predicted in vehicle networks, if effective, may reduce the price of cars. Expensive crash mitigation components, like airbags, may become unnecessary, and may be superseded by more effective crash avoidance components.

To achieve network latencies far faster traditional ones, a new set of protocols was developed and a new spectrum assigned specifically for vehicles. The Federal Communications Commission (FCC) dedicated a 75 MHz spectrum in the 5.9 GHz band for Dedicated Short Range Communications. DSRC features two distinct network/transport layer protocols. The first is called WAVE, Wireless Access for Vehicular Environments, which features WSMP, WAVE Short Message Protocol, which would typically be used in V2V safety applications. The second, IPv6/ TCP/UDP typically would be used in V2I (vehicle-to-infrastructure), especially when accessing infotainment or enabling safety services which use SAE J2735 message protocols that include a message type called a Basic Safety Message, BSM, also referred to as a heartbeat message.




Safety
1. Collision Avoidance (BSM)
10. Undeanonymizability (Silent Period)
4. Untrackability (ID Switching)
5. Untraceability (ID Switching)
2. Authentication (Digital Signature)
6. Accountability (TTP)
7. Revocability (CRL)
9. Decentralization (Blind Signature)
Trust
3. Pseudonymity (PseudoID)
8. Anonymity (Protocol)
Privacy

Property 1: Collision Avoidance. BSMs Must Maximize Safety.

To achieve collision avoidance, vehicles inform each other of whether or not they are on a trajectory to collide. In VANETs, collision avoidance is accomplished using BSMs. The BSM is the fundamental building block of VANET safety systems, and the fundamental privacy vulnerability addressed by the privacy protocols evaluated in this research. In the United States, the Society of Automotive Engineers (SAE) has established a standard, J2735, the DSRC Message Set Dictionary, which specifies the message sets and data elements for VANET communications. SAE J2945 specifies the minimum performance requirements; SAE J2945-1 specifies the requirements for the BSM (SAE International 2009). This exposition uses only the American terminology, though similar standards are set by European Telecommunications Standards Institute (ETSI) and the European Committee for

Standardization (CEN), and by the Japanese Association of Radio Industries and Businesses (ARIB).

BSMs must be fast, frequent, and localized, and include the transmitting vehicle's precise position, velocity and direction at precise times. For BSMs, *fast* means ultra low-latency wireless communication using 802.11p which omits BSS and associated MAC sublayer overhead (Kenney 2011). An average-sized 320-byte message sent over a standard 6 Mbps channel would take 4.27 ms to transmit; no handshaking or acknowledgement is required. *Frequent* means each vehicle transmits every 100 to 300 ms, depending on localized range and vehicle density. *Localized* means a vehicle transmits within a 15 m to 300 m radius (Raya and Hubaux 2005), depending upon vehicle speed and the number of other vehicles in the vicinity, which may contribute to wireless network congestion. *Precise position* is measured by the error range of the location reported by a vehicle's global positioning system (GPS).

Velocity and *direction* can be computed from multiple precise positions. Vehicular positional accuracy continues to be an active area of research. Methods have been proposed to sharpen the accuracy of GPS using ancillary systems, such as RFID-assisted localization systems (Lee et. al. 2009). Since GPS can suffer from interruption or interference, non-GPS solutions (Bohlooli 2012) have also been suggested. IEEE 1609.4 specifies GPS as an effective source of *precise time*. The preciseness and frequency of BSMs are absolutely required for safety applications, but also represent fundamental vulnerabilities to privacy.

Privacy protocols diminish safety when they include a *silent period*, i.e. a time span when no BSMs are transmitted. The importance of availability of the safety system can be illustrated using rough figures from government and NGO reports. According to the NHTSA Fatality Analysis Reporting System (fars.nhtsa.dot.gov) in 2011 there were 1.10 fatalities per 100 million vehicle miles traveled. There were 3 trillion miles traveled according to the Federal Highway Administration (www.fhwa.dot.gov), which equates to roughly 33,000 fatalities. If we estimate 82% of fatalities could be eliminated using BSMs without silent periods (Kenney 2011), then the system would save 27,060 lives. If a privacy protocol required all vehicles to activate a BSM silent period 10% of the time, then the cost to safety would be 2,706 lives per year. These are ~~a~~ naïve estimates, of course, but they serve to illustrate the impact of unavailability of BSMs. We hope more accurate estimates will be published in future research.

Property 2: Authentication. BSMs Must Be Trustworthy.

Vehicles must be able to rely on each other's BSMs. It is assumed that non-malicious vehicles will attempt to transmit accurate BSMs because each vehicle's safety depends on its neighboring vehicles knowing its precise position, and vice versa. However, inaccurate BSMs may be transmitted, either accidentally or maliciously. It is possible that a vehicle's damaged or defective GPS could accidentally generate inaccurate location readings. Cryptographic authentication would not solve the inaccuracy problem because the messages would be authentic. Authentication is designed to protect against attackers who may maliciously transmit false BSMs, degrading traffic flow or perhaps even inducing vehicle collisions.

To achieve authentication, BSMs must include valid *digital signatures*. Because digital signatures provide quick, low-overhead authentication, they have long been accepted as the most effective mechanism to ensure authentication and message integrity in vehicular environments where nodes may often be in transmission range of each other for only a brief time (Raya and Hubaux 2005).

Digital signatures do not provide confidentiality, but confidentiality is not important. BSM data do not contain secret information. In fact, BSM data are designed to be transparent for safety reasons, so confidentiality would be counterproductive. However, this lack of confidentiality, coupled with verification of credentials provided by the digital signature, also makes BSMs vulnerable to privacy attacks. Privacy protocols diminish authentication effectiveness when they introduce vulnerabilities or impair performance in processing or accessing digital signatures.

Some proposals discuss a tradeoff between storage, computation time and transmission time of digital signatures. In our evaluation, we do not consider dollar costs of private key storage. We only consider availability, confidentiality and integrity of the keys and digital certificates. To protect against attackers who might attempt to steal private keys used for digital signatures, some have proposed that such secret information be stored in a *tamper proof device* (TPD) in each vehicle. Such a component in a privacy system would tend to bolster its rating. To protect against both accidental and malicious BSMs, it has been proposed that neighboring vehicles' BSM data be confirmed by other sensors, such as short range radar or cameras similar to those used by self-parking cars, or that neighboring vehicles be measured for their trustworthiness by reputation-based trust systems (Raya and Hubaux 2005). This and any other method which explicitly improves the trustworthiness of digital signatures would receive a commensurately improved authentication rating.

By our definition, the property of *authentication* is separate from the properties of *accountability* and *revocability*, described below. A privacy protocol which requires only one single certificate authority, for example, may centralize key distribution which may simplify certificate revocation (revocability) or may help law enforcers identify culpable vehicles involved in traffic accidents (accountability), but it may also introduce a single point of attack and a performance bottleneck, a threat to the availability and integrity of the digital signatures (authentication). In such a case, we would not classify the privacy protocol as high-trust though we might rate it highly in other respects.

Property 3: Pseudonymity. ~~Bsms~~ Must Not Reveal Real Identities of Vehicles or Owners.

To achieve pseudonymity, a type of *identity privacy*, BSMs must use pseudonyms, or *pseudoIDs*, each of which having a corresponding digital certificate. Except in circumstances requiring *accountability* or *revocability*, described below, pseudoIDs and their certificates are *unlinkable* to the vehicle identification number (VIN) of the vehicle and to the personally identifiable information (personal information) of the vehicle owner.

In the literature, the term *unlinkability* may refer to the inability to correlate vehicle identities with pseudoIDs, but it may also refer to the inability to correlate between multiple pseudoIDs of a particular vehicle. To avoid ambiguity, we refer to the former as *pseudonymity* and to the latter as either *untrackability* or *untraceability*, defined below.

Privacy protocols diminish pseudonymity when they risk linkage between pseudoID and VIN or personal information. There is a natural tradeoff between authentication by digital signature and pseudonymity by pseudoID. For authentication to be fast and efficient, the fewer the pseudoIDs and certificates the better; for pseudonymity, the more the better.

Property 4: Untrackability. Pseudoids in Currently Transmitted BSMs Must Not Be Linkable to Pseudoids in Immediately Preceding BSMs from the Same Vehicle.

If a vehicle were identified (*marked*), and its pseudoID linked to personal information even a single time, then the vehicle could be tracked as long as its BSM used that same pseudoID.

To achieve *untrackability*, a type of *location privacy*, BSMs must use multiple pseudoIDs, rotating between them frequently, on average every 5-10 minutes. A single vehicle may contain several, or even thousands of pseudoIDs, each with its own digital certificate. By periodically changing among many pseudoIDs, theoretically a vehicle could only be tracked while a particular pseudoID was in use subsequent to the vehicle being marked (Kenney 2011). Privacy protocols diminish untrackability when they risk linkage between current pseudoIDs and their immediately preceding pseudoIDs. There is a natural tradeoff between authentication by digital signature and untrackability. For authentication to be fast and efficient, the fewer the pseudoIDs and certificates the better; for untrackability, the more the better. If a pseudoID switching technique in a privacy protocol includes a silent period when no BSMs are transmitted, there could also be a tradeoff between collision avoidance and untrackability.

Property 5: Untraceability. Pseudoids in Current or Past BSMs Must Not Be Linkable to Other Pseudoids from the Same Vehicle, Except by Proper Authorities.

To achieve *untraceability*, another type of *location privacy*, sometimes called *historical location privacy*, BSMs must use multiple pseudoIDs, switching between them, as in untrackability, above. However, the property of untraceability is distinct from untrackability. By our definitions, tracking a vehicle would be performed in real-time, while the vehicle is in motion. Tracing the vehicle would be a matter of historical investigation, to determine what vehicle was at what location at what time. This sort of evidence-gathering has been used by proper authorities, such as courts of law (see *accountability*, below).

But tracing could also be used by stalkers or paparazzi for gathering background information on vehicles to establish locations at specific times or to establish transportation patterns of people under unauthorized surveillance. Sometimes in the literature definitions of the terms untrackability and untraceability are interchanged. Sometimes they are used as synonyms. The properties are similar but not exactly the same. We offer our definitions as standard terminology to distinguish between protocols protecting real-time location privacy (untrackability) and historical location privacy (untraceability).

Privacy protocols diminish untraceability when they risk linkage between pseudoIDs and preceding pseudoIDs for a given vehicle. There is a natural tradeoff between authentication by digital signature and untraceability. For authentication to be fast and efficient, the fewer the pseudoIDs and certificates the better; for untraceability, the more the better. If a pseudoID switching technique in a privacy protocol includes a silent period when no BSMs are transmitted, there could also be a tradeoff between collision avoidance and untraceability.

Property 6: Accountability. Pseudoids Must Be Linkable to personal information by Proper Authorities.

Sometimes it is beneficial to link a vehicle to its owner's identity and/or its location, such as when a vehicle may have been used in a crime or involved in an accident. It may be argued that a privacy protocol without the property of accountability would introduce more risk to the public by concealing criminals than it would introduce security to the public by protecting people's privacy. To achieve accountability, a certificate authority

(CA) or other trusted third party must protect vehicle and owner identity and location while maintaining the capability to link this information with pseudoIDs if requested by a proper authority. This capability to link is sometimes referred to as *conditional privacy*. Privacy protocols diminish accountability when they do not provide a secure mechanism for linkage between pseudoIDs and vehicle/owner identity and location. There is a natural tradeoff between trust and privacy, and this balance must be struck in the property of accountability. The TTP must be able to determine the circumstances under which a proper authority may circumvent a privacy protocol and reveal the true identity associated with a pseudoID.

Property 7: Revocability. Pseudoids and Digital Certificates Must Be Rescindable.

It is possible that valid digital certificates could be stolen and used maliciously. Stolen certificates should be revoked. To achieve revocability, a CA or other TTP must provide valid digital certificates for pseudoIDs while maintaining the capability of rescinding certificates by updating and distributing a *certificate revocation list* (CRL) if requested by a proper authority. Privacy protocols diminish revocability when they impair the distribution of CRLs securely, quickly and broadly (Haas 2011). For authentication to be fast and efficient, the smaller the CRLs, the better; for effective revocability, some protocols indicate large CRLs. There is a natural tradeoff between trust and privacy, and this balance must be struck in the property of revocability. The TTP must be able to determine the circumstances under which more harm than good comes from BSMs bearing a particular pseudoID and that the benefit of revoking that pseudoID's digital certificate exceeds its cost.

Property 8: Anonymity. Privacy Models Must Maximize Indistinguishability Between Pseudoids.

Privacy protocols can be evaluated by *anonymity*, which we define as the quantifiable amount of privacy the vehicle's pseudoID enjoys by using the protocol. Anonymity could measure identity privacy or location privacy. The pseudoID is the mechanism which protects identity privacy; therefore, it follows that pseudonym anonymity could measure identity privacy protection. But what about location privacy?

Shokri enumerates four methods of preserving location privacy: obfuscation, hiding events, adding dummy events, and anonymization (Shokri, Freudiger and Hubaux 2010). For BSMs, obfuscation is not possible because precision is required for safety applications. Hiding BSMs is not possible because safety applications depend on the detectability of BSMs. Adding dummy BSMs may threaten safety by inducing vehicles to react to nonexistent vehicles; in fact digital signatures are used to reduce the possibility of malicious fake BSMs. The only remaining method is anonymization. Since the identities used in BSM transmissions are pseudonyms, pseudoIDs, the only way to protect privacy in VANETs is by *pseudonym anonymity*, or as we call it, anonymity.

It is necessary to make fine distinctions between the terms, anonymous and pseudonymous, to clarify computational privacy in vehicular network contexts. The dictionary definition of *anonymous* is, "not named or identified" (Merriam Webster), a definition which cannot apply in vehicular networks that require identifiers. When referring to computers, the term anonymous sometimes means using a pseudonym which is unlinkable to a person's true identity, as in an *anonymous post* on a blog. This definition introduces ambiguity with the term anonymity when used as in *anonymity set*, defined below. We use the dictionary definition of *pseudonymous*, "bearing or using a fictitious name" (Merriam Webster) to indicate unlinkability to personal information. Vehicle networks use pseudoIDs, which achieve the property of *pseudonymity*. In this paper, we use the term anonymity as it is used in set theory, as in an anonymity set (Dıaz 2002). Thus we can define two distinct privacy properties, pseudonymity and pseudonym anonymity.

To achieve anonymity, privacy models must maximize the *anonymity set size* of each pseudoID. An anonymity set (AS) is a group of entities indistinguishable from one another. Anonymity set size, $|AS|$, is the number of entities in that group. To achieve pseudonym anonymity, privacy models must create conditions where $|AS| > 1$ for BSM pseudoIDs. Pseudonym anonymity is difficult to achieve because of the rapid repetition of BSMs containing the same pseudoID, and because BSMs contain data fields which may uniquely or partially identify a vehicle.

Incidentally, one application of the anonymity set concept is *k-anonymity*, which requires that in the results of a database query each entity must be indistinguishable from k - 1 other entities (Sweeney 2006). Gruteser and Grunwald apply k-anonymity to vehicular location privacy (Gruteser 2003), but in the context of database queries. Some researchers have challenged the appropriateness of k-anonymity, which depends on a centralized anonymity server (CAS) to obfuscate queries, as a valid metric for location privacy (Shokri, Freudiger and Hubaux 2010). In the context of BSMs, under our threat model, there is no way to enforce k-anonymity. If an antenna were set up at a particular intersection, it could be used to record and log the BSMs of vehicles. The logs of multiple antennae could be used to track or trace any or all vehicles in their vicinities by following the pseudoIDs contained in BSMs. Whoever sets up the antennae would control the logs, and anyone could set up antennae. No one could enforce obfuscation in queries.

BSMs with the same pseudoID are transmitted frequently, every 100 ms, which at 60 mph (100 kph) is every 8.8 ft per 100 ms (2.8 m per 100 ms). A Ford F-150 pickup truck is about 18 feet (5.5 m) long, so even at relatively high speeds, if one were to take top view snapshots of an F-150 the instant it transmitted BSMs and superimpose the images on a map, the snapshots would overlap even if the vehicle were traveling at highway speeds. Rapid repetition of BSMs simplifies tracking for attackers because even if all vehicles in an area were suddenly to change their pseudoIDs an eavesdropper could tell which preceding pseudoIDs correlated to their successors by comparing the nearest most likely previous positions.

Privacy protocols address the BSM tracking and tracing problems by periodically changing pseudoIDs, as discussed above. Privacy protocols address the BSM pseudoID anonymity problem using protocol-specific pseudonym-changing techniques, usually involving mix zones, silent periods and/or group signatures. *Mix zones* are geographical areas where vehicles cannot be detected. In mix zones multiple vehicles are in close proximity. Vehicles may or may not change the direction of their trajectories, but they will change their pseudoIDs (Beresford 2003). Due to the precision of BSM location data, mix zones without silent periods provide little or no effectiveness in achieving anonymity. *Silent periods*, as mentioned above, are time spans when no BSMs are transmitted and when vehicles change pseudoIDs. A silent period without a mix zone (area of multiple vehicles) is of minimal use, because if there is only one vehicle in an area, changing pseudoIDs will not fool anyone.

Group signatures are authentications of BSMs made by a key shared by multiple vehicles. In the *group model*, vehicles travel in clusters, all using the same group temporary identifier, and authenticating messages using the same group signature (Guo 2007). In simulations the group model has been shown to be effective in achieving anonymity in VANET communications but it is less effective the lower the vehicle density, since anonymity level depends on the number of vehicles in each group. The group model introduces inefficiency by additional overhead for group setup and group join/verify processes. Some researchers suggest the group model is infeasible due to limitations of bandwidth and computation power, since pseudoID schemes create large *certificate revocation lists*, CRLs, and associated *checking costs*, network overhead necessary to verify that certificates have not been revoked (Sun 2010).

Silent periods may impair the safety property. Group signatures may impair digital signature efficiency, the trust property. The method used by the protocol to achieve anonymity was not a factor in our evaluation of a protocol's anonymity property, though it may have affected our evaluations of other properties. Anonymity has been measured using a range of metrics, not all of which apply to our evaluation. Because we assume no central capability to obfuscate data, a wide range of database query metrics do not apply, including k-anonymity (Sweeney 2002), l-diversity (Machanavajjha-la, Kifer, Gehrke and Venkitasubramaniam 2007), t-closeness (Li 2007), L1 similarity (Coull 2008), m-invariance (Xiao 2007)(Dewri 2010), and ε-differential privacy (Dwork 2006). We also set aside network metrics such as combinatorial anonymity degree (CAD), zone-based receiver k-anonymity (ZRK) and evidence theory anonymity (Kelly 2008). These are all valuable measures of anonymity but they do not measure pseudonym anonymity in the context of BSMs. We evaluate pseudoID anonymity of privacy protocols using one or more of the following metrics: pseudonym anonymity set size, $|AS|$; entropy of the anonymity set size, $H(|AS|)$, also called entropy anonymity degree (EAD) (Kelly 2008); and tracking probability, Pt.

Privacy protocols diminish anonymity when they do not provide high levels of indistinguishability between pseudoIDs, as measured by $|AS|$, $H(|AS|)$ and Pt.

Property 9: Decentralization. BSMs Must Not Be Traceable by a Single TTP.

To achieve decentralization, BSMs must include *blind signatures*, which require a traceable anonymous certificate (TAC). A TAC is a certificate issued by two TTPs, a Blind Issuer (BI) and an Anonymity Issuer (AI). Therefore, decentralized protocols by our definition must call for multiple independent TTPs. As far as authentication is concerned the certificate works the same as any digital signature, but to trace back to the personal information of the vehicle requires the cooperation of multiple TTPs (Park and Kent 2009, Chaum 1983).

The purpose of decentralization is to prevent a single TTP from being able to trace or track vehicles. Several researchers have referred to traceability as Big Brother, so we include in this property of decentralization the concept of Big Brother defensibility. Technically, if both the BI and the AI were controlled by the same administrative authority, it would be possible to have decentralization without Big Brother defensibility. Privacy protocols diminish decentralization when they fail to call for multiple independent TTPs.

Property 10. Map Database Undeanonymizability. Privacy Models Must Minimize Map Database Undeanonymizability.

LBS applications such as Google Maps API or US Census TIGER database may be used to convert longitude and latitude coordinates into postal addresses, that is, if a vehicle attempted to send fake coordinates to an LBS, the LBS could to *deanonymize* the data. Protected data must be, to the extent possible, un-deanonymizable. Undeanonymizability may require application level defenses, such as EPZ, endpoint protection zones (Corser et al. 2013).

CONCLUSION

Wireless network communication systems currently being developed specifically for transportation systems will likely both increase passenger safety and decrease passenger privacy. As autonomous cars emerge as fundamental components of transportation systems, they will likely rely on vehicle networks for an increasingly large range and volume of data, perhaps creating a new array of privacy vulnerabilities.

Researchers continue to investigate methods for maximizing privacy, but they do not always consider the full range of issues involved in implementing protocols, nor do they always use terminology consistently. This paper presents terminology to describe ten desired properties of VANETs relevant to location privacy preservation which the authors hope will be helpful to future researchers in examining and evaluating proposed VANET privacy protocols.

REFERENCES

Anonymous. Merriam-Webster.com. https://www.merriam-webster.com/dictionary/anonymous.

Beresford, A. R., and Stajano, F. 2003. "Location Privacy in Pervasive Computing." *IEEE Pervasive Computing*, 2(1): 46–55.

Bohlooli, A., & Jamshidi, K. 2012. "A GPS-Free Method for Vehicle Future Movement Directions Prediction Using SOM for VANET." *Applied Intelligence*, 36(3): 685–697.

Bush, I. 2013, Feb 25. GM, AT&T readying in-vehicle wi-fi. http://philadelphia.cbslocal.com/2013/02/25/gm-att-readying-in-vehicle-wi-fi/

Chaum, D. 1983. "Blind Signatures for Untraceable Payments." In *Advances in Cryptology: Proceedings of Crypto 82*, ed. David Chaum, 199-203. New York: Springer Science + Business Media.

Corser, G., Fu, H., Shu, T., D'Errico, P., Ma, W. 2013. "Endpoint Protection Zone (EPZ): Protecting LBS User Location Privacy Against Deanonymization and Collusion in Vehicular Networks." *The 2nd International Conference on Connected Vehicles & Expo, Las Vegas* (ICCVE 2013).

Coull, S. E., Wright, C. V., Keromytis, A. D., Monrose, F., & Reiter, M. K. 2008. "Taming the Devil: Techniques for Evaluating Anonymized Network Data." In *Proceedings of Network and*

Distributed System Security Symposium 2008: February 10-13, 2008, San Diego, CA, 125-135. Internet Society.

Dewri, R., Ray, I., & Whitley, D. 2010, May. "Query M-Invariance: Preventing Query Disclosures in Continuous Location-Based Services." In *Mobile Data Management (MDM), 2010 Eleventh International Conference*, Kansas City, MO, 95-104. IEEE Xplore Digital Library.

Dwork, C. 2006. "Differential Privacy." In Automata, Languages and Programming: Proceedings of 33rd International Colloquium, ICALP 2006, Venice, Italy, July 10-14, 2006, 1–12. Berlin Heidelberg: Springer.

Diaz, C., Claessens, J., Seys, S., & Preneel, B. 2002. "Information Theory and Anonymity." In *Proceedings of the 23rd Symposium on Information Theory in the Benelux*, 179-186.

Fatality Analysis Reporting System (FARS). National Highway Traffic Safety Administration (NHTSA), www.nhtsa.gov/FARS.

Gruteser, M., & Grunwald, D. 2003. "Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking." In *Proceedings of the 1st International Conference on Mobile Systems, Applications and Services* 31-42. New York: Association for Computing Machinery.

Guo, J., Baugh, J. P., & Wang, S. 2007, May. "A Group Signature Based Secure and Privacy-Preserving Vehicular Communication Framework." In *Proceedings of Mobile Networking for Vehicular Environments*, 103-108. IEEE Catalog CFP0738D.

Haas, J. J., Hu, Y. C., & Laberteaux, K. P. 2011. "Efficient Certificate Revocation List Organization and Distribution." Selected Areas in Communications, *IEEE Journal* 29(3), 595–604.

IEEE Standard for Wireless Access in Vehicular Environments Security Services for Applications and Management Messages," IEEE Std 1609-2.

Johnson, L. 2012, Oct 31. Location-Based Services to Bring in $4b Revenue in 2012: Study. http://www.mobilemarketer.com/cms/news/research/14115.htmlhttp://www.mobilemarketer.com/cms/news/research/14115.html

Kelly, D. J., Raines, R. A., Grimaila, M. R., Baldwin, R. O., & Mullins, B. E. 2008, October. "A Survey of State-of-the-Art in Anonymity Metrics." In *Proceedings of the 1st ACM Workshop on Network Data Anonymization*, 31-40. New York: Association for Computing Machinery.

Kenney, J. B. 2011. "Dedicated short-range communications (DSRC) standards in the United States." *Proceedings of the IEEE* 99.7: 1162–1182.

Koslowski, T. 2012, Jan. 3. "Your connected vehicle is arriving." http://www.technologyreview.com/news/426523/your-connected-vehicle-is-arriving/

Lee, E. K., Yang, S., Oh, S. Y., & Gerla, M. 2009. "RF-GPS: RFID Assisted Localization in VANETs." In *Mobile Adhoc and Sensor Systems*, 2009. MASS'09. IEEE 6th International Conference 621-626. IEEE.

Li, N., Li, T., & Venkatasubramanian, S. 2007. "T-Closeness: Privacy Beyond K-Anonymity and L-Diversity." In *Data Engineering, 2007*. ICDE 2007. IEEE 23rd International Conference 106-115. IEEE.

Lu, H., Jensen, C. S., & Yiu, M. L. 2008, June. "PAD: Privacy-Area Aware, Dummy-Based Location Privacy in Mobile Services." In *Proceedings of the Seventh ACM International Workshop on Data Engineering for Wireless and Mobile Access*, 16-23. New York: Association For Computing Machinery.

Machanavajjhala, A., Kifer, D., Gehrke, J., & Venkitasubramaniam, M. 2007. "L-Diversity: Privacy Beyond k-anonymity." ACM *Transactions on Knowledge Discovery from Data* (TKDD), 1(1): 3.

Park, H., & Kent, S. 2009. Traceable anonymous certificate. Pseudonymous. Merriam-Webster.com. https://www.merriam-webster.com/dictionary/pseudonymous.

Raya, M., & Hubaux, J. P. 2005, September. "The Security of VANETs." In *Proceedings of the 2nd ACM International Workshop on Vehicular Ad Hoc Networks* 93-94. New York: Association For Computing Machinery.

Shokri, R., Freudiger, J., & Hubaux, J. P. 2010. "A Unified Framework for Location Privacy" (No. EPFL-REPORT-148708).

Sun, Y., Lu, R., Lin, X., Shen, X., & Su, J. 2010. "An Efficient Pseudonymous Authentication Scheme with Strong Privacy Preservation for Vehicular Communications." *Vehicular Technology*, IEEE Transactions 59(7): 3589–3603.

Sweeney, L. "K-Anonymity: A Model for Protecting Privacy." 2002. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 10.05: 557–570.

Xiao, X., & Tao, Y. 2007, June. "M-Invariance: Towards Privacy Preserving Re-Publication of Dynamic Datasets." In *Proceedings of the 2007 ACM SIGMOD International Conference on Management of Data* 689-700. New York: Association for Computing Machinery.

Queries for maca-44-03-03

This manuscript/text has been typeset from the submitted material. Please check this proof carefully to make sure there have been no font conversion errors or inadvertent formatting errors. Allen Press.