Pseudonym-on-demand: a new pseudonym refill strategy for vehicular communications

Zhendong Ma, Frank Kargl, and Michael Weber Institute of Media Informatics, Ulm University, Germany {zhendong.ma|frank.kargl|michael.weber}@uni-ulm.de

Abstract—Pseudonyms are pseudonymous certificates, which are regarded as a silver bullet to meet the security and privacy requirements of vehicular communications. Most works so far assume that pseudonyms are readily available when they are needed. In this paper, we identify and compare two strategies to refill pseudonyms. We then propose the pseudonym-on-demand scheme, which is an implementation of one of the strategies. We show that our approach supports the functionalities of pseudonyms in terms of secure and privacy-preserved vehicular communications. Furthermore, our proposed scheme serves as a platform, into which many features to enhance security and privacy can be integrated.

I. INTRODUCTION

The emerging vehicular communications (VC) and vehicular networks (VN) are envisioned to greatly improve the road safety, traffic efficiency, and driver's comfort in the near future. However, there are challenges which need to be addressed before the deployment in the real world. Security and privacy are two of the challenges in VC which have drawn a lot of attentions in the research community in recent years [1] [2].

The open and wireless nature of VN makes it the target of various attacks. An attacker might insert false or modified messages into VN, or locate or track vehicles by their communications. The consequence of such attacks ranges from disruption of normal functionalities of VN to serious damages to the public safety on the road. A set of requirements on VC has been identified, such as authentication, integrity, nonrepudiation of the messages, as well as adequate privacy protection of the driver and passengers against identification and location profiling [3].

Most proposals on securing VC are converging into a *digital signature* based approach, which assumes the existence of a Public Key Infrastructure (PKI). Nevertheless, conventional certificates contain information which can reveal the owners' identities and thus deteriorate the privacy level of the vehicles in VC. Many researchers propose to use pseudonyms, which are pseudonymous certificates with all identifiable information removed, to provide privacy protection while still meeting the security requirements on VC [4] [5] [6]. Vehicles are supposed to change pseudonyms often to avoid being tracked on the same pseudonyms. Consequently, a vehicle needs a on-board *pseudonym pool* to support the pseudonym change.

Although the pseudonym based approach has attained a steady development in recent years, there are still open issues.

One of the issues is to find a strategy for vehicles to obtain new sets of pseudonyms when their stored pseudonyms are used up. Is it better to load a large amount of pseudonyms at one time or to load a small amount of pseudonyms at several times? The answer to this question can lead to big differences in the organization, architecture, and communications, as well as the level of security and privacy in the future VN.

In this paper, we propose a solution for pseudonym refill. Our investigation identifies two strategies for obtaining new sets of pseudonyms. After evaluating their pros and cons, we propose a pseudonym refill scheme as an implementation of one of the strategies which we name as *pseudonym-ondemand* (POD). We also develop a mechanism which can be easily integrated into the POD scheme to enhance privacy of individual vehicles. By evaluating the POD scheme by a set of criteria, we show that it is a practical and efficient solution for enhancing security and privacy in VC.

Sec. II describes the system and threat model we assume in this paper. We go on to discuss pros and cons of two strategies for pseudonym refill in Sec. III. We propose the basic and a privacy-enhanced scheme to realize POD in Sec. IV. We evaluate the schemes and show that they can enhance security and privacy in Sec. V. Sec. VI briefly reviews the state-of-theart, followed by the conclusion and outlook in Sec. VII.

II. BACKGROUND

A. System model

Within VN, vehicles can communicate to each other (Vehicle-to-Vehicle) and to roadside units (RSUs) (Vehicle-to-Infrastructure) via short range wireless Dedicated Short Range Communications (DSRC) [7] technology. A Road Authority (RA) is the entity in charge of the administration of the road network (e.g., road safety and traffic control) in a certain geographic region. The RA can use RSUs to communicate with vehicles driving within its region. Nodes in VN rely on digital signature to authenticate and verify the validity of received messages. To do so, a node needs certificates, which are certified public keys issued by a Certificate Authority (CA) in the infrastructure network. When sending a message, the node uses one of its private keys corresponding to a certificate and calculates a digital signature over the message. The node also appends the certificate to the message so a receiving node can verify the signature. The RA can communicate with the CA via the infrastructure network.

Part of this work has been supported by the European Commission through the SEVECOM project (IST-027795).

B. Threat model

The potential attacks on VC have been extensively studied in recent years. In this paper, we assume that an attacker targets the security of communications among the nodes in VN and the privacy of individual vehicles in VC. We adapt the attacker model identified in [5], which describes attackers as *local vs. global, active vs. passive,* and *insider vs. outsider.* Due to space limitations, we refer our readers to the reference for a detailed description of attacks on VN.

III. HOW TO REFILL PSEUDONYMS?

To avoid being tracked on the same pseudonym, a vehicle uses a pseudonym only on a sequence of messages and switch to a new one during communications. Therefore, a vehicle always needs *adequate* fresh pseudonyms in its pseudonym pool. If there are few fresh pseudonyms left, it needs to request new pseudonyms from a *pseudonym provider* (PP), an entity in the infrastructure which can issue pseudonymous certificates and is trusted by all nodes in the network. This process is called *pseudonym refill*. The strategy which a vehicle can apply to refill pseudonym ranges from one pseudonym at a time to once for a long time (e.g., a quarter or a year). For our analysis, we choose two strategies within the range, which represent two opposite directions of the refill strategy.

- *Strategy 1*: request and refill as many pseudonyms as possible at a time, store them on-board, and use them over a long period of time;
- *Strategy 2*: request and refill only a few pseudonyms at a time, repeat the process whenever pseudonyms are needed.

Most prior art implicitly assumes the first strategy. The advantage is that vehicles need to only occasionally establish connections with the PP, so the drivers have the convenience to choose time and place to carry out pseudonym refill. The disadvantage of loading a large amount of pseudonyms is that it increases the demand on storing and securing pseudonyms in the on-board unit (OBU). Despite the lifetime, pseudonyms can only be used on a sequence of messages within the validity time. Thus the usage of pseudonyms is related to the usage of the vehicle, and pseudonyms are supposed to have a relatively short lifetime to limit the effect of key compromise, e.g., valid for one day [8]. Thus it is very difficult for a driver to estimate the number of pseudonyms needed for future usage at the time of refill. A more serious problem is the impact on pseudonym revocation. Pseudonym revocation is an inherited problem from PKI, where for reason like a key compromise, the certificate needs to be revoked before its expiration date [9]. Compared to the wirelined Internet where PKI is originally indented for, the propagation of information on pseudonym revocation becomes more challenging due to the large scale and loose connectivities among nodes in VN. To aggravate, instead of having only one to a few certificates like most entities in PKI do, vehicles can acquire and store multiple pseudonyms well before their valid and expiration date. The size of revocation lists grows exponentially when the network scales up.

On the other hand, the second strategy allows vehicles to refill pseudonyms only when they need them. To enable such strategy, a vehicle needs to establish secure connections to the PP often, which increases the requirement on connectivities and the communication overhead. The advantage is that a vehicle can much better estimate the pseudonym consumption in a short time window from the time of refill. So the resource on computing and storing pseudonyms can be optimized. Furthermore, pseudonyms are *freshly issued*, *immediately used*, and quickly expired in the second strategy, the need for pseudonym revocation can be minimized (or even eliminated if the time to disseminate revocation lists is longer than the pseudonym lifetime). Given the short time window, even a compromised pseudonym will have very limited effect.

Table I summarizes the pros and cons of the two strategies. Despite the relatively higher cost for deployment, strategy 2 offers many benefits and provides some of the solutions to the problems in VN, which cannot be easily solved under strategy 1. The overhead of deployment can be reduced if it can utilize the existing Intelligent Transportation System infrastructure. Therefore, in this paper we propose schemes to implement the second strategy, which we call *pseudonym-on-demand* (POD). We will show in the following sections that the implementation leads to the creation of a platform, into which more features to enhance security and privacy can be easily integrated.

 TABLE I

 Comparing two pseudonym refill strategies

| Criteria | Strategy 1 | Strategy 2 |
|----------------------------|----------------------|------------------------|
| Connectivities to PP | occasional | very often |
| Communication overhead | low | slightly higher |
| No. of pseudonym to refill | difficult to predict | easy to estimate |
| Pseudonym storage | big storage | very small storage |
| Vulnerability time window | big | very small |
| Pseudonym revocation | no good solution | minimized / not needed |
| Deployment cost | low | relatively high |

IV. PSEUDONYM-ON-DEMAND

This section first introduces the basic scheme of POD, and then a privacy-enhancing feature based on the basic scheme.

A. Basic scheme

In our scheme, a vehicle V has a long-term certificate $Cert_V$ which is issued by the CA and can be used to uniquely identify the vehicle (e.g., by binding the vehicle's registration number with the certificate). Besides, the vehicle has two sets of pseudonyms: a set of long-term pseudonyms (LTPs) and a set of short-term pseudonyms (STPs). LTPs are issued by a PP and valid for a long period of time (e.g., months). To obtain LTPs, a vehicle uses the certificate $Cert_V$ to authenticate itself to the PP for a set of LTPs. The CA timely updates the PP with the latest Certificate Revocation Lists (CRLs). Since the PP has the information on the revoked certificates, it can decide whether a certificate from the vehicle is valid

at the time of authentication. STPs are issued by RAs, and are valid for a very short time since issuing (e.g., tens of minutes). The RAs maintain connections to the PPs via the infrastructure network, so it can download and cache the upto-date pseudonym revocation lists containing revoked LTPs published by the PP. A vehicles can only obtain STPs from a RA *on-site* (i.e., within the geographic region covered by the RA) by authenticating itself to the RA using a valid LTP. $Cert_V$, LTPs, and STPs contain public keys certified by the CA, the PP, and the RA, respectively. Each of the public keys have their corresponding private keys, which are kept in the Tamper Resistant Module of the vehicle's on-board system.

A vehicle refills pseudonyms by communicating with a RA through one of its RSUs. The request and response process to refill STPs involves the following message exchanges:

1) The RA periodically broadcasts the service announcement through one of its RSUs. $I_{service}$ is the *service information* which include the service offered, the identity of the responsible RA, the geographic boundaries of the RA, and the location of the RSU. It can also include the information on the position or distance to the next possible RSU down the road. With such information, a vehicle can estimate the number of STPs needed at the time of refill. T_{RA} is a time stamp. The RA calculates a signature $\sigma_{SK_{RA}}(I_{service}, T_{RA})$ by using its private key SK_{RA} and appends the corresponding certificate.

2) When an approaching vehicle V receives the announcement, it validates the message by its signature. It can also check the consistence of the alleged location of the RSU by comparing the location of the RSU with its own on-board positioning system. V pre-calculates a set of asymmetric key pairs, keeps the private keys, and provides the public keys to the RA. When sending the public keys, V extracts the RA's public key PK_{RA} from the certificate $Cert_{RA}$, uses it to encrypt the whole message, which includes *i* public keys PK_V^1, \ldots, PK_V^i , a time stamp T_V , and a signature calculated by the private key SK_V^j corresponding to its *jth* LTP, and the LTP_V^j . Notice that a vehicle can choose to use the same LTP for all refill processes within the region of the same RA.

3) After the RA receives the message through the RSU, it decrypts the message with its private key SK_{RA} . Then it validate the message by first checking the cached lists for revoked LTPs and then verifying the signature. Next, the RA calculates *i* STPs for *V* by signing the public keys from *V* and sends them back to *V*. To prevent a third party knowing the content of the message to *V*, before sending the message, the RA extracts the public key PK_V^j from LTP_V^j and uses it to encrypt the message. To facilitate the target vehicle to pick up the message, the RA inserts the time stamp from *V* in front of the ciphertext. So *V* knows that the message is sent to it before it tries to decrypt the ciphertext.

4) Upon receiving the message, V uses the *jth* private key SK_V^j to decrypt the message and obtains the set of STPs. V sends an acknowledgment to the RA. The acknowledgment contains the same time stamp T_V and a signature of the time stamp $\sigma_{SK_V^j}(T_V)$ by the same private key corresponding to LTP_V^j . The message is encrypted with the RA's public key PK_{RA} to prevent a third party from knowing the content.

When the RA received the acknowledgment, it archives the STPs together with the LTP. In case of a liability investigation, an authorized party can request the RA to reveal the link between a STP to the LTP. V uses the newly acquired STPs in its communications until it reaches another refill RSU down the road. Depending on the status of storage of unused STPs, V decides whether to start the refill process again.

B. Privacy-enhanced scheme

Pseudonyms achieve unlinkability between the certificates used in communications and a vehicle's long-term identity. However, it is possible that pseudonyms from the same vehicle can be linked by their spacial and temporal correlations. The aforementioned basic scheme cannot guarantee to prevent a global passive attacker (i.e., an attacker passively eavesdrops all communications in VN) from linking and building location profiles of the vehicles (i.e., to form the tracks of the vehicle's daily movement). Based on the location profile, an attacker might be able to infer or discover the vehicle's long-term identity. Intuitively, there should be mechanisms to hide a vehicle's communications from such attacker.

Our privacy-enhanced scheme uses a symmetric group key distributed by RSUs to create a virtual *mix zone*, in which vehicles encrypt their communications so an attacker cannot read the content of the messages while the vehicles switch to the newly-refilled STPs. A mix zone [10] is a spatial region where users' positions cannot be located. Users update their identities within the mix zone, and make it harder for an attacker to correlate the new identities at the exit of the mix zone to the old ones at the entrance of the mix zone.

In our scheme, mix zones are created downstream after the refill RSUs. For this reason, the RA generates a symmetric encryption key K_e together with the information on the mix zone I_{mix} , and adds them in the encrypted part of the response from the RA to the vehicle:

$$RA \xrightarrow{RSU} V : T_V, \{K_e, I_{mix}, STP_V^1, \dots, STP_V^i\}_{PK_{i}^j}$$

A vehicles receiving the response is able to communicate in two modes: *private mode* and *public mode*. In private mode, the vehicle encrypts outgoing messages with the symmetric encryption key K_e . In public mode, messages are sent without encryption. I_{mix} specifies the border lines of the mix zone, at where a vehicle may start and stop encrypting its communications. Notice that the boarder lines should start and end at least one-hop (e.g., 300 meters) away from the RSUs, so the approaching vehicles within one-hop to the mix zone can receive the same symmetric key K_e from the RSUs to decrypt the messages originated from the mix zone. Based on the geographic features of the road network, a RA can design mix zones which conveniently use the road network to achieve maximum unlinkabilities of changing pseudonyms. Fig. 1 shows a simple example of creating a mix zone in a segment of a straight road between two RSUs. Vehicles passing the two RSUs receive the *same* encryption key to encrypt their communications. Vehicles passing the RSU are informed of the board lines of the mix zone, P_{start} and P_{end} . So they can encrypt their communications accordingly.



Fig. 1. An example of a virtual mix zone created in a segment of the road by using two RSUs to distribute the same encryption key.

If all vehicles switch back to public mode when they reach P_{end} , an attacker might still be able to link most of the pseudonyms by the strong correlation of the vehicles in and out of the mix zone. To avoid this, the vehicles on the same direction are designed to have a fixed P_{start} but a variable P_{end} , which means, after switching to the private mode at P_{start} , a vehicle chooses randomly when to switch back to public mode before it reaches the position specified by P_{end} .

The privacy-enhanced scheme is still vulnerable to an attacker in possession of one or more legitimate vehicles. Imagine an attacker uses a legitimate vehicle to obtain the symmetric encryption key, the attacker can then read all encrypted communications in the mix zone. Therefore, the encryption key needs to be reset often. To make sure that vehicles in and approaching the mix zone have the same encryption key so important messages can reach all legitimate users, the key is only changed when the mix zone and its vicinity is empty. The RA can dynamically adjust the parameters of the mix zone (P_{start} and P_{end}) in I_{mix} to facilitate key reset.

V. EVALUATION

Being a special form of certificates, pseudonyms provide security functionalities which meet a set of requirements on vehicular communications such as authentication, message integrity, and non-repudiation etc. The very short lifetime of STPs, usually in the time frame of minutes, exposes a very short vulnerability window to potential attacks on the cryptographic primitives. The very short lifetime also minimizes the need for pseudonym revocation. Thus the security level of VC is higher because the security level of the pseudonyms is strengthened. On the other hand, vehicles can only obtain STPs by providing legitimate LTPs to the RAs, who have the up-to-date information on the revoked LTPs. Furthermore, the system can further enhance security by enforcing all vehicles to use and trust only STPs issued by local RAs, which requires a vehicle to provide a valid LTP to obtain STPs from a RA to participate in the local communications.



Fig. 2. Hierarchy of identities and their responsible entities.

Our privacy evaluation focuses on two aspects: the linkability of pseudonyms to the long-term identity (identification) and linkability of a series of pseudonyms from the same vehicle (tracking). Fig. 2 shows the hierarchy of identities and their corresponding organizational entities as described in Sec. IV-A. The identity hierarchy has a downward one-to-nmapping relation, e.g., the certificate bound to the vehicle's long-term identity is mapped to m LTPs and one of the LTPs is mapped to n STPs. The RA might have partial knowledge of the usages of STPs, i.e., the location at where a specific STP appears, since it can overhear messages originated from vehicles though its RSUs. However, since the LTP to STP mapping is kept local by the RA, i.e., the RA does not inform the PP based on which LTP it issues the STPs, it is not possible to link a given STP upwards in the hierarchy without involving both the PP and the RA at the same time. The same applies to the LTP to the certificate mapping between the PP and the CA. Therefore, by organizationally keeping entities separate and enforcing privacy policies on their practices, privacy threats on identification from both outside and inside attackers are mitigated.

Although a global attacker is theoretically possible, a global coverage is difficult and expensive to achieve in practice, especially in vehicular networks where communications are based on short-range wireless technology. However, in the following, we will show that our scheme is robust even under a strong attack model (i.e., a global passive attacker).



Fig. 3. An example of a vehicle's track.

A track of a vehicle can be regarded as a path through a series of areas covered by different RAs and a cascade of mix zones created by the RAs. Consider the example illustrated in Fig. 3, a vehicle's path traverses three RAs and four mix zones (gray areas). Although a global attacker can intercept all communications along the path, the mix zones make it very difficult to build an end-to-end track which consists of all pseudonyms from the same vehicle. Using the quantitative

model developed in [10], the level of privacy provided by the system can thus be calculated as the level of anonymity introduced by the mix zones, measured by entropy in *bits*. The entropy of a mix zone is calculated as $h = -\sum p_i logp_i$,

where p_i is the probability of mapping a pseudonym at the egress of a mix zone to a pseudonym at the ingress of the mix zone, and *i* is the number of such possible mappings. The higher the entropy, the higher the number of vehicles that are indistinguishable from each other. In the context of our scheme, parameters influencing the value of h include the vehicle arrival rate to the mix zone, the number of exits of the mix zone, and the attacker's prior knowledge of the mix zone (e.g., turning probabilities at the intersection and characteristics of vehicle mobility). The overall entropy of a track is the sum of entropies from each mix zone on the track. In the example in Fig. 3, if each mix zone contributes 1 bit of entropy to the track, i.e., a pseudonym at the egress of the mix zone can be mapped to two equally possible pseudonyms at the ingress, the overall entropy will be 4 bits, which means from the attacker's point of view, LTP^k has the probability of $\frac{1}{2^4} = \frac{1}{8}$ to be linked to LTP^j . It can be imagine that the linking of pseudonyms becomes virtually useless if the entropy reaches a certain level, e.g., 20 bits. This shows that under the assumption of a strong (i.e., global passive) attacker, our privacy-enhanced scheme can still provide a lower bound on the vehicle's level of privacy.

VI. RELATED WORK

Most works on security and privacy in VC focus on how to utilize the security and privacy features provided by pseudonyms in the communications. Many assume that pseudonyms will be readily available when communications need them. [4] proposes an architecture which touches issues related to pseudonyms like organizational structure, format, and the need for periodic refill. Recently, [11] proposes an interesting approach which enables vehicles to generate their own pseudonyms using group signature. However, there are still open research questions on group signature (e.g., dynamic group formation, efficiency, and revocation etc.). This approach is orthogonal to ours. The revocation of pseudonyms poses a challenge to the security in VN. Our approach tries to minimize the need for pseudonym revocation. However, in case the revocation of some of the pseudonyms becomes necessary, the protocols introduced in [12] can be applied, which detect and evict malicious nodes from VN. A similar approach to use cryptographic means to protect vehicle privacy is also proposed in [13], which uses a symmetric encryption based protocol to create vehicle mix zones at road intersections. The effectiveness of using intersections as mix zones has been studied in [14]. Alternatively, Gerlach [15] proposes to use context-aware pseudonym change to create mix zones among vehicles in VANETs. In contrast to the overall public-key cryptography based approaches, [16] proposes a symmetrickey based scheme for security and privacy in VC. Key distribution is burdensome in symmetric cryptographic system. The schemes proposed in this paper can be complementary for the symmetric key distribution.

VII. CONCLUSION AND FUTURE WORK

In this paper, we identify two pseudonym refill strategies for secure and private vehicular communications – the traditional one of refilling pseudonyms occasionally and a new strategy which allows vehicles to frequently refill pseudonyms with short lifetime. Our comparison shows that depending on the particular deployment scenario in future vehicular networks, the new strategy can minimize the need for pseudonym revocation and enhance security and privacy in the vehicular communications. To implement this refill strategy we propose a basic pseudonym-on-demand (POD) scheme and the derived privacy-enhanced scheme. We show that such schemes are able to strengthen the functionalities of pseudonyms in terms of security and privacy in vehicular communications.

Our next step will be to optimize and enhance the bandwidth efficiency and robustness of the POD scheme, and provide a rigorous evaluation of the privacy level of the system. Another ongoing work is to develop further *plug-in* features to enhance security and privacy in vehicular communications based on the POD scheme.

REFERENCES

- B. Parno and A. Perrig, "Challenges in securing vehicular networks," in *HotNets-IV*, Nov. 2005.
- [2] J.-P. Hubaux, S. Čapkun, and J. Luo, "The security and privacy of smart vehicles," *IEEE Security and Privacy*, vol. 4, no. 3, pp. 49–55, 2004.
- [3] Frank Kargl, Zhendong Ma, and Elmar Schoch, "Security engineering for VANETs," in ESCAR'06, November 2006.
- [4] P. Papadimitratos, L. Buttyan, J-P. Hubaux, F. Kargl, A. Kung, and M. Raya, "Architecture for secure and private vehicular communications," in *ITST2007*, June 2007.
- [5] M. Raya and J.-P. Hubaux, "Securing vehicular ad hoc networks," *Journal of Computer Security, Special Issue on Security of Ad Hoc and Sensor Networks*, vol. 15, no. 1, pp. 39 – 68, 2007.
- [6] E. Fonseca, A. Festag, R. Baldessari, and R. Aguiar, "Support of anonymity in vanets - putting pseudonymity into practice," in WCNC 2007, Hong Kong, March 2007.
- [7] "DSRC (Dedicated Short Range Communications)," http://www. leearmstrong.com/DSRC/DSRCHomeset.htm.
- [8] M. Raya and J.-P. Hubaux, "The security of vehicular ad hoc networks," in SASN 2005, Alexandria, USA, Nov. 2005.
- [9] ITU-T, "ITU-T Recommendation X.509."
- [10] A. Beresford and F. Stajano, "Mix zones: user privacy in location-aware services," in *PerCom Workshops*, 2004, pp. 127–131.
- [11] G. Calandriello, P. Papadimitratos, J.-P. Hubaux, and A. Lioy, "Efficient and robust pseudonymous authentication in vanet," in VANET '07, NY, USA, September 2007, pp. 19–28.
- [12] Maxim Raya, Panagiotis Papadimitratos, Imad Aadz, Daniel Jungels, and Jean-Pierre Hubaux, "Eviction of misbehaving and faulty nodes in vehicular networks," *IEEE JSAC, Special Issue on Vehicular Networks*, vol. 25, no. 8, pp. 1557 – 1568, 2007.
- [13] Julien Freudiger, Maxim Raya, Mark Felegyhazi, Panos Papadimitratos and Jean-Pierre Hubaux, "Mix-zones for location privacy in vehicular networks," in *WiN-ITS*, 2007.
- [14] Levente Buttyan, Tamas Holczer, and Istvan Vajda, "On the effectiveness of changing pseudonyms to provide location privacy in vanets," in ESAS 2007, July 2007.
- [15] Matthias Gerlach, "Assessing and improving privacy in VANETs," in ESCAR'06, November 2006.
- [16] J. Y. Choi, M. Jakobsson, and S. Wetzel, "Balancing auditability and privacy in vehicular networks," in *Q2SWinet* '05, 2005, pp. 79–87.