Strong and Affordable Location Privacy in VANETs: Identity Diffusion Using Time-Slots and Swapping

David Eckhoff*, Christoph Sommer*, Tobias Gansen†, Reinhard German* and Falko Dressler*
*Computer Networks and Communication Systems, Dept. of Computer Science, University of Erlangen, Germany
†Audi Electronics Venture GmbH, Gaimersheim, Germany
{eckhoff,csommer,german,dressler}@informatik.uni-erlangen.de, tobias.gansen@audi.de

***Abstract*—Public acceptance, and thus the economical success of Vehicular Ad Hoc Networks (VANETs), is highly dependent on the quality of deployed privacy mechanisms. Neither users nor operators should be able to track a given individual. One approach to facilitate this is the usage of pseudonym pools, which allow vehicles to autonomously switch between different identities. We extend this scheme with that of a time-slotted pseudonym pool of static size, reducing the storage and computation needs of the envisioned Intelligent Transportation System (ITS) while further improving users' privacy. In addition, we allow the exchange of pseudonyms between nodes, eliminating the mapping between vehicles and pseudonyms even for operators of the VANET. Here, we support the exchange of both the currently used pseudonym and those of future time-slots, further enhancing users' privacy. We evaluate the feasibility of our approach and back up privacy claims by performing a simulative study of the system using the entropy of nodes' anonymity sets as the primary metric.**

I. Introduction

Intelligent Transportation Systems (ITSs) offer a wide range of services based on wireless communication. Many of these, such as location-based services or driving assistance, require the exchange of positions and identifiers of cars in the vicinity to operate. A commonly used method in VANET is that vehicles send periodic beacon messages to inform other entities about their current state and position. An adversary is thus able to track a single entity throughout the system just by overhearing communication, collecting and then aggregating this data [1]. This can severely compromise the privacy of users, because a car is usually only driven by very few different drivers [2]. However, even if the location is not included in these messages, the position of a sending node can be determined with sufficient precision by other vehicles and Roadside Units (RSUs), using triangulation or simple range estimations. This allows an operator or any other user to create accurate traces of all participants if the number of observations is high enough [3].

A common approach to avoid this problem is the use of multiple pseudonyms instead of static identifiers [4], [5]. Nodes can then autonomously change their identifier, which is also used for directed communication, to complicate tracking of their positions. This approach, however, has several weaknesses. First, simply changing the pseudonym regardless of nodes' current status – meaning their current positions, speeds, headings and the amount of cars in transmission range – has been shown to only yield small improvements in terms of location privacy [6]. The second problem is the feasibility in a real environment. Systems usually employ an additional base identity in order to restrict the participation of a vehicle in the network. Without a mechanism like this, it is difficult to prevent freeloaders or known adversaries from participating in the system. The base identity is then used to request pseudonyms from a central authority. This is commonly realized by cryptography, i.e. the use of a Public Key Infrastructure (PKI). Identities and pseudonyms are certificates that are only valid if they are signed by a root Certificate Authority (CA) and only for a limited time.

From this it follows that each node in the network has to be equipped with a large number of pseudonyms, so that if (a) the CA is not reachable due to lack of connectivity or (b) the car was not used for a longer time period, the vehicle can still send messages until the CA supplies new pseudonyms. This can, of course, still not guarantee that the car can instantly participate in the Vehicular Ad Hoc Network (VANET): After long disconnection times, there might simply be no more valid pseudonyms available. A larger number of pseudonyms stored on each node can therefore decrease the possibility of a car not being able to transmit messages, but the required disk space, transfer volume, and management costs will also significantly increase. Moreover, while frequent pseudonym changes, for example one change every 60 s, will increase the privacy enjoyed by a node, this will require the node to request a huge number of pseudonyms from the CA. If the network grows there will be a noticeable computational and network overhead just to equip all nodes with a sufficient number of pseudonyms.

Aside from this, the PKI approach also enables the CA to resolve any pseudonym to the base identity with which it was requested. This means that every node could be tracked throughout the network by the operator of the PKI. This is a major violation of drivers' privacy and, depending on the trustworthiness of the operator, can be highly undesirable. Though split knowledge dual control schemes exist to counter this drawback, their application is, by design, optional with no way of allowing users to check whether their policies are actually enforced [7], [8].

We contribute to the state of the art by developing a scheme that offers both low-bandwidth pseudonym management and unlinkability of pseudonyms, thus, by design, providing strong privacy for all participants in the VANET. In order to achieve this, we employ time-slotted pseudonym pools, which signifi-

978-1-4244-9524-5/10/$26.00 ©2010 IEEE

cantly reduce network and computational load for the operator, and introduce static upper-bounds for disk space usage and communication overhead between vehicles and CA. In addition, we combine this approach with the concept of pseudonym exchange of both the currently used pseudonym and those of future time-slots to further improve the level of privacy enjoyed by drivers and to counter the ability of system providers to map pseudonyms to unique base identifiers (Section II).

We present a communication protocol followed by a discussion of problems and possible attacks (Section III). We evaluated the offered privacy using nodes' entropy. As can be seen from the results, the achieved entropy is much higher than in related approaches. We show that our pseudonym exchange scheme is a feasible approach for VANETs (Section IV).

II. Time-slotted Pseudonym Pools and Pseudonym Swapping

Instead of having a very large amount of pseudonyms, every node has a time-slotted pseudonym pool with slot length t, so that $\frac{p}{t}$ time-slots cover the total period length p. For each time-slot, there is exactly one assigned pseudonym, resulting in $\frac{p}{t}$ pseudonyms per car, and only one valid pseudonym for every arbitrary point in time. When a time-slot has passed, each node will change its pseudonym. This can be achieved by clocks roughly synchronized with the GPS signal.

While the use non-overlapping pseudonyms, as also proposed in [9], is very similar to time-slots, nodes in our scenario will reuse pseudonyms. When the last $\frac{p}{t}$th time-slot has passed, time-slot 1 will become active again, meaning that the time period will simply restart from the beginning.

A straightforward choice for those values, t = ten minutes and p = one week, results in a pseudonym being valid for, e.g., Monday from 6:00 a.m. till 6:10 a.m. Note that this pseudonym is then, in fact, valid on *every* Monday for said ten minutes. It can be seen that location privacy with a time-slotted pool alone relies on the time-slot length t, which determines how often a node changes its pseudonym, but also on the extent to which a node exhibits periodic behavior, e.g., starting the commute to work every Monday at 6:00 a.m.

Li et al. have shown that the exchange of pseudonyms can increase privacy in Mobile Ad Hoc Networks (MANETs) and complicate tracking for an adversary [10]. If nodes are able to exchange their pseudonyms in secrecy by using encryption and to keep third parties from tracking which nodes have swapped pseudonyms, a possible mapping at an authority will also become invalid. Due to the time-slotted pseudonym scheme, only pseudonyms valid for a specific time-slot can be exchanged, otherwise it cannot be guaranteed that every vehicle has exactly one pseudonym per time-slot. This means that, I_n being the pseudonym valid for time-slot n, vehicles v and v' must only exchange pseudonyms I_n with I'_n.

When it comes to swapping the currently used pseudonym, it is a difficult task to choose partners in a way that will benefit the privacy of the participants. For example, two cars passing, each one going in a different direction, will most likely not increase their anonymity by swapping pseudonyms because this action could be easily detected due to the unlikeliness of both cars having turned around at the same time. Gerlach and Guettler have therefore proposed to take context information of a node into account [11]. This means that a node evaluates its environment (such as number, speed and heading of its neighbors) and then decides if changing its pseudonym is profitable, so an adversary cannot simply infer the nodes' pseudonyms after the exchange by extrapolating their expected position based on their last known heading and speed [6].

Algorithm 1 Request exchange of a pseudonym

Require: Beacon Message of v' received **and** not ignored
  ignore beacons of v' for 20 s
  if v_{speed}, $v_{heading} \approx v'_{speed}$, $v'_{heading}$ **and** $swap = true$ **then**
    request exchange of current $I_n = I_{now}$
    $swap \leftarrow false$ for 60 s if swap(v') was true
  else if traceable pseudonyms in pool **then**
    request exchange of random $I_n \in$ traceable
    $I_n \leftarrow$ not traceable if successful
  else
    request exchange of random $I_n \neq I_{now}$
  end if

In our approach, we use the speeds, headings, and positions of other vehicles to determine whether a node v will ask a node v' in its vicinity to swap the currently valid pseudonym. In the scope of this paper, we refer to all nodes meeting these requirements as *candidates*.

By carefully choosing bounds for similarity, we increase the likelihood of both exchange partners being indistinguishable in terms of position. An adversary can then never be sure whether a pseudonym exchange has taken place or not. The efficiency of this scheme, of course, is highly dependent on the frequency and positional accuracy of the beacons each car emits. The privacy achieved by this approach could thus be amplified by using random silent periods [12], meaning both cars will not send beacons for a certain amount of time after a possible exchange.

However, one problem remains: If vehicles only exchange currently valid pseudonyms, that is, their current identifier, each vehicle will start using the same pseudonym every $\frac{p}{t}$ slots, because once a new slot $n+1$ has begun, the pseudonym last used in slot n will not be touched or exchanged again until slot n will be active again. This way an attacker, or the authority, is able to link two locations to one node: the present one (e.g., this Monday 6:00:00 a.m.) and the one from the last time the time-slot was active (e.g., last Monday 6:09:59 a.m.). Furthermore, each time a car enters a time-slot for the first time, which will happen $\frac{p}{t}$ times after being equipped with the ITS device, the operator can link the first location in this time-slot to a car. It has been shown that accumulated information about vehicles can be used to create traces and profiles for a user [13].

Therefore, cars have to be able to exchange these pseudonyms *before* actually using them. To achieve this, each time a time-slot ends, the last used pseudonym is marked as *traceable*.

Algorithm 2 Respond to pseudonym exchange request

Require: v' requests exchange of I_n
  ignore beacons of v' for 20 s
  if $I_n \neq I_{now}$ **then**
    exchange pseudonym I_n
    $I_n \leftarrow$ not traceable
  else if rand() < 0.5 **and** $swap = true$ **then**
    exchange pseudonym I_{now}
    $swap \leftarrow false$ for 60 s
  else
    $swap \leftarrow false$ for 60 s if $swap$ was true
    exchange random pseudonym from pool
  end if

Similarly, all pseudonyms that are freshly obtained from the CA are marked *traceable*. When a node encounters another node, it may then either exchange the current pseudonym, or one marked *traceable*, removing the flag if successful. As the exchange does not affect the current pseudonym of a node and the attacker cannot see which pseudonyms were exchanged, constraints like speed or heading of node v' can be ignored and the acceptance of the request does not need to be probabilistic but is defined to be always positive. Algorithm 1 shows the simplified working principle for requesting a pseudonym exchange. As can be seen, a node decides dependent on the speed and heading of another node whether to request the exchange of the currently used pseudonym or another (preferably traceable) pseudonym.

However, Schoch et al. have shown that too frequent pseudonym changes have a negative impact on geographic routing in VANETs [14]. We therefore set the Boolean variable $swap$ to $false$ to suppress exchanges of the current pseudonym, but only when node v' was also ready to exchange the current pseudonym, that is, variable $swap$ at node v' was $true$. The state of swap(v') prior to the request is communicated in an response packet. After 60 s the value is reset to $true$, which means that a node is again able to request exchange of the currently valid pseudonym. In addition, to avoid overloading the network, a node v must only contact v' every 20 s.

To ensure secure communication for private key data and to counter overhearing, nodes will build up a secure channel with their current pseudonyms using cryptography as proposed in WAVE [15]. To erase predictability, we introduce a probability whether a node will send a positive response to a request for the current pseudonym. This probability is set to 50 %, meaning if a node v asks for pseudonym exchange, node v' will accept the request as often as it rejects it, when it was allowed to exchange the current identifier, that is, if $swap$ was $true$. Even if the request is rejected, node v' and v will exchange another pseudonym so that an attacker cannot determine if the nodes have swapped their current pseudonyms from the overheard transfer. Algorithm 2 shows the response of a node if another node requests the exchange of a pseudonym.

Figure 1 depicts possible flows of the pseudonym exchange process. Vehicle A requests an exchange of the currently valid

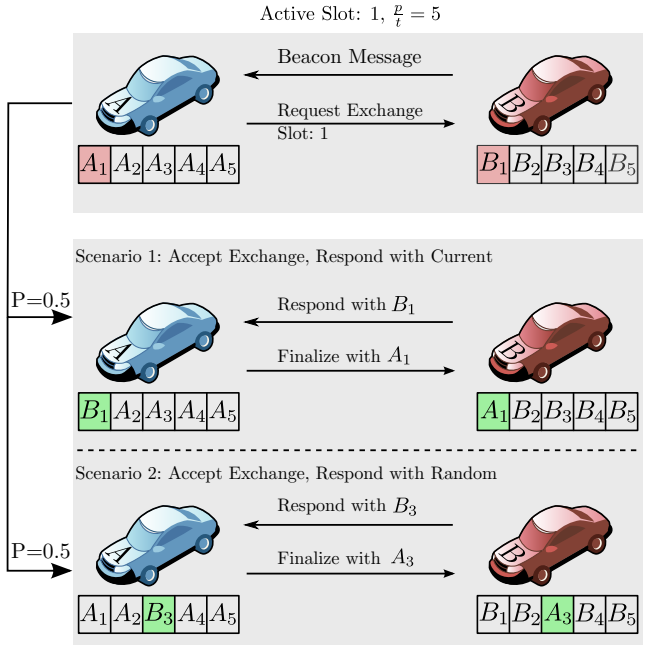


Figure 1. Pseudonym exchange between two cars: The currently valid pseudonym is requested and confirmed in scenario 1 (resulting in the change of the current pseudonym), but rejected and answered with a random pseudonym from the pool in scenario 2

pseudonym from vehicle B, because both vehicles happen to have similar values for heading and speed. In 50 % of all cases B will respond with the currently active pseudonym and A will finalize the exchange process by handing over its current pseudonym as well. The vehicles will then use the new pseudonyms. In the other 50 % of cases, vehicle B will not exchange its currently valid identifier but rather respond with another one from its pseudonym pool, preferably one marked as traceable. Vehicle A will accept this, and answer with the corresponding pseudonym from its own pool. Both vehicles will replace their old pseudonym for the given slot with the one from the other node.

III. Benefits and Limitations

An advantage of the time-slotted approach over huge pseudonym pools is its property to ensure that, ideally, a vehicle always has a pseudonym to participate in the ITS as long as it has received its $\frac{p}{t}$ pseudonyms in the setup phase. Even if the CA is not reachable or the car was not used for a longer time period the vehicle will not run out of pseudonyms because it can reuse the old ones.

In addition our scheme introduces upper limits for disk space and, more importantly, traffic volume. This simplifies the design of on-board units and also reduces the communication costs, making the deployment of an ITS more affordable. The pseudonym pool size is reduced to a constant value of $\frac{p}{t} \times s$ bytes and, more importantly, the workload at the CA is no longer dependent on the number of nodes actually participating in the network but rather on the ones joining it.

Using time-slots and synchronized clocks, every node

will change its pseudonym at the same time. Depending on penetration rate and traffic density, this can increase drivers' privacy, as we will show in our evaluation. By further applying a pseudonym exchange scheme, the privacy of users is significantly increased. Allowing the exchange of current and future pseudonyms eliminates the mapping at an authority and allows nodes to start new time slots already anonymous.

Accountability in pseudonym exchange environments remains an open problem. Therefore, the use of our scheme should be limited to non-safety critical messages to avoid misuse. The class of 'critical safety messages' includes messages such as accident (e.g., triggered by airbag release) and emergency break messages. We argue that for non-critical service messages, but also for periodic beaconing, preservation of unlinkability and privacy is more important than accountability.

Without further installing a mechanism for revocation in pseudonym exchange environments, it is not possible to revoke all pseudonyms of a vehicle to fully keep it from participating without relying on additional information such as camera footage or license plate snapshots. If one is willing to give up full unlinkability and enable third parties to cooperatively resolve pseudonyms to base identities, our scheme could be extended: Vehicles periodically upload a log file that contains every pseudonym exchange to a server. Combining information from many log files, misbehaving vehicles can then be identified with high likelihood. This will be the focus of future research.

While, by design, in our scheme every node has only one valid pseudonym for any point in time, the use of tamper proof devices is crucial. Tampered on-board units could be configured not to delete old pseudonyms after exchanging them with another node, allowing an adversary to build up a pool of many pseudonyms, all valid for the same time-slot.

IV. Evaluation

There exist different metrics to measure the level of location privacy enjoyed by an individual in a network [16]. Anonymity, in our case the precondition for location privacy, is interpreted by Pfitzmann and Hansen as the "state of being not identifiable within a set of subjects, the anonymity set" [17]. The anonymity set A hence contains all nodes in the network that could possibly be a targeted individual P. However, in our network, not all members of A are equally likely to be this individual P, meaning the cardinality of the anonymity set $|A|$ alone is not a sufficient metric to measure the location privacy enjoyed by P. Instead we use the entropy $\mathcal{H}_p$ of the anonymity set A, which can be seen as the uncertainty in determining the current identifier of individual P [18].

In order to calculate the entropy, let p_i be the probability of node i to be the target P, $\forall i \in A$ and let further

$$\sum_{i=1}^{|A|} p_i = 1 \ .$$

The entropy $\mathcal{H}_p$ of identifying a target P in the anonymity set is then defined to be

$$\mathcal{H}_p = -\sum_{i=1}^{|A|} p_i \times \log_2 p_i \ . \tag{1}$$

Based on this, the upper limit of $\mathcal{H}_p$, the maximum value of entropy, can be calculated as

$$\mathcal{H}_{p_{max}} = -\sum_{i=1}^{|A|} p_i \times \log_2 p_i \ = \ \log_2 |A| \quad \text{if } \forall i : p_i = \frac{1}{|A|} \ . \tag{2}$$

From this it follows that for $\mathcal{H}_p = \mathcal{H}_{p_{max}}$ all entities $i \in A$ have to be equally likely to be individual P. However, this is almost impossible to achieve in a VANET, because nodes may contact a large number of other nodes and the relation v *has met* v' is rarely transitive.

A simple example shows how to interpret entropy values for a given individual P. Assume an attacker is not sure whether P uses identifier v or v' and that both nodes are equally likely to be driven by P, then the anonymity set for P is $A = \{0.5, 0.5\}$. The entropy $\mathcal{H}_p$ for P is thus 1.

On the other hand, if P is more likely to be the driver of v than v' with a certainty of 80 %, then the anonymity set would be $A = \{0.8, 0.2\}$ and the resulting entropy $\mathcal{H}_p \approx 0.72$.

If we were to consider three nodes, each equally likely to be the target, the anonymity set is $A = \{0.33, 0.33, 0.33\}$ and the entropy $\mathcal{H}_p = 1.5$.

A. Attacker Model

The evaluated level of location privacy enjoyed by an individual is always relative to the power of an attacker trying to track this person in the network. In our simulations, we assume a global passive attacker, that is, an attacker that is able to overhear *every* message sent in the network. The attacker is further able to evaluate the content of all broadcast beacon messages (which we assume to include the speed, position and heading of a node). As the attacker is, of course, well aware of the protocol, it is able to conclude which nodes might exchange of their current pseudonyms.

The attacker is, however, not able to actually follow the pseudonym exchange, as all of these messages are encrypted using public key cryptography. All the attacker can gather from observing transmissions in the network is the fact that pseudonym requests and replies have been exchanged.

Furthermore, our attacker model is based on the strong assumption that at the beginning of the lifetime of node v, the attacker has full knowledge of all mappings between vehicles v and individuals I. If this was not the case, the individual would already be anonymous from the start and could only be exposed through origin/destination pairs if tracking throughout the network was successful.

When modeling an attacker using tracking algorithms, the apparent strength of the attacker is heavily dependent on the used mobility and driver model. If, for example, nodes do not change lanes or drive in a very predictable manner, tracking algorithms will perform significantly better. Therefore we choose to use a probabilistic attacker model:

Figure 2. Region of Interest for the urban scenario

As we have shown, the entropy is based on $p_i \forall i \in A$. However, the distribution of p_i is directly dependent on the attacker strength α. The attacker strength $\alpha \in [0.5, 1]$ is defined as the probability with which an attacker is able to follow a pseudonym exchange between two nodes. The weakest possible attacker in our scenario would thus be an attacker with $\alpha = 0.5$, meaning after a pseudonym exchange between v and v', the individual I, previously known to be the driver of v, is now equally likely to be the driver of v or v'. The strongest possible attacker is an attacker with $\alpha = 1$. This attacker is not confused by pseudonym exchanges and is therefore able to track every entity throughout the network. Obviously, the entropy $\mathcal{H}_p$ for each individual in the network would then be zero.

The attacker strength α also affects by how much the level of privacy is increased when a new slot in the slotted pseudonym pool becomes active, that is, when all nodes will start using new pseudonyms. If we assume that two nodes very close to each other can confuse an attacker by exchanging their pseudonyms (the extent being dependent on its strength), this attacker will also be confused when these two nodes both switch to a new pseudonym simultaneously. From this we follow that the level of confusion is based on the amount of candidates directly neighboring a node. NB: Not all cars within transmission range of v are considered *candidates*, but only those with similar speed, heading and position. A node will never request the exchange of the currently valid pseudonym from a non-neighbor.

In our simulation experiments, we used the following metric to determine the probability φ_v of an attacker successfully tracking a node v beyond a slot change, with $n(v)$ being the amount of candidates near v:

$$\varphi_v = \frac{\alpha}{\alpha + (n(v) \times (1 - \alpha))} \tag{3}$$

From this it follows that when three nodes v_1, v_2, and v_3 are all very close to one another and a new slot begins, the weakest attacker with strength $\alpha = 0.5$ will be able to determine with probability $\varphi_{v_1} = \varphi_{v_2} = \varphi_{v_3} = 0.33$ which node was which. Consequently, the strongest attacker with strength $\alpha = 0.95$ can link new identifiers to old ones with $\varphi_v \approx 0.90$ in this particular case.

B. Simulation Setup

We investigated our scheme with the help of our *Veins*[1] simulation environment [19], [20], which is based on two simulation toolkits, both well established in their respective domain. Highly detailed vehicular mobility models, in particular with regard to intersection management, were provided by SUMO, a dedicated traffic micro simulation toolkit from the domain of traffic engineering. We further implemented the presented protocol for pseudonym exchange in the network simulator OMNeT++ using its INET Framework extension to simulate wireless transmissions.

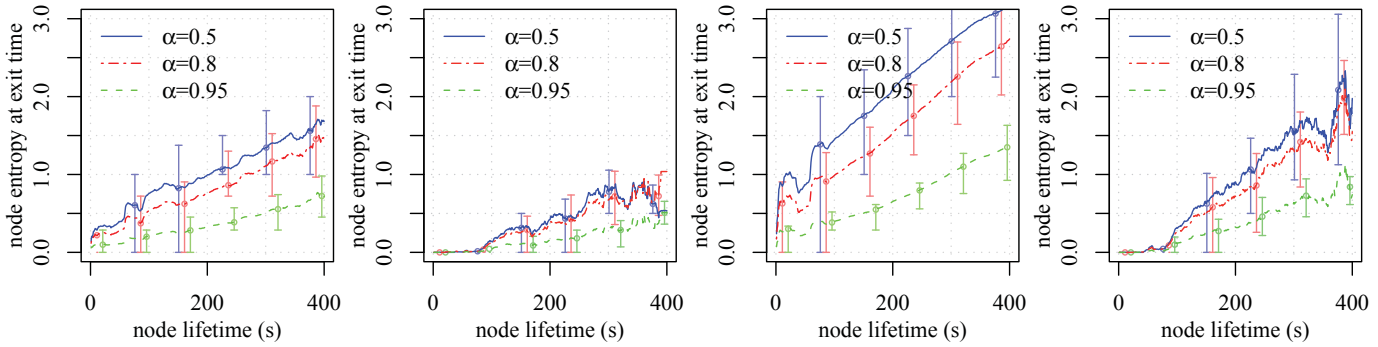
For the evaluation, we chose the following protocol parameters: A node may not change its current pseudonym more often than once every 60 s. Each node will only contact an already contacted node if 20 s have passed. The pseudonym pool length p is set to 1 week, the slot length 10 min. Cars are considered to be eligible for exchange of the current pseudonym, or *candidate*, when their speed difference is at most $10\,\mathrm{km/h}$, the difference in heading is at most $15°$ and their distance is at most $30\,\mathrm{m}$. The beacon frequency for each node is 5 s. NB: This does not affect the achieved level of privacy in our simulation as we used a stochastic attacker model. We simulated over 350 h of traffic with a total of over 1 500 000 cars until the margin of error was low enough. We evaluated the proposed scheme in an urban scenario as well as in a freeway setup, the latter one is only briefly described in the paper.

The urban scenario models traffic in the city of Ingolstadt. The road network itself was based on data by the Open Street Map project[2], adapted to reflect realistic intersection management. Traffic was created by randomly generating O/D pairs and iteratively applying dynamic user assignment [21], as implemented in SUMO, until the algorithm reported a stable, optimal distribution of flows. In the evaluation, we focus on the 4 km^2 Region of Interest (ROI) shown in Figure 2, which contains a typical mixture of high- and low-capacity roads, traffic lights, and unregulated intersections, as well as high- and low-density areas. While traffic is simulated in the whole city of Ingolstadt, we apply our privacy scheme only to nodes within the ROI.

We compared the measured vehicle densities with values provided by the local authorities. It turned out that even the high density scenarios match to sparse real traffic patterns.

To calculate the communication overhead caused by our privacy scheme, we base the amount of data needed for pseudonym exchange on the proposed algorithms and certificate lengths in the IEEE WAVE draft [15]. To setup a secure channel a slightly modified version of the NIST 800-38C [22] is used. We assume a certificate length of 288 B with asymmetric key length of 1024 bit and a symmetric key length of 128 bit for the *aes_128_ccm* scheme. From this, we conclude that the traffic needed for the exchange of a pseudonym, including IP overhead is roughly 1 KiB, that is, 0.5 KiB per node. Note that we disregard beacon messages in these calculations, since we

[1] http://www7.informatik.uni-erlangen.de/veins/
[2] http://www.openstreetmap.org/



(a) 16 cars/km^2 slotted exchange (b) 16 cars/km^2 random change (c) 100 cars/km^2 slotted exchange (d) 100 cars/km^2 random change

Figure 3. Comparison of node entropy in the 4 km^2 urban scenario with 25 % and 75 % quartiles

consider them to be part of the ITS service provision, not of the privacy mechanism.

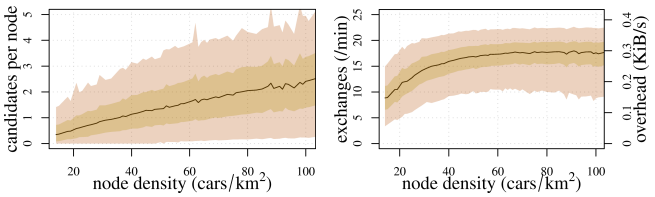
C. Results

1) Urban Scenario: The results for the simulation of the urban scenario are depicted in Figure 3. We compared our scheme with an approach that uses random pseudonym change with cooperative silent periods [12]. The random pseudonym change does not offer unlinkability between pseudonym and base identity. Cars will randomly change their pseudonyms and the effect on the gained privacy is dependent on nearby cars.

We observe nodes moving through the ROI and calculate the entropy resulting from pseudonym exchanges and slot changes. For both densities we observe that with a medium-strength attacker the privacy enjoyed by drivers is not considerably smaller than with the weak attacker. However, it takes about 60 s longer to reach the same level of privacy. This is exactly the time a pseudonym has to stay active. We therefore conclude that in order for users to gain approximately the same level of anonymity for the $\alpha = 0.8$ attacker as for the $\alpha = 0.5$ attacker, one additional pseudonym exchange is needed.

Comparing results for the strong attacker vs. the medium-strength attacker, we see that the difference in entropy is significantly bigger. In contrast to the high density scenario (Figure 3d), the mean entropy for nodes does not exceed 1 in the lower density setup (Figure 3a). For the weaker attackers the entropy reaches a level that can be considered to provide anonymity after about 300 s. In both scenarios, the pseudonym exchange approach performs considerably better than random pseudonym changing, with additionally providing unlinkability between pseudonyms and base identities.

The bumps at about 40 s and 90 s can be explained by the topology of our region of interest. Two highly frequented roads, one in the lower left and one in the middle, cut the ROI. It took nodes about 40 s and 90 s, respectively, to pass these roads. The set of cars with these lifetimes therefore includes a significant amount of cars with higher privacy levels, since on busy roads nodes will find potential partners for pseudonym exchanges more easily.



(a) Median candidates per node in the urban scenario. Overlaid are the 25 % and 75 % quartiles and 5 % and 95 % quantiles, respectively

(b) Exchanges of non-current pseudonyms per node in the urban scenario with resulting traffic overhead. Overlaid are the 25 % and 75 % quartiles and 5 % and 95 % quantiles, respectively

We conclude that in a low-density environment trips shorter than 400 s can be tracked by attackers if they are able to follow pseudonym exchanges with high probability and if the mapping of individual I to node v is known at the time of departure. In high-density scenarios with weak or medium-strength attackers, drivers will enjoy a sufficient level of anonymity after only 1-2 min, because the probability of finding a suitable node for pseudonym exchange will rise with the node density. We conclude that trips exceeding 5 min cannot be tracked.

We measured the number of nodes suitable for exchange of the current pseudonym, the *candidates* of a node, according to our simulation setup parameters (speed difference $\leq 10\,\mathrm{km/h}$, heading difference $\leq 15°$, distance $\leq 30\,\mathrm{m}$). As can be seen in Figure 4a in scenarios with density $\leq 40\,\mathrm{cars/km^2}$ most of the nodes are only very infrequently able to find one or more candidate. As expected, the number of candidates rises with the density. With 70 cars/km^2, 75 % of all nodes frequently have one or more nodes suitable for pseudonym exchange nearby. The 5 % quantile is still very low for the 100 cars/km^2 scenario, because there are always nodes traveling on infrequently-used streets, e.g., in residential neighborhoods. It should be pointed out that, even though finding a suitable node for pseudonym exchange was already very likely in higher density scenarios, it will even be more likely in real world scenarios, which frequently exhibit even higher node densities.

Figure 4b shows the number of exchanged non-current pseudonyms per minute, that is, pseudonyms for slots other

than the currently active one. One might expect that with higher density, the amount of pseudonym exchanges also rises. However, as can be seen, exchanges only marginally rise for scenarios with densities higher than $60\,\mathrm{cars/km^2}$. With a beacon frequency of only 0.2 Hz and one allowed connection between two nodes in 20 s the slope is significantly lower than expected. The reasons for this are twofold: First, with more nodes in the network, the concurrency of nodes reacting to a beacon message will also increase. That is, new nodes do not only offer more possibilities to exchange a pseudonym, but also compete for requesting exchange from other nodes. Secondly, the more significant reason is that cars preferably exchange their current pseudonym than pseudonyms from other slots. With higher node densities, nodes will find suitable partners for exchanging their current identifier more easily as previously shown in Figure 4a. This also explains the slightly declining 5 % quantile in higher density scenarios.

As expected, the traffic overhead caused in the wireless network by our scheme is insignificantly lower (Figure 4b). It did not exceed an average of 0.5 KiB/s and can therefore be deployed in VANETs without restriction.

Extrapolating our results, the observed pseudonym exchange rates meet a rate of 1200 pseudonyms/h. Assuming that, in a worst case scenario, traceable pseudonyms are only exchanged when the node carrying it initiates the pseudonym exchange, it would take less than 2 h to exchange the whole pseudonym pool. After this time period a node would only carry untraceable pseudonyms and already be completely anonymous when a new slot begins.

2) Freeway Scenario: We found that in a freeway scenario the entropy of nodes increases almost linearly with the lifetime of cars on the freeway. This is caused by nodes almost immediately finding a suitable candidate for pseudonym exchange on freeways. Our findings suggest that after 10 min on a freeway, even in sparse scenarios with a strong attacker, pseudonym exchanging cars have reached a sufficient level of privacy (data not shown).

V. Related Work

There are several approaches to protect the privacy of users in a MANET. Li et al. present a user-centric approach for updating pseudonyms based on velocity and direction changes with respect to nodes in their neighborhood to complicate tracking [10]. They also propose *Swap*, a scheme for exchanging pseudonyms among nodes in mobile networks, however, their approach has not been evaluated for VANETs. They assume that after a successful exchange nodes are always indistinguishable from each other. They combine both approaches and conclude that their scheme offers better privacy than random pseudonym updates.

Buttyan et al. examine the effect of frequent pseudonym changes on location privacy in mix zones [23], which were first proposed in [24]. For the evaluation of their approach they generate traffic with realistic parameters on non-trivial road maps and assume that attackers position antennas in the network to overhear communication.

Dötzer also uses a CA to equip vehicles with signed pseudonyms. To exclude misbehaving or malfunctioning nodes from the network he employs Certificate Revocation Lists (CRLs) that contain pseudonyms of these nodes [5]. There have also been approaches in which nodes that already have a signed base certificate do not request pseudonyms from a base authority, but rather create them autonomously [25].

Fischer et al. deal with the problem that a single CA might be able to resolve vehicles' identities and proposes the separation of a CA into a *Privacy Authority* and *Identity Authority*, with both sharing just parts of the identities. Only when both authorities cooperate they are able to resolve an identity [7]. This idea of *separation of concerns* is pushed further by Schaub et al., where the used pseudonyms itself contain all necessary information to resolve an identity if multiple authorities cooperate, thus greatly enhancing the scalability of the system [8].

Chaurasia and Verma, among others, use entropy and anonymity sets to measure privacy in VANETs [26]. They have shown that changing pseudonyms with regard to the state of other vehicles in the transmission range maximizes the size of anonymity sets and therefore maximizes the anonymity for each vehicle.

Huang et al. introduce random silent periods to increase location privacy: nodes updating their identifier do not send any messages for a given time period to complicate tracking for an adversary. They show that these periods can significantly improve the privacy of users in mobile networks [12].

Leinmueller et al. introduce a series of heuristics to detect fraudulent vehicles that broadcast fake positions [27]. Their approach can also be used to detect Sybil attacks in VANETs. They use a random waypoint model to simulate urban traffic and conclude that their model will drastically reduce the possibility of position forgery.

Wiedersheim et al. have shown that simple pseudonym change is not enough [3]. They show that using Multi-Hypothesis Tracking and Kalman filtering can lead to very high tracking success rates. However, high density spots, mostly found at intersections and traffic lights, can cause the tracking algorithm to fail. They also note that beaconing intervals have significant effects on the tracking accuracy and find that beacon intervals upward of 2 s complicate tracking for adversaries.

We extended the described pseudonym-based solutions by, first, providing a time-slotted scheme that features a fixed-size pseudonym pool and frequently changed identities. Secondly, our approach fosters the exchange of single pseudonyms among vehicles to further enhance the users' privacy.

VI. Conclusion

We presented a novel approach to increase the level of location privacy enjoyed by users in a VANET and to eliminate the mapping between pseudonyms and base identities at a central authority. We make use of a time-slotted pseudonym pool, in which for every time-slot there exists exactly one pseudonym. By using this method, the workload at the CA is

much less dependent on the number of nodes participating in the network, but rather on the rate of nodes joining it.

The synchronous change of identifiers increases the privacy of users that are close to other nodes in the network. To further increase anonymity and to keep a central authority from resolving pseudonyms to real identities of users, nodes exchange pseudonyms between one another. We showed the applicability of identifier exchange in vehicular environments and measured the resulting degree of privacy, using the entropy of nodes' anonymity sets.

Nodes exchange their currently active pseudonym if a node suitable for exchange can be found. Only nodes with similar speed, position and heading are considered suitable. We evaluated our approach in realistic scenarios with probabilistic attacker models and showed that suitable nodes can be found frequently in urban scenarios. We showed that even with attackers that can follow pseudonym exchanges with high probability (95 %) nodes can become sufficiently anonymous when moving in a high-density environment for about 5 min. Our approach works in both urban and freeway environments and scales with the lifetime of a node in the network.

Furthermore, the exchange of pseudonyms other than the current one makes it impossible for a central authority to resolve pseudonyms to identities even when a new time-slot just became active. We showed that with very low communication overhead nodes can exchange a sufficient amount of pseudonyms to swap all traceable pseudonyms for anonymous ones in short time periods.

Future work in this area will cover the combination of group building and pseudonym exchange to further increase the level of location privacy in VANETs and balancing this with accountability concerns in systems where identities can be swapped.

References

[1] P. Papadimitratos, A. Kung, J. Hubaux, and F. Kargl, "Privacy and Identity Management for Vehicular Communication Systems: A Position Paper," in *1st Workshop on Standards for Privacy in User-Centric Identity Management*, Zurich, Switzerland, July 2006.

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

[5] F. Dötzer, "Privacy Issues in Vehicular Ad Hoc Networks," in *5th International Workshop on Privacy Enhancing Technologies (PET 2005)*, Cavtat, Croatia, May 2005, pp. 197–209.

[6] K. Sampigethaya, L. Huang, M. Li, R. Poovendran, K. Matsuura, and K. Sezaki, "CARAVAN: Providing location privacy for VANET," in *Embedded Security in Cars (ESCAR 2005)*, Tallinn, Estonia, July 2005.

[7] L. Fischer, A. Aijaz, C. Eckert, and D. Vogt, "Secure Revocable Anonymous Authenticated Inter-Vehicle Communication (SRAAC)," in *4th Conference on Embedded Security in Cars (ESCAR 2006)*, Berlin, Germany, November 2006.

[8] F. Schaub, F. Kargl, Z. Ma, and M. Weber, "V-tokens for Conditional Pseudonymity in VANETs," in *IEEE Wireless Communications and Networking Conference (WCNC 2010)*. Sydney, Australia: IEEE, April 2010.

[9] M. Raya, R. Shokri, and J. Hubaux, "On the tradeoff between trust and privacy in wireless ad hoc networks," in *Proceedings of the third ACM conference on Wireless network security*. Hoboken, NJ, USA: ACM, May 2010, pp. 75–80.

[10] M. Li, K. Sampigethaya, L. Huang, and R. Poovendran, "Swing & swap: user-centric approaches towards maximizing location privacy," in *5th ACM Workshop On Privacy In The Electronic Society*. Alexandria, VA: ACM, October 2006, pp. 19–28.

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

[12] 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.

[13] Z. Ma, F. Kargl, and M. Weber, "Measuring location privacy in V2X communication systems with accumulated information," in *6th IEEE International Conference on Mobile Ad Hoc and Sensor Systems (MASS 2009)*, Macau SAR, China, October 2009.

[14] E. Schoch, F. Kargl, T. Leinmuller, S. Schlott, and P. Papadimitratos, "Impact of pseudonym changes on geographic routing in vanets," in *3rd European Workshop on Security and Privacy in Ad hoc and Sensor Networks (ESAS 2006)*, Hamburg, Germany, September 2006.

[15] "Wireless Access for Vehicular Environments," IEEE, Draft Standard 802.11p-D4.0, March 2008.

[16] C. Diaz, "Anonymity Metrics Revisited," in *Dagstuhl Seminar on Anonymous Communication and its Applications*, Dagstuhl, Germany, October 2005.

[17] A. Pfitzmann and M. Hansen, "Anonymity, Unlinkability, Unobservability, Pseudonymity, and Identity Management - A Consolidated Proposal for Terminology," TU Dresden, TR v0.28, May 2006.

[18] A. Serjantov and G. Danezis, "Towards an Information Theoretic Metric for Anonymity," in *2nd International Workshop on Privacy Enhancing Technologies (PET 2002)*, San Francisco, CA, April 2002, pp. 259–263.

[19] C. Sommer, Z. Yao, R. German, and F. Dressler, "On the Need for Bidirectional Coupling of Road Traffic Microsimulation and Network Simulation," in *1st ACM International Workshop on Mobility Models for Networking Research (MobilityModels 2008)*. Hong Kong, China: ACM, May 2008, pp. 41–48.

[20] C. Sommer, R. German, and F. Dressler, "Bidirectionally Coupled Network and Road Traffic Simulation for Improved IVC Analysis," *IEEE Transactions on Mobile Computing*, 2010, available online: 10.1109/TMC.2010.133.

[21] C. Gawron, "Simulation-Based Traffic Assignment – Computing User Equilibria in Large Street Networks," PhD Thesis, University of Cologne, 1999.

[22] M. Dworkin, "Recommendation for Block Cipher Modes of Operation: The CCM Mode For Authentication and Confidentiality," NIST Special Publication, Tech. Rep. 800-38c, 2003.

[23] L. Buttyán, T. Holczer, and I. Vajda, "On the Effectiveness of Changing Pseudonyms to Provide Location Privacy in VANETs," in *4th European Workshop on Security and Privacy in Ad hoc and Sensor Networks (ESAS 2007)*, Cambridge, UK, July 2007.

[24] A. R. Beresford and F. Stajano, "Location Privacy in Pervasive Computing," *IEEE Pervasive Computing*, vol. 2, no. 1, pp. 46–55, January 2003.

[25] G. Calandriello, P. Papadimitratos, J. P. Hubaux, and A. Lioy, "Efficient and robust pseudonymous authentication in VANET," in *4th ACM International Workshop on Vehicular Ad Hoc Networks (VANET 2007)*, Montréal, Canada, September 2007, pp. 19–28.

[26] B. Chaurasia and S. Verma, "Maximizing anonymity of a vehicle through pseudonym updation," in *4th International Conference on Wireless Internet (WICON 2008)*, Maui, HI, November 2008.

[27] T. Leinmüller, C. Maihöfer, E. Schoch, and F. Kargl, "Improved security in geographic ad hoc routing through autonomous position verification," in *3rd ACM International Workshop on Vehicular Ad Hoc Networks (VANET 2006)*, Los Angeles, CA, September 2006, pp. 57–66.