ARTICLE IN PRESS

Computer Networks xxx (2011) xxx–xxx

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

On the difficulty of achieving anonymity for Vehicle-2-X communication

Carmela Troncoso [a,*], Enrique Costa-Montenegro [b], Claudia Diaz [a], Stefan Schiffner [a]

[a] K.U. Leuven/IBBT, ESAT/SCD-COSIC, 3001 Heverlee-Leuven, Belgium
[b] Departamento de Enxeñería Telemática - Universidade de Vigo ETSE Telecomunicación, 36310 Vigo, Spain

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Privacy
Anonymity
Vehicle-2-X communications
IntelliDrive

ABSTRACT

Vehicle-2-X communications are hailed as the future to improve safety on the roads. Ensuring that messages sent by vehicles contain correct information is crucial to fulfill this objective, as misleading information could disrupt traffic and create potentially dangerous situations. Thus, Vehicle-2-X communication requires authentication to ensure that messages come from legitimate vehicles, and to identify vehicles that send misleading information. If a unique public key certificate per vehicle is used to authenticate messages, then the identification of misbehaving (or malfunctioning) vehicles is straightforward, and so is the revocation of their credentials. This solution however, offers no privacy protection to drivers, as the tracking of all the vehicles' movements is equally trivial. A privacy-preserving alternative is to authenticate messages using (unlinkable) one-time pseudonyms, but these protocols are computationally expensive and their certificate revocation process is more complex. Intermediate solutions that trade off privacy and efficiency are based on multiple certificates per vehicle, which may or may not be unique, that are reused to authenticate messages. In this work we analyze two such intermediate solutions that have been proposed by IntelliDrive, US Department of Transportation (DoT). We show that by exploiting the reuse of pseudonyms and spatio-temporal constraints the service provider is capable of tracking a large percentage of vehicles. Furthermore, we find that one of the schemes fails to provide privacy even if the adversary does not control the service provider and only listens to the communications of vehicles.

1. Introduction

IntelliDrive [3] is an initiative of the USA DoT supported by a coalition of Federal, state and local transportation agencies, trade associations, and vehicle manufacturers – such as BMW, Chrysler, Ford, Honda, Mercedes-Benz, etc. The IntelliDrive program is ultimately focused on deployment, and its results are used by several Federal associations (e.g. the National Highway Traffic Safety Administration) to assess the safety and effectiveness of V2X applications in order to decide whether to pursue a rulemaking process to require or encourage this technology on some or all vehicles [23].

IntelliDrive has designed a framework in which vehicles use dedicated short-range communications (such as the IEEE 1609 standard family [2]) to send messages to other vehicles and to the Roadside Equipment (RSE). One of the main concerns related to the deployment of V2X communication networks is to ensure that vehicles send accurate information. This is crucial to protect the physical safety of drivers, as misleading information on the status of the road may prove fatal. Therefore, the system must be capable of isolating misbehaving users (who have sent misleading messages), and must prevent those users from sending further messages by revoking their authentication credentials. A second concern relates to the protection of the drivers' privacy. The fact that vehicles interact with their environment may allow the service provider, or even passive eavesdroppers, to track vehicles and thus infer private

* Corresponding author. Tel.: +32 16321045.
E-mail addresses: Carmela.Troncoso@esat.kuleuven.be (C. Troncoso), kike@det.uvigo.es (E. Costa-Montenegro).

1389-1286/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.comnet.2011.05.004

Please cite this article in press as: C. Troncoso et al., On the difficulty of achieving anonymity for Vehicle-2-X communication, Comput. Netw. (2011), doi:10.1016/j.comnet.2011.05.004

information about their drivers [20,18,14,16]. A discussion on the consequences of losing location privacy can be found in [7].

The IntelliDrive framework attempts to reconcile authentication and privacy requirements by using anonymous certificates that allow vehicles to send authenticated messages without revealing their actual identity [10,11,26]. The goal of these anonymous certificates, optimized for scalability, easy certificate management and revocation, is to ensure that the IntelliDrive service provider (or other entities) cannot link the message received at a particular location to the vehicle that sent it, and thus to preserve the anonymity of drivers and prevent vehicle tracking. Moreover, in case of repeated misbehavior, malicious vehicles lose their ability to renew anonymous certificates and thus to send messages (once all their certificates have been revoked).

In this work we show that both schemes [10,11,26] fail to prevent the service provider from tracking vehicles and re-identifying their drivers. We present two attacks that exploit the time and location where anonymous certificates are (re-)used. We have tested the effectiveness of our attacks through software simulations, discovering that even in heavy traffic conditions the service provider succeeds in reconstructing most of the trajectories followed by vehicles. Further, we have also considered a weaker adversary, who does not know how sets of certificates have been distributed to vehicles (i.e., does not control the service provider and can only eavesdrop on the communications of vehicles). We find that one of the schemes, based on shared certificates [10,26], is also vulnerable to this adversary model. The attack reveals the sets of certificates associated with most vehicles, and recovers the trajectories that they have followed.

It has been shown that an individual can often be uniquely identified by just obtaining the approximate locations of her home and work place [14], and that these locations can be found by studying the movement patterns of vehicles [20] (e.g., vehicles are likely to spend the night at their home location). Thus, our results imply that drivers' privacy would be compromised if any of the two anonymous certificate schemes in [10,11,26] were deployed. The key feature that is exploited by our attacks is the reuse of certificates, suggesting that one-time pseudonyms are necessary to achieve a reasonable level of location privacy.

The rest of this paper is organized as follows. Section 2 discusses related work. Section 3 describes the IntelliDrive model and the studied anonymous certificate schemes. The adversary model and attack algorithms are explained in Section 4. We describe our experimental setup in Section 5 and the results of our evaluation in Section 6. We discuss in Section 7 the implications of our experimental results for real traffic scenarios and offer our conclusions in Section 8.

2. Related work

A large body of research [13,15,21,19] focuses on achieving location privacy by hiding the actual location of users from the service provider. In these approaches users communicate with the service provider through a trusted proxy. This proxy is continuously aware of the location of all users in the system, and uses this knowledge to build *cloaking regions* in which they are k-anonymous [25] (i.e., indistinguishable from at least k other users). However, these solutions are not suitable for the IntelliDrive framework, as its main goal is to improve road safety and mobility, and thus it requires knowledge of the exact location of vehicles.

A second family of solutions uses *mix zones* [5,12] (regions in which users do not communicate with the environment), inside of which users change their credentials. Therefore, it is not possible to perform tracking when several users traverse a mix zone simultaneously. The model in [5,12] considers that every time users authenticate, the use a different certificate. Thus, the only information available to the attacker is the time and location of messages. The IntelliDrive certificate management schemes reuse certificates, and thus leak additional information that can be exploited to improve tracking. This information allows us to perform tracking across multiple mix zones, contrary to [12] where the use of one-show certificates ensures that after traversing several mix zones the probability of following vehicles becomes negligible. The analysis in [5], on the other hand, focuses only on an isolated mix zone.

A theoretical analysis of shared certificates was presented in [17], which studies the tradeoffs between anonymity towards the certification authorities and ease of revocation, given that n certificates can be linked. The paper finds that if the parameters of the certificate scheme are adjusted to provide high levels of privacy (i.e., each certificate is shared by many vehicles), then the revocation mechanism performs poorly, as it revokes all the certificates of many well-behaved vehicles (which lose their ability to send messages). If on the other hand the scheme is tuned to ensure a good revocation performance (i.e., certificates are shared by few vehicles and thus revocation only affects malicious vehicles) then the levels of privacy are low. The main difference between our study and the one by Haas et al. [17] is that we do not assume that certificates are linked but instead provide methods to do so. Yet, our results corroborate the theoretical analysis in [17]. Further, we show that shared certificates do not protect users' privacy even if the service provider is honest and certificate distribution information is not available to the adversary.

3. The IntelliDrive model

In the IntelliDrive model [10,26,11], vehicles communicate with Road Side Equipment (RSE). Each vehicle has a set of anonymous certificates $\mathcal{G} = \{c_1, \ldots, c_M\}$ that are used to authenticate messages. The certificates are designed to prevent vehicles from being identified and tracked – i.e., no entity should be able to associate the certificate used to authenticate a message with the vehicle that sent the message, or to link messages as being sent by the same vehicle. For this purpose the IntelliDrive uses Certificate Authority Partitioning, and distributes the issuing of

Please cite this article in press as: C. Troncoso et al., On the difficulty of achieving anonymity for Vehicle-2-X communication, Comput. Netw. (2011), doi:10.1016/j.comnet.2011.05.004

certificates between two authorities. The Authorizing Certificate Authority (ATA) issues long term vehicle identifying certificates. The Assigning Certificate Authority (ASA) issues anonymous certificates to vehicles. The ATA has no knowledge of the vehicle's anonymous certificates, while the ASA does not know the long term vehicle identity. Vehicles who persistently send malicious messages can be identified because they have a higher rate of anonymous certificate revocations and corresponding certificate requests than the other (honest) vehicles.

The IntelliDrive Consortium has proposed two methods [10,26,11] for anonymous certificate management. Both schemes have four parts: (i) key generation, (ii) key distribution, (iii) key usage and management, and (iv) key revocation and update. In this work we are only concerned with key distribution, management and usage. For further details on the other aspects of the schemes we refer the reader to [10,26,11].

3.1. Vehicle segment certificate management using short-lived, unlinked certificate schemes

The first approach to achieve anonymity is described in [11]. Vehicles receive short-lived certificates that are unique and valid only for a short period of time (e.g., one or two weeks) after which they expire. Upon expiration or revocation the vehicle must request new certificates to the ASA. It proceeds as follows (see Fig. 1):

1. The vehicle sends a request $(R, E_{ATA}(Id))$ to the ASA, where $E_{ATA}(Id)$ is an encryption of the vehicle's identity under the public key of the ATA.
2. The ASA forwards $E_{ATA}(Id)$ to the ATA, who decrypts the message and verifies whether or not the vehicle should obtain new certificates (the decision is based on whether or not the vehicle is suspected of sending malicious messages). The ATA sends the result of the verification to the ASA.
3. If the response from the ATA is positive, the ASA issues new anonymous certificates to the vehicle.

The scheme in [11] assumes that the ATA is trusted not to reveal the vehicle's identity to the ASA. The ASA on the other hand, is considered potentially adversarial. The security goal of the scheme is to prevent the ASA from tracking and re-identifying vehicles. To achieve this, the ASA must not know the full set $\mathcal{G}$ of anonymous certificates assigned to a car, and thus vehicles cannot renew all their certificates at the same time. One possibility would be to make an individual request per certificate. To improve efficiency by reducing the number of requests, it is suggested that vehicles renew their anonymous certificates in small batches [11].

Let $\mathcal{G}$ denote the full set of certificates of a vehicle, and M be the number of certificates it contains, i.e., $M = |\mathcal{G}|$. Let b be the size of the batch, with $M = \beta b$, where β is a positive integer. When vehicles request new certificates to the ASA, they obtain a batch $g_i = \{c_{i,1}, \ldots, c_{i,b}\}$ of b certificates, i.e., $b = |g|$. At any time, vehicles have M certificates grouped in β independent batches of b elements, such that $\mathcal{G} = \bigcup_{i=1}^{\beta} g_i$.

Note that in this scheme the ASA knows how batches of certificates have been issued; i.e., given two certificates c_v and c_w, the ASA knows whether or not they belong to the same batch. Thus, the security of the scheme relies on the ASA not being able to link certificates from different batches as belonging to the same vehicle.

3.2. Vehicle segment certificate management using shared certificate schemes

The second approach is presented in [10,26]. In this scheme, shown in Fig. 1(b), the ASA creates a pool of N distinct anonymous certificates and then provides each vehicle with a set $\mathcal{G}$ of M certificates, which are randomly selected from the pool. Thus, in a population $\mathcal{V}$ of V vehicles, each certificate is shared (on average) by k vehicles, with $k = \frac{VM}{N}$.

The scheme in [10,26] is designed to protect vehicles from being tracked by the ASA, assuming that the ASA knows how certificates are grouped in sets (each corresponding to a vehicle), and that it can see the messages received by the RSE. The security of the scheme relies on k-anonymity [25]. The key idea is that given a message authenticated using certificate c_v, the ASA cannot distinguish which among the k vehicles who received c_v signed the message.

When a certificate is used to authenticate a malicious message, the ASA will revoke it. Vehicles sharing this certificate will need to request a new valid certificate, as in the case of short-lived certificates. Note that in this scheme new certificates are requested one-by-one, and that the ASA updates the sets $\mathcal{G}$ that contained the revoked certificate with the new certificate.

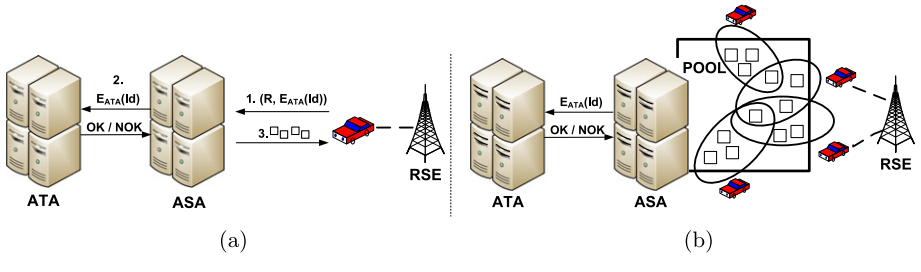

Fig. 1. (a) Short-lived certificates scheme, and (b) shared certificates scheme ($M = 4$). The □ symbol represents an anonymous certificate.

Please cite this article in press as: C. Troncoso et al., On the difficulty of achieving anonymity for Vehicle-2-X communication, Comput. Netw. (2011), doi:10.1016/j.comnet.2011.05.004

4. Attacking anonymous Vehicle-2-X communications in the IntelliDrive framework

We have developed two attacks on the schemes for vehicular communications presented in the previous section. The first attack is based on a brute-force search of possible vehicle routes, taking into account that the vehicles use shared certificates as described in Section 3.2. The second attack is based on heuristic clustering techniques and it applies to the short-lived certificates described in Section 3.1. Note that both attacks operate on information collected at the application layer, and are oblivious of issues at the network or transport layers.

4.1. Attacker model

Assuming that the ATA is honest and does not reveal the identity of the vehicles nor learn their location, the goal of the schemes in [11,10,26] is to provide privacy towards the ASA and the RSE. We note that this threat model is common in the literature [5,22], and that it is described in Annex F of the 1609. 2 WAVE standard [1]. In this paper, we evaluate the security of [11,10,26] against this *strong attacker*. Moreover, we consider a *weak attacker* who just observes interactions between the vehicles and the RSE, as shown in Fig. 2.

For each message, the adversary collects a tuple $\mathcal{T}_i = (c_v, l_a, t_i)$, where c_v is the anonymous certificate used by the vehicle, l_a is the location of the antenna, and t_i is the time when the communication took place. In addition, the attacker estimates the time that vehicles take to travel between each pair of adjacent RSE devices in different traffic conditions. This can be done, for instance, by covering the routes several times and collecting samples of the time it takes to travel between RSE locations. Given a pair of locations l_a and l_b, we denote as $\Pr_{a,b}[t]$ the probability that a vehicle travels from l_a to l_b in time less or equal than t (i.e., $\Pr_{a,b}[t]$ is a cumulative distribution function). $\Pr_{a,b}[t]$ can be used to decide whether or not two tuples $\mathcal{T}_i$ and $\mathcal{T}_j$, collected at locations l_a and l_b respectively, could possibly correspond to the same vehicle. Let γ be a parameter such that $0 < \gamma \leqslant 1$. We consider that $\mathcal{T}_i$ and $\mathcal{T}_j$ might refer to the same vehicle if $\Pr_{a,b}[t_j - t_i] < \gamma$.

In the example shown in Fig. 2(b), the adversary obtains a tuple $\mathcal{T}_1 = (c_1, l_a, t_1)$ from a message received at location l_a. After some time, two messages are received at location l_b, with associated tuples $\mathcal{T}_2 = (c_2, l_b, t_2)$ and $\mathcal{T}_3 = (c_3, l_b, t_3)$. As we can see in the upper part of the figure, $\Pr_{a,b}[t_2 - t_1] < \gamma$, while $\Pr_{a,b}[t_3 - t_1] > \gamma$. The adversary concludes that $\mathcal{T}_3$ cannot belong to the same trajectory as $\mathcal{T}_1$ because they are too far apart in time. On the other hand, $\mathcal{T}_2$ was received within the expected time frame, and thus $\mathcal{T}_1$ and $\mathcal{T}_2$ may originate from the same vehicle.

Moreover, we assume the attacker knows the city layout (i.e., the allowed and forbidden directions and turns) and uses this information when deciding whether two tuples may or may not correspond to a vehicle's trajectory.

Attacker's goal. The goal of the adversary is to track vehicles. For this she tries to link the observed tuples so that they reveal the vehicles' trajectories. The adversary exploits two types of constraints: first, space–time constraints (e.g., a vehicle cannot travel one kilometer in one second), and second, constraints imposed by the anonymous certificate schemes (e.g., a vehicle has M different certificates). It has been shown that tracking enables driver re-identification [14], and the inference of other private information [18]. To recover trajectories *weak* adversaries only use the information extracted from the tuples (i.e., certificate, time, and location). *Strong* adversaries also use the certificate distribution information available to the ASA. In the case of shared certificates, the adversary knows the complete set $\mathcal{G}$ of M certificates assigned to each vehicle. When short-lived certificates are used, the adversary only knows the grouping in batches g_i of b certificates.

Attacker's success. We consider that the attack succeeds when the adversary learns *all* the certificates belonging to a vehicle and recovers the vehicle's *full* trajectory. This is a strong definition of success, as private information could be compromised even if the attacker only learns partial trajectories or an incomplete set of certificates.

4.2. Brute-force clustering

In this section we present a brute-force clustering algorithm to reconstruct vehicle trajectories when shared certificates are implemented. The key idea is to explore

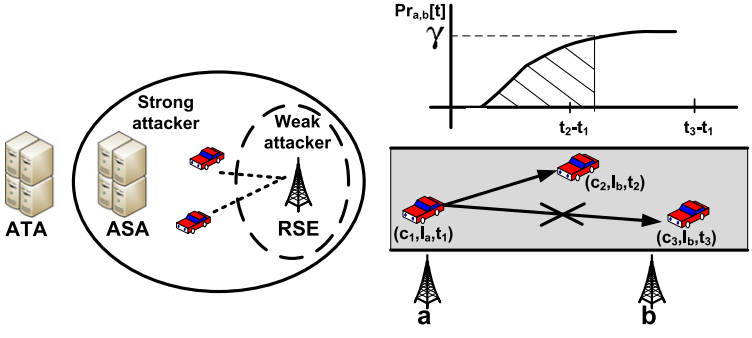

Fig. 2. Weak and strong adversary models (a) and cumulative distribution function describing the time needed to go from l_a to l_b (b).

Please cite this article in press as: C. Troncoso et al., On the difficulty of achieving anonymity for Vehicle-2-X communication, Comput. Netw. (2011), doi:10.1016/j.comnet.2011.05.004

possible trajectories and discard those which are inconsistent with the constraints. The attack succeeds if the actual trajectory of a vehicle is the only one that satisfies all the constraints.

The algorithm takes as input the messages sent by vehicles to the RSE. Each of these messages is represented by a tuple $\mathcal{T}_i = (c_v, l_a, t_i)$, as defined in the previous section. The adversary orders the list of tuples chronologically.

The attack starts by taking the first tuple, $\mathcal{T}_1$, and constructing a tree that has $\mathcal{T}_1$ as its root node. The tree is built in a breadth-first manner. To find the next point in the trajectory of the vehicle that sent the message represented by $\mathcal{T}_1$, the adversary compiles an list of candidate children nodes. These candidates are the tuples $\mathcal{T}_j = (c_w, l_b, t_j)$ that satisfy the following conditions:

1. Location l_b is adjacent to location l_a.
2. $\Pr_{a,b}[t_j - t_1] < \gamma$.

The attacker then checks that candidate trajectories (branches of the tree) are consistent with the parameters of the shared certificate scheme. A branch is consistent if the path between the leaf and the root node contains at most M different certificates, and inconsistent if adding it contains $M+1$ certificates (including the leaf's certificate c_w). The attacker discards branches which do not pass this consistency check. The algorithm is iterated to find candidate children of the remaining tree leafs, until no more tuples can be added to the tree. At this point, if only one branch remains, the adversary considers that the tuples in the branch describe a vehicle's trajectory. Those tuples are removed from the list, as they cannot belong to any other vehicle, and the algorithm proceeds starting from the earliest remaining tuple. If on the other hand the tree has more than one branch left, the adversary considers that the attack has failed. The root tuple $\mathcal{T}_1$ is discarded and a new tree is constructed starting from tuple $\mathcal{T}_2$.

Let us illustrate this process with the example shown in Fig. 3. There are two vehicles in this scenario (represented as ○ and □, respectively). Each of them has $M = 2$ anonymous certificates and they choose randomly which certificate to use for each communication. The vehicles send in total five messages to four RSE devices, located in l_a, l_b, l_c, and l_d. Thus, the observation of the adversary is represented by the list of tuples $\{\mathcal{T}_1 = \{c_1, l_a, t_1\},\ \mathcal{T}_2 = \{c_2, l_b, t_2\},\ \mathcal{T}_3 = \{c_3, l_b, t_3\},\ \mathcal{T}_4 = \{c_4, l_d, t_4\},\ \mathcal{T}_5 = \{c_2, l_c, t_5\}\}$.

The adversary selects $\mathcal{T}_1 = (c_1, l_a, t_1)$ as root node for the tree, and searches for possible successors. Let us consider that in this example both $\mathcal{T}_2$ and $\mathcal{T}_3$ are possible successors (i.e., $\Pr_{a,b}[t_2 - t_1] < \gamma$ and $\Pr_{a,b}[t_3 - t_1] < \gamma$). The adversary now performs a consistency check taking into account that $M = 2$. Both branches pass the consistency check, as at this stage both contain no more than two different certificates.

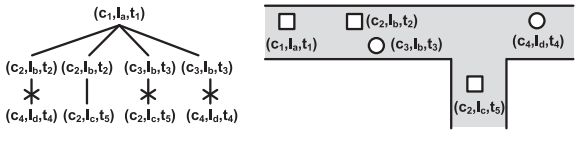

Fig. 3. Brute-force clustering.

In the next step, the adversary searches for possible successors to tuple $\mathcal{T}_2$. We consider that $\Pr_{b,d}[t_4 - t_2] < \gamma$ and $\Pr_{b,c}[t_5 - t_2] < \gamma$, and thus both $\mathcal{T}_4$ and $\mathcal{T}_5$ are candidate successors to $\mathcal{T}_2$. The adversary then checks the consistency of the branches with the shared certificate scheme parameters. As we can see, the branch formed by $\{\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_4\}$ cannot represent the trajectory of a vehicle, as it contains three different certificates $\{c_1, c_2, c_4\}$. The branch formed by $\{\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_5\}$ on the other hand does pass the test, as it only contains two different certificates $\{c_1, c_2\}$.

The attacker repeats the algorithm to find the candidate successors of tuple $\mathcal{T}_3$. We consider that $\Pr_{b,d}[t_4 - t_3] < \gamma$ and $\Pr_{b,c}[t_5 - t_3] < \gamma$. None of the two resulting branches ($\{\mathcal{T}_1, \mathcal{T}_3, \mathcal{T}_4\}$ and $\{\mathcal{T}_1, \mathcal{T}_3, \mathcal{T}_5\}$) passes the certificate consistency check though. In both cases, branches contain three different certificates, and thus they cannot describe the trajectory of the vehicle that starts its trip at location l_a.

At this point no more tuples can be added to the tree, which has only one branch left (formed by tuples $\{\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_5\}$). Thus, the adversary considers that the algorithm has succeeded in tracking the vehicle with certificates $\{c_1, c_2\}$, which passed by l_a at t_1, by l_b at t_2, and by l_c at t_5. Tuples $\{\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_5\}$ are removed from the list, as they cannot belong to another trajectory. The adversary proceeds to create a new tree starting from the first tuple in the new list (i.e., $\mathcal{T}_3$) in order to track another vehicle.

4.3. Heuristic clustering

In this section we present a heuristic clustering algorithm to reconstruct vehicle trajectories when short-lived certificates are implemented. We recall that short-lived certificates are unique (i.e., they are not shared by various vehicles). Therefore, tracking a vehicle v_i is trivial once the adversary learns its full set of certificates $\mathcal{G}_i$. As in the previous case, the algorithm takes as input the tuples $\mathcal{T}_i = (c_v, l_a, t_i)$ associated with the messages received by the RSE. The attack proceeds in three phases.

In the *first phase* the attacker constructs a weighted graph. Nodes in the graph represent certificates and the weight of the edge between two nodes expresses the likelihood that the two certificates belong to the same vehicle. To compute these weights, we observe the frequency with which two certificates are seen at adjacent locations within a certain time frame. The intuition is that two certificates that belong to the same vehicle appear consecutively more often two certificates belonging to different vehicles. This approach follows the spirit of the Statistical Disclosure Attack (SDA). The SDA was proposed in [9] as a method to uncover long-term sender-recipient relationships in mix-based anonymous communications networks.

Let us consider two certificates c_v and c_w. The weight of the edge between nodes c_v and c_w is incremented by one for each pair of tuples $\{\mathcal{T}_i = (c_v, l_a, t_i), \mathcal{T}_j = (c_w, l_b, t_j)\}$ that satisfies the following two conditions:

1. Locations l_a and l_b are adjacent.
2. $\Pr_{a,b}[t_j - t_i] < \gamma$.

Please cite this article in press as: C. Troncoso et al., On the difficulty of achieving anonymity for Vehicle-2-X communication, Comput. Netw. (2011), doi:10.1016/j.comnet.2011.05.004

Furthermore, impossible certificate pairs are "blacklisted". Let us assume that the minimum time to travel between two arbitrary locations l_x and l_y is $t_{min(x,y)}$. If there exists a pair of tuples $\{\mathcal{T}_i = (c_v, l_x, t_i), \mathcal{T}_j = (c_w, l_y, t_j)\}$ such that $t_j - t_i < t_{min(x,y)}$, then the edge between c_v and c_w is set to $-\infty$, as those two certificates cannot belong to the same vehicle.

We consider a scenario in which a total of $V = 3$ vehicles have $M = 3$ certificates each. Fig. 4 (left) shows an example graph constructed using this methodology. In the shown example, certificates c_4 and c_6 have been seen five times consecutively at adjacent locations, and thus the edge between them has weight five. On the other hand, certificate c_9 was used at location l_x at the same time as c_6 was used at (far-away) location l_y. Thus the weight of the edge between c_9 and c_6 is set to $-\infty$. If two certificates have neither been blacklisted, nor seen at adjacent locations as part of a possible vehicle trajectory, the edge between them has weight zero. For simplicity, these edges are not shown in Fig. 4.

We assume that the adversary knows the total number V of vehicles. Each vehicle v_i has a set $\mathcal{G}_i$ of M short-lived anonymous certificates, where M is a known parameter of the system. The full set of certificates in the system is denoted by $\mathcal{C}$ and has size $|\mathcal{C}| = VM$. Let $\mathcal{P}$ be the partition of $\mathcal{C}$ into the subsets $\mathcal{G}_i$; i.e., $\mathcal{P} = \{\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_V\}$. The goal of the attacker is to correctly reconstruct as many subsets $\mathcal{G}_i$ as possible, and ideally recover the full partition $\mathcal{P}$.

The adversary estimates $\mathcal{P}$ using the graph obtained in the first phase of the attack. The key idea is to partition the graph into V clusters of M nodes, such that the sum of the weights of the edges within the clusters is maximized. The adversary's estimation of $\mathcal{P}$ is defined as $\hat{\mathcal{P}} = \{\hat{\mathcal{G}}_1, \hat{\mathcal{G}}_2, \ldots, \hat{\mathcal{G}}_V\}$, where each subset $\hat{\mathcal{G}}_i$ represents a cluster in the graph (i.e., contains the same certificates as the cluster).

In order to compute $\hat{\mathcal{P}}$, we need to solve a graph clustering problem with known cluster sizes. We first create a starting solution $\mathcal{P}_0$ and then apply a heuristic optimization algorithm. To accelerate the convergence of this algorithm, it is important to choose a good starting solution $\mathcal{P}_0$. We construct $\mathcal{P}_0$ as follows.

1. We randomly choose a certificate $c_v \in \mathcal{C}$ and assign it to the cluster corresponding to $\hat{\mathcal{G}}_1$.
2. We complete this cluster with the $M - 1$ nodes with which c_v has edges of highest weight.
3. We remove from $\mathcal{C}$ the M certificates assigned to the cluster, and go to step 1.

In the toy example shown in Fig. 4 (top left), we have $V = 3$ vehicles with $M = 3$ certificates each. Thus, we need to cluster the certificates in three subsets containing three elements:

- $\hat{\mathcal{G}}_1$: c_6 is randomly chosen at the beginning of the algorithm. Then, c_4 is added to the cluster, as the weight of $\{c_6, c_4\}$ (5) is the largest. The cluster is completed with c_2 (chosen at random between c_2 and c_5, both with weight 1). All three nodes are removed from the list of candidates.
- $\hat{\mathcal{G}}_2$: c_3 is randomly chosen to start this cluster. We then add c_1 (linked to c_3 by an edge of weight 1). The third node is chosen at random between the remaining nodes linked to c_3 by an edge of weight zero (i.e., between c_5, c_8, and c_9). We assume c_9 is chosen.
- $\hat{\mathcal{G}}_3$: The remaining nodes (c_5,c_7,c_8) are assigned to the last cluster.

The third phase consists of applying a heuristic algorithm that finds the optimal solution $\hat{\mathcal{P}}$. The attack succeeds in identifying and tracking vehicle v_i if one of the subsets in partition $\hat{\mathcal{P}}$ contains the same certificates as $\mathcal{G}_i$. If $\hat{\mathcal{P}}$ is equivalent to $\mathcal{P}$, then the adversary succeeds in identifying and tracking all vehicles.

The adversary must find the graph clustering solution that maximizes the objective function (sum of the weights of edges within clusters), Given the starting solution $\mathcal{P}_0$, our algorithm proceeds as follows. First, it calculates the objective function of $\mathcal{P}_0$ (in the example in Fig. 4 this is $(5 + 1) + (1 + 0) + (0 + 0) = 7$). Then it randomly selects two nodes from different clusters, swaps them, and computes the objective function of the resulting solution. If the resulting objective function has a lower value, then the algorithm rejects the swap, and randomly chooses a new pair of nodes. If on the other hand the objective function increases with the swap, then the algorithm accepts the swap of nodes, updates the starting point with the new solution, and proceeds to select a new random pair of nodes. The algorithm stops when many consecutive iterations have resulted in a rejection of the swap.

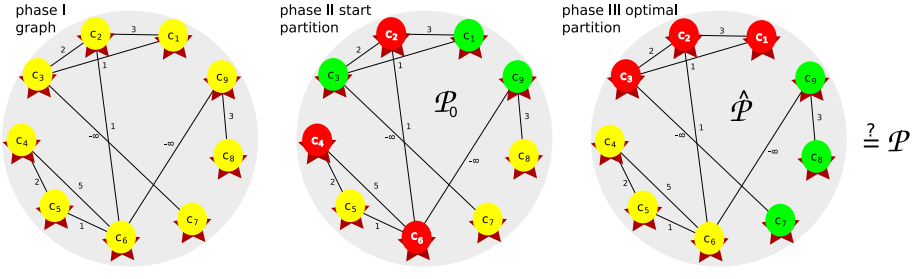

Fig. 4. Overview of the heuristic clustering algorithm. The different colours represent the different subsets in a partition $\hat{\mathcal{P}}$. Edges with weight zero do not appear in the graphs.

Please cite this article in press as: C. Troncoso et al., On the difficulty of achieving anonymity for Vehicle-2-X communication, Comput. Netw. (2011), doi:10.1016/j.comnet.2011.05.004

5. Experimental setup

Similarly to [12], we consider the square-shaped reticular city. Streets in this city have one or two lanes and are unidirectional. These streets divide the city in identical square blocks of 100 m wide (i.e., the total area is 1 km^2), and there is a ring road surrounding the city. In each of the intersections, we place traffic lights and RSE equipment. Vehicles send a message to the RSE every time they cross an intersection, and they are out of the RSE communication range while they drive on a street. We assume that vehicles do not reveal their speed or direction in their communication. The information collected for the purpose of the attack is the time of the communication, the certificate used to authenticate it, and the location of the RSE device that received the message.

At the beginning of the simulation each vehicle is randomly assigned a *home* and a *work* address. Every day, all vehicles leave at roughly the same time their *home* to go to their *work*, where they spend a deterministic time before heading back (i.e., there are two peak hours per day). Vehicles always choose the shortest path between *home* and *work*. They drive as fast as possible (with a maximum speed of $v = 50$ km/h) while respecting traffic lights and avoiding collisions with other vehicles. We simulate this scenario using the Netlogo[1] programmable modeling environment and collect data for the equivalent of two weeks.

Our choice of the parameters for the anonymous communications management schemes follows the examples provided in the schemes under study. For the shared certificate proposal [10,26], we consider a pool of $N \in \{5000, 10{,}000, 15{,}000\}$ certificates and we test scenarios in which vehicles are assigned $M \in \{4, 5, 6\}$ certificates. In the case of short-lived certificates [11], we consider vehicles have $M \in \{100, 1000\}$ certificates requested in batches of $b = 10$.

When vehicles communicate with the RSE one of their anonymous certificates must be chosen. In both schemes [10,26,11] certificate selection algorithms are mentioned but not specified. In this work we consider three selection algorithms: (i)*One-trip:* vehicles use one certificate for all the communications with the RSE during a trip, and the certificate is changed every time the car engine is started; (ii) *Sequential:* vehicles choose an order for their certificates and rotate them sequentially for each communication with the RSE; and *Random* vehicles select uniformly at random one of their certificates for each communication with the RSE. We note that, although the choice of these algorithms is arbitrary we believe they are representative of the possible strategies. In particular, the Random selection algorithm represents the worst-case scenario for the attacker as no information can be extracted from the order in which certificates are selected.

We define a set of standard parameters: $M = 5$ and $N = 10{,}000$ for shared certificates, and $M = 100$ and $b = 10$ for short-lived certificates. We have performed experiments in various traffic conditions, varying the total number of vehicles in the city between $V = 50$ and $V = 1000$ (traffic becomes heavy when $V = 500$) and studied the impact of the certificate scheme parameters on the effectiveness of the attacks. For values of V smaller than 500 we set $\gamma = 1$, while for $V = 500$ and $V = 1000$ we considered $\gamma = 0.9$ to reduce the computation complexity (lower values of γ reduce the number of candidate successors for each tuple). In Section 7 we discuss how the results obtained in this experimental setup can be extrapolated to more realistic scenarios.

[1] NetLogo itself: Wilensky, U. 1999. NetLogo. http://ccl.northwestern.edu/netlogo/. Center for Connected Learning and Computer-Based Modeling, Northwestern University. Evanston, IL.

6. Evaluation results

In this section we present the results of our experiments. We first discuss the protection provided by shared certificates against the attack described in Section 4.2. Then we proceed to evaluate the short-lived certificate scheme against the heuristic clustering attack explained in Section 4.3. We analyze the robustness of both systems against the strong and weak attackers presented in Section 4. Further, we evaluate the influence that the observation time (i.e., number of days for which the adversary has collected data) has on the attacks' success rate. We recall that we consider the attack is successful if the adversary recovers *all* the certificates of a vehicle *and* recovers the vehicle's *full* trajectory. Once a trajectory is known to the adversary, it is possible to recover private information about the driver: her *home* and *work* locations [20] and even her identity [14].

6.1. Shared certificates

6.1.1. Strong attacker model

We first evaluate the effectiveness in tracking and re-identifying vehicles of a strong attacker. The attacker knows the list of tuples $\mathcal{T}_i$ associated with the messages received by the RSE (each tuple contains the certificate used to authenticate the message, the time when it was received, and the location of the RSE device that received it), and additionally controls the ASA (i.e., knows the complete set $\mathcal{G}$ of M certificates assigned to each vehicle).

Our first experiment compares the level of privacy provided by the certificate selection algorithms explained in the previous section in conditions of normal traffic ($V = 100$ vehicles) and heavy traffic ($V = 500$ vehicles). The results are shown in Table 1 (left). We can see that Random selection provides better protection than the Sequential and One-trip strategies. Nevertheless, even in heavy traffic conditions, the adversary is able to recover all the information of 96% of the vehicles. For this reason, in the remaining experiments we only consider the Random selection algorithm.

Our second experiment evaluates the impact of the parameter M on the success of the attack. Given a fixed pool size N, this parameter varies the number of certificates k shared by vehicles. We carry out experiments in which vehicles are assigned $M = \{4, 5, 6\}$ certificates from a pool of $N = 10{,}000$. The results are presented in Table 1 (right). As expected, the tracking algorithm performance decreases for larger values of M (note that a larger M makes

Please cite this article in press as: C. Troncoso et al., On the difficulty of achieving anonymity for Vehicle-2-X communication, Comput. Netw. (2011), doi:10.1016/j.comnet.2011.05.004

Table 1
Strong attacker: success rate for shared certificates.

Certificate selection algorithm				Number of certificates per vehicle			
	One-trip (%)	Sequential (%)	Random (%)	M=	4 (%)	5 (%)	6 (%)
V = 100	100	100	94.7	V = 100	98.7	94.7	93.3
V = 500	99.7	99.8	95.8	V = 500	97.9	95.8	92.1
(N = 10,000, M = 5)				(N = 10,000, Random selection)			

Table 2
Strong attacker: success rate for shared certificates.

Pool size				Number vehicles in the system				
N=	5000 (%)	10,000 (%)	15,000 (%)	V = 50	100 (%)	250 (%)	500 (%)	1000 (%)
V = 100	95	94	91	96	94	94	95	95
V = 500	96	95	93					
(M = 5, Random selection)				(N = 10,000, M = 5, Random selection)				

the revocation of malicious vehicles less efficient). Intuitively, as M grows vehicles share more certificates and take longer to reuse them. Nevertheless, we note that even for M = 6 with V = 500 vehicles the attack success rate is 92%.

Next we studied the influence of the total number of certificates generated by the ASA, and considered pool sizes $N \in \{5000, 10{,}000, 15{,}000\}$. In Table 2 (left) we can see that this parameter has little impact on the success of the adversary. This is because with any of the three pool sizes considered vehicles share a very small fraction of their certificates. Thus, the tracking algorithm is rarely mistaken and the attacker success is high: more than 90% of vehicles have their private information compromised. These results confirm the findings in [17]: When vehicles do not share a high fraction of their certificates, their routes are easily traceable by an attacker with knowledge of all groups $\mathcal{G}$.

Last, we assessed the performance of the shared certificates scheme in different traffic conditions, considering that in the simulated town there are $V \in \{50, 100, 250, 500, 1000\}$ vehicles. In Table 2 (right) we observe that, even when there is a traffic jam (V = 1000), 95% of the vehicles can be tracked.

6.1.2. Weak attacker model

In the previous section we considered an adversary who controls the ASA, and thus has information on how certificates have been distributed to vehicles. In this section we are concerned with a weaker adversary, who only knows the tuples $\mathcal{T}$ collected by the RSE.

Table 3
Shared certificates: strong vs weak attacker

	Strong (%)	Weak (%)
V = 100	94	92
V = 500	95	86
(N = 10,000, M = 5, Random selection)		

Table 3 shows that the success rate of a weak attacker decreases with respect to the strong attacker. However, in normal traffic conditions, the attacker can track up to 92% of vehicles, and even in heavy traffic conditions 86% of vehicles are traceable.

6.1.3. Influence of the observation time

Our last experiment studies the effect of the observation time on the adversary's success. We simulated two scenarios in which V = 100 and V = 500 vehicles are given M = 5 certificates from a pool of N = 10,000. The adversary observes the system for d days, and collects a tuple for each communication between vehicles and the RSE. Fig. 5 shows the evolution of the attack's success rate with the number of days of observation (d ranges from 1 to 14 days).

As expected, the success rate of the attacker increases as more information is available. When the groups $\mathcal{G}$ are known (strong attacker), after two days the tuples that

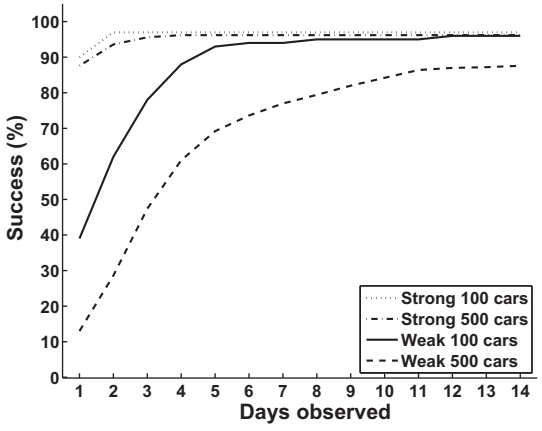

Fig. 5. Strong vs weak attacker: Influence of observation time on the success rate of the attack (N = 10,000, M = 5, Random selection).

Please cite this article in press as: C. Troncoso et al., On the difficulty of achieving anonymity for Vehicle-2-X communication, Comput. Netw. (2011), doi:10.1016/j.comnet.2011.05.004

Table 4
Strong attacker: success rate for short-lived certificates.

Number vehicles in the system						Number of certificates M		
V=	50 (%)	100 (%)	250 (%)	500 (%)	1000 (%)	M=	100 (%)	1000
Full	97	96	89	88	49	V = 100	96	0
Partial	1	2	4	5	23	V = 500	93	0
(N = 10,000, M = 100, Random selection)						(N = 10,000, Random selection)		

the adversary is able to cluster together are enough to learn the *home* and *work* locations of most vehicles (and thus identify their drivers [17]), and to track their movements in the city. The small percentage (around 3%) of vehicles that are not fully tracked by the adversary correspond to those who cross a vehicle with which they do share a certificate.

The observation time, however, has great impact on the success of the weak attacker. In light traffic conditions, we can see that on the first day the attacker only succeeds compromises 39% of the vehicles (learning their full set of certificates, their *home* and *work*, and their exact trajectories). This percentage increases up to 94% after one week, and reaches the same level of success as the strong attacker by the end of the observation period. With $V = 500$ vehicles, the attacker compromises 77% of the vehicles after one week, and 87% after two weeks.

6.2. Short-lived certificates

6.2.1. Strong attacker model

In the case of short-lived certificates, an adversary controlling the ASA knows which certificates were issued in the same batch. This information can be used to improve the clustering process described in Section 4.3, as certificates in the same batch can be directly classified as belonging to the same vehicle. Thus the attacker does not need to cluster M certificates to identify one vehicle, but only $\beta = M/b$ batches.

Our first experiment studies the influence of traffic density on the adversary's success. We carry out the attack when $V \in \{50, 100, 250, 500, 1000\}$ vehicles drive in the city, assigning $M = 100$ certificates to each vehicle in batches of $b = 10$. In the first row of Table 4 (left) we can see how increasing the number of vehicles decreases the success of the attack. If traffic is heavy, more vehicles meet at intersections. Thus, each time a certificate in seen in a location, it has many candidate successors to form a trajectory. This introduces noise in the graph created in the first phase of the attack and worsens its effectiveness. Nevertheless, even in a very congested situation ($V = 1000$) the attacker is able to fully compromise nearly half of the vehicles. The second row in Table 4 (left) represents the percentage of vehicles for which the attacker obtains partial information (defined as more than half of the vehicle's certificates).

In a second experiment we evaluate the influence of M on the performance of the attack, considering batches of $b = 10$ certificates. Table 4 (right) shows the success rate of the attack when vehicles possess $M = 100$ and $M = 1000$ certificates. We can see that increasing the number of certificates completely stops the attack, as pairs of certificates from the same vehicle do not appear consecutively very often. As noted in [17] however, increasing M affects both the communication overhead (the Certificate Revocation List becomes very large) and the effectiveness of the revocation mechanism (a large number of revocations is needed to fully remove a malicious user from the system).

6.2.2. Weak attacker model

The weak attacker does not succeed in tracking vehicles for $M = 100$ in our experimental setting (note that this experiment is equivalent to considering a strong attacker with $M = 1000$ and $b = 10$, as shown in Table 4 (right)). We note that the short-lived certificate scheme in [11] is designed to provide privacy towards the strong attacker, and thus our attack still succeeds in compromising the security of the system in its own threat model.

6.2.3. Influence of the observation time

Following the indications in [11] we considered that certificates have a lifetime of two weeks. In this simulation vehicles only travel during weekdays (a conservative estimate), and thus the adversary has data for only ten days.

Fig. 6 shows the percentage of vehicles whose full set of certificates is compromised as a function of time. In normal traffic conditions ($V = 100$), data from one day is sufficient to fully compromise 40% of the vehicles. After two days this percentage grows up to 74%. After six days, the adversary knows the full set of certificates of every vehicle.

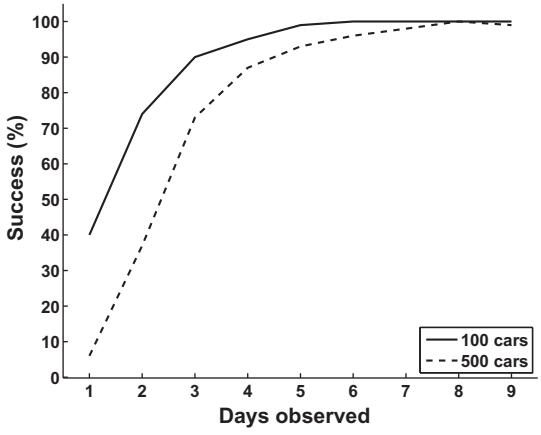

Fig. 6. Strong attack: success rate for short-lived certificates. Influence of observation time ($M = 100$, $b = 10$).

Please cite this article in press as: C. Troncoso et al., On the difficulty of achieving anonymity for Vehicle-2-X communication, Comput. Netw. (2011), doi:10.1016/j.comnet.2011.05.004

When $V = 500$, the adversary needs two days to compromise 40% of the vehicles, and three days to compromise 73%. By the time the certificates expire, the full set of certificates of every car is known to the attacker. Note that this allows the adversary to fully trace the movements of the vehicles for the full two weeks (the adversary can go back to the observations and link each message to the vehicle that sent it).

7. Discussion

We have made some simplifying assumptions in the simulation scenario chosen for the evaluation. In this section we review these assumptions and discuss the implications of our results for real world deployments. We also indicate ways in which the basic attacks presented in this paper could be extended to further improve their effectiveness.

We have considered a scenario in which there are more intersections than one would expect in the average town. We note that the higher the number of intersections, the harder it is to track vehicles, as they can change direction more frequently and mix with more vehicles. We have placed RSE on each intersection, and assumed that vehicles always send a message to the RSE when they are in its range of communication. Existing real world deployments may not have such an exhaustive coverage of all intersections, although we can expect the tendency to be towards greater coverage. A lower density of RSE naturally decreases the effectiveness of attacks, although they can still be deployed. The attacks presented in this paper could be trivially adapted to scenarios with lower RSE density by computing the probability distribution of the time needed to travel between each pair of RSE devices. The exact impact of reducing the number of RSE's in the system on the attacker's capability to track vehicles is left as subject for future work.

With respect to vehicle behavior, we have considered a very simple model: vehicles make two trips a day, from home to work and back; and the number of vehicles in the town is constant, with no vehicles traveling to or from outside the considered area of a square kilometer, which represents a small town. In real world scenarios, some vehicles make many more trips per day, and travel to a variety of destinations. Some other vehicles on the other hand, are used only occasionally.

In a country-wide deployment of the IntelliDrive framework there would be a larger number of vehicles than the ones considered in this work. However, an attacker is only concerned about vehicles driving in the region under monitoring. The *weak* adversary uses only the local messages collected by the RSE, and thus the performance of the attacks would not be affected by the existence of large numbers of vehicles in other areas. The information on certificate distribution available to a *strong* attacker who controls the ASA, would however contain vehicles outside of the area of interest, and thus the advantage offered by prior knowledge with respect to the weak adversary may decrease as more vehicles exist outside of the area of interest.

We note that our adversary does not make any assumptions or has any prior knowledge of vehicle behavior (i.e., it does not assume that cars will repeatedly commute between home and work). The adversary only uses the observations of messages sent by vehicles and knowledge of the anonymous certificate scheme used. Therefore, the attacks are applicable to real scenarios in which vehicles behave differently. In fact, our experimental setup simulates a worst case scenario, in which vehicles leave home and work every day roughly at the same time (peak hour). This means that every day the traffic is similar, and thus the adversary does not gain much additional knowledge with each extra day of observation. In real scenarios, we expect the effectiveness of the attacks to vary with vehicle behavior, and to be higher in scenarios in which vehicles make more and longer trips, since these provide more samples to the attacker. Vehicles that travel at night or during hours of very low traffic density are easy to identify and track.

The less regular the vehicle behavior (variation in destinations and timing of the trip), the easier it is to identify vehicles. If a set of vehicles travel together, their certificates frequently appear associated to each other, and it is hard to distinguish which certificates belong to which vehicles. If we consider instead a scenario in which vehicles move randomly, the certificates of a vehicle will appear associated to each other with a much higher frequency than to other vehicles' certificates.

Other factors present in real scenarios, such as short-term parking or vehicles entering and leaving the controlled area, would introduce some noise in the observations that may attenuate the effectiveness of the attacks. However, we note that our attacks use only a subset of the available information. Additional sources of information, such as speed or direction, or the location and timing of requests for anonymous certificates to the ASA, could be exploited to strengthen the attacks. Given a sufficient number of samples, noise can be removed with additional processing by applying consistency checks to the recovered sets of certificates and trajectories.

Our current attack algorithms can also be improved using more sophisticated traffic flow models. Consider a target vehicle is seen in a certain location l_1 at time t_1, and let t_{min} and t_{max} be, respectively, the minimum and maximum amount of time needed to travel between locations l_1 and l_2 given the average traffic density. The attacks in this paper treat all vehicles seen in l_2 between $t_1 + t_{min}$ and $t_1 + t_{max}$ as equally likely candidates to be the target vehicle. However, given the traffic density between l_1 and l_2 and the exact timing of the interactions between the RSE and the vehicles, it is possible to define a more refined probabilistic model that determines which of the vehicles seen in l_2 are more likely than the others to be the target.

The results presented in this paper were obtained by applying the attack algorithm once to the set of samples. This gives as result a first set of samples, certificates and trajectories for which the attack has succeeded and a second set for which the attack has failed (i.e., the adversary has not been able to associate those samples and certificates to a vehicle). One way to extend the attack would be to remove from the total set of samples those for which the attack succeeded, and iterate the analysis algorithm over the remaining samples.

Finally, the analysis of real world deployments should take into account both side channel attacks such as

Please cite this article in press as: C. Troncoso et al., On the difficulty of achieving anonymity for Vehicle-2-X communication, Comput. Netw. (2011), doi:10.1016/j.comnet.2011.05.004

wireless device fingerprinting, as well as cryptographic attacks on the certificate protocols.

8. Conclusions

The IntelliDrive Consortium has proposed two anonymous certificate management schemes [11,10,26] for Vehicle-2-X communications, aimed at enabling revocation while protecting drivers' privacy. Our results demonstrate that the schemes, optimized for scalability and easy certificate management, fail to achieve the privacy-preserving properties they are designed to provide with respect to a potentially adversarial service provider. If implemented, these *anonymous* certificates would enable the re-identification and tracking of a large percentage of vehicles.

The design of privacy-preserving Vehicle-2-X communications remains a challenge, specially when the detection and identification of misbehaving vehicles is required. Our results indicate that privacy-preserving solutions should be based on one-time pseudonyms, as the reusing of certificates is the key feature that enables our attacks. Furthermore, one-time pseudonyms would provide forward (and backward) security properties: even if a vehicle is tracked in a trip (e.g., because it is traveling alone in the road and does not cross any other vehicles), one-time pseudonyms would not provide any useful information for tracking the past or future trips of that vehicle.

Anonymous credentials [8] are one way of implementing one-time pseudonyms with optional anonymity revocation. Anonymous authentication protocols were initially considered not suitable for V2X due to efficiency reasons. However, the advances in the field [6,24,4] suggest that anonymous credentials will soon be suitable for low-latency authentication.

The analysis we have presented in this work demonstrates the potential of traffic analysis attacks to compromise privacy in vehicular communications. However, our methods cannot be considered a benchmark against which improved versions of the certificate management schemes should be tested. The attacks presented in this paper succeed using a very simple algorithm. Attacks based on more sophisticated traffic flow models and traffic analysis methods are left as subject of future work.

Acknowledgments

This research was conducted during a research visit by Costa-Montenegro to the COSIC group (K.U. Leuven, Belgium) during 2008, funded by the Programa de Axudas á Mobilidade dos Investigadores from the University of Vigo, Spain. This work was supported in part by the IWT SBO SPION project, and the IAP Programme P6/26 BCRYPT. Troncoso and Diaz are funded by the Fund for Scientific Research in Flanders (FWO).

References

[1] IEEE P1609.2 Trial-Use Standard for Wireless Access in Vehicular Environments – Security Services for Applications and Management Messages. Available from: <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=01653011>.
[2] IEEE P1609 Family of Standards for Wireless Access in Vehicular Environments (WAVE). Available from: <http://www.standards.its.dot.gov/fact_sheet.asp?f=80>, 2009.
[3] IntelliDrive USA. Available from: <http://www.intellidriveusa.org/>, 2009.
[4] M. Belenkiy, M. Chase, C. Erway, J. Jannotti, A. Küpçü, A. Lysyanskaya, E. Rachlin, Making P2P accountable without losing privacy, in: ACM Workshop on Privacy in the Electronic Society, ACM, 2007, pp. 31–40.
[5] A.R. Beresford, F. Stajano, Location privacy in pervasive computing, IEEE Pervasive Computing 2 (1) (2003) 46–55.
[6] P. Bichsel, J. Camenisch, T. Groß, V. Shoup, Anonymous credentials on a standard Java Card, in: 16th ACM Conference on Computer and Communications Security, ACM Press, 2009, pp. 600–610.
[7] A.J. Blumberg, P. Eckersley, On locational privacy, and how to avoid losing it forever, Technical report, Electronic Frontier Foundation, 2009.
[8] J. Camenisch, A. Lysyanskaya, An efficient system for non-transferable anonymous credentials with optional anonymity revocation, in: EUROCRYPT, LNCS, vol. 2045, Springer, 2001, pp. 93–118.
[9] G. Danezis, Statistical disclosure attacks, in: 18th International Conference on Information Security IFIP TC11, IFIP Conference Proceedings, vol. 250, Kluwer, 2003, pp. 421–426.
[10] G. Di Crescenzo, S. Pietrowicz, E. Van Den Berg, R.G. White, Tao Zhang, Vehicle segment certificate management using shared certificate schemes, 2008.
[11] G. Di Crescenzo, S. Pietrowicz, R.G. White, T. Zhang, Vehicle segment certificate management using short-lived, unlinked certificate schemes, 2008.
[12] J. Freudiger, M. Raya, M. Felegyhazi, P. Papadimitratos, J.P. Hubaux, Mix-Zones for Location Privacy in Vehicular Networks, in: ACM Workshop on Wireless Networking for Intelligent Transportation Systems (WiN-ITS), 2007.
[13] B. Gedik, L. Liu, Location privacy in mobile systems: a personalized anonymization model, in: ICDCS, IEEE C. Society, 2005, pp. 620–629.
[14] P. Golle, K. Partridge, On the anonymity of home/work location pairs, in: Pervasive, LNCS, vol. 5538, Springer, 2009, pp. 390–397.
[15] M. Gruteser, D. Grunwald, Anonymous usage of location-based services through spatial and temporal cloaking, in: 1st International Conference on Mobile Systems, Applications, and Services, USENIX, 2003, pp. 31–42.
[16] M. Gruteser, X. Liu, Protecting privacy in continuous location-tracking applications, IEEE Security & Privacy 2 (2) (2004) 28–34.
[17] J. Haas, Y.C. Hu, K. Laberteaux, The impact of key assignment on VANET privacy, Security and Communication Networks (2009).
[18] M. Iqbal, S. Lim, An automated real-world privacy assessment of GPS tracking and profiling, in: 2nd Workshop on Social Implications of National Security: From Dataveillance to Uberveillance, 2007, pp. 225–240.
[19] P. Kalnis, G. Ghinita, K. Mouratidis, D. Papadias, Preventing location-based identity inference in anonymous spatial queries, IEEE Transactions on Knowledge and Data Engineering 19 (12) (2007) 1719–1733.
[20] J. Krumm, Inference attacks on location tracks, in: Pervasive Computing, 5th International Conference, LNCS, vol. 4480, Springer, 2007, pp. 127–143.
[21] M. Mokbel, C.Y. Chow, W.G. Aref, The new casper: a privacy-aware location-based database server, in: 23rd International Conference on Data Engineering, IEEE, 2007, pp. 1499–1500.
[22] P. Papadimitratos, L. Buttyan, J.P. Hubaux, F. Kargl, A. Kung, M. Raya, Architecture for Secure and Private Vehicular Communications, in: IEEE International Conference on ITS Telecommunications, 2007, pp. 1–6.
[23] S.J. Row, A detailed overview of the U.S. DOTs IntelliDrive initiative. Available from: <http://mobilesynergetics.com/overview-of-usdot-intellidrive/>, 2010.
[24] M. Sterckx, B. Gierlichs, B. Preneel, I. Verbauwhede, Efficient implementation of anonymous credentials on java card smart cards, in: 1st IEEE International Workshop on Information Forensics and Security, IEEE, 2009, pp. 106–110.
[25] L. Sweeney, k-anonymity: a model for protecting privacy, International Journal Uncertainty Fuzziness Knowledge-Based Systems 10 (5) (2002) 557–570.
[26] R. White, S. Pietrowicz, E. van den Berg, G. Di Crescenzo, D. Mok, R. Ferrer, T. Zhang, H. Shim, Privacy and scalability analysis of vehicular combinatorial certificate schemes, in: 6th IEEE Conference on Consumer Communications and Networking Conference, IEEE Press, 2009, pp. 624–628.

Carmela Troncoso holds a Master in Telecommunications Engineering from the University of Vigo (2006). Since then, she is a PhD candidate in the Katholieke Universiteit Leuven in 2006. Her research is focused on Privacy Enhancing Technologies, amongst them anonymous communications, anonymity metrics and location privacy.

Enrique Costa-Montenegro received the M.Sc. degree in 2000 and the Ph.D. degree in 2007 from the University of Vigo, Spain; both in Telecommunication Engineering. He is associate professor since 2002 at the Department of Telematic Engineering, at the ETSE de Telecomunicacion (Higher Technical School of Telecommunication Engineering) in the University of Vigo. His research interests include wireless networks, car to car communication technologies, multi-agent systems and peer-to-peer systems. He is author of several articles in international journals and conferences and has participated in diverse R&D projects in these areas.

Claudia Diaz is an Assistant Professor in Privacy Technologies at the Department of Electrical Engineering of the Katholieke Universiteit Leuven. She holds a Master in Telecommunications Engineering from the University of Vigo (2000) and a Ph.D. in Engineering from the Katholieke Universiteit Leuven (2005). Her expertise is on privacy enhancing technologies, an area in which she has been active in the last ten years. Her research interests include the formalization and analysis of anonymity, location privacy, privacy-preserving social networking, the application of privacy by design to systems engineering, and interdisciplinary aspects of privacy.

Stefan Schiffner received the degree of Diplom Informatiker in 2006 from Technische Universität Dresden, Germany during his studies he worked for Siemens Dematic (Offenbach am Main). He worked for the data security group of TU Dresden until he started his Ph.D. at COSIC, Katholieke Universiteit Leuven in 2006. His research interests include privacy friendly technologies, reputation systems and peer to peer systems.

Please cite this article in press as: C. Troncoso et al., On the difficulty of achieving anonymity for Vehicle-2-X communication, Comput. Netw. (2011), doi:10.1016/j.comnet.2011.05.004