2011 Third International Conference on Intelligent Networking and Collaborative Systems

Distributed Privacy-Preserving Secure Aggregation in Vehicular Communication

Bo Qin[1,2], Qianhong Wu[1,3]
[1]Department of Computer Engineering and Mathematics
Universitat Rovira i Virgili, Tarragona, Spain
[2] Department of Mathematics, School of Science
Xi'an University of Technology, China
Email: {bo.qin,qianhong.wu,josep.domingo}@urv.cat

Josep Domingo-Ferrer[1], Willy Susilo[4]
[3] School of Computer, Wuhan University, China
[4] Centre for Computer and Information Security Research
School of Computer Science and Software Engineering
University of Wollongong, Wollongong, Australia
Email: wsusilo@uow.edu.au

***Abstract*—Vehicular *ad hoc* networks (VANETs), formed by computers embedded in vehicles and the traffic infrastructure, are expected to develop in the near future to improve traffic safety and efficiency. To this end, VANETs should be designed to be resistant against various abuses and attacks. In this paper, we first review the existing proposals to provide security, privacy, and data aggregation in vehicle-to-vehicle communication. We then address the fundamental issue of achieving these conflicting properties in a unified solution, having observed that separate efforts cannot fulfill the VANET design objectives. A set of new mechanisms are suggested for efficiently managing identities and securely compressing cryptographic witnesses, which are among the major obstacles to the deployment of strong security mechanisms in VANETs.**

***Keywords*-Authentication; vehicle privacy; data aggregation; VANETs.**

I. Introduction

With the fast advancement and pervasive deployment of information and wireless communication technologies, vehicular *ad hoc* networks (VANET) are expected to develop in the near future [1]. A VANET consists of on-board units (OBUs) embedded in vehicles serving as mobile nodes and road-side units (RSUs) working as the information infrastructure located in the critical points of the road. OBUs and RSUs are equipped with built-in sensory, data processing, and wireless communication modules. These modules allow vehicles and road-side infrastructure units to communicate with each other over single or multiple hops to exchange and share information about the routine driving status reports of vehicles and the driving environment changes. With these mechanisms, the OBUs and RSUs form a self-organized network which is the first commercial version of mobile *ad hoc* networks.

VANETs have various potential applications. The main thrust behind VANETs is applications related to traffic safety. According to the World Health Organization [2], approximately 1.3 million people die each year on the world's roads, and between 20 and 50 million sustain non-fatal injuries. Many traffic accidents come from the lack of cooperation between drivers. By giving more information about possible conflicts, most life-endangering accidents can be averted. VANETs also facilitate traffic optimization. Indeed, vehicles can collect data about traffic jams, weather or road surface conditions, construction zones, highway or rail intersections, emergency vehicle signal preemption, etc., and become information sources by sending those data to other vehicles in the VANET. These mechanisms enable transportation administration authorities to guide vehicles and manage them electronically (*e.g.* speed control, permits, etc.), which is much more efficient than traditional manual administration. Finally, in addition to safety-related applications, value-added services can be provided via VANETs. By implementing advanced electronic payment protocols in VANETs, one can expect to pass a toll collection station without having to reduce speed, wait in line, look for some coins and so on. As GPS systems have become available in many vehicles, it is also possible to realize location-based services in VANETs, for instance, finding the closest fuel station, restaurant and hotel. Other kinds of services include infotainment, vehicle-based electronic commerce and so on. All these services lead to a more comfortable driving experience for drivers.

Having realized the great commercial opportunities of VANETs, many academic and industrial organizations are committed to developing them. In the USA, the Dedicated Short Range Communications (DSRC, [3]) standard is being developed to support wireless communications for vehicles and road-side infrastructure. In Europe, the Car2Car Communication Consortium deals with vehicular communication standardization. With OBUs having a wireless connection to Internet, some infotainment can be provided, *e.g.*, Microsoft Corp.'s MSN TV and KVH Industries Inc. have introduced an automotive vehicle Internet access system named TracNet, which can bring Internet service to any in-car video screen. It is predicted that the market of VANETs can be up to billions of Euros in the near future.

Despite the potentially great benefits derived from VANETs, there are many challenges, especially conflicting security, privacy and storage requirements [4]. Among these concerns, security is the primary focus [5], [6], [7], [8]. VANETs aim at providing a safer and more comfortable driving environment by allowing vehicles to periodically disseminate messages to other vehicles in their vicinity. However, selfish vehicles can

978-0-7695-4579-0/11 $26.00 © 2011 IEEE
DOI 10.1109/INCoS.2011.78

also exploit this mechanism to send fraudulent messages for their own profit. Malicious vehicles may impersonate innocent ones to launch attacks without being caught. Driving privacy or vehicle anonymity is another critical concern in VANETs [9], [10]. Usually, to achieve security, vehicle-generated messages must be signed so that the receiving vehicles can verify that these messages have been originated by authentic sources and have not been modified during transmission [11]. However, with these signatures, it is possible for attackers to identify who generated a vehicular message containing speed, location, direction, time and other driving information. A lot of private information on the driver can be inferred if the driving pattern of his/her car can be tracked. Furthermore, the signed vehicle-generated messages have to be stored by the receiving vehicles for possible liability investigation: if some signed messages are later found false and to have misguided other vehicles into accident, the message generators and endorsers should be traceable. However, vehicular messages, especially their appended signatures, grow linearly with time while the storage capacity of OBUs in the vehicles is limited. Therefore, security and privacy of vehicle-to-vehicle (V2V) communications need to be conciliated with data aggregation/compression.

In this paper, we first review the state of the art of security, privacy and data aggregation in VANETs. Threats to be resisted include eavesdropping, impersonation, source spoofing, message modification and replay, identity theft, and privacy compromise. We note that the existing work treats security, privacy and data aggregation separately, which is a flawed approach because they are *not* independent issues. For instance, security requires the vehicular messages to be stored for a long period but the storage space of OBUs in vehicles is limited; data aggregation can alleviate storage consumption in vehicles but it may potentially weaken the security of VANETs; privacy requires the vehicular messages to be anonymous and this may potentially degrade the trustworthiness of V2V communications because the generators of messages are indistinguishable and cannot be identified for liability; and most privacy mechanisms aggravate the storage overhead at vehicles. These observations imply that a unified solution is needed to achieve the conflicting goals of security, privacy and data aggregation in VANETs. We suggest a set of new mechanisms for efficiently managing identities and securely compressing cryptographic witnesses, which are among the major obstacles to the deployment of strong security in VANETs.

Sections II, III and IV review, respectively, security, privacy and aggregation solutions in VANETs. Section V describes an approach achieving simultaneously security, privacy and aggregation. Section VI concludes the paper.

II. Security Challenges and Countermeasures

There are various categories of attacks against the security of VANETs. The first category is vehicular message source cheating, a.k.a. spoofing or impersonation: an attacking vehicle generates messages to cheat receiving vehicles by impersonating other vehicles [12]. This is attractive for an attacker to generate bogus messages to misguide target vehicles to his/her own advantage without being caught later when the messages are found to be false. A related attack is the Sybil attack [13]: the attacker generates several copies of the same cheating message by impersonating several source vehicles to deceive the receivers into believing that there are many vehicular sensors sensing the same driving environment; this increases the likelihood of receivers believing the content of the cheating message. Such an attack is powerful especially when receivers use some threshold/voting mechanism to decide whether they accept a message as true. The second category of attacks against security is to modify/replay messages generated by vehicles. If there is no security mechanism, the receiving vehicles cannot distinguish whether the received messages are intact/fresh or not. Therefore, they may be misguided and led into making wrong decisions, which may benefit the attackers or result in traffic accidents. The third category of attacks against security consist of generating incorrect messages to cheat other vehicles while writing a random identity in the identity field of the message. As a result, the receiving vehicles might be fooled but the attacker can deny having originated these bogus messages, because the random identity cannot be linked with the attacking vehicle.

Attacks against security can be thwarted by providing authentication, integrity and non-repudiation in vehicular communications. Authentication implies that, if a message is verified to be valid, then it comes from the claimed source. Integrity means that the received message has not been modified during transmission, and non-repudiation states that, if Alice generated a message, she cannot later deny authorship of the message. Authentication, integrity and non-repudiation can be simultaneously achieved using a cryptographic signature. That is, a vehicle generates a message with a timestamp, it signs the message with its private key and then it sends the message together with the signature on the message to other vehicles. The message will be trusted only if its signature is verified to be valid.

There are two ways to implement signatures in VANETs. The most commonly suggested one is to employ a conventional public key infrastructure (PKI) [3]. This approach requires the existence of a trusted third party called certification authority (CA). A CA can be materialized by some transportation administration office (TAO). A vehicle can register to a CA with its public key by showing that it holds the secret key corresponding to the public key. Then the CA generates a certificate for the public key. Finally, the registered vehicle can sign any message with its secret key. The receiving vehicle accepts the received message only if both the certificate of the public key and the signature on the message are valid. Traditional PKI-based signatures incur a heavy overhead to generate, distribute, store, verify, and revoke certificates of a huge number of vehicles. An alternative is to implement identity-based signatures in VANETs [14], [15]. In this approach, the public key of a vehicle is just its recognizable identity, *e.g.* the licence plate number; the private key is generated by a private key generator (PKG), *e.g.* the transportation administration office, by taking as inputs the PKG master secret key and the vehicle's unique identity. Then the registered vehicle can sign messages similarly to the PKI setting, but it does not require any certificate because the vehicle identity as a public key is recognizable and the vehicle can only have the secret signing key after registration. Hence, using ID-based signatures eliminates the requirement of certificates in VANETs.

The signature-based authentication of vehicular messages can be strengthened by a voting/threshold mechanism. With this mechanism [12], [16], [17], a message is trusted only if it has been endorsed by at least a threshold number of vehicles. The motivation behind the threshold mechanism is the fact that a message with a valid signature is not necessarily correct or truthful. For instance, an attacking vehicle (*e.g.*, a stolen car) may sign a wrong message to cheat other vehicles. Note that, if an event is sensed and endorsed by more vehicles, then the corresponding message is more trustable. Hence, the threshold mechanism can be used to thwart the above attacks. Voting can take place at either the message-generator side or the message-verifier side. In the former case, a threshold of sensing vehicles cooperatively sign the same message and a verifier is convinced if the jointly generated signature is valid. In the latter case, each sensing vehicle independently signs the same message and the verifier is convinced if it receives at least a threshold number of distinct signatures on the same message. Due to the high mobility of vehicles and the volatility of connections between them, verifier-side voting is easier to implement.

Authenticating vehicular communications using signatures also raises a number of practical challenges. Firstly, it raises the concern of vehicle privacy. To authenticate a vehicle-generated message, a signature must be included together with a public key or an identity which can be used to uniquely identify the message-generating vehicle. Note that the vehicle-generated message contains many sensitive data. By collecting and linking messages with the vehicle identity, an attacker can infer much private information about the driver such as her mood, personality and life style. Secondly, the security solutions raise the challenge of checking against the list of compromised vehicles. This is especially serious when PKI-based signatures are employed [18], [19], [20], [21], because, whenever a vehicle verifies a signature, it also has to check against a list of revoked certificates which grows linearly with time. Distributing, storing and checking this revocation list becomes a heavy burden after a secure VANET has been up and running for some time. This problem also appears in VANETs using ID-based signatures. In this case, if a vehicle is compromised due to the leakage of secret signing keys, the corresponding identities should be revoked. Hence, one faces the problem of checking revocation lists as in the case of VANETs using PKI-based signatures. Thirdly, establishing liability raises the challenge of storing numerous vehicular messages together with their signatures, the corresponding public keys and their certificates. Each vehicle can be assumed to periodically send messages over a single hop every 300ms within a distance of 10s travel time [3], which means a distance range between 10m and 300m. Assume that an OBU is on for two hours every day (tax and bus may run longer than that every day) on average and the vehicle is in a 40m road of 80 vehicles/km^2. Suppose that each message has to be stored for at least one year. Then a vehicle has to store about 8.4×10^8 signatures plus their corresponding public keys and certificates as cryptographic witnesses. Assume that an elliptical curve cryptosystem is used which is very efficient in bandwidth consumption, *e.g.*, about 22 bytes, 44 bytes and 22 bytes for each signature, public key and certificate for the available shortest signature, respectively. A vehicle has to store about 7.4×10^{10} bytes as cryptographic witnesses. Thus, the storage of these data is a heavy burden for vehicles whose OBUs have usually very limited storage capacity.

III. Privacy Challenges and Countermeasures

Since the security mechanisms in VANETs alone do not guarantee vehicle privacy, a number of efforts have been directed at tackling the privacy threats in VANETs. Generally, two approaches are employed, one based on pseudonyms and the other on group signatures.

Pseudonyms in VANETs allow both security and (conditional) privacy to be achieved [3], [22], [23]. With this approach, a vehicle has many public keys as pseudonyms, and the CA authenticates each public key with a certificate of short duration, say, several minutes. That is, the CA knows the real identity of the owner of the public keys. Then the vehicle can authenticate or endorse vehicular messages as in the case of regular signatures. Once the lifetime of a public key expires, the key will never be used again. Since each pseudonym is used only a few times and/or for a very short period, an attacker cannot link the pseudonym with the real identity of the vehicle. However, using this approach: i) each vehicle needs to pre-load a huge pool of pseudonyms, *i.e.*, the public keys and their certificates; and ii) the CA also needs to maintain all the anonymous certificates of all the vehicles to trace the originator of a malicious traffic report message endorsed with some pseudonym, which results in substantial certificate management burden. To relieve this overhead, identity-based signatures with many fake identities for each vehicle can also be resorted to (see [15]). However, similarly to what we discussed in Section II, using ID-based signatures does not reduce the overhead of checking revocation lists, although it eliminates the requirement of certificates inherent to the PKI setting.

Group signatures are an alternative to achieve privacy in VANETs [24], [16]. In a group signature, there is a group manager who maintains the group; members may join or leave the group dynamically. To join the group, a member holding a private signing key can register her public key with the group manager. The manager generates a secret certificate for the member. Then the member can anonymously sign any message on behalf of the group with her secret signing key and secret group certificate. A verifier can verify the group signature with only the group public key but cannot know which registered vehicle is the message generator. However, if necessary, the group manager can reveal the originator of any group signature. Group signatures can be implemented in VANETs to achieve vehicular communications authentication and vehicle privacy, by letting the transportation office play the role of group manager and vehicles the role of group members. The main merit of the group signature based technique over the pseudonym approach is that the former overcomes the limitation of pre-storing a large number of anonymous certificates.

Authentication of vehicular communications based on group signatures is conceptually simple. However, it also raises a number of challenges for practical deployment. Although group signatures eliminate the requirements of pre-loading a large number of pseudonyms and managing PKI-based certificates, a list has to be maintained to record the compromised vehicles. Whenever verifying a group signature, the verifier must also check whether the message generator has been revoked. Observe that the revocation list grows as time passes. The system performance may greatly degrade after the system has been deployed for some time. Furthermore, group signatures are usually much longer than regular signatures and the existing secure group signatures do not allow aggregation. If a message is endorsed by a number of anonymous vehicles, then the same number of group signatures has to be appended. This causes a heavy communication overhead for message relay and an expensive storage load for saving witnesses for the purpose of accident investigation.

IV. Performance Challenges and Aggregation

Although data aggregation is important in VANETs, only few proposals have addressed this issue. In [25], it is suggested to let vehicles send unsigned semantically aggregated information. This solution saves bandwidth and does not require receiving vehicles to verify any signatures or certificates. However, since no authentication mechanism is employed, this solution does not provide security and the received vehicle-generated message cannot be used as evidence in court. Also, privacy is not guaranteed, which is a major problem in VANETs.

The goal of aggregation can be partially realized with probabilistic verification and storage. A probabilistic verification scheme is proposed in [26] whereby vehicles cooperatively work to probabilistically verify only a small percentage of these message signatures based on their own computational capacity. Only the selectively verified signatures are stored as witnesses. Conceptually, this approach does not use data aggregation, but it achieves the same goal of compressing evidence. However, since many signatures will not be verified/used, this approach is actually based on a tradeoff between security and storage costs. Besides, not much attention is paid to vehicle privacy. Hence, the proposal [26] needs to be improved for practical deployment.

Cryptographic aggregation has recently been suggested in VANETs. In [27], aggregate signatures are applied to authenticate emergency messages. In an aggregate signature scheme, signatures on different messages are aggregated into one regular signature which can convince a verifier that the signers agree upon their respective messages. The verification of an aggregate signature is similar to that of the underlying signatures. The proposal [27] is implemented in the PKI setting and the certificate aggregation only allows one CA; hence, this proposal suffers from the so-called single-point of failure weakness. Although this scheme enables signature aggregation and certificate aggregation, the public keys of vehicles cannot be aggregated. The resulting cryptographic witnesses still grow linearly with time.

V. Simultaneous Security, Privacy and Aggregation in VANETs

As discussed above, the existing authentication schemes for V2V communications suffer from heavy overhead introduced by signature relay, verification and storage, or they cannot meet the goals of security and privacy. In this section, we propose a solution simultaneously achieving the following goals:

- *Security*. Authentication, integrity and non-repudiation are provided, so that explicit liability can be established in VANETs. This allows V2V communication to be used for accident reconstruction and liability investigation.
- *Conditional privacy*. Privacy is guaranteed to honest vehicles. However, if a vehicle behaves maliciously, it can be traced by a trusted third party.
- *Data aggregation*. Vehicular messages consist of non-cryptographic fields (payload, time stamps, etc.) and cryptographic fields (signatures) necessary to verify the validity of the former. Non-cryptographic fields contain much redundant information and can be compressed using general-purpose compression algorithms (*e.g.* Lempel-Ziv-Welch, [28]). However, cryptographic fields look random and cannot be compressed using general compression algorithms. We focus on securely compressing cryptographic fields in VANETs.

We observe that ID-based aggregate signatures are especially suitable for securing VANETs. In an ID-based aggregate signature scheme, a trusted private key generator (PKG) has a master secret-public key pair. PKG takes as input his master secret key and each user's identity to output the user's private signing key. Since a user can only get a secret signing key after registering to PKG, certificates are unnecessary for users in an ID-based aggregate signature scheme. A registered user can sign any message using her private signing key. The signature can be verified with PKG's master public key and the user's identity. The distinguishing feature of ID-based aggregate signatures is that independently generated signatures can be aggregated into a single one, no matter whether these independent signatures are on the same message and from the same signer or not. The validity of all signatures can be verified by merely validating the aggregated signature.

The aggregatability of ID-based aggregate signatures matches both the security and efficiency requirements of VANETs. The public key of PKG can be pre-loaded onto the OBU of each vehicle. After receiving thousands of message-signature pairs from vehicles in the permitted communication range, the verifier can aggregate all the signatures into a single one, namely an aggregate signature, and then verify it with one verification operation for all the received signatures. If the verification shows that all the signatures are valid, then the vehicle can make its driving decision.

A. High level description of the scheme

The above direct implementation of ID-based aggregate signatures in VANETs does not guarantee privacy for vehicles because the vehicle-generated message contains the identity of the message generator. To obtain an efficient privacy-preserving solution, we suggest that each RSU managed by the transportation administration department play the role of PKG. As shown in Figure 1, for each RSU we define the management coverage as the area reached by the neighboring RSUs' communication range. Hence, an RSU's management coverage is usually much larger than its communication range and neighboring RSUs' management domains may overlap. If the communication coverage of two neighboring RSUs does not intersect, then vehicles on the road may relay the signed timestamp till it reaches the communication coverage of the neighboring RSUs. By using relaying vehicles, the management domains of RSUs can fully cover all the road sections, even if the RSUs are sparse at the initial deployment stage of VANETs. This implies that our system is very scalable.

A vehicle first registers to the transportation office and obtains a secret token regarding its public identity. We assume that each RSU has two secret keys, one master secret pseudonym key for generating pseudonyms and one master secret signing key for generating private signing keys for vehicles. When passing an RSU, the vehicle shows to the RSU that it holds a secret token regarding its identity. Then the RSU generates a pseudonym for the vehicle by taking as input its master secret pseudonym key, its name, a timestamp and the vehicle's identity. The RSU further generates a private signing key for the vehicle by taking as input the pseudonym and its master secret signing key. The RSU stores the timestamp and the vehicle's identity in its local database and securely forwards the pseudonym and the private signing key.

In the management domain of an RSU, any vehicle can anonymously produce a signature on any vehicular message which can be verified by other vehicles in the same management domain. A receiving vehicle accepts the vehicle-generated message only if (1) the timestamp contained in the sending vehicle's pseudonym is consistent with the current time and, (2) the signature can be verified with the current pseudonym and the RSU's master public key corresponding to its master secret signing key. With RSU-aided on-the-fly pseudonyms, no certificates or revocation lists are required to preserve security and privacy in the proposed system. The format of privacy-preserving vehicle-generated messages is illustrated in Table I. For each vehicle-generated packet, the payload field may include the information on the vehicle's position, direction and speed, as well as traffic events encountered, event time and so on. According to DSRC, the payload of a message is 100 bytes. A timestamp is used to specify the signature generation time, which is employed to prevent replay attacks. We suggest to use the ID-based aggregate signature scheme by Herranz [29], which contains two group elements each one about 22 bytes long.

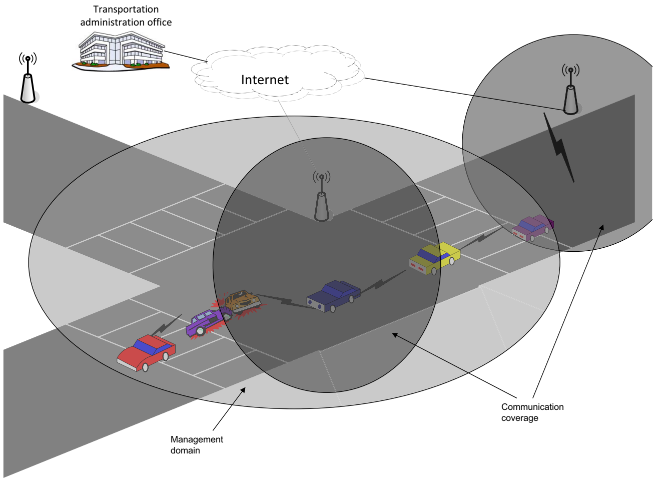

Fig. 1. System architecture

TABLE I
FORMAT OF PRIVACY-PRESERVING VEHICULAR MESSAGES

Payload	Timestamp	Pseudonym	Signature
100 bytes	4 bytes	22 bytes	22 bytes

B. The Scheme

We briefly review bilinear maps underlying our vehicular authentication protocols. Following the notations of [30], we use PGen to represent an algorithm which, on input a security parameter 1^λ, outputs a tuple $\gamma = (p, G_1, G_2, G_T, g_1, g_2, e)$, where finite cyclic groups $G_1 = \langle g_1 \rangle$ and $G_2 = \langle g_2 \rangle$ have the same prime order p, and $e : G_1 \times G_2 \rightarrow G_T$ is an efficient non-degenerate bilinear map such that $e(g_1, g_2) \neq 1$ and for all $g_1 \in G_1$, $g_2 \in G_2$ and $u, v \in \mathbb{Z}$, $e(g_1^u, g_2^v) = e(g_1, g_2)^{uv}$.

Global System Parameters: The algorithm PGen takes as input a security parameter 1^λ, and outputs a tuple $\gamma = (p, G_1, G_2, G_T, g_1, g_2, e)$. Here, $G_1 = G_2 = G$ and $g_1 = g_2 = g$. Three hash functions are selected: $H : \{0,1\}^* \rightarrow \mathbb{Z}_q$, $H_1 : \{0,1\}^* \rightarrow \mathbb{Z}_q$ and $H_2 : \{0,1\}^* \rightarrow G$. The system parameters $\pi =< \lambda, \gamma, H_1, H_2 >$ are embedded into TAO, RSUs and OBUs.

TAO Key Generation: TAO randomly chooses its secret key $sk_{TAO} \in \mathbb{Z}_p^*$ and computes its public key $PK_{TAO} = g^{sk_{TAO}}$. The TAO public key PK_{TAO} is also embedded into RSUs and OBUs, but the corresponding secret key $sk_{TAO} \in \mathbb{Z}_q$ is only known by TAO.

RSU Key Generation: Each RSU has two secret keys, one master secret pseudonym key for generating pseudonyms and one master secret signing key for generating private signing keys for vehicles. An RSU randomly chooses $x, y \in \mathbb{Z}_p^*$ and computes $X = g^x, Y = g^y$ as the RSU's public key, where x and y are the RSU's secret keys for generating private signing keys and pseudonyms of vehicles, respectively. Also, the RSU's public key (X, Y) needs to be certified by TAO with a certificate $Cert_{RSU} = H_2(ID||X||Y||PK_{TAO})^{sk_{TAO}}$, where ID is the RSU's identity containing a timestamp specifying the time when the RSU public key is certified. Each RSU's public key (X, Y) as well its certificate can be pre-verified and embedded into TAO, RSUs and OBUs. For a large-scale VANET up to one thousand RSUs, this storage cost is only about 66K Bytes, which is affordable in practice.

Vehicle Key Generation: A vehicle randomly picks $sk_v \in \mathbb{Z}_p^*$ as its long-term secret key and computes $PK_v = g^{sk_v}$ as its public key. The vehicle's public key PK_v also needs to be certified by TAO with a certificate $Cert_v = H_2(V||PK_v||PK_{TAO})^{sk_{TAO}}$ through a secure (*i.e.*, confidential and authenticated) channel, where V is the vehicle's identity containing a timestamp specifying the time when the vehicle's public key is certified.

Pseudonym and Signing Key Generation: This interactive procedure consists of two steps.

First, when a vehicle passes an RSU, it securely proves to the RSU that it is a registered vehicular user. Let the vehicle V be at location L_v, the current time be T_v, and the RSU have identity ID. The vehicle randomly selects $\gamma \in \mathbb{Z}_p^*$, and computes g^γ, a piece of registration argument $argument = H_2(V, L_v, T_v, ID)^{sk_v} Cert_v$ and the ciphertext $Enc_{key}(V, L_v, T_v, PK_v, Argument)$ using any secure symmetric encryption algorithm $Enc(\cdot)$, where $key = Y^\gamma$. The ciphertext and g^γ are sent to the RSU who can extract $V, L_v, T_v, PK_v, Argument$ with $key = (g^\gamma)^y$. Then the RSU checks that (1) L_v and T_v are sound, and (2) $e(argument, g) = e(H_2(V, L_v, T_v, ID),\ PK_V)e(H_2(V|| PK_v||PK_{TAO}), PK_{TAO})$. Only if the check passes, the RSU proceeds to the second step.

Second, the RSU returns an on-the-fly pseudonym and a corresponding signing key to the vehicle both of which will expire if the vehicle leaves the RSU's management domain. To do this, the RSU computes $r = H(y||V, L_v, T_v, PK_v||ID, T_{ID})$ and $R = g^r$, where R serves as a pseudonym of the vehicle and T_{ID} is the timestamp at which the pseudonym is issued by the RSU with ID. The RSU further computes the value $\sigma = r + xH_1(ID, T_{ID}, R) \mod p$ which later serves as a signing key of the requesting vehicle. Then the RSU returns $Enc_k(R, \sigma, T_{ID})$ to the vehicle. Finally, the RSU adds $(V, L_v, T_v, T_{ID}, PK_v)$ to its local records, appended by $Argument' \leftarrow Argument \times Argument' \in G$ where $Argument'$ is initialized to 1. One may note that $(V, L_v, T_v, T_{ID}, PK_v)$ can be significantly compressed because the regular location and time information contains much redundancy while the long-term identity V and public key PK_v are identical for the same vehicle.

Clearly, the vehicle can extract the pseudonym and the signing key. Their correctness can be validated by checking whether $g^\sigma \overset{?}{=} RX^{H_1(ID, T_{ID}, R)}$.

Signing Vehicular Reports: Let a vehicle receive a pseudonym R_i and a signing key σ_i with a timestamp T_{ID_i} from the RSU with identity ID_i. When it wants to sign a report m_i, it computes $\theta = H_2(m_i, ID_i, T_{ID_i}, R_i)^{\sigma_i}$ as the resulting signature on m_i. The vehicle sends the authenticated vehicular message $M_i = (m_i, ID_i, T_{ID_i}, R_i, \theta_i)$ to other vehicles nearby.

Aggregation and Verification: Assume that a vehicle receives n vehicular messages $M_i = (m_i, ID_i, T_{ID_i}, R_i, \theta_i)$ to be verified in an interval, where one can have $ID_i = ID_j$ while $m_i \neq m_j$. The vehicle computes the aggregate $\theta_A = \prod_{i=1}^n \theta_i$ for reports $m_1, \cdots, m_n$. The verifying vehicle accepts $(m_1, \cdots, m_n)$ if the contained location and time information is inconsistent and $e(\theta_A, g) = \prod_{i=1}^n e(H_2(m_i, ID_i, T_{ID_i}, R_i), R_i \cdot X^{H_1(ID_i, T_{ID_i}, R_i)})$.

Finally, the verifying vehicle adds $M_i = (m_1|| \cdots ||m_{i-1}; (ID_1,\ T_{ID_1})|| \cdots ||(ID_{i-1}, T_{ID_{i-1}});\ \theta_1 \cdots \theta_{i-1})$ to its local database $\mathbb{M}$.

One may note that the last field is always of constant size, *i.e.*, 22 bytes in the above proposal, while the rest of fields contain redundant non-cryptographic data which can be greatly compressed. The saving in storage cost is significant.

C. Tracing Authorship of Questionable Reports

Even if the vehicular reports are signed and verified to be authentic, they may contain untruthful information. If a vehicle was misguided by an untruthful report, the original vehicle signing this report can be identified by an RSU. Let the authenticated report m be indeed untruthful. The receiving vehicle can later find the corresponding (ID, T_{ID}) item in its local records and report it to the RSU with identity ID. With T_{ID} the RSU can trace the real identity of the dishonest vehicle V (Refer to the procedure of Pseudonym and Signing Key Generation).

D. Secret Key Updates

One advantage of the above scheme is easy secret key updating, if the secret key SK_{ID} is compromised before its expiry date, for example, because of an attack or an accidental exposure. Compared to most other ID-based schemes, our scheme need not change the secret key x of the TAO or the public hash function H_1 or the secret keys of other vehicles and RSUs to update the secret key of an individual vehicle. This is because the ID-based signature scheme we use is probabilistic: the secret key for an identity ID is a Schnorr (probabilistic) signature $SK_{ID} = (R, \sigma)$ obtained from TAO by choosing at random $a \in Z_q$ and then computing $R = g^r$ and $\sigma = a + xH_1(ID, T_{ID}, R)$. If this secret key of a vehicle is compromised, TAO can compute and distribute a new secret pair (R', σ') for this vehicle by just choosing a different value r'. In this way, the rest of the vehicles and RSUs can keep their secret keys, because the parameters of PKG remain unchanged.

E. Storage, Verification and Relay of Vehicular Reports

In a VANET, a vehicle periodically receives a large number of messages with appended signatures for verification. Usually, these signatures are verified separately and the verification

TABLE II
FORMAT OF THE STORED AGGREGATE VEHICULAR MESSAGE (BYTES)

Payload	Timestamp	RSU ID	Pseudonym	Signature
$100n$	$4n$	$\leq 10N$	0	22 bytes

TABLE III
FORMAT OF FORWARDED VEHICULAR MESSAGES (BYTES)

Payload	Timestamp	RSU ID	Pseudonym	Signature
$100n$	$4n+4$	$22n$	$22n$	22+22

is time-consuming. Some recent proposals suggest batch verification to speed up the verification process. To enable a *provably secure* batch verification [16], a linear number of additional exponentiations is required. Hence, the batch verification approach can only partially relieve the verification overhead. By employing aggregate signatures in VANETs, the received signatures can be aggregated into a single signature and then be verified as a normal signature. The aggregation operation needs only n multiplications rather than n exponentiations and, for security in a cryptographic sense, the involved multiplications are usually two orders of magnitude more efficient than the corresponding exponentiations. Hence, aggregate verification allows vehicles to quickly respond to driving environment challenges.

Vehicular messages which have been verified to be valid should be stored for possible liability investigation. However, as time passes, the vehicular messages received by each vehicle grow linearly but the OBU's storage capacity is limited. This conflict can be mitigated by aggregate signatures: the receiving vehicle only needs to save the aggregate signature. If new message-signature pairs are received, the new signatures and the stored aggregate signature can be re-aggregated into a new aggregate signature to be stored. Hence, a vehicle needs to store only one signature, no matter how many message-signature pairs are received. We also suggest that the receiving vehicle generate an aggregate signature to show that it has verified the validity of the stored messages. As for the pseudonym domain, since the pseudonym in each vehicular message can be reconstructed by the issuer (*i.e.*, an RSU) of the pseudonym with the timestamp and the RSU's location contained in the message, the receiving vehicle does not need to store the pseudonyms for each message. The format of the stored aggregate vehicular message is illustrated in Table II; note that the pseudonyms do not need to be stored and N is the total number of the RSUs. In the table, we do not specify the compression effect on non-cryptographic data which, as mentioned above, can be compressed by using general-purpose compression algorithms.

Some valid messages might need to be forwarded to the management domains of other RSUs. The verifying vehicle can remove the original signatures and endorse them with a new signature. The endorser's pseudonym needs to be included so that its anonymous signature can be verified by other vehicles. Also, the endorsing vehicle needs to add a new timestamp to reflect the endorsement time. The non-cryptographic data contained in the payload and timestamp fields can be greatly compressed before being forwarded, as they contain much redundancy.

F. Security and Performance

In this section, we briefly show that the security, privacy and aggregation goals are achieved in the above proposal.

Firstly, the unforgeability of ID-based aggregate signatures guarantees authentication, integrity and non-repudiation in vehicular communications. Unforgeability implies that only a registered vehicle can sign and endorse vehicular messages and, if there is any modification on any signature or any signed message, the aggregate verification procedure will signal the modified message-signature pair as invalid. Hence, if a verifying vehicle receives a batch of vehicular messages and the aggregate verification shows that the aggregate signature is valid, then these messages must come from authentic sources and have not been tampered with since they were originated; accordingly, the message generators cannot deny that they have agreed on these messages. If one or more messages in the batch are invalid, the aggregate verification fails. In this case, it is known that an efficient binary algorithm can be employed to find the invalid messages, similarly to the batch verification approaches in existing proposals.

The proposed solution preserves conditional privacy. On the one hand, whenever a vehicle generates a message-signature pair, only a pseudonym is specified in the message. Also, the pseudonym is different for different vehicles. For the same vehicle, the same pseudonym will be used only as long as the vehicle remains in the management domain of the RSU which generated that pseudonym. Next time the same vehicle enters the same management domain, its pseudonym will change as the pseudonym is also determined by a timestamp. Hence, for an attacker eavesdropping vehicular communications (see Table I), the pseudonym cannot be linked to the real identity of the message-generating vehicle. On the other hand, if some signed messages were later found untruthful and harmful to some vehicle, the data (see Table II) stored in the harmed vehicle and other witness vehicles can be used to trace the message generators and establish liability. From the payload of each message, the location of the pseudonym-issuing RSU is retrieved and this (trusted) RSU can recover the pseudonym of the message generator with its master secret key combined with the timestamp in the message. The pseudonym is uniquely bound with the real identity of some registered vehicle. Since the aggregate signature can be verified to be valid, the originator of the message cannot deny authorship of the message due to the non-repudiability of the aggregate signature. Hence, the vehicles signing cheating messages can be caught by using aggregate cryptographic fields.

As to performance, let us examine storage and computation. Table II shows that the storage saving for cryptographic fields (*i.e.*, pseudonyms and signatures) is very impressive: pseudonyms do not need to be stored and signatures are compressed into constant length, *i.e.*, 22 bytes, no matter how many message-signature pairs are received. Without the above mechanisms, the size of both the stored pseudonyms and signatures would be linear with time and might be up to the order of 10^{11} bytes (See II and note that many pseudonyms are also stored as evidences). Regarding computation, aggregate verification costs much less than separate verifications. Aggregate verification is also more efficient than batch verification, which requires a linear number of exponentiations. Finally, the proposed solution allows aggregate message relay without requiring to forward a large number of signatures appended to the original messages. Instead, only one new signature is generated and forwarded, at a cost of a new 4-byte timestamp and a 22-byte pseudonym of the endorser.

VI. Conclusion

In this paper, we have reviewed the state of the art to achieve security, privacy, and data aggregation in vehicular communications. We have argued that targeting those three properties separately cannot meet the requirements of VANETs. We have presented a comprehensive solution to attain those three conflicting goals. A set of new mechanisms have been described for efficiently managing identities and securely compressing cryptographic witnesses, two major problems in strong VANET security deployment. Our analysis shows that security, privacy, and data aggregation can be well conciliated and simultaneously achieved in VANETs.

Acknowledgments

This paper is partly supported by the EU 7FP through project "DwB", the Spanish Government through projects CTV-09-634, PTA2009-2738-E, TSI-020302-2010-153, PT-430000-2010-31, TIN2009-11689, CONSOLIDER INGENIO 2010 "ARES" CSD2007-0004 and TSI2007-65406-C03-01, by the Catalonia Government under grant SGR2009-1135, and by the NSF of China under grants 60970114, 60970115, 91018008, 60970116, 61173154, 61003214 and 61100173. We also acknowledge support by the Fundamental Research Funds for the Central Universities of China through Project 3103004, Beijing Municipal Natural Science Foundation through Project 4112052, and Shaanxi Provincial Education Department through Scientific Research Program 2010JK727. The third author is partially supported as an ICREA-Acadèmia researcher by the Catalan Government. The fourth author is supported by the ARC Future Fellowship FT0991397. The first three authors are with the UNESCO Chair in Data Privacy, but this paper does not necessarily reflect the position of UNESCO nor does it commit that organization.

References

[1] J. Blau, *Car Talk*, IEEE Spectrum, 45(10): 16-16, 2008.

[2] http://whqlibdoc.who.int/publications/2009/9789241563840_eng.pdf.

[3] http://www.iteris.com/itsarch/html/standard/dsrc5ghz.htm.

[4] D. Ma, G. Tsudik, *Security and Privacy in Emerging Wireless Networks*, IEEE Wireless Communications, 17(5): 12-21, 2010.

[5] M. Raya, P. Papadimitratos, I. Aad, D. Jungels and J.-P. Hubaux, *Eviction of Misbehaving and Faulty Nodes in Vehicular Networks*, IEEE Journal on Selected Areas in Communications, 25(8): 1557-1568, 2007.

[6] S. Kaza, J. Xu, B. Marshall and H. Chen, *Topological Analysis of Criminal Activity Networks: Enhancing Transportation Security*, IEEE Transactions on Intelligent Transportation Systems, 10(1): 83-91, 2009.

[7] Y. Jiang, M. Shi, X. Shen and C. Lin, *BAT: A Robust Signature Scheme for Vehicular Networks Using Binary Authentication Tree*, IEEE Transactions on Wireless Communications, 8(4): 1974-1983, 2009.

[8] T. Alpcan and S. Buchegger, *Security Games for Vehicular Networks*, IEEE Transactions on Mobile Computing, 10(2): 280-290, 2011.

[9] K. Sampigethaya, M. Li, L. Huang and R. Poovendran, *AMOEBA: Robust Location Privacy Scheme for VANET*, IEEE Journal on Selected Areas in Communications, 25(8): 1569-1589, 2007.

[10] J. Sun, C. Zhang, Y. Zhang and Y. Fang, *An Identity-Based Security System for User Privacy in Vehicular Ad Hoc Networks*, IEEE Transactions on Parallel and Distributed Systems, 21(9): 1227-1239, 2010.

[11] L. Zhang, Q. Wu, A. Solanas and J. Domingo-Ferrer, *A Scalable Robust Authentication Protocol for Secure Vehicular Communications*, IEEE Transactions on Vehicular Technology, 59(4): 1606-1617, 2010.

[12] V. Daza, J. Domingo-Ferrer, F. Sebé and A. Viejo, *Trustworthy Privacy-Preserving Car-generated Announcements in Vehicular Ad-Hoc Networks*, IEEE Transaction on Vehehicluar Technology, 58(4): 1876-1886, 2009.

[13] T. Zhou, R.-R. Choudhury, P. Ning and K. Chakrabarty, *P^2DAP – Sybil Attacks Detection in Vehicular Ad Hoc Networks*, IEEE Journal on Selected Areas in Communications, 29(3): 582-594, 2011.

[14] J. Sun, C. Zhang and Y. Fang, *An ID-based Framework Achieving Privacy and Non-repudiation in Vehicular Ad Hoc Networks*, MILCOM 2007, pp. 1-7, 2008.

[15] Y. Jiang, M. Shi, X. Shen, C. Lin, *BAT: a Robust Signature Scheme for Vehicular Networks Using Binary Authentication Trees*, IEEE Trans. Wireless Communications, 8(4): 1974-1983, 2009.

[16] Q. Wu, J. Domingo-Ferrer, and U. Gonzalez-Nicolas, *Balanced Trustworthiness, Safety, and Privacy in Vehicle-to-vehicle Communications*, IEEE Transactions on Vehicular Technology, 59(2): 559-573, 2010.

[17] L. Chen, S.-L. Ng, and G. Wang, *Threshold Anonymous Announcement in VANETs*, IEEE Journal on Selected Areas in Communications, 29(3): 605-615, 2011.

[18] A. Wasef and X. Shen, *EDR: Efficient Decentralized Revocation Protocol for Vehicular Ad Hoc Networks*, IEEE Transactions on Vehicular Technology, 58(9): 5214-5224, 2009.

[19] A. Wasef, R. Lu, X. Lin and X. Shen, *Complementing Public Key Infrastructure to Secure Vehicular Ad Hoc Networks*, IEEE Wireless Communications, 17(5): 22-28, 2010.

[20] A. Wasef, Y. Jiang and X. Shen, *DCS: An Efficient Distributed-Certificate-Service Scheme for Vehicular Networks*, IEEE Transactions on Vehicular Technology, 59(2): 533-549, 2010.

[21] J.-J. Haas, Y.-C. Hu and K.-P. Laberteaux, *Efficient Certificate Revocation List Organization and Distribution*, IEEE Journal on Selected Areas in Communications, 29(3): 595-604, 2011.

[22] M. Raya and J. Hubaux, *The Security of Vehicular Ad Hoc Networks*, In 3rd ACM Workshop on Security of Ad hoc and Sensor Networks-SASN 05, pp. 11-21, 2005.

[23] Y. Sun, R. Lu, X. Lin, X. Shen and J. Su, *An Efficient Pseudonymous Authentication Scheme With Strong Privacy Preservation for Vehicular Communications*, IEEE Transactions on Vehicular Technology, 59(7): 3589-3603, 2010.

[24] X. Lin, X. Sun, P.-H. Ho, and X. Shen, *GSIS: A Secure and Privacy Preserving Protocol for Vehicular Communications*, IEEE Transactions on Vehicular Technology, 56(6): 3442-3456, 2007.

[25] F. Picconi, N. Ravi, M. Gruteser and L. Iftode, *Probabilistic Validation of Aggregated Data in Vehicular Ad Hoc Networks*, Proc. of VANET'06, 2006.

[26] C. Zhang, X. Lin, R. Lu, P. Ho, and X. Shen, *An Efficient Message Authentication Scheme for Vehicular Communications*, IEEE Transactions on Vehicular Technology, 57(6): 3357-3368, 2008.

[27] H. Zhu, X. Lin, R. Lu, P. Ho and X. Shen, *AEMA: An Aggregated Emergency Message Authentication Scheme for Enhancing the Security of Vehicular Ad Hoc Networks*, IEEE-ICC'08, pp. 1436-1440, 2008.

[28] T. A. Welch, *A Technique for High-performance Data Compression*, Computer, 17(6): 8-19, 1984.

[29] J. Herranz, *Deterministic Identity-based Signatures for Partial Aggregation*, The Computer Journal, 49(3): 322-330, 2006.

[30] Q. Wu, Y. Mu, W. Susilo, B. Qin and J. Domingo-Ferrer, *Asymmetric Group Key Agreement*, Eurocrypt'09, LNCS 5479, pp. 153-170, 2009.