An ETSI look at the State of the Art of pseudonym schemes in Vehicle-to-Everything (V2X) communication

Anonymous
Some Research Group
Some Institution
Some Email Addresses

***Abstract*—While Intelligent Transportation Systems can contribute to increased road safety, they also allow tracking their user's location.**
This survey combines an overview of the ETSI ITS standard with a look at the state-of-the-art of pseudonym schemes for V2X communication, evaluating their applicability for protecting against location tracking and the possibility of combination of different approaches. Thereby it focuses on the middle layers of the used network stack.

***Index Terms*—Networks, Intelligent transportation systems, Security, Mesh networks, Privacy**

I. Introduction

In recent years, traffic got safer and safer. Improved safety technologies in our vehicles have contributed a lot to that development. But so far safety assistant systems are mostly working on their own while trying to evaluate the situation around them.
Intelligent Transportation Systems aim to create an ecosystem of networked vehicles and their infrastructure, collaborating with other vehicles and road infrastructure to improve safety and additionally providing new services to users. This step will be crucial for achieving the *vision zero* of no deaths caused by traffic worldwide.

While being an important step for traffic safety, Intelligent Transportation System (ITS) can pose a danger for user's privacy as always connected vehicles sending their positional data around in computer networks might allow tracking the users and creating location profiles.
Multiple solutions have been proposed so far to tackle this issue, protecting the human right of privacy. There already are some surveys giving an overview about the usage of different *pseudonym schemes* for preserving privacy in ITSs. But often the cutting-edge research is far ahead of standardization attempts, while the latter are deciding how future practical implementations might work while the former can provide valuable inspirations and introduce new technologies to the stack.

This survey combines the current status of the European standardization efforts for ITSs by the European Telecommunications Standards Institute (ETSI) with state-of-the-art approaches from newer research. Thereby it takes a look at how the middle layers of the ETSI ITS standard architecture are affected by the threat against privacy and what can be done about this.

In Section II I describe the background knowledge needed to judge the functionality of ETSI ITS networks by giving an overview of their architecture. Afterwards I describe the protocols involved in the middle layers of the networking stack and single out potential identifiers usable for the tracking of users.

In Section III I describe the pseudonym scheme proposed in the ETSI standard, emphasize the importance of pseudonym change strategies and present some further cutting edge pseudonym schemes not covered by standards so far.

Section IV defines attacker models, uses them to evaluate the privacy gained by the ETSI pseudonym scheme and looks at the feasability of that approach from a performance perspective.

II. Background

A. ITS Architecture

This section gives a brief overview of the ETSI architecture for Intelligent Transport Systems. It isn't meant to be elaborate but has a focus on identifiers and other message contents allowing linkability of messages.

Vehicular Ad-Hoc Networks (VANETs) have some special requirements: Due to many nodes being constantly on the move at higher speeds, tolerance for quickly changing topologies and low-latency communication are important points. Multi-hop mesh-networking is an important ability to keep the network functional in areas without designated infrastructure.

A VANET consists of different kinds of ITS stations:
On-Board Units (OBUs) residing inside vehicles can be divided into the communication and Communication & Control Unit (CCU), managing the ITS specific network communication over the car's wireless interfaces, and Application Units (AUs) utilizing the network services provided by the CCU to communicate transparently over a standard IPv6 stack.
On the stationary infrastructure side, Road-Site Units (RSUs) can either just provide some special local services

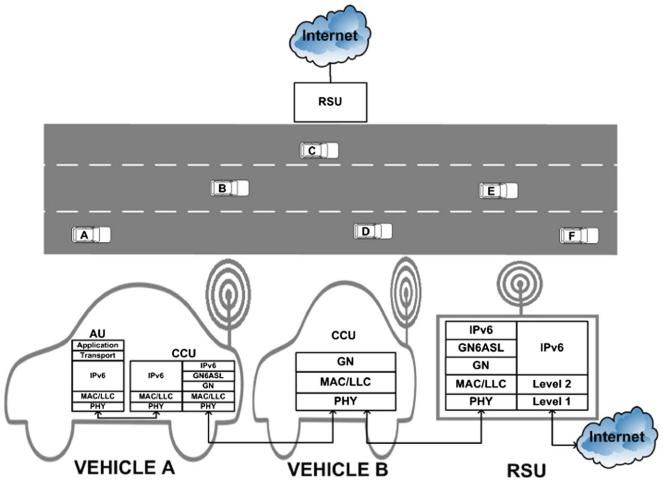

Fig. 1. Components of an ITS network, communicating with the internet; source: [1]

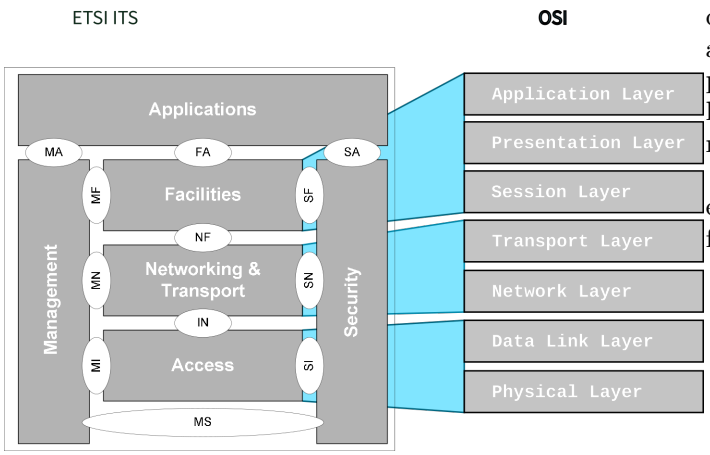

Fig. 2. The ETSI ITS-station reference architecture, based on [2]

or even be connected to a network operator's infrastructure and thus provide an uplink to the Internet (see Fig. **??**).

The protocol architecture of ITS stations according to the ETSI reference architecture [2] is mostly based on the well-known Open Systems Interconnection (OSI) layer model.

OSI layers 1 and 2 are combined into the *Access* layer, OSI layers 3 and 4 into the *Networking & Transport* layer and OSI layers 5, 6 and 7 are put into the *Facilities* layer (see Fig. 2).
The two vertical *Management* and *Security* layers provide supporting functionality throughout the whole stack. *Applications* make use of the ITS-station services and thus sit on top of it all.

Designed for modularity, the ETSI ITS architecture allows for a big number of access protocols. Similarly, a great variety of applications can run on top of the stack. Because of that variety, access and application layer are considered out-of-scope of this survey.

The **Networking & Transport** layer takes care of addressing and routing of messages within the ITS network and multiplexing them to higher-level services. Similarly to the OSI model, the groundwork of this functionality is provided by various networking protocols:
ETSI explicitly mentions the usage of Internet Protocol version 6 (IPv6) (possibly equipped with mobility support), the CALM FAST protocol [3] and the GeoNetworking (GN) protocol, which can also be used to encapsulate IPv6 packets.

CALM FAST [3] is a non-IP port-mapper protocol designed for single-hop communication between ITS stations and extensible with additional features. Due to a lack of proper access to the standard document, this protocol is considered out-of-scope of this survey.

1) GeoNetworking: GeoNetworking (GN) ([4] et seq.) is an ETSI-standardized networking protocol for routing and forwarding packets through VANETs based on geographical information. It sits between the link and network layer and provides its services to other networking and transport protocols. The background section of [1] gives a good high-level overview of the GN networking architecture and the rationale behind it.

Every GN node has to know its geographical position, e.g. through Global Navigation Satellite Systems (GNSSs), for the routing to work. The services provided by GN are:

- geo-unicast: routing a packet to a single node at a specific location
- geo-multicast: first routing a packet to a specified destination area, then flooding it to all nodes within that area
- topology-scoped broadcast: broadcast of packet within a certain number of neighbour hops
- single-hop broadcast: sending packets to all neighbouring nodes
- geo-anycast: routing packet to an arbitrary node within a specified geographical area

For this to work, each node maintains a GeoNetworking Location Table (LT) with the positions of its direct neighbours. This LT is populated with information from periodically-sent beaconing messages. These beacons advertise a node's position, GN address, its speed, station type and heading (see II-B1). This information is also included in all other sent GN packets. LT entries have a lifetime attached, after which they expire if not refreshed periodically.

Security properties of GN messages are ensured by signing (authenticity), encrypting (confidentiality) the messages and checking their plausibility and consistency. The necessary information for that is given in a security header [5].

2) IPv6: IPv6 [6] specifies the 6th version of the Internet Protocol, the routing protocol used in the networking layer of the Internet. Relevant details for VANETs are the addressing using 128 bit long IP addresses [8] with the first up to 64 bits specifying the network part and the last

64 bits specifying the interface ID (node ID) within that subnetwork. Additionally to the globally unique routable IPv6 address, nodes are also addressable with their link-local address. This special address is only valid in the scope of the same OSI layer 2 link and is automatically derived from lower-layer identifiers. Together with the huge number of globally unique IPv6 addresses, this new property makes it usable for vehicular ad-hoc networks. Another improvement in IPv6 is *neighbour discovery* [9] using link-local multicast. One application of that is the *Router Advertisement (RA)*, where routers just periodically announce their parameters so clients are able to derive an address themselves without further negotiation.

3) IPv6 over GeoNetworking: Transparently exposing IP networking to higher layers allows re-using existing services based on the classical Internet TCP/IP stack without modification. The GeoNetworking to IPv6 Adaptation Sub-Layer (GN6ASL) [10] specifies a mechanism for sending IPv6 packets over the GN protocol by using it as a sub-IP coupling layer. GN takes care of encapsulating and routing the IP packets to their final destination node, so that the whole underlying VANET looks like a flat layer 2 network to IP services. GN6ASL specifies how to derive a GN address from an IPv6 address and extends IPv6 with some GeoNetworking specific extensions like geographic multicast, Geographically Scoped stateless Address Configuration or (un)reachability detection.

4) BTP: The transport layer protocol above GeoNetworking is the Basic Transport Protocol (BTP) [11]. It provides a connectionless multiplexing/ demultiplexing of datagrams to the layers above, adding minimal overhead while providing an unreliable packet transport comparable to UDP.

If IPv6 over GN is used at the network layer, transport protocols like TCP and UDP from the standard Internet protocol suite can of course be used, too.

The **Facilities Layer** unifies the three upper OSI layers (application, presentation, session layer) and provides different support tasks to services and applications like time management, position management, database management and session management. It is also responsible to manage service priorities when passing down data to the Network and Transport Layer.

The **Security Layer** is a vertical layer providing security functionality like identity, key and certificate management to all other layers. It also contains all cryptographic functions like encryption or verification of data.

The **Management Layer** takes care of software changes like updates and installation of additional components and is considered out-of-scope of this survey.

B. Identifiers

There are many different addresses, IDs or other identifying information scattered around the network layers. This sections gives a list of relevant identifiers and the information encoded in them. Media-dependent, that means

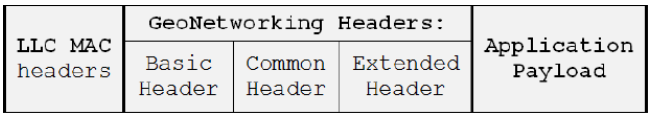

Fig. 3. Structure of an unsecured GN packet, source: [12]

Fig. 4. Structure of a secured GN packet, source: [12]

bound to a certain physical or data link layer technology, additional identifiers are considered out-of-scope.

1) GeoNetworking: Each GN node is identified by a 64bit GN_ADDR address [5], containing information about the ITS station type (passenger car, cyclist, pedestrian, RSU, …) and 48bit derived from the link-layer address. In case of a pseudonym change, only the latter part is supposed to change.

As shown in Fig. 3, GN packets have a basic, a common and an optional extended header. The *basic header* contains information like the packet's maximum lifetime and the remaining hop limit. These information are non-critical for identification. The *common header* also doesn't contain identifying information, only the flag indicating a mobile or stationary ITS station could slightly reduce the anonymity set. The *extended header* fields depend on the actual GN package type and can contain information like the sequence number (initialized with 0) and position vectors.

The LT is populated with information from beaconing messages and all other messages received by the ITS node. GeoNetworking Location Table entries also contain identifying data: Additionally to the GN_ADDR, station type and link-layer address of the peer node it contains a timestamped geographical position (including accuracy), its current speed and its heading.

Parts of GN packets can be secured by wrapping them into security headers as defined in [13] and shown in Fig. 4. This service is provided by the vertical security layer in the ETSI ITS architecture and secures all parts shown in Fig. 4 between security header and trailer according to the chosen security profile. The standard defines security profiles for encrypted, signed, externally signed, and signed encrypted messages.

The certificates used contain information about signer subject (name, type, keys), validity restrictions and the actual certificate signature from the Certificate Authority (CA). The signer information can be given in form of a digest, certificate or certificate chain.

The security trailer contains a signature for verifying authenticity and integrity of the message.

2) BTP: The BTP header as defined in [11] is only 4 bytes long and has a quite simple structure.
There are 2 modes of operation for BTP: *interactive packet*

transport using the BTP-A header, meant for services requiring replies to their messages, and *non-interactive packet transport* using the BTP-B header. The BTP-A header consists out of 2 16bit numbers denoting the source and destination ports. The BTP-B header contains the 16bit long destination port and 16bit for optional destination port information (depending on the service). Some of the facility layer services have well-known ports assigned in [14], so the destination port might identify the service used.

3) IPv6: While each IPv6-capable network interface can have multiple addresses, it has at least one link-local address with the interface ID (the lower 64bits) uniquely derived from its data-link layer address. The mapping of IPv6 link-local address and GN_ADDR is straightforward, as both addresses are deterministically derived from the same 48bit link layer address. Additionally to the IPv6 address, the IPv6 header can also contain a 20bit *flow label* [15] which could lead to partial linkability of packets even after an address change: Although a flow shall be identified by the triplet of flow label, source and destination address, an equal flow label could indicate the resumption of a connection even after an address change.

There exists a static mapping between IPv6 multicast groups and geographical areas (relative to the station). That means it is possible to contact IPv6-based services within a node's surrounding. But as this mapping is static and relative, it shouldn't help reidentifying hosts. Geographical Virtual Links (GVLs) are another important concept for understanding the visibility scope of IPv6 packets to other nodes. These virtual links are defined as non-overlapping, restricted geographical areas wherein all IPv6 multicasts within the same subnet are forwarded via GN to all nodes of that GVL. Usually this is a zone around a specific RSU serving as an Internet uplink and thus managing the whole subnet and its addresses. Globally routable IPv6 addresses are usually obtained via the stateless autoconfiguration with the help of RAs. So changing the GVL means getting another IPv6 prefix announced via RA and thus implies a change in the node's global IPv6 address.

4) Facilities Layer: The Facilities layer introduces a *StationID*, an integer identifying the ITS system. The standard document [16] already mentions that this ID may be a pseudonym.

Some further identifiers might be introduced in real-world implementations, e.g. for realising certain service over their dedicated protocols.

III. Pseudonym Schemes

As shown in the previous section, ITS communication contains many identifiers potentially allowing linking vehicle communication even over longer periods of time and thus tracking and creating movement profiles of vehicles.

This is a clear threat to the vehicle user's privacy, more precisely the *location privacy*. Complete anonymity of all network participants is no viable countermeasure, as security critical systems like these require certain levels of authenticity of data and accountability of the participants. Furthermore, request-response message schemes require at least short-term linkability of messages to establish a mutual session. This is needed e.g. for requesting data from infrastructure or managing automatical payment at car chargers.

A widely chosen approach for restoring user privacy is the usage of temporary pseudonyms for identification in the network. This section will look at the usage and kinds of pseudonym schemes in the ETSI standards, explore other approaches outside of the standardized ETSI world and look at the issue of when to change pseudonyms to minimize long-term linkability of nodes.

A. Pseudonym Schemes for ETSI ITS Systems

1) Pseudonym Management: The ETSI standard on trust and privacy management [18] mentions the goal of pseudonymity and unlinkability of ITS nodes and their messages as the way to achieve ITS privacy. This privacy goal is subdivided into two dimensions:

The **privacy** of ITS registration and authorization shall be achieved by limiting the knowledge of a node's canonical (fixed) identifier to a limited number of authorities. Furthermore, the responsibility for verifying the validity of a canonical identifier is given to an Enrolment Authority (EA) and split from the authorization to services by the Authorization Authority (AA). Both these authorities are parts of the needed Public Key Infrastructure (PKI) and need to be operated in different areas of control to achieve a surplus of privacy.

During manufacture, the following data is to be stored in an ITS node using a physically secure process:

- a globally unique canonical identifier
- contact addresses + public keys of an EA andAA,
- a set of trusted EA and AA certificates

The EA has to hold the following information about a node: The permanent canonical identifier, its enrollment credentials, its public key and a link to further profile information. ITS nodes can now request an enrolment certificate with their enrolment credentials from the EA. The task of the EA is to verify that an ITS node can be trusted to function correctly as the EA must only know the credentials of certified ITS nodes. Credentials of compromised nodes have to be revoked. With the enrollment request being encrypted and signed by the enrolling node and the response being encrypted as well, only the EA knows the mapping between the enrollment certificate and the requesting identity. The enrollment certificate contains a pseudonymous identifier being signed with a certificate chain leading back to the originating EA. This enrollment certificate can then be used to get Authorization Tickets (ATs) from an AA. These ATs too are certificates denoting the permissions a node has. Authorization ticket certificates may be stored in a Hardware Security Module (HSM)

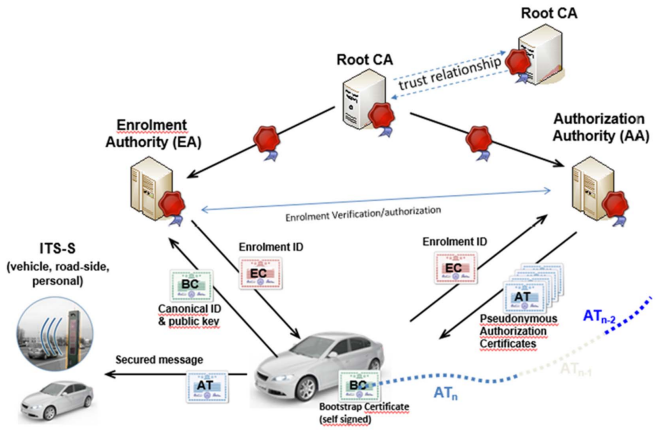

Fig. 5. ETSI ITS PKI trust model; source: [19]

to prevent direct unregulated access to the cryptographic keys, at least the security service Specification [17] offers such an option.
All authority responses are encrypted and signed in a way verifiable for the node. Certificate requests include a start and end time as well as a *challenge* [17], a random string encrypted with the public key of the receiver. These two measures prevent against message replay attacks. Enrolment credentials and ATs can also be updated if needed over similar mechanisms.
The overall trust model is sketched in Figure 5.

The second dimension of privacy covers the communication between ITS stations. The obtained authorization tickets serve as pseudonyms for authenticating and signing messages with other ITS services and nodes. ITS stations have to check the validity of the AT certificates included in every message and can check the permissions for the message's action (e.g. sending messages to certain broadcast domains) or access to certain services. These pseudonyms are to be regularly changed to preserve the privacy of the node's user by achieving long-term unlinkability of messages by the ITS node. According to [5] the AT may even be used to derive a GN_ADDR from.
There are different kinds of ATs: Those used by official role vehicles (e.g. state authorities) and ITS infrastructure don't always need to preserve the node's privacy and thus can contain a long-lived identifier for the official role they are fulfilling. ATs of personal user nodes can contain further personal identifying information if required for service usage, but then shall only be sent to already authorized nodes over encrypted channels. For broadcasting, first contact and all other uses, personal user nodes shall only use minimal pseudonymous ATs which then can be sent even over non-encrypted channels.

The ETSI standard [17] mentions the retaining of an audit log of incoming messages as the way of holding nodes **accountable** in case of misbehaviour. This only helps though if the EA retains a mapping of enrollment certificates to the canonical identifiers they were given to and the AA does the some for ATs and enrolment certificates. The legal and organisational framework for making sure that the information from the EA and AA are only combined for legitimate cases is crucial for maintaining user privacy, but is left out-of-scope of this survey.

For **revocation** of node access to the ITS network, e.g. in case of misbehaviour, there exist multiple mechanisms: The EA can be told to revoke the node's enrollment credentials to prevent it from updating its enrollment certificate and thus acquiring further ATs. Additionally, the EA revokes the validity of the enrollment certificate and the AA does the same for the authorization tickets. As ITS nodes are expected to check the validity of certificates using Certificate Revocation Lists (CRLs) and Certificate Trust Lists (CTLs) [19], messages of the revoked node are not accepted anymore.

2) Pseudonym Change for IPv6 ITS Networking: Section 11 of the ETSI standard on IPv6 usage over GN [10] covers the support for pseudonyms and their change of that protocol stack. The binding of a GVL's prefix to a distinct geographical area can be a threat to users' location privacy as a static interface identifier part of the IPv6 address would allow singling out a node over multiple GVL networks and track their location by the GVL prefix and its associated geographical region.
The proposed countermeasure is again the adoption and regular change of pseudonyms. In this case the affected identifier is the interface identifier part of IPv6 address. As this identifier is derived from the link-layer address, this also implies a change of the link-layer identifier address (MAC address). The same is true for the GN_ADDR thus it also changes accordingly with the changed link-layer address. All existing IPv6 connections have to be terminated as a clear cut between the old and new pseudonym IP address has to be made to prevent correlation of the old and new pseudonym during migration. A possible countermeasure against the interruption is the usage of *Network Mobility support* [20]. As this mobility support requires a home agent where all traffic flows through, this home agent needs to be trusted as it still has the possibility of location tracking by GVL.

B. Pseudonym Change Strategies

A crucial parameter of pseudonym schemes has been left out so far: How and when pseudonyms are actually changed. To show why that is so important, let us imagine a lone car on a street in the countryside: If a single car just changes pseudonyms there, immediately continuing its broadcasts under the new pseudonym, linkage of both pseudonyms is trivial for an observer.
Another example: Let us look at a traffic jam with 10 cars standing within reception range of an observer. Now there are multiple cars around making the mapping of pseudonyms to cars not totally trivial. But if we assume that each car only changes pseudonyms every 24 hours and

does this at an arbitrary time, the probability that only 1 vehicle changes pseudonyms within a short time range is very high, making linkage of pseudonyms easy again.
A last example so far: Focusing on one vehicle, let us assume it changes its pseudonym in a perfectly ambiguously way which can't be linked to the old one reliably. But after the pseudonym change, an already enqueued packet is sent, containing identifiers linkable to the previous pseudonyms.

These examples already show important points to take care of when changing pseudonyms: There needs to be some ambiguity regarding which node changed to which pseudonym – there shall be other nodes present within the reception range, coordination and frequency of change matter, and all identifiers need to be changed simultaneously with buffers being flushed or discarded. The position needs to be updated during pseudonym change, too, to prevent re-identification through stale position coordinates included in GN packets.

The ETSI ITS working group gathers a number of concepts for pseudonym change strategies in a technical report [19]: The parameters deciding about a pseudonym change (e.g. time period or way length) shall be randomized to prevent linkability by analyzing the periodicity of changes. After changing pseudonyms, random-length *silent periods* shall be abided in which nodes stop sending any packages. When using a *vehicle-centric* strategy, pseudonym change time, its frequency and duration of silent periods are influenced by the vehicle's mobility and trajectory to make linkage of pseudonyms based on broadcasted movement parameters harder. When using a density-based approach, pseudonyms are changed only if enough other vehicles are around to avoid unnecessary unambiguous pseudonym changes.

Mix-zones are geographical areas where no messages of location-aware services are exchanged. This concept is supposed to make linkage of in-going and outgoing vehicles from the zone difficult. These zones are especially effective in high-density and high-fluctuation areas like intersections or parking spots.
Within these zones, vehicles could collaboratively change pseudonyms by first announcing it via broadcast messages and then changing synchronously. The efficiency of that approach depends heavily on the density of the situation. A special variant are *cryptographic mix-zones*: Within these zones with a size limited to the radio coverage of an RSU, no identifying data is sent in plaintext but everything is encrypted with the same symmetric key provided by the RSU. This allows the usage of location-aware collision detection messages while preventing an outsider from eavesdropping, without having to switch off important safety features.

An alternative to just changing from one pseudonym to the next one from a node's internal storage is swapping pseudonyms randomly between nearby vehicles. This approach is limited though by the inclusion of vehicle-specific data into messages and legal requirements demanding the possibility of an identity resolution for law enforcement.

The ETSI survey [19] also gives an overview of used strategies in existing standards or projects. These include some interesting further approaches:
The SCOOP@F project proposes a timeslot-based round-robin pseudonym selection. The interesting thing about this is that reuse of pseudonyms from the local pool is explicitly allowed as the selection mechanism makes sure they are not always re-used in the same order. This is a useful approach against the problem of pseudonym refill (acquiring new pseudonyms) not always being possible.
The strategy proposed by the Car-2-Car Communication Consortium is dividing each trip into at least 3 segments: The first one from the start of the trip to a middle segment, the middle segment being common to a number of people and unassociated to certain origins and destinations, and the last segment to the intended destination of the trip. This shall achieve that locations significant to a user can neither be linked together nor to the user and thus preventing individual movement profiles. The values for changing pseudonyms have been statistically obtained with the outcome of changing pseudonyms at the beginning of a trip, then randomly after 0.8-1.5 km, and from then on randomly at least every 0.8 km or 2-6 minutes.

Some safety requirements of the ETSI standard affect pseudonym change: In critical situations when a receiving station would need to take immediate action in response to received safety information, pseudonyms have to be locked. The reason behind that is that cooperational collision avoidance depends on all vehicles broadcasting their location and trajectory. Vehicles in a silent period due to a pseudonym change wouldn't be taken into account, and vehicles changing pseudonyms without silent period could appear as duplicate or ghosting vehicles hindering collision evasion. Recognizing such critical situations and initiating the pseudonym locking is done by the receiving ITS vehicle, which decreases the risk of an attacker trying to deliberately lock pseudonyms without a critical situation being present.

C. Further Pseudonym Scheme Techniques

Petit et al. made an extensive survey [21] of cryptographic approaches for pseudonym schemes and defined a representative pseudonym life-cycle for comparing the different approaches.

1) Certificate-based Pseudonyms: The ETSI standardized pseudonym scheme is one instance of the ones categorized as *asymmetric cryptography schemes* in that survey. The class of these schemes is characterized by the use of asymmetric cryptography based on hierarchical certificates acquired from a PKI. This PKI has to be divided into at least 2 different administrative and legal control domains to make sure pseudonym resolution using the retained pseudonym-to-identity escrow mapping informa-

tion only happens under specific legal circumstances. Important parameters of these kinds of pseudonym schemes are the number of available pseudonyms acquired and available at a time, their lifetime, the used way of acquiring new pseudonyms (*pseudonym refill*) and the number of collaborating different authorities to resolve the split information for pseudonym resolution.

Some approaches covered don't require contact to an external PKI for pseudonym refill, but allow pseudonym self-issuance: Armknecht et al. [22] propose the self-issuance of pseudonym certificates with the node's own master keys. Verification of these pseudonyms utilizes zero-knowledge proofs and bilinear pairings while revocation of certificates works via changing the cryptographic system's parameters.
Calandriello et al. [23] combine the classical certificate scheme with *group signature schemes* (see III-C3) for pseudonym generation with individual private keys, and verification with the public common group key.

When it comes to enhancing the privacy of pseudonym resolution, several approaches of further splitting and distributing identity mapping information over several authorities utilizing blind signature schemes or group signature schemes are mentioned.

The IFAL protocol [24] introduces a mechanism tackling the issue of pseudonym refill: Pseudonym certificates can be distributed in big numbers already well in advance, as they are in principal valid in the future, but only if activated with periodically distributed activation codes. This is possible even over bad connections, SMS messages or via broadcasts as the codes are not confidential, but requires more storage space for the unactivated certificates.

The clear advantage of this class of schemes is the applicability to existing Vehicle-to-Everything (V2X) standards, as all major V2X Specifications use some kind of certificates.

These certificates have to be included into each message though and their storage and verification requires notable resources. Furthermore is the maintenance of the PKI system quite complicated, both regarding infrastructure requirements and legal and organisational frameworks. Because of these disadvantaged, I now take a look at other cryptographic pseudonym schemes.

2) Identity-based Cryptographic Pseudonyms: Identity-based cryptography is a form of asymmetric cryptography where a node's identifier (i.e. network interface and protocol address) serves as a nodes public key. A private key has to be derived from that public-key-id, this is usually done by a central Trusted Authority (TA) which has additional secret parameters to prevent that any node would be able to do this derivation. Some of the parameters are published and required for verifying message signatures. This TA can then also retain identity-mapping information, but doesn't distribute these mappings over multiple authorities. Revocation of pseudonyms can work similarly to the classical certificate-based scheme by revoking the canonical registration identifier of a node. The lifetime of pseudonyms can also be limited by adding an additional timestamp to the identifier string before deriving the private key from it. In theory revocation of certain pseudonyms could also be done by distributing revocation lists, but this has the same scalability issues like it has with certificates (see evaluation in IV).
When it comes to pseudonym change, the same strategies as for certificate-based pseudonyms apply. As the network interface identifiers are equivalent with the public key, especially the strategies for changing the network identifiers are relevant.

As the public key is directly derivable from the destination address of messages, a Man-in-the-Middle (MITM) relay-interception is prevented. Not having to include the certificate into each message and the smaller size of pseudonyms reduce the needed storage resources of ITS nodes. This though has to be compensated by the higher computational requirements of the used *bilinear mappings*, which are the basis for most of these schemes.

With the TA being involved in deriving the public key, pseudonym refill always requires a connection to this authority node. Another downside of this scheme is the required high trust into the TA which retains all the mapping information and needs to be directly exposed to the ITS network, thus being an exposed and valuable attack target. Some promising attempts for approaching this downside are mentioned in the survey [21] though.

3) Group Signature Scheme based Pseudonyms: The idea behind group signature schemes is that all nodes of a group are using the same shared public key for signing their messages, but have individual private keys for creating these signatures. As every group member could have created the signature validated with that shared public key, all nodes of the group are using the same pseudonym and this are anonymous within the anonymity set of the group. Two messages of the same vehicle are not linkable to each other as they're not distinguishable from two messages of different vehicles which are members of the same group.

Groups require a setup, during which the members of the group are determined and individual private keys are assigned to them by the *group leader*. The group manager is an entity that determines the system parameters including the public group key, creates and assigns private keys based on them to members and may revoke pseudonymity for certain members. This role could be assigned to any node of the group, but as it allows certain privileged actions the process of group manager election needs to be concisely designed. Proposals include using RSUs as regional group managers, which gives infrastructure operators even more powerful potential tracking abilities.

Pseudonyms are only changed to manage group dynamics, i.e. change of members of the group. Then the group manager generates new system parameters and issues new keys. When this happens, already mentioned strategies

like silent periods may be used. But individual network interface addresses still need to be unique per node and thus still have to change regularly like in other pseudonym schemes.

As an advantage of these schemes, nodes don't have to deal with generating, issuing and storing many pseudonym certificates.

Revocation is more complicated in group signature schemes: As all group nodes are indistinguishable by their exposed pseudonym identifiers, it's not possible to distribute revocation lists. A re-setup of the group by changing system parameters can exclude certain nodes, but has a big overhead as all group members are required to change their keys. A proposed solution for that circumvents the problem by remote-controlling the HSM to remove the keys from its memory.

The keys from group signature schemes are not directly usable for public key encryption of messages due to the special relationship of one public and multiple private keys. They can be used though to authenticate key-exchange protocols like Diffie-Hellman which are unauthenticated by themselves.

A special kind of group signature schemes not requiring setup and being more dynamic are *ring signature schemes*. Their usage is only briefly covered in [21].

4) Pseudonyms using Symmetric Cryptography: There are also pseudonym schemes utilizing symmetric cryptography authentication using Message Authentication Codes. Symmetric crypto algorithms are often computationally more efficient which would fit the requirements of near-realtime processing in VANETs.
The big issue with these schemes is that creation and verification of signatures uses the same key. Thus every node having the key for verification purpose can also create valid signatures in the name of another node pseudonym. Thus signature verification can't be done by each node themselves. After a node got a vehicle-ID from an EA, it creates several pseudonyms from it by hashing and combining with seed and counter values. These values then serve as pseudonym identifiers for connecting to an RSU and jointly creating a symmetric signature key. The RSU retains a mapping of key and pseudonym identifier.
For verification a node has to send the message (or a hash of it, depending on the MAC scheme) and the supposed sender pseudonym to the RSU. That station then verifies the signature using the retained mapping and sends the result back to the requesting node.

Thus symmetric pseudonym signature schemes heavily rely on infrastructure for signature verification and introduce additional delays due to the needed round trips. These issues make them hardly usable in practice.

There are some attempts of getting rid of these issues. The TESLA protocol [25] for example manages to reduce the infrastructure dependence by revealing previous signature keys using beaconing messages. This approach still suffers from high latency times though.

IV. Evaluation

This section evaluates the security of the proposed pseudonym schemes with an emphasis on the goals of privacy and anonymity, and the pseudonym schemes proposed in the ETSI standards. I also look at how much the pseudonym schemes influence the general functionality of the ITS system.

A. Attacker Model

In a security system for a network so ubiquitous like ITS networks will be in our world with omnipresent nodes, users and infrastructure, we can have a wide range of different adversaries with different capabilities and interests. So let's try to categorize the possibilities:

We now consider the **reach** of an attacker: Is the attacker limited to a single position, do they have a set of access points or do they even have a nearly global view on the network and their participants? Are they accessing the network over wireless interfaces or are they part of the backbone infrastructure or internet?

Is the attacker **active**ly trying to create, forge, block, modify, ...messages like a *Dolev-Yao adversary* [26] or just **passive**ly eavesdropping?

Is the attacker an **insider** - i.e. can it successfully authenticate at least with parts of the network – or an **outsider**?

So let us combine some of these characteristics to common attacker models and take them as a basis for evaluation:
Our first attacker is a *multi-point passive outsider* which we then further extend to a *global passive outsider*.
For our third attacker we look at the power of *attackers in the infrastructure*.

The trust assumptions of the ETSI ITS security services architecture are layed out in section 6.2.5 of [17].

B. Resilience against Attacks

I assume our attacker to be a multi-point passive outsider eavesdropping on the wireless communication and our ITS network to use the pseudonym scheme proposed in the ETSI standards.
As all communication to the AA and EA is securely encrypted, we can't get any information about the exchanged certificates and IDs from the eavesdropped communication to the PKI even if it happens to occur in our range of reception. Assuming that all identifiers are changed simultaneously, we now can only threaten a node's location privacy by managing to link its pseudonyms to each other. The change strategy proposed by the Car-2-Car Communication Consortium defined in III-B is deliberately designed with our chosen adversary in mind: Way lengths of segments are chosen big enough to prevent a single radio station tracking multiple segments including the pseudonym change itself while the middle-segment change interval time is chosen short enough to prevent multiple stations tracking the same pseudonym at multiple points.

So unless the adversary is lucky enough to have enough stations located at the correct points, we don't even need cooperative pseudonym change strategies so far.
When it comes to a global passive outsider though, the presence of other nodes and a cooperative pseudonym change strategy are necessary for reducing the linkability of pseudonyms well enough. Cooperative dynamic pseudonym change reduces the probability of correctly linking pseudonyms together with each change and with the number of cooperating vehicles. Silent periods in mix zones even improve the improbability as now projecting the last broadcasted trajectory into the the future includes too much entropy to reliably link pseudonyms. As we are dealing with an outsider we can even choose the concept of a cryptographic mix zone to keep safety features working. This changes though as we move to an insider attacker: As all authenticated ITS nodes get dealt the same symmetric key, our attacker can decrypt the broadcasted messages of all nodes, too, rendering this measure useless compared with a real silent period. Other cryptographic measures like using a group signature scheme within the mix zone might help with the indistinguishability of nodes, though correlations of the actual beaconing messages including positions and trajectories can still help with the linkage of pseudonyms. Additionally this can introduce other attack vectors like the *Sybil attack* described later in this section.

Authority vehicles shall only use their non-anonymized privileged tickets when they clearly want to exhibit this privileged status. Ambulances or firefighter trucks using these non-anonymized ATs can be recognized immediately and are granted special privileges. Nevertheless there needs to be an additional mechanism of utilizing these privileges while being pseudonymous and not appearing as an authority node to everyone. Police cars need a possibility of being undercover without passive outsider adversaries just recognizing them as the authority they are, otherwise avoiding police cars without even seeing them becomes much easier. For executing their privileges they can authenticate themselves as a privileged authority over an encrypted connection, similar to the personal ATs.

Other active insider attackers can attempt a *pseudonym depletion attack* by initiating so many pseudonym changes that the victim node runs out of pseudonyms and has to keep the same pseudonym although a change would be due. One possibility for this can be deliberately creating colliding network interface identifiers e.g. on the link layer. As many identifiers are derived from the node's link layer address, such a collision breaks several functionality throughout the stack, one of them e.g. GN. To evade this collision and restore functionality again, the victim node changes its network identifiers, triggering a pseudonym change.
For this to work, pseudonym refill needs to be obstructed, e.g. by preventing the connection to an AA. A connection might fail due to bad network connectivity, possibly made worse by active jamming of the attacker, a denial-of-service attack to the AA itself rendering it unusable or by collaboration of parts of the infrastructure (e.g. the RSUs) as our third attacker type suggests. The SCOOP@F change strategy (see III-B) allows pseudonym reuse and thus prevents pseudonym depletion. But this again can open an attack vector for *Sybil attacks*.

If the attacker has access to infrastructure components the issues with cryptographic mix zones already mentioned arise, too. As all RSUs are connected to the internet, they can even collaborate to track all changes in (cryptographic) mix zones to become a long-term global active insider adversary. Only frequent cooperative pseudonym change with silent periods introduces enough entropy to obstruct reliable pseudonym linkage.
Thanks to router advertisement and stateless autoconfiguration node's IPv6 addresses can't be linked to each other by the RSU serving as the subnet router, as nodes don't have to request an IPv6 address but just construct it themselves using the announced prefix and their own interface identifier. Thus also arbitrary IPv6 peers in the internet can't link the IPv6 addresses to recognize ITS clients again.

Personal ATs sent to already authenticated ITS stations can include additional personal data. This might be necessary for some kinds of services (e.g. payment information for charging services) but allows limited loaction tracking, especially if multiple stations of this kind and the same operator are located at different positions. They might exchange information about a node being close to them over the internet. As countermeasures it needs to be ensured that such personal identifying data is only included if it's really necessary. Additionally this data must be only sent to the service nodes when they're actually used, not just because they're within reception range.

If an insider active attacker node has access to multiple pseudonyms at once and can change between these at will, it can create the impression of additional spoofed ITS nodes in the surrounding area, tricking victim nodes into assuming being surrounded by many other vehicles and doing an ineffective pseudonym change. This so called *Sybil attack* can be prevented by limiting the number of available pseudonyms at a time, e.g by not exposing the pseudonym key material directly by storing it inside a HSM.

C. Influence of Pseudonyms on Performance

Preserving user's privacy through the use of pseudonym schemes is an additional requirement likely to add additional overhead to ITS networks. So we need to ask ourselves: Is this additional overhead still reasonable?

As shown in the previous section, frequent pseudonym change is needed at least each few minutes to prevent linkability of pseudonyms. This requires all network identifiers to change with the same frequency, too, interrupting existing long-standing connections. Applications either need

to tolerate this or adopt countermeasures like the usage of a NEMO mobile IP home agent. [20]

To prevent old identifiers being sent after pseudonym changes in packets already queued before the pseudonym change it is recommended to flush or drop all packet buffers before the change. This isn't necessary if one can be sure that there is no node identifying data in the queued packets. That is true for the GN packet forwarding queue, as nodes don't add their own source address when forwarding packages. The same is true for GeoNetworking Location Service (LS) packets. The source address included in there is the address of the original requesting node and though gives no reliable information about the address of the packet's sender as that node can also just be forwarding the package.

Active pseudonym certificate revocation turns out quite problematic in pseudonym schemes using asymmetric certificates and a PKI: CRLs or CTLs can quickly grow so big that they don't propagate through the network in reasonable times. Additionally checking each message against them quickly becomes too much for the limited computational resources of the node. So instead of active revocation, passive revocation by preventing misbehaving nodes from refilling their short-lived pseudonyms is the approach to choose.

V. Summary

The European Telecommunications Standards Institute (ETSI) Intelligent Transportation System (ITS) standard architecture contains many identifiers throughout the stack allowing to recognize and track the movement of vehicles and their communication behaviour.

To counter this threat for a user's location privacy, various pseudonym schemes have been proposed. The one proposed for usage with the ETSI standards uses asymmetric cryptography and a Public Key Infrastructure (PKI), but lacks a proper definition of important aspects like a detailed pseudonym change strategy, pseudonym resolution resilient against authority misuse or the usage of more advanced cryptographic schemes. But combined with technologies from other research the scheme is feasible to protect user privacy against several proposed attackers.

As many advanced cryptographic schemes are not compatible with the standards proposed by ETSI so far, future work should evaluate whether the standard could be changed to utilize some of these more modern approaches to counter current drawbacks.

VI. Glossary

AA	Authorization Authority
AT	Authorization Ticket
AU	Application Unit
BTP	Basic Transport Protocol
CA	Certificate Authority
CCU	Communication & Control Unit
CRL	Certificate Revocation List
CTL	Certificate Trust List
EA	Enrolment Authority
ETSI	European Telecommunications Standards Institute
GN6ASL	GeoNetworking to IPv6 Adaptation Sub-Layer
GN	GeoNetworking
GNSS	Global Navigation Satellite System
GVL	Geographical Virtual Link
HSM	Hardware Security Module
IPv6	Internet Protocol version 6
ITS	Intelligent Transportation System
LLC	Logical Link Control
LS	GeoNetworking Location Service
LT	GeoNetworking Location Table
MAC	Medium Access Control
MITM	Man-in-the-Middle
OBU	On-Board Unit
OSI	Open Systems Interconnection
PKI	Public Key Infrastructure
RA	Router Advertisement
RSU	Road-Site Unit
TA	Trusted Authority
V2X	Vehicle-to-Everything
VANET	Vehicular Ad-Hoc Network

References

[1] V. Sandonis, I. Soto, M. Calderon, and M. Urueña, "Vehicle to Internet communications using the ETSI ITS GeoNetworking protocol: V. Sandonis et al." *Transactions on Emerging Telecommunications Technologies*, vol. 27, no. 3, pp. 373–391, Mar. 2016.

[2] European Telecommunications Standards Institute (ETSI), "ETSI EN 302 665 V1.1.1; Intelligent Transport Systems (ITS); Communicatons Architecture," Sep. 2010.

[3] I. . S. intelligents de transport ISO/TC 204 Systeme für Verkehrsbeeinflussung und -information , ISO/TC 204 Intelligent transport systems, "ISO 29281-1," 2013.

[4] European Telecommunications Standards Institute (ETSI), "ETSI EN 302 636-1 V1.2.1; Intelligent Transport Systems (ITS); Vehicular Communications; GeoNetworking; Part 1: Requirements," Apr. 2014.

[5] ——, "ETSI EN 302 636-4-1 V1.3.1; Intelligent Transport Systems (ITS); Vehicular Communications; GeoNetworking; Part 4: Geographical addressing and forwarding for point-to-point and point-to-multipoint communications; Sub-part 1: Media-Independent Functionality," Aug. 2017.

[6] S. Deering and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification," Internet Requests for Comments, RFC Editor, STD 86, Jul. 2017.

[7] A.-L. Baecker and C. Schrimpe, "RFCE014: IPv6."

[8] R. Hinden and S. Deering, "IP Version 6 Addressing Architecture," Internet Requests for Comments, RFC Editor, RFC 4291, Feb. 2006, http://www.rfc-editor.org/rfc/rfc4291.txt.

[9] T. Narten, E. Nordmark, W. Simpson, and H. Soliman, "Neighbor Discovery for IP version 6 (IPv6)," Internet Requests for Comments, RFC Editor, RFC 4861, Sep. 2007, http://www.rfc-editor.org/rfc/rfc4861.txt.

[10] European Telecommunications Standards Institute (ETSI), "ETSI EN 302 636-6-1 V1.2.1; Intelligent Transport Systems (ITS); Vehicular Communications; GeoNetworking; Part 6: Internet Integration; Sub-part 1: Transmission of IPv6 Packets over GeoNetworking Protocols," May 2014.

[11] ——, "ETSI EN 302 636-5-1 V2.1.1; Intelligent Transport Systems (ITS); Vehicular Communications; GeoNetworking; Part 5: Transport Protocols; Sub-part 1: Basic Transport Protocol," Aug. 2017.
[12] E. Hamida, H. Noura, and W. Znaidi, "Security of Cooperative Intelligent Transport Systems: Standards, Threats Analysis and Cryptographic Countermeasures," *Electronics*, vol. 4, no. 3, pp. 380–423, Jul. 2015.
[13] European Telecommunications Standards Institute (ETSI), "ETSI TS 103 097 V1.3.1; Intelligent Transport Systems (ITS); Security; Security header and certificate formats," Oct. 2017.
[14] ——, "ETSI TS 103 248 V1.1.1 Intelligent Transport Systems (ITS); GeoNetworking; Port Numbers for the Basic Transport Protocol (BTP)," Nov. 2016.
[15] S. Amante, B. Carpenter, S. Jiang, and J. Rajahalme, "IPv6 Flow Label Specification," Internet Requests for Comments, RFC Editor, RFC 6437, Nov. 2011.
[16] European Telecommunications Standards Institute (ETSI), "ETSI TS 102 894-2 V1.2.1 Intelligent Transport Systems (ITS); Users and applications requirements; Part 2: Applications and facilities layer common data dictionary," Sep. 2014.
[17] ——, "ETSI TS 102 731 V1.1.1; Intelligent Transport Systems (ITS); Security; Security Services and Architecture," Sep. 2010.
[18] ——, "ETSI TS 102 941 V1.1.1; Intelligent Transport Systems (ITS); Security; Trust and Privacy Management," Jun. 2012.
[19] ——, "ETSI TR 103 415 V1.1.1; Intelligent Transport Systems (ITS); Security; Pre-standardization study on pseudonym change management," Apr. 2018.
[20] V. Devarapalli, R. Wakikawa, A. Petrescu, and P. Thubert, "Network Mobility (NEMO) Basic Support Protocol," Internet Requests for Comments, RFC Editor, RFC 3963, Jan. 2005.
[21] J. Petit, F. Schaub, M. Feiri, and F. Kargl, "Pseudonym Schemes in Vehicular Networks: A Survey," p. 33, 2015.
[22] F. Armknecht, A. Festag, D. Westho, and K. Zeng, "Cross-layer Privacy Enhancement and Non-repudiation in Vehicular Communication," Jan. 2007.
[23] G. Calandriello, P. Papadimitratos, J.-P. Hubaux, and A. Lioy, "Efficient and Robust Pseudonymous Authentication in VANET," in *Proceedings of the Fourth ACM International Workshop on Vehicular Ad Hoc Networks*, ser. VANET '07. New York, NY, USA: ACM, 2007, pp. 19–28.
[24] E. R. Verheul, "Issue First Activate Later Certificates for V2X," p. 28.
[25] A. Perrig, R. Canetti, J. D. Tygar, and D. Song, "The TESLA Broadcast Authentication Protocol$_*$," p. 12.
[26] D. Dolev and A. Yao, "On the security of public key protocols," *IEEE Transactions on Information Theory*, vol. 29, no. 2, pp. 198–208, Mar. 1983.