MARMARA UNIVERSITY
INSTITUTE FOR GRADUATE STUDIES
IN PURE AND APPLIED SCIENCES

# A New Method For Protecting Location Privacy In Vehicular Ad-hoc Networks

MUHAMMED HANNY SABBAGH

**MASTER THESIS**
Department of Computer Engineering

**Thesis Supervisor**
Assoc. Prof. Mujdat SOYTURK

ISTANBUL, 2022

# ACKNOWLEDGMENT

I would like to thank my advisor, Assoc. Prof. Mujdat Soyturk, for his continuous

supervision and support for me during this thesis work.

I would like to also thank my family members for supporting me throughout my entire master's degree.

September, 2022 Muhammed Hanny Sabbagh

# TABLE OF CONTENTS

# ÖZET

V2X iletişimine katılan araçların, konumlarını, hızlarını ve araçla ilgili diğer temel verileri içeren ortak farkındalık mesajlarını (CAM'ler) periyodik olarak yayınlamaları zorunludur. Bu yaklaşım, aracın yakınında bulunan herkesin bu mesajları almasını sağlar ve analiz etme fırsatı sunar. Her araç, gerçek kimliğini açığa çıkarmamak için bu mesajları yayınlama sırasında bir dizi geçici kimlik veya "takma ad" kullanır.

Ancak araştırmalar, bu geçici adlar periyodik olarak değiştirilseler bile, CAM mesajlarında yer alan aracın genişliği ve uzunluğu gibi ek veriler nedeniyle bu takma adları birbirine bağlamanın mümkün olduğunu ve dolayısıyla aracın gizliliği ihlal ettiğini göstermiştir. Örneğin, kötü niyetli kişiler aynı araca ait takma adları birbirine bağlayarak aracın tüm geçmiş konum geçmişini ortaya çıkarabilecektir.

Bu çalışmada, daha iyi gizlilik koruması sağlayan (1) mevcut yöntemlerde iyileştirmeler ve (2) CAM mesajlarındaki ek tanımlanabilir verileri hesaba katarak araçlar için gelişmiş özgün bir yaklaşım sunuyoruz. Ek olarak, yöntemimizin performansını Avrupa C-ITS takma isim değiştirme stratejisiyle karşılaştırıp ve sınırlarını araştırıyoruz. Buna yönelik olarak, gizlilik performansı sonuçları gerçek dünya alanı ve araç hareketlilik modelleri ile bir V2X iletişimi dijital ikizi kullanılarak gösterilmekte ve bu çalışmanın bulguları, desteklenmektedir.

# ABSTRACT

It is mandatory for vehicles participating in V2X communications to periodically broadcast cooperative awareness messages (CAMs) containing their location, speed and other basic data about the vehicle. This enables anyone near the vehicle to receive those messages and provides an opportunity analyze them. Each vehicle gets a set of temporary IDs or "pseudonyms" to broadcast these messages in order to avoid revealing their real identity.

However, research has shown that linking these pseudonyms together, even when they are periodically changed, is possible because of additional data in CAM messages such as the width and length of vehicle, and hence, breaks privacy. For instance, by linking pseudonyms that belong to the same vehicle together, the adversary will be able to reveal the entire past location history of the vehicle.

In this work, we introduce (1) enhancements to existing methods and (2) a novel approach that takes the additional identifiable data in CAM messages into account and provides an improved privacy protection for vehicles. Additionally, we compare the performance of our method with the European C-ITS pseudonym-changing strategy, and explore its limitations. To this end, the findings of the study are supported by showing privacy performance results based on a digital twin of a V2X communication with real-world area and vehicle mobility patterns.

# SYMBOLS

**DTT** : Dictionary of average travel times between observer points

$\mathbf{D_{ij}}$ : Average travel time from observer *i* to observer *j*.

$L_{dummy}$ : Reported dummy length value

$L_k$ : List of matched pseudonyms

$L_{real}$ : Reported real length value

$N$ : Normal distribution

$O_k$ : List of pseudonyms collected under observer points

$P_{ij}$ : List of pseudonyms *j* which belong to vehicle i

$R_k$ : List of pseudonyms which were not matched completely yet

$U_{ij}$ : List of used pseudonyms *j* by vehicle *i* during a trip.

$W_{dummy}$ : Reported dummy width value

$W_{real}$ : Reported real width value

$X_B$ : Broadcasted value of quantity *x*

$\mu_x$ : Mean of quantity *x*

$\sigma^2$ : The variance

$\sigma$ : The standard deviation

# ABBREVIATIONS

**C2C** : Car-to-car

**C-ITS** : Cooperative Intelligent Transport Systems

**CAM** : Cooperative awareness message

**DMIX** : Double Mix Zones.

**ETSI** : European Telecommunications Standards Institute

**LuST** : Luxembourg SUMO Traffic

**MLPS** : Multilevel location privacy scheme

**OBU** : On-board unit

**PKI** : Public Key Infrastructure

**RSU** : Road-side Unit

**SAE** : Society of Automobile Engineers

**SDS** : Smart Dummy Strategy

**SLOW** : Silence at low speeds

**TA** : Trusted authority

**TROPHY** : Trustworthy VANET ROuting with grouP autHentication keYs

**UPCS** : Urban pseudonym-changing strategy

**V2V** : Vehicle-to-vehicle

**V2I** : Vehicle-to-infrastructure

**V2X** : Vehicle-to-everything

**VANET** : Vehicular ad-hoc network

**VeNIT Lab** : Vehicular Networking and Intelligent Transportation Systems Research Lab

**VLPZ** : Vehicular location privacy zone

**VPBC** : Velocity-based pseudonym-changing strategy

# LIST OF FIGURES

# LIST OF TABLES

# 1. INTRODUCTION

The vehicle connectivity provided by V2X technology offers many promising benefits for safety and traffic efficiency of vehicles and road users in traffic. With the onboard units (OBUs) in the vehicles and the roadside units (RSUs) in the infrastructure, vehicles can communicate with each other and with the infrastructure. In this way, many new applications and services with enhanced content can be developed and used in the field.

Vehicle connectivity became one of the most important components of the Cooperative Intelligent Transportation Systems and Connected Autonomous Vehicles (CAVs). Vehicles access information directly and increase situational awareness, thereby improving safety and traffic efficiency. In V2X communication, vehicles broadcast cooperative awareness messages (CAM) frequently which include vehicle-related speed, position, acceleration, and other basic status data. The ETSI standard specifies that CAM messages in V2X communication should be broadcasted periodically at intervals of not more than 1000 milliseconds [15]. Anyone within the vehicle broadcasting range can receive, read and process these messages because although they are signed to verify their originality, the content of the messages remains unencrypted to be used for various purposes, like traffic safety and collision avoidance.

CAM messages are very important to increase drivers' situational awareness and ensure safety. By using the data shared in CAM messages, possible accidents can be prevented, damages can be reduced in unavoidable collisions, and casualties can be mitigated e.g. by activating safety airbags. Since these messages are also used to increase traffic efficiency and provide driving comfort, there is a dependency on CAM messages. However, there is a trade-off between safety and security/privacy, as sending these CAM messages open can lead to security and privacy breaches.

Each vehicle that uses V2X technology has a unique registered ID and a set of given temporary IDs or “pseudonym” certificates that are used to sign those messages before broadcasting them. Only the trusted authority (TA) - which manages the certificates in ITS (Intelligent Transport System) infrastructure - can know the real unique ID of a vehicle, because it is never used to sign these messages. Instead, only pseudonym

certificates are used to sign these messages in order to maintain a level of privacy. For enhancing security and privacy, each vehicle is assigned a batch of verified pseudonyms by the TA to be used over different periods of time. Other vehicles that receive the signed CAM messages can verify the integrity of the messages by checking the pseudonym certificate of the sender. Vehicles may periodically replace their currently used pseudonym with a new one from a set of pseudonyms based on predetermined conditions or rules.

However, since vehicles can send a large number of CAM messages in a very short time, it would be quite possible for a global or local adversary to collect these messages over a specific geographic area and track a vehicle's past position in order to figure out what places this vehicle had visited (while using the same pseudonym). More importantly, the adversary may try to match changing pseudonyms belonging to the same vehicle to reveal the vehicle's entire past location history and follow it wherever it goes. This, in turn, can compromise the privacy of the vehicle (and the driver), as it allows the adversary to reveal critical and confidential information about him/her without them knowing.

As a countermeasure, the research community has responded to this problem by introducing many pseudonym-changing methods which tried to weaken an adversary's ability to establish this linkability between different pseudonyms [20]. There is also a recommended pseudonym-changing strategy for the European C-ITS platform, which was introduced in 2017 [25]. However, recent research [11] has shown that regardless of the applied pseudonym-changing strategy, the adversary will still be able to establish the linkability between pseudonyms because of additional vehicle-specific unique data in CAM messages, which are the vehicle's width and length.

As each vehicle broadcasts its own actual width and length, even if a pseudonym change occurs for a vehicle in a particular area, the adversary can still link the new pseudonym to the old one by matching the width and length data. It is shown that for vehicles traveling in a city (in [11], the city is Luxembourg) a simple adversary model can link more than 70% of entire journeys.

# 2. BACKGROUND

In this section, we provide an overview of pseudonym-changing approaches in order to protect a vehicle's location. Furthermore, we examine the related work in the literature along with the recommended European C-ITS method.

## 2.1. Pseudonym Change to Defeat Linkability

In a typical scenario, the adversaries aims to reveal the entire journey route of a vehicle (or group of vehicles) they have interest in. This way they can carry out further attacks, or target the driver of the vehicle in different modes of harmful activities.

These adversaries can be categorized as global or local based on the scope and the coverage [20]. A global adversary means that he has full coverage on the area he is monitoring and can collect any CAM messages broadcasted. A local adversary means that he can only monitor small "observer" points that he manually places in different parts of the city or area he is monitoring, and does not have global coverage. The adversaries can be categorized as well as internal or external based on his capabilities to achieve attacks from inside or outside of the system, respectively. An external adversary means that he can only monitor and collect CAM messages, without cooperating with any compromised vehicle on the road, while an internal one can additionally gain more data from compromised or malicious vehicles. Fig. 2.1 illustrates an example of observer points in which an adversary collects CAM messages from the vehicles. In this scenario, an adversary which is classified as local is able to collect all CAM messages inside the shared areas but is not able to collect anything from outside of the shaded areas.

In order to avoid a situation where an attacker could track a vehicle's exact past location history, a method is used in V2X communications which is very similar to the method used in cellular networks and other communications: Instead of signing messages (e.g. CAM messages) with their unique ID (or unique private key), they are signed using pseudonym certificates that change constantly at regular intervals. In this way, due to the changing pseudonyms, the adversary will not be able to trace the entire movement of the vehicle, as it will get messages signed with different names.

**Figure 2.1.** Two example observer points in an urban area, which can collect broadcasted CAM messages under the highlighted red range.

However, research has shown that a simple pseudonym change will not be enough to destroy this linkability, as the adversary will still be able to correctly match and link the new pseudonym to the old one, by processing data coming from the continuously-broadcasted CAM messages [31]. For example, the adversary can detect that only one vehicle in a group of vehicles has changed its pseudonym, so it can instantly link the old one to the new. Additionally, if the pseudonym-changing process happens in a specific rate, then the adversary can estimate the pseudonym-changing times for the vehicles which he is able to monitor, and hence, he can try to match them together based on the data in CAM messages (location, width, length, speed, acceleration, etc.).

At the same time, it is important to provide privacy whilst not harming the safety aspects of CAM messages, as there is a trade-off between safety and privacy; one, for example, could easily turn off CAM messages altogether or under specific conditions as a defense method to deal with this problem. This approach will help achieve 100% protection against the adversary, but this will destroy everything related to safety applications in V2X communications, which is why it is not preferred and recommended in the literature [15]. Therefore, neither simple pseudonym change nor

turning off CAM messages are good enough, and a new, perhaps smarter method is needed to protect the vehicles' privacy against adversaries.

## 2.2. Related Work

This subsection provides a comprehensive review of related studies/approaches in the literature on protecting vehicle privacy.

### 2.2.1. Categorization of Defense Techniques

There is not just one approach to protect location privacy in V2X communications, instead, there is a number of possible defense techniques that can be employed. Below we provide a categorization of these techniques:

- **Cryptographic methods:** Those are methods which depend on encrypting the V2X communication in order to provide privacy. Such as encrypting the broadcasted CAM messages so that no external adversary can access them.
- **Mix-zones methods:** Which are methods that apply a pseudonym-changing process depending on the geographic area where the vehicle is currently located. Such as applying that process at road intersections, shopping malls and/or other places under different conditions.
- **Context-based methods:** Which are methods that apply the pseudonym-changing process depending on the current driving conditions of the vehicle (its current speed, number of surrounding vehicles etc.) so that it gets triggered whenever those conditions are met.
- **Swapping methods:** Which make vehicles change their current pseudonym by swapping it with the current pseudonym of another vehicle in the same vicinity, under specific conditions related to the vehicle.
- **Obfuscation methods:** Methods which do not mainly aim to change the currently used pseudonym of the vehicle, but instead, obfuscate the contents of CAM message in order to make them useless for a possible adversary. Such as obfuscating the location data to provide false vehicle location.

This classification is not absolute, however; one can find many methods which use one

or more of these defense techniques combined in order to provide location privacy in V2X communications. Such methods can be classified as “hybrid” methods.

### 2.2.2. Other Methods in Literature

Benarous et al. [1] proposed the “Alloyed pseudonym-changing strategy”. The pseudonym-changing process can be triggered in two ways, depending on the value of *k*, where *k* is the number of other vehicles in the vicinity near the vehicle. If *k* is bigger than a certain desirable threshold, then a normal pseudonym change process takes place and a new pseudonym is assigned instantly, because there are many vehicles in the vicinity which allows confusing the attacker. If *k* is less than the desirable threshold, then both the location and speed data of the vehicle will be obfuscated to confuse the attacker. The vehicle speed will be set to 0, and the vehicle location will be set to a fixed spot (the location just before the process takes place). This will go on for a short determined time *t*, after which the vehicle can get a new pseudonym. In case of a detected emergency on the road, the vehicle will stop the obfuscation process and restore the location and speed data to real values in order to avoid any accidents.

Boualouache et al. [3] proposed the VLPZ method. It requires deploying “privacy zones” at the entrances of common vehicle areas (like gas stations, traffic lights and shopping malls, etc.). A “router” device is used at the entrance of the privacy zone that requests each vehicle to change its position according to a specific way (e.g change the current lane it is standing in to another lane), and an “aggregator” at the exit of the privacy zone is used to randomize the waiting time before the exit of each vehicle. In between these two, synchronized pseudonym-changing process happens for vehicles in the privacy zone. The order of the entering/exiting vehicles should not be the same in the zone. To ensure that the adversary can not collect CAM messages, CAM messages broadcasting is turned off while the vehicles are in the privacy zone.

Boualouache et al. [2] proposed the UPCS method, which stands for Urban Pseudonym-changing Strategy. In this method, vehicles waiting at a traffic light can choose between two pseudonymization approaches: Either change their currently used pseudonym by them selves, or exchange it with another waiting vehicle's pseudonym. During waiting at the red light, CAM messages broadcasting is turned off until the green light is turned

on.

Bouchelaghem et al. [5] proposed a method that utilizes encryption of CAM messages between vehicles in order to provide privacy against a passive global adversary. RSUs and TA will be responsible for managing the overall of keys management, while vehicles will be able to verify the integrity of CAM messages by checking the public key of the corresponding vehicle.

Buttyan et al. proposed SLOW [6], a method in which vehicles stop broadcasting safety channel messages while they are running at slow speeds such as 30 km/h. While SLOW is good in principle to avoid local and global adversaries, turning off the safety channel messages may not be an option in dense populated cities, because it might affect their safety as we described earlier. Authors argued that only 5% of road accidents cause death below 30 km/h, and hence, they believe it is negligible.

Cirne et al. proposed TROPHY [8], considering a global active adversary, which is a complex cryptography-based method that utilizes the encryption of CAM messages. There is a human-controlled "KDC" (Key distribution center), which can issue encryption keys for vehicles in the VANET. Each vehicle already has the public key of the KDC. Vehicles can decrypt CAM messages using the keys distributed to them by the KDC, and in case of a compromised vehicle or encryption key, then the KDC sends a "refreshment message" to all RSUs in order to distribute the new keys. Additionally, vehicles also help in distribution of the new keys which they got from the RSUs to other vehicles they meet on the roads, so that they don't have to pass by a RSU to get the new ones. The other vehicles can verify that these new keys are given by the KDC, because they are signed using its private key (For which, they already have the corresponding public key).

Eckhoff et al. [10] introduced a pseudonym-changing strategy that depends on swapping pseudonyms between two vehicles if they are located in the same traffic lane whenever the needed conditions are met. In this way, even if the adversary links the pseudonyms together, then he will not guarantee that the same pseudonym wasn't used beforehand by a different vehicle.

Li et al. proposed PAPU [21], considering a global passive adversary, which is a

pseudonym-changing strategy that utilizes pseudonym swapping between vehicles. Researchers have acknowledged that law regulators' need to unveil the real unique ID of any vehicle in VANET in case of a malicious attack or misbehavior, which is why pseudonym swaps are not done by the vehicles themselves. Instead, when a vehicle desires to get a new pseudonym, it needs to be near a RSU so that it can be grouped with other vehicles wishing to get a new pseudonym. Different weights are assigned for each vehicle depending on the trajectory, location and speed parameters, and the more likely two vehicles share the same driving conditions, the more likely their pseudonyms are going to be swapped. The RSU then creates a pool containing the pseudonyms for all vehicles in its range which are wishing to change their pseudonyms, and then swaps them between vehicles based on the previously assigned weight values and an exponential utility calculation to maximize the differential privacy.

Lu et al. [22] introduced Social Spots. In Social Spots, vehicles change their pseudonyms when waiting at traffic lights, gas stations, marketing shops and similar areas which are typically crowded. Vehicles achieve better privacy results when they change their pseudonyms together in these areas.

Sampigethaya et al. [24] introduced CARAVAN. The method depends on creating groups for the moving vehicles on roads on the condition that every vehicle can receive CAM messages from all the other group members. A "group leader" is selected for the vehicles which takes broadcasting on behalf of other vehicles, which go to silence period depending on the time for which they remain in the group. The silence period time is randomized to avoid the adversary's ability to expect it at a fixed rate.

Ullah et al. introduced VBPC. [27] A method which depends on grouping moving vehicles on the roads to different groups based on their current velocity, and then, change their pseudonyms together according to conditions related to travel distance or travel time that they have driven together. No silence period is applied, but the synchronized pseudonym change for the vehicle group makes it harder for the adversary to reveal their real identity.

Ullah et al. Introduced MLPS. [26] In this method, three possible levels of location data obfuscation are possible depending on the number of surrounding vehicles (vehicle

density). The first level is applied when the vehicle density is high, and there are two vehicles before and after the current vehicle. 4 random locations will be selected on the area between the vehicles, and they will be reported as if they are real vehicles but with the dummy location that has been determined for each one of them and a different pseudonym for each, fooling the adversary to think that they are real vehicles. The second level is applied when there is only one surrounding vehicle in the transmission range of the current vehicle. Again 4 location points in the area between the two vehicles are going to be selected, and they will be reported as of they correspond to real vehicles on the road to fool the adversary. The third level is applied when there are no surrounding vehicles, and in this level, one random location points will be selected in the ranges of 0-100 meters, 101-200 meters and 201-300 meters respectively (three in total). They will be reported as if they were additional vehicles driving nearby the main vehicle to confuse the adversary.

Zhou et al. [32] introduced DMIX, or "double mix-zones". Their method depends on grouping vehicles in the same lane direction, and then changing their pseudonyms synchronously. In this way, there would be "double" privacy zone in an area, and not just one. Additionally, a silence period is applied for a short time, and wrong/virtual location data is provided instead of the real one to confuse the adversary trying to guess the location of the exiting vehicles. Real and correct location data is gradually restored after a short period of time.

Fig. 2.2 shows the categorization of these methods and to which defense technique category they belong to.

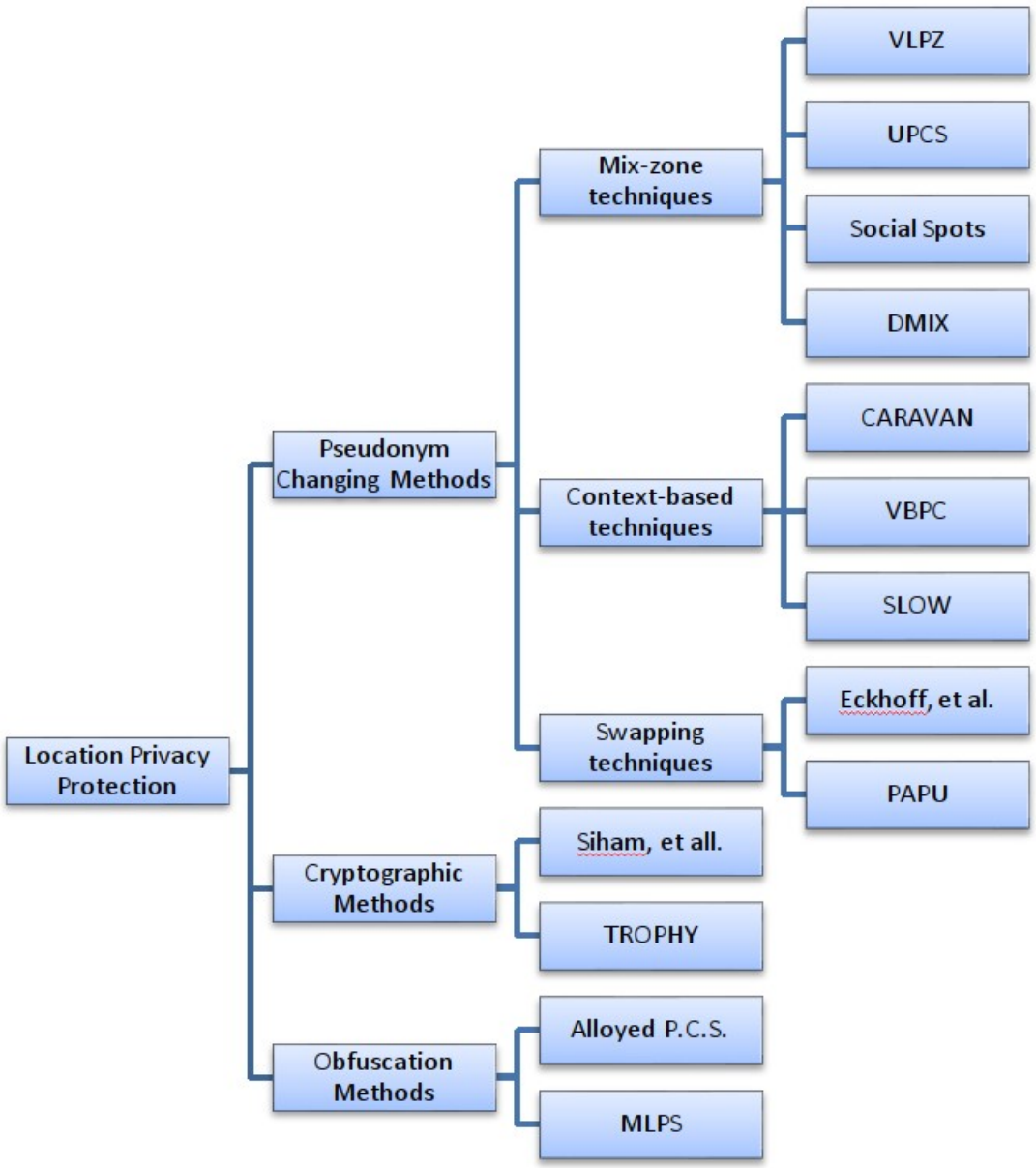


**Figure 2.2.** Categorization of different privacy protection methods in V2X communications.

There are studies ([4], [18] and [20]) in the literature which survey the pseudonym-changing methods, and provide an additional categorization for them.

Many of these mentioned methods include using the concept of silence periods to enhance privacy, and it is indeed one of the finest techniques to apply (without considering other factors like safety) [19], but this usage is discouraged, even in the pre-

standardization study on pseudonym-changing strategies by ETSI [15], because it affects safety applications. However, it may not be possible to protect privacy at all against a global passive adversary (GPA) without the usage of silence periods, because the adversary is capable of collecting every CAM message anywhere and at any time.

Cryptographic methods, which depend on encrypting the entire communication between V2I and V2X devices in different ways in order to maintain privacy without depending on pseudonym-changing, indeed do provide better privacy but at the cost of too much overhead, and the continuous need to defend the network stack against outsiders who may retrieve the encryption keys either in part or in full.

Obfuscation methods are good in theory, but to what data and how should the obfuscation be applied? So far, most methods suggested obfuscating vehicle location and speed data and broadcasting false ones (to everyone in the vicinity, both the possible adversary and other road users) in order to protect privacy, but again, this affects the safety aspect of V2X and may endanger people lives.

As for pseudonym swapping method, the ETSI did not recommend using this approach if it would lead to identity mismanagement and the inability to detect misbehavioring vehicles (because if you swap the pseudonyms of two vehicle, and one of them breaks the law afterwards, then the previous owner of the pseudonym may be in trouble). This was an approach recommended by [10], but a better approach was recommended by [21] which updates pseudonyms ownership whenever it is swapped between two vehicles if they were near RSUs. However, it adds an overhead work to RSUs, and is not applicable if few RSUs are available in the vicinity or none of them, which makes it limited.

Finally, both mix-zones methods and context-based methods have different ways of application and they provide different levels of privacy. But the issue with the first is that once the vehicle leaves the privacy zone, it becomes vulnerable. Moreover, the privacy zones them selves, or their entrances and exit points, will become an attack target for the adversary, because it would be more beneficial for him to monitor them to get a higher success rate than random places in the city. The latter, on the other hand, while good in theory (being vicinity-aware and applying the strategy only smartly under

some conditions is a good plus), but could not introduce an approach that does not affect safety applications so far in V2X communications.

ETSI recommends to use a pseudonym-changing strategy which is described in the following subsection.

## 2.3. European C-ITS Pseudonym-changing Strategy

Due to the mentioned privacy concerns, the C2C Communication Consortium recommended a pseudonym-changing strategy [23], which was later adapted by the European Commission and recommended in [25] to be implemented in the deployment and operation of European C-ITS.

In this strategy, the pseudonyms are changed based on the following rules:

- Initial pseudonym: On engine start (if turned off for 10 minutes).
- 1st change: After 800 to 1500 meters in travel distance.
- 2nd change: After 800 meters of travel distance AND an additional 2 to 6 minutes in driving time.
- 3rd change: After an additional 10 to 20 kilometers in travel distance.
- 4th and every further change: After an additional 25 to 35 kilometers in travel distance.

However, it is discussed in [11] that this strategy does not protect privacy well in the urban area, and the adversary can achieve a great attack success rate. In this study, the researchers presented how well a local adversary can link the changing pseudonyms when the European C-ITS pseudonym-changing strategy is used. Their work comprises the city of Luxembourg using the LuST scenario [9], which contained around 214,000 moving vehicles. They have found that although vehicles changed pseudonyms many times during their trips, an adversary can succeed in linking the full trips of more than 70% of the vehicles (with 200 observer points). They achieved this success rate by utilizing the vehicle width and length information contained in the CAM messages.

According to the ETSI standard [17], broadcasted CAM messages have to contain the real width and length of the vehicle. This is important for safety applications to

precisely estimate movements of surrounding vehicles, and to avoid possible accidents and collisions. However, it allows the adversary to greatly increase his success rate in linking different pseudonyms for the same vehicle to each other, because the probability that there is another vehicle with the exact same width and length combination in the vicinity is so small. Hence, the width and length data are helpful information for the adversary in matching the pseudonyms.

It is concluded that it may not be possible to achieve a good level of location privacy in V2X communications, regardless of the applied pseudonym-changing strategy. The use of real width and length values in CAM messages allows the adversary to link the pseudonyms and reveal the traces of the vehicles. However, the previous works in the literature related to pseudonym-changing did not take this weakness into account when introducing their new protection strategies, so this situation emerges as a new problem in the literature.

## 2.4.System Model

ETSI described and specified the use of certificates (as well their distribution) in the standards; ITS Communication Security Architecture and Security Management in [14], and Trust and Privacy Management in [16]. Certificates are published and distributed (mainly from the TA, and later distributed from RSUs for all vehicles) in order to be used for security and privacy. In this case, pseudonyms or pseudonym certificates are going to be used for signing the CAM messages.

Fig. 2.3. shows a simple vehicle enrollment process in a C-ITS system. The vehicle must first register itself with the central authority before becoming able to communication with other vehicles or RSUs on the road. After this, each vehicle will be given a unique registered ID and a set of given temporary IDs or "pseudonym" certificates that can be used to sign CAM messages before broadcasting them. "Signing" a message means sending it with the sender's signature such that any receiver can verify the sender as well as the integrity of the message.

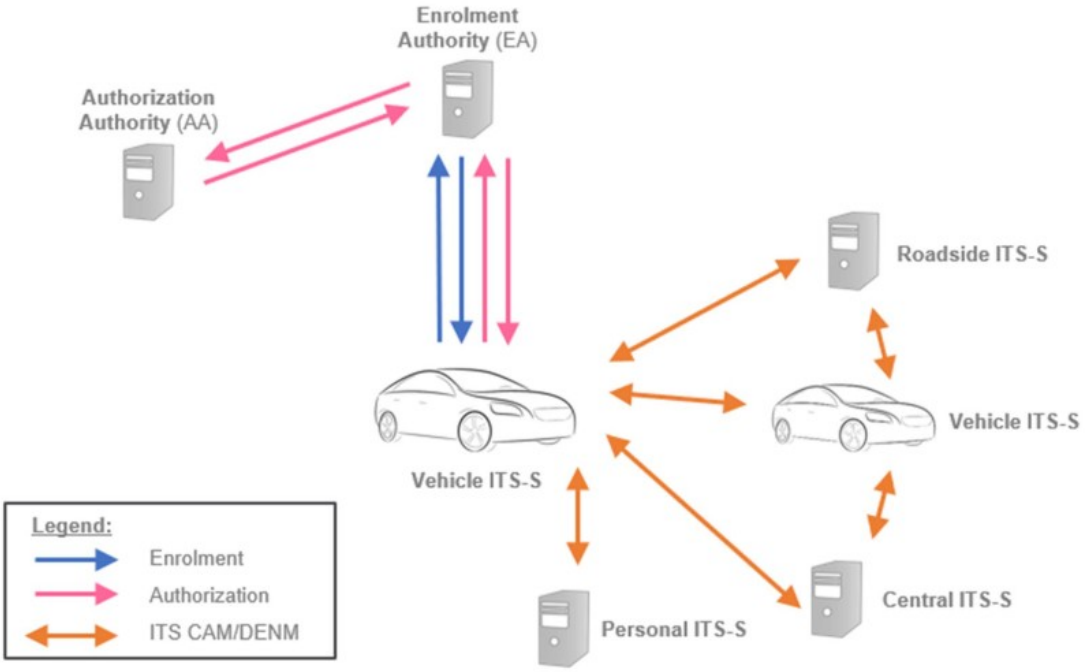


**Figure 2.3.** Vehicle enrollment and registration in a C-ITS system, according to the ETSI standard [16].

# 3. METHODOLOGY & PROPOSED APPROACH

This section presents the methodology used in our research, as well as the description of our proposed method. It also explains the experiment scenario we apply to measure the results.

## 3.1. General Methodology

In our work, we consider a local adversary which can install multiple observer points over a specific geographic area that he wants to monitor. It is unlikely that the adversary will succeed in achieving a global coverage, as the cost, time and efforts needed to carry out such attacks are quite massive and can not be done in most cases.

Additionally, we mainly aim to provide a useful method that can be used in urban areas, where the travel distance is shorter than the travel distance in highways. We believe that the majority of people in most cities will be driving their cars from home to work, or home to other places in the same city most of the time, so we focus on the shorter trips instead of highways or city-to-city travels in this work. Further, our studies can easily

be extended to cover larger regions and longer trips as needed.

The adversary will install observer points over the area he wants to monitor, start collecting CAM messages and then start doing some statistics on the traffic flow (average travel time from one point to another, average traffic lights waiting time, rush hour versus non-rush-hour statistics, etc.). Then, once he has enough information about the traffic movement, he can start collecting CAM messages in order to carry out his attack to link different pseudonyms of the same vehicle together based on different factors. There are many possible models for the adversary, ranging from simple to complex depending on the effort provided to implement them.

If there are many vehicles in the vicinity which have the same width and length, then the adversary is more likely to do a false-positive match. But if it is the other way around (which is more often), then the adversary can easily match the pseudonyms which was indicated in the literature for the first time in [11]. In our work, we also consider the width and length data fingerprinting as a possible matching mechanism the adversary may be using.

However, it is important to note that in [11] the researchers are seeming to use millimeter measurement unit (mm) for vehicle width and length data in their experiment. We believe that this is a huge bias in favor of the adversary, because it would help him in identifying the vehicles even further since the probability of having another vehicle with the exact millimeter measurement in the vicinity is extremely unlikely. The ETSI standard of the EU for reporting vehicle width and length data is using decimeters, not millimeters [17]. And the SAE standard, which is used in the US, is using centimeters [28]. This means that the adversary will face additional difficulty in his matching process, because widths and lengths are not reported in a more identifying fashion like millimeters. In this work, we also compare the performance of our method using the width and length measures in both the ETSI and SAE standards.

**Table 3.1.** Minimum and maximum width and length values for each vehicle category, as well as the difference between them (millimeters). [7]

| Category Name | Min-Max Width | Width Difference | Min-Max Length | Length Difference |
|---|---|---|---|---|
| City Cars | 1615-1683 | 68 | 3466-3686 | 220 |
| Small Cars | 1665-1848 | 183 | 3700-4088 | 388 |
| Compact Cars | 1765-1860 | 95 | 4108-4400 | 292 |
| Family Cars | 1760-1958 | 198 | 4419-4697 | 278 |
| Executive Cars | 1820-1966 | 146 | 4709-4989 | 280 |
| Luxury Cars | 1890-1978 | 88 | 4851-5391 | 540 |
| Sports Cars | 1735-2098 | 363 | 3915-5034 | 1119 |
| Estate Cars | 1732-1890 | 158 | 4262-4986 | 724 |
| MPV Cars | 1751-1916 | 165 | 4242-4857 | 615 |
| Small SUV | 1622-1848 | 226 | 3700-4195 | 495 |
| Compact SUV | 1757-1904 | 147 | 4205-4510 | 305 |
| Mid-size SUV | 1784-1922 | 138 | 4509-4726 | 217 |
| Large SUV | 1839-2016 | 188 | 4708-5207 | 499 |

The market contains many vehicles from different companies, and they come in different width and length combinations (as in Table 3.1). We inferred that in [11] researchers used at least 80 different combinations of widths and lengths (taking 2700 vehicles as an average per combination out of 214,000), therefore, we are using the same number in our work in order to be able to reproduce similar results.

Vehicles which have a unique width and length combination in the city can always be tracked, but as mentioned in [11], their percentage does not exceed 0.03% of total vehicles. We have omitted these vehicles from our experiment, and assume that there are no vehicles with unique width and length combination.

In our experiments, we are using a pool size of 60 pseudonyms, as recommended by the C2C Communication Consortium [15]. All of these pseudonyms are unique to every vehicle (no sharing), and re-use is allowed.

## 3.2. Proposed New Method

We propose a new method based on the recommended European C-ITS pseudonym-changing strategy with some modifications on: (1) the pseudonym-changing frequency and intervals, (2) the reported width and length values in the CAM messages.

We observed that in the European C-ITS pseudonym-changing strategy, vehicles do not change their pseudonym so frequently which allows the adversary to trace the vehicles. For example, we observed that most vehicles only change their pseudonym once or twice during short trips, which is not good for a small urban area like a city, town, or village to provide privacy. In our experiments, we have also observed that the 3rd pseudonym change may not happen at all if the total travel distance is less than 10 km, and similarly, the 4th pseudonym change only occurs on additional 25 km of travel distance, in total after around 35 km in travel distance.

It is inferred that this original strategy is designed for vehicles on highways or vehicles at high speeds. We consider that pseudonym changes should be made more frequently in small areas or areas with slow-moving vehicles (for example, due to traffic density or congestion). Otherwise, the adversary will be able to obtain the same pseudonym in many areas (observer points), which would harm location privacy and enable the adversary to trace the vehicles.

It is essential for the vehicle to change its pseudonym more frequently considering the issues described above. For this purpose, we modified the steps in the strategy with some relaxations which enforce the vehicle to change its pseudonym more frequently to provide privacy.

In these relaxations:

- The condition for 2nd pseudonym change has been changed from "AND" to "OR".
- 3rd and further pseudonym changes apply the same rule given above.

The use of "OR" approach in the first relaxation enforces slow-speed vehicles to change their pseudonym in the same area, and also enforces fast-speed vehicles to change their pseudonyms more frequently to avoid the use of the same pseudonym over many areas

(observer points). The second relaxation removes the 10-25 km and 25-35 km travel distance conditions and avoids the use of the same pseudonym for a long time and over many areas. Which greatly enhances location privacy.

Since width and length data are responsible for the privacy threat, this has led us to believe that obfuscation of length and width data could be a possible solution for this problem. We propose a novel method called "Smart Dummy Strategy (SDS)" to prevent an attacker from utilizing such data. By using this method, adversaries will not be able to process this data in their attacks, because instead of finding the original correct data, they will find changed width and length data which overlap between many vehicles, rendering the data useless in such attacks whilst not harming safety.

While at the first instance, it might be seen as unreasonable and perhaps risky to obfuscate the width and length data, we argue that changing the width and length data slightly than the real values will not be a problem in most - if not all - cases. Moreover, the SAE and the ETSI approaches do not use the exact measurements in millimeters either. They instead use the truncated values to centimeters and decimeters respectively.

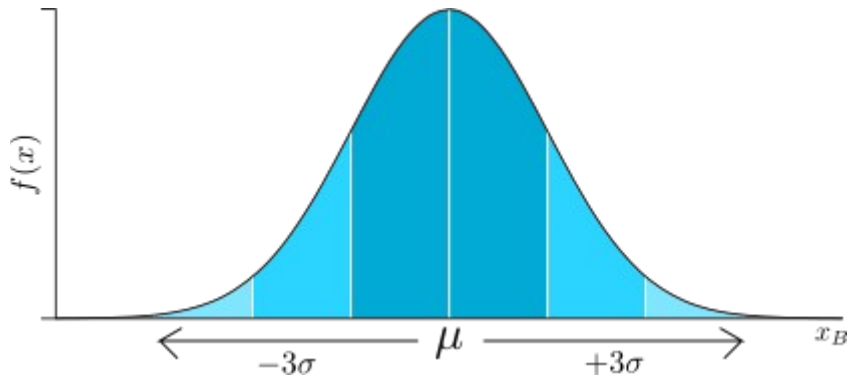


**Figure 3.4.** An example normal distribution showing that more than 99% of broadcasted values for length or width ($x_B$) will fall between ±3σ from the mean ($\mu$), which is the real value for the width or length.

Our proposed method for obfuscating width and length data is as follows: When transmitting their length and width in CAM messages, instead of the actual length and width, the vehicles transmit dummy but realistic values that are randomly generated for each CAM message. Various probability distributions can be used for random value

generation, but it is important to generate a random value and keep the length/width values constantly changing at each CAM message. For example, by using the Normal Distribution, realistic values can be produced more often and unrealistic values can be avoided. In our case, each vehicle, instead of reporting its real width ( $W_{real}$) and real length ($L_{real}$) in every CAM message, should report a random (dummy) width ( $W_{dummy}$) and a random (dummy) length ($L_{dummy}$) generated from Normal Distribution $N(\mu, \sigma^2)$ with the real width and length as the mean, and a standard deviation (σ) of 5 centimeters.

Based on the Empirical Rule (and properties of Normal Distribution), 68.26% of broadcasted $W_{dummy}$ and $L_{dummy}$ values for width or length will be in the range of $\mu \pm \sigma$. 95.44% of values will be in the range of $\mu \pm 2\sigma$, and 99.76% of the possible probabilities are going to be in the range of $\mu \pm 3\sigma$, which will not exceed ±15 cm in the case of our proposed approach, as can be seen in Fig. 3.1. This is because 95.44% of the dummy values will be at most 10 cm deviated from the mean, we believe that this small offset from the real values (which the will be truncated to decimeter or centimeter before insertion into the CAM based on the ETSI and SAE, respectively) will not cause a problem for safety applications of V2X communications.

In cases where the exact vehicle width and length values would be important for specific safety scenarios e.g. parking and stopping areas, the real values can be reported in CAM messages. This, however, does not affect the remaining journey of the vehicle in terms of privacy.

The SDS method greatly destroys the adversary's ability to match and link vehicles pseudonyms with the vehicles' width and length data since all vehicles are reporting varying width and length values in every single broadcasted CAM message. Therefore, the adversary will not be sure whether the new pseudonym is linked to the old one since the all vehicles will be using similar width and length combinations generated from random distribution with its own $N(\mu_{length}, \sigma^2_{length})$, $N(\mu_{width}, \sigma^2_{width})$.

## 3.3. Adversary Model

We use an adversary model similar to the one used in [11] but with slight changes, which was originally introduced as part of a framework by [12].

The goal of the adversary is to reveal the full trip of each vehicle. Assume that each vehicle $i$ is allocated a set of pseudonyms $P_{i,j}$ ($j$ is the index for the pseudonym), to use for privacy and security purposes, and assume that $U_{i,j}$, a subset of $P_{i,j}$ are used during the trip of vehicle $i$. In the adversary model, a set of $O_k$ pseudonyms will be collected by the adversary via broadcasted CAM messages under the observer points. The adversary model takes $O_k$ and aims to link the pseudonyms by building the set of $L_k$, which are linked pseudonyms. $R_k$ is the set of pseudonyms which are not linked yet by the adversary. The output of the adversary model will be a set of linked pseudonyms which belong to the same vehicle, with the corresponding observer points under which they were detected (and eventually the location traces observed at those observer points).

**Algorithm 1** Estimation method for the adversary model, based on the average travel time and width/length data.

**Require:** *DTT* (Dictionary of Travel Time): is a key-value store for average travel times from current observer *i* to all remaining observers, where *i* is index for the observer of the current selected CAM message/pseudonym where it was detected (if there is no travel occurrence between *i* and an observer point, then $d_{ij}$ will be 0).

**Require:** $R_k$: Remaining pseudonyms together with other attributes in CAM messages.

```
for each Pseudonym in R_k do:
  for each observer and travel_time in DTT do:
    expected_time = Pseudonym.time + observer.traveltime
    error_margin = x
    candidate_list = []
    for each P in [R_k – Pseudonym] do:
      if expected_time is between P_time ± error_margin then:
        Add P to candidate_list
      end if
    end for
    for P_candidate in candidate_list do:
      If P_candidate_width = Pseudonym_width and P_candidate.length = Pseudonym_length
      then:
        Consider it a successful hit and break the current loop.
      end if
    end for
  end for
end for
```

There are three possible scenarios for linking pseudonyms by the adversary:

- The vehicle is using the same pseudonym under different observer points, allowing the adversary to easily see where the vehicle is moving.
- The vehicle does the pseudonym change under an area covered by an observer point, which instantly enables linking both pseudonyms together and adding them to $L_k$.
- For the remaining pseudonyms (and other attributes), $R_k$, which were not matched yet, the adversary will build an estimation model to try to match them based on average travel time from one observer point to another, as well as vehicle width and length.

In the adversary model, the adversary is able to collect all CAM messages which were broadcasted under his observer points. Unlike [11], we do not define enter events and exit events for an adversary observer point. Instead, we define "occurrences" which tell the adversary that a vehicle was caught at a specific observer point at a specific time with the vehicle's metadata (CAM message information).

For pseudonym changes which occurred under the observer points, the adversary will match the new pseudonyms with the old ones instantly (and put them in $L_k$), because he is capable of detecting the change at any time and location under his observer points. For pseudonyms and occurrences which were not matched so far, $R_k$, the adversary will run an estimation model that estimates whether a specific vehicle that appeared in any time is the same one detected in a new observer point. To do that, the adversary must have already gone through a learning phase, where he collects the average travel time from each of the observer points to all other observer points. In this way, the adversary can expect that a vehicle that just left the $X$ observer point should be in the $Y$ observer point after $M$ seconds. Further, in order to increase his success rate, the adversary will match the detected width and length of the vehicle with the new vehicles that appeared after the expected travel time, and if they are the same, the adversary will consider it a match. Fig. 3.2 presents the estimation method applied for the third scenario of the adversary model when trying to match two detected occurrences. The adversary also allows a small window period before and after the expected arrival time from one

observer point to another, in case of possible delays that could happen for the vehicle.

Finally, in order to evaluate the results of the adversary model, each vehicle will reveal its traces (the partial traces at the observer points it had traveled) using its unique IDs, as well as the pseudonyms $P_{i,j}$ that belong to it. The matched pseudonym list by the adversary will be compared to each vehicle's $P_{i,j}$, as well as the observer points they appeared under. If they completely match, then it would be considered a success.

# 4. RESULTS & EVALUATION

This section contains performance results for all the pseudonym-changing methods applied under a realistic environment, including our novel approach and the incremental changes applied on the European C-ITS pseudonym-changing strategy.

## 4.1. Experiment Setup

We have evaluated the performance of the proposed approaches with a comprehensive and realistic platform. We have used the digital twin platform "vTRIx" [30] built with the support of VeNIT Lab [29] for V2X communication and mobility. vTRIx uses the real V2X protocol stack (ITS G5 [13]) and includes various components for connected cars and V2X communications. It allows (a) large-scale demonstration (b) with realistic scenarios (c) with real data exchanged in real V2X messages (d) with the vehicles running on real map of the operation area.

**Table 4.2.** Average travel time and travel distance for vehicles in the simulation.

| | Average Travel Time (Seconds) | Average Travel Distance (Meters) |
|---|---|---|
| 175 Vehicles (Low density) | 432 | 4668 |
| 350 Vehicles (Medium density) | 504 | 4800 |
| 542 Vehicles (High density) | 532 | 4743 |

**Table 4.3.** Experiment parameters.

| Parameter | Value |
|---|---|
| Area size | 13 km |
| Message types | CAM |
| Protocol stack | ITS G5, WSMP |
| Vehicle density | low (175 vehicles),<br>medium (350 vehicles)<br>high (542 vehicles) |
| Vehicle speed (average) | 37.8 km/h |
| CAM message period (message frequency) | 100-1000 ms (1-10 Hz) |
| Experiment duration | 20 minutes |
| Number of runs | 25 runs |
| Methods Evaluated | 1. European C-ITS strategy "original"<br>2. European C-ITS strategy with "no trip segmentation"<br>3. European C-ITS strategy with "OR method"<br>4. SDS method<br>5. Original strategy with SDS method |

We conduct our experiment on a real geographical area with a size of around 13 km which is located at Kadikoy district of Istanbul - Turkey. We defined a realistic scenario for vehicle mobility considering the mobility pattern of the vehicles in that region. The traces for 542 vehicles for 20 minutes are generated using the "vTRIx" tool with real CAM messages generated and exchanged among the vehicles. The transmitted and received CAM messages are recorded for post-processing and applying the adversary model with the described methods.

In the experiments, various vehicle densities are used to observe the effect of traffic density on the adversary success rate and its impact on the privacy. The parameters in the experiments are given in Table 4.1 and Table 4.2.

We consider that the adversary is capable of collecting CAM messages at observers located at some points in the experiment area. In a real-world implementation, the adversary would be able to analyze collected CAM messages in real-time. Similarly, we

apply a post-processing approach analogous with the adversary model for the proposed pseudonym-changing methods and for their performance comparison. Vehicles in the experiment may pass through multiple observer points, which will allow the adversary to build the vehicle traces and threaten their privacy.

In our experiment, vehicles pass through 4 observer points at least. The reason for using 4 observer points (at least) is related to the success rate of the adversary (the strength of the adversary model) and the strength of the proposed novel pseudonym-changing approach. The more observer points the adversary needs to detect the vehicle under in order to achieve a success hit, the more difficult and unbiased the experiment would be for him. (e.g matching the trace of a vehicle passing through just 2 observer points in its journey is much easier than 3 or 4 or 5, etc.).

The initialization time of vehicles in the simulation is given in Fig. 4.1, which shows that all vehicles are initialized (Deployed on the road) by time 800 seconds (13.33 minutes).

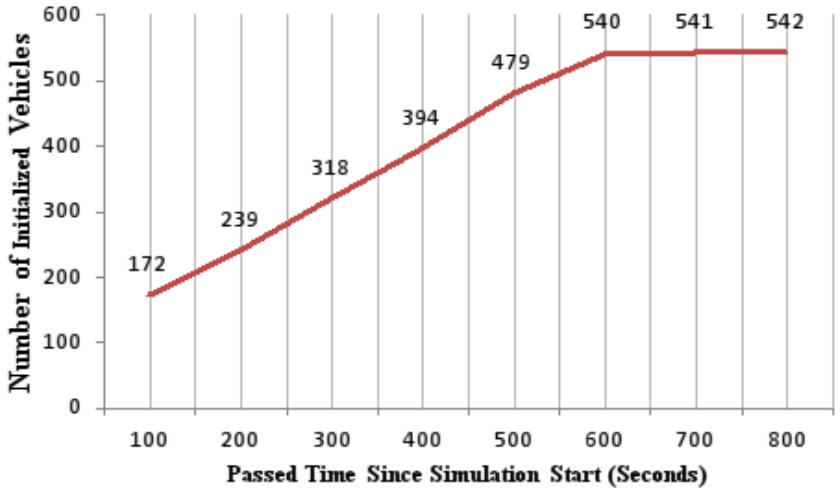


**Figure 4.6.** Initialization of vehicles in the simulation according to passed time (seconds).

## 4.2. Observers Setup

**Figure 4.7.** Highlighted simulation area in Kadikoy, Istanbul - Turkey. Circular (i) blue boxes are the placed observer points.

We consider a local adversary which can collect the broadcasted CAM messages under 100 meters of his manually-placed observer points, which is the same range which was defined by [11]. We are using 30 observer points which are placed uniformly over the operation area 400 meters distance from each other at minimum. We also aimed to place them at traffic junctions and road intersections, but they can be found in other places as well. Those places are more likely to contain more vehicles, and hence, can give the adversary more chances to capture higher number of CAM packets. The observer points are dispersed over the experiment area as can be seen in Fig. 4.2.

## 4.3. Performance Results

The following subsection presents the results of experiments on pseudonym-changing strategies and methods. All results are the averages of 25 runs.

In our studies, we have found that indeed as suggested [11], using the width and length data as additional parameters to link different pseudonyms of the vehicles will dramatically increase the success rate of the adversary. However, the work in [11] does not consider ETSI and SAE standards for CAM messaging format. The work in used exact width and length values in millimeters to be reported in CAM messages, but the

standards (SAE and ETSI) are using centimeters and decimeters units respectively for width and length reporting in CAM messages. We focused on the standards in our work, therefore, this thesis presents the results based on the standards and the applied pseudonym-changing strategies.

There are 5 methods which were tested and evaluated in our experiments:

- **Method 1** - Original European C-ITS pseudonym-changing strategy is applied without any modification.
- **Method 2** - European C-ITS pseudonym-changing strategy with "no trip segmentation": Is a slight modification to the original strategy. This corresponds to the second relaxation where the 2nd pseudonym change rule is used regardless of travel distance. We named this method as C-ITS strategy with "no trip segmentation".
- **Method 3** - European C-ITS pseudonym-changing strategy with "OR method": In addition to the 2nd method above, the condition for 2nd pseudonym change has been changed from "AND" to "OR". We named this method as C-ITS strategy with "OR method".
- **Method 4** - SDS method: Is the novel approach we proposed. In performance results, SDS method is evaluated by considering the modifications of methods 2 and 3 above.
- **Method 5** - Original strategy with "SDS method": In this method, we use the original European C-ITS pseudonym-changing strategy together with our SDS method.

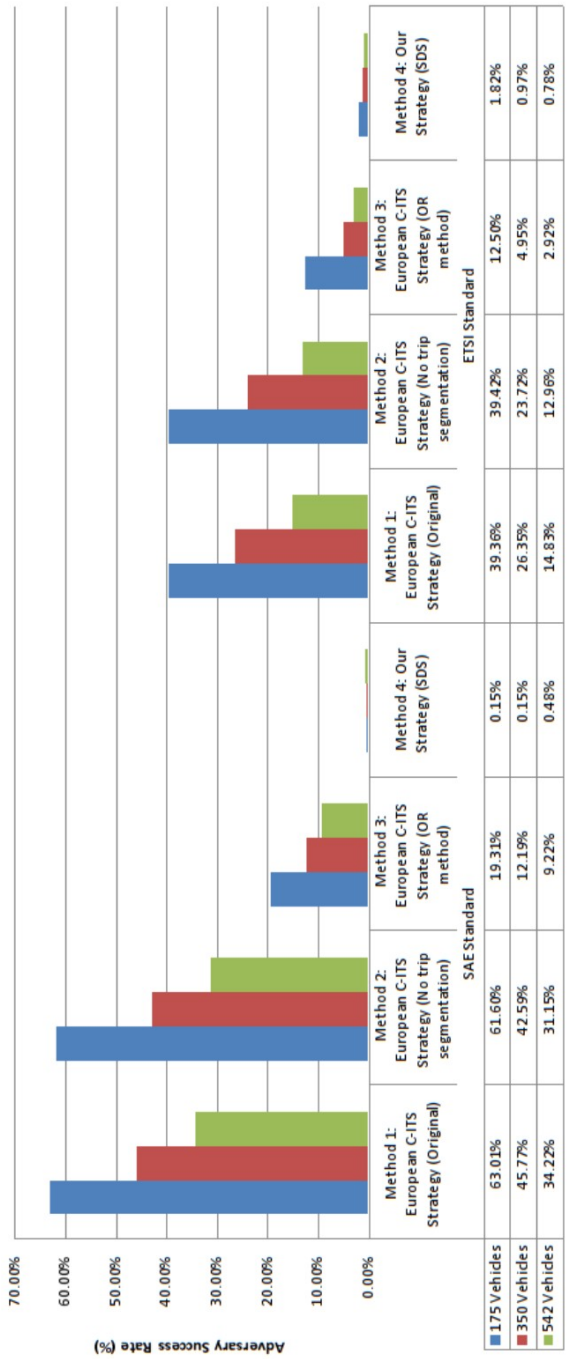


**Figure 4.8.** Adversary success rate in matching different pseudonyms which belong to the same vehicle with different traffic densities in the experiment, according to different standards used for reporting width and length values, and also the applied pseudonym-changing strategy.

The performance results of these methods for SAE and ETSI standards are presented in Figures 4.3-4.5. In Fig. 4.3, adversary success rate for various vehicle densities is presented for the first four methods. It is seen that as the vehicle density increases, the adversary success rate is decreasing for each method and standard. It is much more easier for the adversary to link the pseudonyms when the vehicle density in an area is low. As the vehicle density increases, more number of pseudonyms will be received by the adversary making the linking more difficult.

When the number of vehicles increased from 175 to 542, adversary success rate is reduced from 63% to 34.22% for the original European C-ITS pseudonym-changing strategy (Method 1) with SAE standard implementation. This means that the threat size previously thought was blown out of proportion when it came to matching the width and length data from CAM messages, as we did not get the same results of [11].

A similar decrease observed with ETSI standard from 39.36% to 14.83%. When the Method 2 is applied, the adversary success rate decreased slightly (approximately 3%) for all methods (Methods 1-4) with SAE and ETSI standards. The reason of this improvement is related to the use of segmentation, in other words, removing the different segments (the 10-20 km and 25-35 km conditions in travel distance) for the pseudonym change approaches (Methods 1-4) allowed triggering pseudonym-changing slightly more often than before, but the trigger conditions were still restricted by the "AND" condition relating to the traveled distance and time (that is, 800 meters of additional travel distance AND 2 to 6 minutes of travel time). When the "AND" condition is replaced with the "OR" condition (Method 3) it is seen that there is great decrease in the adversary success rate for all vehicle densities and at both standards. The main reason for this improvement is (as mentioned earlier) is that vehicles change their pseudonyms either at time-based or distance-based rules, whichever comes first, and not both of them together as was done before. This ensures more frequent pseudonym changes and makes the adversary's task more difficult.

When the "OR method" (Method 3) is compared with the original method (Method 1), it is seen that adversary success rate has decreased from 63.01% to 19.31% in low density (175 vehicles), from 45.77% to 12.19% in medium density (350 vehicles) and from

34.22% to 9.22% in high density (542 vehicles) with the SAE standard. Similar enhancements are seen in ETSI standard, too. It is seen that adversary success rate decreases from 39.36% to 12.50% in low density (175 vehicles), from 26.35% to 4.95% in medium density (350 vehicles) and from 14.83% to 2.92% in high density (542 vehicles) with the ETSI standard.

It is seen that, in general, the adversary has more success rate using the SAE standard than the ETSI standard, because in the SAE standard vehicles are reporting more unique lengths and widths compared to the ETSI standard (centimeters instead of decimeters). The use of more unique width and length values allows the adversary to link the changed pseudonyms more easily. For example, the adversary will have a harder time in matching width values like 1.9 than 1.93 meters, because the vehicles in the vicinity will be reporting overlapping values when using decimeters unit instead of centimeters.

**Table 4.4.** Number of pseudonym changes when using different pseudonym-changing strategies on high density (542 Vehicles).

| Method Name | Number of Pseudonym Changes per Vehicle (mutually exclusive) | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Method 1: European C-ITS strategy (original) | 133 | 409 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Method 2: European C-ITS strategy (no trip segmentation) | 133 | 258 | 127 | 22 | 1 | 0 | 0 | 0 | 0 | 0 |
| Method 3: European C-ITS strategy (or method) | 1 | 6 | 46 | 129 | 161 | 113 | 56 | 19 | 8 | 1 |
| Method 4: Our strategy (SDS) | 1 | 6 | 47 | 126 | 161 | 112 | 59 | 18 | 8 | 1 |
| Method 5: Original European C-ITS strategy + SDS | 133 | 409 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

The use of relaxations previously mentioned (corresponding to Method 2 and Method 3)

has an impact on the privacy protection. These relaxation approaches cause the vehicles to change their pseudonyms more frequently, thus, they are introducing a challenge for the adversary. The number of pseudonyms used in these methods are presented in Table 4.3. In the original European C-ITS pseudonym-changing method (Method 1), all vehicle use at most 2 pseudonyms during their travel (133 of 542 vehicles use only one pseudonym, the rest use 2 pseudonyms). The improvements in Method 2 and Method 3 against the success of the adversary are totally related to use of more pseudonyms in these two methods. It can be seen that in Method 3, for example, 99% of vehicles switch their pseudonym three times or more, unlike the original European C-ITS strategy (Method 1), where 75% of vehicles just switch their pseudonym two times only, giving the adversary a better chance to track them.

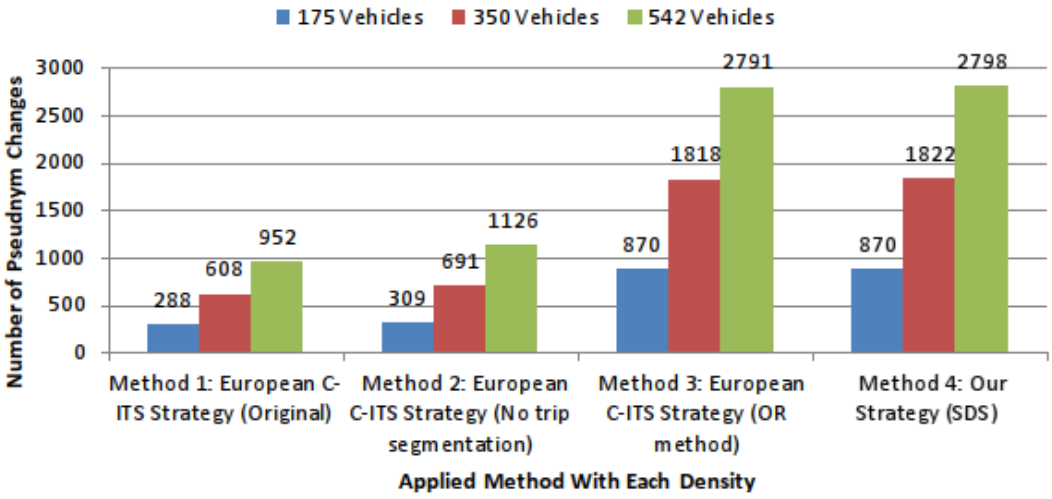


**Figure 4.9.** Number of pseudonym changes for each applied pseudonym-changing strategy, per different traffic densities.

The results show how frequent-pseudonym-change negatively impacts the success rate of the adversary. However, despite the use of so many different pseudonyms, the adversary is still able to link the pseudonyms. Our novel method, SDS, aims to solve this issue. Method 4, which is SDS together with the previously-discussed relaxations, successfully reduce the adversary's success rate to nearly 0% (Fig. 4.3). The significant additional gain in privacy is due to the obfuscation in SDS; each vehicle is reporting random width and length values (closer to the real values) in each CAM message rather than reporting its real width and length. As a result, the adversary is no longer capable

of distinguishing the vehicles from each other based on these data, and consequently, the adversary can no longer benefit width and length information for linking.

Number of total pseudonym changes in each method (Method 1 - 4) is given in Fig. 4.4. Method 3 and 4 use more pseudonyms while reducing the adversary success rate. It is seen that there is trade-off between the privacy (reducing the adversary success rate) and the load introduced due to use of more pseudonyms. Using more pseudonyms at each vehicle requires them have a large set of pseudonyms for this purpose. There could be another discussion on the number of pseudonyms allocated to each vehicle and how to use them for security and privacy purposes.

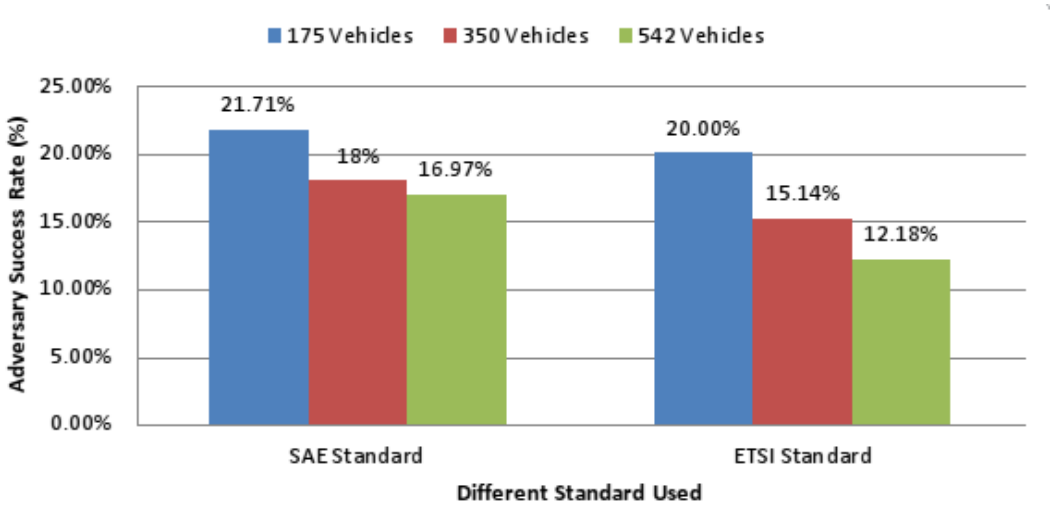


**Figure 4.10.** Adversary success rate when applying the original European C-ITS method with our SDS method, for both ETSI and SAE standards.

Protecting the privacy with the use of more pseudonyms may introduce a challenge to the nodes which have limited capabilities e.g., memory, processing. The SDS Method can be used to address this issue as well. In Method 5, SDS method has been used together with the the original European C-ITS pseudonym strategy, meaning that the vehicles use original European C-ITS pseudonym strategy while using the SDS method in reporting the width and length values. This approach causes the use of minimum number of pseudonyms while preserving privacy at the same time.

The results of Method 5 (for SAE and ETSI) are presented in Figure 4.5. It is found that

the adversary's success rate dropped from 34.22% to 21.71% (SAE) and from 14.83% to 9.62% (ETSI), as can be seen in Fig. 4.5.

This means that our SDS method can be used with any pseudonym-changing strategy (European C-ITS method, or others) to obtain further privacy protection, regardless of that method's details. This could be useful in scenarios where the number of pseudonym changes need to be controlled and limited, but there is also a desire for providing additional privacy protection without further changes.

# 5. CONCLUSION

The earlier research on the threat of location privacy due to width and length data in CAM messages was exaggerated. Using the real-world SAE and ETSI standards, the adversary's success rate did not exceed 63.01% and 39.36% respectively when using the original European C-ITS pseudonym-changing strategy even in the worst case of low density. With high density, these numbers drop to 34.22% and 14.83% respectively, which is too low of a motivation for possible adversaries to carry their attacks.

Still, the adversary model we used (which was also used by other researchers) is simple and does not employ the usage of traffic lanes and directions information. Hence, a more powerful, complex and vicinity-aware adversary model may achieve a higher success rate.

Using our proposed SDS method and enhancements, however, we eliminated the threat factor of any possible adversary model. The proposed method ac hived width and length data obfuscation, which greatly reduced any adversary's ability to use these data to establish linkability between different pseudonyms of the same vehicle.

In general, we have observed that the ETSI standard for reporting width and length data resulted in much lower adversary success rate in different scenarios, sometimes even by around 70%. We recommend to modify the SAE standard for V2X communications regarding width and length data reporting towards the direction of ETSI standard. This way, possible attacks on privacy can be reduced without implementing any further changes or defense techniques.

# REFERENCES


[1] Benarous, L., Kadri, B., & Boudjit, S. (2020). Alloyed Pseudonym Change Strategy for Location Privacy in VANETs. 2020 IEEE 17th Annual Consumer Communications & Networking Conference (CCNC), 1–6. https://doi.org/10.1109/CCNC46108.2020.9045740

[2] Boualouache, A., & Moussaoui, S. (2017). Urban pseudonym-changing strategy for location privacy in VANETs (UPCS). International Journal of Ad Hoc and Ubiquitous Computing, 24, 49. https://doi.org/10.1504/IJAHUC.2017.080914

[3] Boualouache, A., Senouci, S.-M., & Moussaoui, S. (2016). VLPZ: The Vehicular Location Privacy Zone. Procedia Computer Science, 83, 369–376. https://doi.org/10.1016/j.procs.2016.04.198

[4] Boualouache, A., Senouci, S.-M., & Moussaoui, S. (2018). A Survey on pseudonym-changing Strategies for Vehicular Ad-Hoc Networks. IEEE Commun. Surv. Tutorials, 20(1), 770–790. https://doi.org/10.1109/COMST.2017.2771522

[5] Bouchelaghem, S., & Omar, M. (2020). Secure and efficient pseudonymization for privacy-preserving vehicular communications in smart cities. Computers & Electrical Engineering, 82, 106557. https://doi.org/10.1016/j.compeleceng.2020.106557

[6] Buttyán, L., Holczer, T., Weimerskirch, A., & Whyte, W. (2009). SLOW: A Practical pseudonym-changing scheme for location privacy in VANETs. 2009 IEEE Vehicular Networking Conference (VNC), 1–8. https://doi.org/10.1109/VNC.2009.5416380

[7] Car dimensions of all makes with size comparison tools. (n.d.). Retrieved June 25, 2022, from https://www.automobiledimension.com/

[8] Cirne, P., Zúquete, A., & Sargento, S. (2018). TROPHY: Trustworthy VANET routing with group authentication keys. Ad Hoc Networks, 71, 45–67. https://doi.org/10.1016/j.adhoc.2017.12.005

[9] Codecá, L., Frank, R., Faye, S., & Engel, T. (2017). Luxembourg SUMO Traffic (LuST) Scenario: Traffic Demand Evaluation. IEEE Intelligent Transportation Systems Magazine, 9(2), 52–63.

[10] Eckhoff, D., Sommer, C., Gansen, T., German, R., & Dressler, F. (2010). Strong and affordable location privacy in VANETs: Identity diffusion using time-slots and swapping. 2010 IEEE Vehicular Networking Conference, 174–181. https://doi.org/10.1109/VNC.2010.5698239

[11] Escher, S., Sontowski, M., Berling, K., Kopsell, S., & Strufe, T. (2021). How well can your car be tracked: Analysis of the European C-ITS pseudonym scheme. 2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring), 1–6. https://doi.org/10.1109/VTC2021-Spring51267.2021.9449078

[12] Förster, D., Kargl, F., & Löhr, H. (2015). A framework for evaluating pseudonym strategies in vehicular ad-hoc networks. Proceedings of the 8th ACM Conference on Security & Privacy in Wireless and Mobile Networks, 1–6. https://doi.org/10.1145/2766498.2766520

[13] Intelligent Transport Systems (ITS); ITS-G5 Access layer specification for Intelligent Transport Systems operating in the 5 GHz frequency band. ETSI EN 302 663—V1.3.1. (2020).

[14] Intelligent Transport Systems (ITS); Security; ITS communications security architecture and security management; Release 2. ETSI TS 102 940. (2021).

[15] Intelligent Transport Systems (ITS); Security; Pre-standardization study on pseudonym change management. ETSI TR 103 415. (2018).

[16] Intelligent Transport Systems (ITS); Security; Trust and Privacy Management; Release 2. ETSI TS 102 941 V2.1.1. (2021).

[17] Intelligent Transport Systems (ITS); Users and applications requirements; Part 2: Applications and facilities layer common data dictionary. ETSI TS 102 894-2 V1.3.1. (2018).

[18] Kalaiarasy, C., Sreenath, N., & Amuthan, A. (2019). Location Privacy

Preservation in VANET using Mix Zones – A survey. 2019 International Conference on Computer Communication and Informatics (ICCCI), 1–5. https://doi.org/10.1109/ICCCI.2019.8822028

[19] Kerkacha, N., Hadj-Said, N., Chaib, N., Adnane, A., & Ali-Pacha, A. (2021). Impact of Silent Periods on Pseudonym Schemes. 2021 18th International Multi-Conference on Systems, Signals & Devices (SSD), 79–85. https://doi.org/10.1109/SSD52085.2021.9429464

[20] Khan, S., Sharma, I., Aslam, M., Khan, M. Z., & Khan, S. (2021). Security Challenges of Location Privacy in VANETs and State-of-the-Art Solutions: A Survey. Future Internet, 13(4). https://doi.org/10.3390/fi13040096

[21] Li, X., Zhang, H., Ren, Y., Ma, S., Luo, B., Weng, J., Ma, J., & Huang, X. (2020). PAPU: Pseudonym Swap With Provable Unlinkability Based on Differential Privacy in VANETs. IEEE Internet of Things Journal, 7(12), 11789–11802. https://doi.org/10.1109/JIOT.2020.3001381

[22] Lu, R., Lin, X., Luan, T. H., Liang, X., & Shen, X. (2012). pseudonym-changing at Social Spots: An Effective Strategy for Location Privacy in VANETs. IEEE Transactions on Vehicular Technology, 61(1), 86–96. https://doi.org/10.1109/TVT.2011.2162864

[23] Position Paper regarding personal data protection aspects in C-ITS, CAR 2 CAR Communication Consortium. (2017). CAR 2 CAR Communication Consortium.

[24] Sampigethaya, K., Huang, L., Li, M., Poovendran, R., Matsuura, K., & Sezaki, K. (2005). CARAVAN: Providing location privacy for VANET. Washington Univ Seattle Dept of Electrical Engineering.

[25] Security Policy & Governance Framework for Deployment and Operation of European Cooperative Intelligent Transport Systems (C-ITS), Release 1. (2017). https://transport.ec.europa.eu/system/files/2018-06/c-its_security_policy_release_1.pdf

[26] Ullah, I., Shah, M. A., Khan, A., & Jeon, G. (2021). Privacy-preserving

multilevel obfuscation scheme for vehicular network. Transactions on Emerging Telecommunications Technologies, 32(2), e4204. https://doi.org/10.1002/ett.4204

[27] Ullah, I., Wahid, A., Shah, M. A., & Waheed, A. (2017). VBPC: Velocity based pseudonym-changing strategy to protect location privacy of vehicles in VANET. 2017 International Conference on Communication Technologies (ComTech), 132–137. https://doi.org/10.1109/COMTECH.2017.8065762

[28] V2X Communications Message Set Dictionary. J2735_202007. (2020). SAE International. https://www.sae.org/standards/content/j2735_202007/

[29] Vehicular Networking and Intelligent Transportation Systems Research Lab (VeNIT Lab). (n.d.). Retrieved August 15, 2022, from https://venit.org/

[30] VTRIx, A Digital Twin Platform for Connected Cars/V2X Communications. (n.d.). Retrieved August 15, 2022, from https://bigtri.net/product/DigitalTwin/

[31] Wiedersheim, B., Ma, Z., Kargl, F., & Papadimitratos, P. (2010). Privacy in inter-vehicular networks: Why simple pseudonym change is not enough. 2010 Seventh International Conference on Wireless On-Demand Network Systems and Services (WONS), 176–183. https://doi.org/0.1109/WONS.2010.5437115

[32] Zhou, Y., & Zhang, D. (2019). Double Mix-Zone for Location Privacy in VANET. Proceedings of the 2019 7th International Conference on Information Technology: IoT and Smart City, 322–327. https://doi.org/10.1145/3377170.3377250