Cooperative Location Privacy in Vehicular Networks: Why Simple Mix-zones are not Enough

Mohammad Khodaei, *Member, IEEE,* and Panos Papadimitratos, *Fellow, IEEE*

***Abstract*---Vehicular communications disclose rich information about the vehicles and their whereabouts. Pseudonymous authentication secures communication while enhancing user privacy. To enhance location privacy, cryptographic mix-zones were proposed to facilitate vehicles covertly transition to new ephemeral credentials. The resilience to (*syntactic* and *semantic*) pseudonym linking (attacks) highly depends on the geometry of the mix-zones, mobility patterns, vehicle density, and arrival rates. We introduce a tracking algorithm for linking pseudonyms before and after a cryptographically protected mix-zone. Our experimental results show that an eavesdropper, leveraging standardized vehicular communication messages and road layout, could successfully link ≈73% of pseudonyms during non-rush hours and ≈62% of pseudonyms during rush hours after vehicles change their pseudonyms in a mix-zone. To mitigate such inference attacks, we present a novel *cooperative mix-zone* scheme that enhances user privacy regardless of the vehicle mobility patterns, vehicle density, and arrival rate to the mix-zone. A subset of vehicles, termed *relaying vehicles*, are selected to be responsible for emulating non-existing vehicles. Such vehicles cooperatively disseminate decoy traffic without affecting safety-critical operations: with 50% of vehicles as relaying vehicles, the probability of linking pseudonyms (for the entire interval) drops from ≈68% to ≈18%. On average, this imposes 28 ms extra computation overhead, per second, on the Road-Side Units (RSUs) and 4.67 ms extra computation overhead, per second, on the (relaying) vehicle side; it also introduces 1.46 KB/sec extra communication overhead by (relaying) vehicles and 45 KB/sec by RSUs for the dissemination of decoy traffic. Thus, user privacy is enhanced at the cost of low computation and communication overhead.**

***Index Terms*---Privacy, Anonymity, Pseudonymity, Location Privacy, Mix Networks, Vehicular Communication, VANETs.**

I. Introduction

Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communications seek to enhance transportation safety and efficiency. It has been well-understood that Vehicular Communication (VC) systems are vulnerable to attacks and that the privacy of their users is at stake. As a result, security and privacy solutions have been developed by standardization bodies (IEEE 1609.2 WG [1] and ETSI [2], [3]), harmonization efforts (Car2Car Communication Consortium (C2C-CC) [4]), and projects (SeVeCom [5], PRESERVE [6], and CAMP [7]). In VC systems, vehicles disseminate Cooperative Awareness Messages (CAMs) and Decentralized Environmental Notification Messages (DENMs) periodically at a high rate. To secure VC systems, a consensus towards using Public Key Cryptography (PKC) to protect V2V/V2I (V2X) communication is reached: a set of Certification Authorities (CAs) constitutes the Vehicular Public-Key Infrastructure (VPKI), e.g., [7], [8], providing multiple anonymous credentials, termed *pseudonyms*, to legitimate vehicles. Vehicles switch from one pseudonym to a non-previously used one towards unlinkability of digitally signed messages, and improved sender privacy for V2V/V2I messages. Pseudonymity is conditional, in the sense that the corresponding long-term vehicle identity (Long Term Certificate (LTC)) can be retrieved by the VPKI entities if needed, e.g., for eviction of a faulty, misbehaving vehicle.

Due to the openness of wireless communication and dissemination of basic safety messages in plaintext (as confidentiality is not needed in VC systems [1], [9]), an external entity could eavesdrop communications, towards inferring vehicle-sensitive information. Although pseudonymous authentication is a promising approach to protect user privacy, an adversary, eavesdropping all traffic in an area, could link successive pseudonymously authenticated messages. An adversary might observe an isolated pseudonym change, and associate the old and the new pseudonymous identifier through *syntactic linking*, e.g., [10], [11], [12], [13]. Alternatively, an adversary could leverage the physical constraints of the road layout [14], together with data in message payloads, e.g., location, velocity, time, acceleration, the length and width[1] of a victim's vehicle, to predict its trajectory towards linking messages *semantically*, e.g., [11], [14], [18], [19]. Such information could be unique, or one of few (locally rare), and thus, it could be easily linked. While appropriate pseudonym provisioning policies alleviate syntactic linking, by issuing time-aligned pseudonyms [8], [20], [21], compromising user privacy by conducting semantic linking attacks is still feasible[2].

Different schemes were proposed, leveraging pseudonymous authentication primitives to mitigate inference, by an adversary, e.g., silent periods [26], [27], [28], [29], silent cascades [30], SLOW [11], and random encryption periods [31]. The common denominator among all *silent period*-based schemes is that vehicles refrain from transmitting CAMs in certain intervals. This would result in diminished situational awareness, e.g., *collision avoidance* [32], thus increased probability of an accident, notably near intersections with congested traffic conditions [33]; thus, the practicality of such schemes is questionable.

The authors are with the Networked Systems Security group at KTH Royal Institute of Technology, Stockholm, Sweden. E-mail: {khodaei, papadim}@kth.se.

[1]Length and width of vehicles are specified with a precision of 10 centimeters [15], [16], [17].

[2]Connecting such anonymous location profiles to real identities of vehicle owners is the final step, e.g., tracing their commutes and identifying home/work locations [22], [23], [24], the information obtained from Vehicular Social Networks (VSNs) [25], or full de-anonymization of vehicles by *honest-but-curious* VPKI entities [8].

Alternatively, vehicles could change their pseudonyms when approaching designated areas, termed mix-zones [34], [35], [36]. The Cryptographic Mix-Zone (CMIX) was initially proposed [37] in the VC systems to establish a cryptographically protected region at appropriate times and places, e.g., at road intersections. All legitimate vehicles within the mix-zone receive a symmetric session key from a Road-Side Unit (RSU), responsible for the initiation of the pseudonym transition process and the symmetric key updates [37]. Vehicles encrypt CAMs and opt in to change their pseudonyms while crossing these regions, towards pseudonym unlinkability (by an external eavesdropper). The achieved privacy protection level for such statically constructed mix-zones highly depends on the geometry of mix-zones, mobility patterns, and vehicle arrival rates. For example, based on the mix-zone geometries [38], or the traffic mobility pattern [39] and vehicle speed [14], [37], [40], [41], an adversary can link successive pseudonyms of a given vehicle by observing the mix-zone entry and exit points. Such schemes are mostly effective when vehicle density and arrival/exit rates of vehicles in/from the mix-zones are uniformly distributed. Moreover, a fraction of non-cooperative vehicles within the mix-zone could affect anonymity by simply not changing their pseudonyms in a mix-zone; this yields a smaller anonymity set size for the mix-zone, compared to the one when all vehicles switch their pseudonyms.

Recently, a *chaff-based* CMIX scheme [42] was proposed: RSUs pre-generate and broadcast chaff CAMs to *relaying* vehicles, responsible to periodically disseminate to emulate a non-existing (chaff) vehicle. This imposes significant computation overhead on RSUs and communication overhead: each RSU needs to sign all chaff CAMs and distribute them to each relaying vehicle. Moreover, the RSU needs to precisely predict trajectories of all vehicles in the neighborhood, to properly construct chaff CAMs for all chaff vehicles; the challenge lies in that traffic conditions are volatile, thus changes vehicles trajectories would invalidate (make relatively easily distinguishable) chaff CAMs. All these issues become clear in our performance evaluation in Sec. VI-F.

Contributions: In this paper, we fundamentally re-design the well-known approach (in different domains, e.g., [43]) of introducing decoy traffic and re-design the VC-specific scheme [42], proposing a fully decentralized system (Sec. IV). We show how to enhance user privacy, notably in low-density areas and non-rush hour periods, and how to mitigate syntactic and semantic linking attacks without affecting the operation of safety applications (Sec. IV). Our scheme (i) enhances user privacy regardless of the vehicle mobility patterns, density, and arrival rate (to the mix-zone(s)); at the same time, (ii) it balances user privacy protection and (communication and computation) overhead based on vehicle density and mobility pattern. Furthermore, our scheme incurs low (computation and communication) overhead and prevents abuse of the mechanism towards diminishing the performance of the system or harming user privacy (Sec. VI). We also (iii) introduce a tracking algorithm towards linking vehicles before and after a cryptographically protected mix-zone (Sec. VI-A). We leverage information in CAMs and the road layout towards linking pseudonyms syntactically and semantically, thus compromising user privacy. To mitigate such inference attacks, we introduce cooperative dissemination of decoy traffic: vehicles and RSUs emulate a non-existing vehicle by broadcasting decoy traffic in order to generate sufficiently many vehicles, thus diminishing the probability of linking two successive pseudonyms by an eavesdropper (Sec. VI). Our scheme efficiently and effectively enhances user privacy and it maintains strong user privacy protection for vehicles upon pseudonym change in a mix-zone in the presence of *honest-but-curious* system entities (Sec. V).

In the rest of the paper, we survey the state-of-the-art research efforts (Sec. II) and describe the system model, adversarial model, and the requirements (Sec. III). We present our novel CMIX scheme with decoy traffic (Sec. IV), followed by a qualitative analysis of security and privacy (Sec. V). We evaluate the performance of our scheme (Sec. VI) before conclusion and future work (Sec. VII).

II. Related Work

Due to the openness of wireless transmissions and dissemination of basic safety messages in plaintext (as confidentiality is not needed in VC systems [1], [2], [9], [17], [44], [45]), an external entity can arbitrarily eavesdrop VC systems [46], [47], [48]. With advances in broadcast technology to extend the transmission range of On-Board Units (OBUs) [49], Vehicular Ad-hoc Network (VANET) messages become increasingly accessible for an attacker. This information allows semantic linking attacks that rely on location and heading information of continuously broadcast CAMs [14]. Prior works, e.g., [37], [42], assume that the system entities that are fully trustworthy, i.e., RSUs and VPKI entities, could link successive pseudonyms belonging to a given vehicle. However, recent revelations of mass surveillance, e.g., [50], [51], show that assuming service providers are fully-trustworthy is no longer a viable approach. Thus, in [35], [37], [52], [53], the VPKI entities can easily link pseudonyms issued for the vehicles, thus tracking them for the entire trip duration. Unlike the chaff-based CMIX scheme [42] that requires vehicles provide their intended trajectory path to the RSUs, our scheme does not provide additional information and maintains strong user privacy protection upon pseudonym change in the presence of *honest-but-curious* system entities.

There are different solutions for location privacy: *K-anonymity* [54] and dummy-based privacy protection schemes, e.g., [55], ensure that a target node is not distinguishable from at least *K-1* nodes within an *anonymity set* with respect to the information each node disseminates. However, safety applications require precise information to operate correctly, e.g., *intersection collision warning* [56]. Alternatively, one can rely on group signature schemes, e.g., [57], [58], [59], [60], [61], to enhance user privacy. However, the performance of safety-related applications could be degraded. For example, leveraging such anonymous authentication schemes by the majority of vehicles results in a 30% increase in cryptographic processing overhead [61]. Moreover, with all vehicle-sensitive information in CAMs and DENMs, e.g., location, velocity, and acceleration, a targeted node could be unique, or one of few, and thus, successive messages could be linked sequentially by an external observer.

Different pseudonym transition strategies, to prevent an attacker from inferring such information, have been proposed. To evade correlation attacks, each vehicle could turn its wireless transmitter off for a randomly chosen interval and change pseudonym within that silent period [26], [62], [27]. Based on the quality of service required for each application, this interval of being silent or being active can be dynamically adjusted [30]. Even though such schemes could improve user privacy, they impose a performance penalty on safety applications [33], thus jeopardizing human safety. To mitigate such a problem, vehicles could become silent and change their pseudonyms when their speed drops below 30 km/h since the risk of a fatal accident at a slow speed is expected to be low [11]. Alternatively, vehicles could change their pseudonyms when refilling fuel at a gas station [63], or each vehicle changes its pseudonyms randomly, e.g., every 5 minutes or every 1-3 KM [12]. However, an adversary can still conduct syntactic linking attacks due to a lack of synchronization among vehicles [13], or track vehicles across pseudonym changes by predicting their trajectories [29]. In general, any individually-determined or user-defined pseudonym changing strategy could act as user fingerprints, thus enabling an adversary to track users, i.e., syntactic linking attack.

Another line of study proposes pseudonym transitions strictly within CMIX [37], which does not impair transportation safety applications. A cryptographic mix-zone was initially proposed [37] in the VC systems to establish a cryptographically protected region at appropriate times and places, e.g., at intersections. When crossing these regions, vehicles change their pseudonyms privately while their communication is encrypted, which prevents syntactic and semantic linking attacks. However, the achieved privacy protection highly depends on the number of vehicles participating in the mix-zone, i.e., user privacy is degraded under low traffic density, e.g., in a highway scenario [64]. Moreover, an attacker could compromise unlinkability within a mix-zone based on the traffic mobility pattern and vehicle speed [41]. To counter this, vehicles could randomly switch lanes and speed prior to entering and/or crossing the mix-zones to confuse an adversary [31], [65]. However, such schemes would not be practical as they could seriously jeopardize human safety. Unlike such schemes, we provide privacy protection without affecting the operation of safety applications and regardless of variations in road layout, vehicle density, and mobility patterns.

Another alternative approach is to participate into a dynamic mix-zone, e.g., [31]: each OBU is provided with a global symmetric key, using it to initiate a pseudonym change process. However, an internal attacker could terminate the encryption period on behalf of any vehicle; this impairs the functionality and operation of the scheme, thus eliminating user privacy protection. A dynamic cooperative location privacy protection scheme was proposed [66]: time-aligned pseudonyms are issued for all vehicles to facilitate synchronous pseudonym changes. Upon reaching a pseudonym transition process, a dynamic mix-zone formation is initiated by a vehicle and all CAMs within each mix-zone is encrypted using a distinct symmetric session key [66]. However, in a low traffic density area where there are very few vehicles to cooperatively change pseudonyms, vehicles could be semantically linkable. Unlike such schemes, our system ensures that user privacy is strongly protected even in situations with inherently low traffic density, e.g., suburban areas, and during low traffic periods.

In order to mitigate inferences by a compromised RSU, an asymmetric group key agreement protocol by leveraging identity-based cryptography (not compatible with the standardization bodies, i.e., IEEE 1609.2 WG [1] and ETSI [2]) was proposed [67]: vehicles cooperatively conduct a group key agreement protocol to derive asymmetric keys. However, it is not clear how to determine the group size and handle dynamic changes of each group to perform group key agreement. Moreover, the pseudonyms (and their corresponding private keys) are generated by a Trusted Authority, i.e., a fully-trusted entity, and the keys are pushed to the vehicles; this raises concerns in terms of accountability and has not been adopted by standardization [1]. Furthermore, the 'Trusted Authority' could trivially link all pseudonyms belonging to a vehicle, and thus the pseudonymously authenticated messages towards tracking it for the entire duration of its presence in the system. Moreover, the scheme lacks an extensive performance evaluation, notably in terms of communication and computation overhead in highly congested traffic conditions.

MobiMix [38], [68] shows that an adversary could infer user-sensitive information based on the vehicle population in a mix-zone, the statistical behavior of the population, and the geometry of a mix-zone. To mitigate such inferences, it is proposed to dynamically adjust the geometry of a mix-zone based on multiple factors, e.g., the statistical behavior and the movement patterns of the users. But, an adversary could still perform semantic linking attacks when the traffic density is sparse [69]. Swing & Swap [70], [71] and MixGroup [53] propose to construct a region in which vehicles exchange their pseudonyms (and the corresponding private keys). But, such schemes do not achieve liability attribution and non-repudiation, which are basic requirements for a secure VC system [1], [5], [9], [72].

III. Model and Objectives

A. System Model and Assumptions

Fig. 1 shows an overview of secure and privacy-preserving VC systems. We assume a VPKI, shown on top of Fig. 1, with distinct entities and roles that registers vehicles in a domain [73] and issues pseudonyms, e.g., [7], [8]. The Root CA (RCA), the highest-level authority, certifies other lower-level authorities; the Long Term CA (LTCA) provides registered vehicles (and RSUs) with a Long Term Certificate (LTC), used to authorize the acquisition of pseudonyms from a Pseudonym CA (PCA) [8]. To facilitate the overall intra-domain and multi-domain operations, a vehicle first finds such information from a Lightweight Directory Access Protocol (LDAP) [74] server. This is carried out without disclosing the real identity of the vehicle. Any CMIX-based scheme requires vehicles change their pseudonyms whenever crossing a mix-zone. As a result, vehicles need several pseudonyms with overlapping lifetimes, compatible with the proposals of standardization bodies, i.e., IEEE 1609.2 WG [1] and ETSI [2]. This allows pseudonym

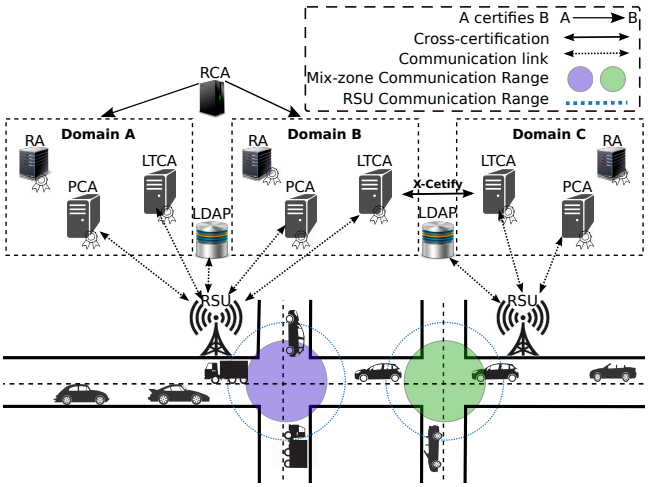


Fig. 1. An overview of secure and privacy-protecting V2X communication.

changes, without pseudonym reuse, to be straightforward, i.e., without extensive prior knowledge on CMIX placement, trip details, pseudonym lifetimes, etc.

Each vehicle triggers pseudonym acquisition process based on various factors [21]. Our scheme requires sparse connectivity to the VPKI, allowing an OBU to be *preloaded* with numerous pseudonyms proactively, covering a longer period, e.g., a week or a month (should the connectivity be relatively scarce). We assume that a state-of-the-art VPKIs, e.g., [7], [8], [75], can provide pseudonyms in a timely manner. Moreover, we assume the VPKI pseudonym lifetime policy is the same for all registered vehicles, so that timing information does not harm user privacy [8], [20].

OBUs and RSUs are equipped with Hardware Security Modules (HSMs). Private keys stored in an HSM cannot be extracted and only one pseudonym can be active at any time. In case of any deviation, pseudonymously authenticated messages can be used by a Resolution Authority (RA) to retrieve the long-term identity of the vehicle [8]. The misbehaving entities should be evicted and revocation information be distributed [76], [77]. The certificates of higher-level authorities are installed on the OBUs and their clocks are (loosely) synchronized with the VPKI servers through Global Navigation Satellite System (GNSS) or other means, e.g., Network Time Protocol (NTP) servers over the Internet.

We assume that each vehicle and RSU have their own location information; RSUs could communicate with VPKI entities and they are aware of the road layout (within their communication ranges). We further assume that appropriate countermeasures are in place to prevent location spoofing, e.g., [78], towards enabling secure neighborhood discovery [79], and facilitating physical position verification [80]. Upon a pseudonym change inside a mix-zone, vehicles change their Media Access Control (MAC) and IP addresses [81] to prevent their old and new pseudonyms from being (trivially) linked based on these interfaces [82], [83]. The impact of pseudonym change on the quality of services, e.g., safety applications [84], or protocols, e.g., geographic routing [85], [86], is orthogonal to our investigation.

The choice of RSUs to establish CMIXs depends on different factors, e.g., desired level of privacy, traffic conditions, physical constraints of the road layout, and efficiency; for example, the more mix-zones are constructed, the higher the frequency of changing pseudonyms becomes. This would result in higher number of unlinkable segments for any journey, thus, enhance user privacy. The frequent change of mix-zone location makes it even harder for an adversary to eavesdrop the communication: an adversary would need to deploy eavesdropping facilities near most, if not all, the RSUs to improve her chance, which seems to be practically infeasible. However, full deployment of mix-zones over all RSUs, i.e., mandating frequent pseudonym changes, could affect the operation of safety application and impose communication overhead. The VPKI system chooses a fraction of RSUs to establish a cryptographically protected mix-zone; in Fig. 1, two mix-zones are constructed, each established by an RSU, to facilitate private pseudonym changes. In our performance evaluation in Sec. VI, the mix-zones locations are fixed and known to an adversary, i.e., the best case for an adversary. Leveraging an optimal placement of mix-zones, e.g., [36], [87], in order to balance the achieved level of privacy and cost, is orthogonal to this investigation.

B. Adversarial Model

We consider the general adversary model in [5], [9] for secure and privacy-preserving VC systems and more specifically the adversarial model assumptions of CMIX schemes [35], [37], [42], [52], [53] that consider external eavesdroppers, possibly with broad or global coverage range. Along these lines, we assume that RSUs and participating users/vehicles are honest (i.e., trustworthy entities). We consider external adversaries with wireless receivers placed near each mix-zone, to eavesdrop VC systems to infer user-sensitive information towards harming user privacy. They passively eavesdrop communication of vehicles entering and exiting the mix-zone, covering all entry and exit points of the mix-zones, towards linking pseudonyms before and after a mix-zone. This is based on information derived from CAMs, e.g., timing, velocity, and location. We do not constrain the choice and design of the inference algorithm, i.e., a tracking algorithm to link two pseudonyms of a vehicle, prior to and after pseudonym change in a mix-zone. Rather, in order to achieve tangible results, we devise a tracking algorithm (see Sec. VI-A), orthogonal to the defense mechanism.

In addition, we explore the consequences of strengthening the adversarial model in Sec. VI-F. In particular, we consider (i) RSUs and VPKI entities that are *honest-but-curious*, i.e., entities complying with security protocols and policies, but motivated to profile users by collecting or inferring user sensitive information based on the execution of the protocols. Moreover, (ii) the collaboration (collusion) of honest-but-curious entities that share information individually inferred by each (see Sec. V for detailed security analysis). Finally, we consider (iii) a set of non-cooperative actions by registered vehicles that can affect the operation (or level) of protection of the scheme (and any CMIX scheme).

Extending the passive eavesdropper model: In this paper, we focus on the effect and improvement of the CMIX approach. The investigation can be extended to the entire network, considering the optimal placement of eavesdroppers,

increasing their coverage, and overall pseudonym usage. The adversarial model can be further strengthened if internal adversaries, including the non-cooperative vehicles joining the mix-zone, report the symmetric keys of the mix-zones and the observed communication to an external adversary (collection point). For example, an RSU could share a transcript of pseudonymously authenticated messages with an honest-but-curious VPKI entity to perform syntactic and semantic linking attacks. However, this adversarial model is beyond the scope of this investigation. Moreover, non-VC mechanisms such as traffic monitoring cameras (with object recognition techniques), Radio Frequency (RF)-based characteristics, e.g., angle-of-arrival [88], [89], physical layer device identification [90], [91], [92], and physical layer localization with additional equipments, e.g., [93], [94], [95], [96], which can localize vehicles based on the physical layer attributes of transmitters or identify decoy traffic from the actual one, are out of scope and warrant a separate investigation. Further, attacks on GNSS are orthogonal to our work. In fact, the consequence of location spoofing would be that mix-zones are not formed properly; however, the effect is more dire for the VC systems to begin with. These extensions of the adversarial model is part of future work.

C. Requirements

Security and privacy requirements for V2X communications have been specified in the literature [9], and additional requirements for VPKI entities in [8]. The security and privacy requirements for the Certificate Revocation List (CRL) distribution are in [76], [77]. Beyond 'conventional' security requirements [9], it is ingrained in CMIX schemes to establish neighbors proximity (physical or communication neighborhood [79], [97], [98], [99], [100]). This can be done in a secure manner, e.g., with RSU/vehicle protocols, or increased protection of the vehicle location information [86]. In the following, we describe the security and privacy, as well as functional and performance, requirements for a privacy-preserving CMIX scheme.

R1. Privacy (anonymity and unlinkability): Vehicles should participate in the VC system *anonymously*, i.e., vehicles should communicate with others without revealing their long-term identifiers and credentials. Anonymity is conditional in the sense that the corresponding long-term identity can be retrieved by the VPKI entities, and accordingly, the long-term credential revoked if vehicles deviate from system policies. In order to achieve *unlinkability*, we need to diminish the inference by an eavesdropper upon pseudonym change, i.e., mitigating syntactic and semantic linking attacks.

R2. Availability: The system should ensure any legitimate vehicle is notified about CMIX parameters, e.g., the location, geometry, and the symmetric key corresponding to an approaching mix-zone, to facilitate their participate in the mix-zone. Moreover, a small fraction of bandwidth should be used for the distribution of mix-zone related material, to not interfere with the safety- and time-critical operations.

R3. Auditability and misbehavior detection: Auditability refers to the ability of a system to audit the processes and operations of the system entities. In case of any deviation, the system should be able to initiate a (resolution) process to identify the misbehaving entity. This essentially allows an RSU to interact with the VPKI system towards detecting misbehavior. Depending on the situation, appropriate actions could be initiated, e.g., de-anonymizing the misbehaving entity, and/or revoking its cryptographic materials and evicting it from further accessing the system. In the context of this work, each RSU monitors the behavior of vehicles when entering and exiting the mix-zone; if a substantial fraction of vehicles exit the mix-zone without changing their pseudonyms, the RSU would increase the percentage of decoy traffic in order to achieve a desired level of privacy protection.

R4. Efficiency and scalability: All mix-zone operations should be efficient and scale with the number of vehicles. The scalability results from fast generation and lightweight dissemination of the credentials, efficient operations, and fault-tolerant design to ensure that the system remains operational in the presence of benign failures or be resilient to resource depletion attacks.

IV. CMIX with Decoy Traffic

A. System Overview

The VPKI system chooses a subset of RSUs, located near intersections where vehicles physically mix [37], to establish a cryptographically protected area and construct a CMIX for private pseudonym changes. RSUs are responsible for the initiation of the pseudonym transition process and maintaining a symmetric key to establish the encrypted region. To mitigate syntactic and semantic linking attacks, we introduce broadcasting decoy traffic at each mix-zone. Such traffic emulates vehicles that do not exist in reality. The RSU at each mix-zone facilitates obtaining *Chaff Pseudonyms (CPs)* in order to generate *chaff CAMs* (or *chaff DENMs*). The purpose is to decrease the probability of linking two pseudonyms of a vehicle prior to and after pseudonym change. In case of sparse traffic (low vehicle density), RSUs could also emulate a chaff vehicle by periodically broadcasting chaff CAMs (signed under the private key of a chaff pseudonym). Our system can be configured so that for each vehicle, multiple *seemingly identical* chaff vehicles could (potentially) appear as if they uniformly exit from different exit points of a mix-zone. As a result, it is hard for an eavesdropper to identify actual traces based on the CAMs attributes, e.g., velocity, acceleration, mix-zone geometry, and time spent in a mix-zone. Each vehicle could request multiple chaff pseudonyms (and the corresponding chaff private keys) from an RSU. For ease of exposition, we assume each vehicle requests one chaff pseudonym in each mix-zone. Extension to multiple chaff pseudonyms and multiple PCAs operating in a domain is straightforward.

Fig. 2 shows three mix-zones: the colored disks indicate the approximate encrypted range of a mix-zone; the blue dotted circles denote the transmission range of RSUs. The coverage range of eavesdroppers denoted by red dotted circles; for mix-zones B and C, the external adversaries eavesdrop all entry and exit points of the RSUs while for mix-zone A, the eavesdropper eavesdrops all entry and exit points of the mix-zone. The RSU coverage range can be either larger or smaller than the local

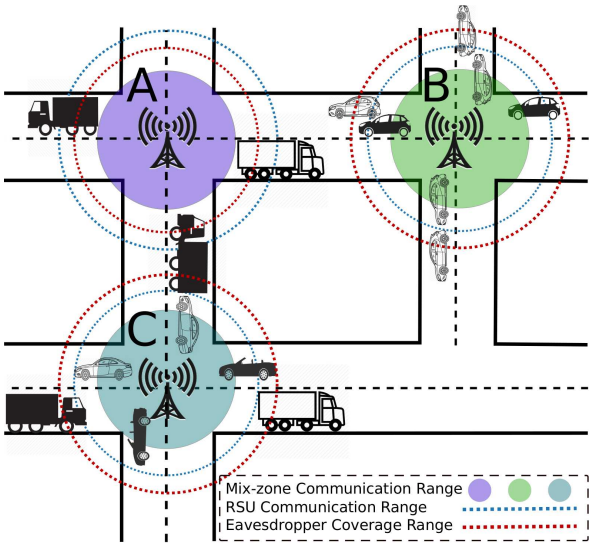


Fig. 2. Mix-zone construction with decoy traffic.

eavesdropper; however, the operation of our scheme does not depend on these ranges. The RSU range needs to always exceed the mix-zone range, simply in order to allow vehicles to execute the CMIX participation protocol, notably obtaining the mix-zone symmetric key. Black vehicles are the real ones while the white ones represent non-exiting vehicles, i.e., the decoy traffic. Once a vehicle enters a mix-zone, it requests to obtain the mix-zone symmetric key. An RSU leverages its knowledge about the road layout and vehicles to determine how many chaff vehicles are required. In the case of sparse traffic density, an RSU generates synthetic CAMs, resembling the traces towards an exit point of the mix-zone. The system can be configured to have RSUs provide and/or emulate one (see mix-zone C in Fig. 2) or multiple (see mix-zone B in Fig. 2) chaff vehicles. In our scheme, each vehicle only provides its length to the RSU; this information is used by an RSU to coordinate with another vehicle in the mix-zone towards disseminating decoy traffic, i.e., generating synthetic CAMs towards resembling a non-existing, *but seemingly identical*, vehicle, exiting from an opposite exit point of the mix-zone.

Each PCA pre-generates a distinct set of chaff public and private keys (chaff pseudonyms) and delivers them to an RSU, responsible for a mix-zone construction. Each vehicle could send a request to the RSU to obtain one chaff pseudonym. The RSU randomly assigns chaff pseudonyms to a subset of vehicles, termed *relaying vehicles*. The VPKI system cannot correlate a vehicle and a chaff pseudonym since the RSU randomly assigns a chaff pseudonym to a requesting vehicle. Note that accountability for chaff CAMs is not paramount as such (chaff) credentials are not valid and they cannot be used for any application. In case of deviation from system protocols, a misbehaving vehicle can still be identified (see Sec. IV-E).

In order to preserve the correct functionality of transportation safety applications, our scheme provides vehicles with information to identify chaff messages. Therefore, each PCA proactively constructs a Cuckoo Filter (CF) [101] by including chaff pseudonyms in a probabilistic data-structure and RSUs distribute these condensed fingerprints of chaff pseudonyms among legitimate vehicles across a region. This facilitates discarding chaff pseudonyms by legitimate vehicles, thus, ensuring the correct operation of safety applications. Similarly to Bloom Filter (BF) [102], [103], CFs provide fast membership tests at the cost of a false positive rate (ρ), but in contrast support dynamic updates of the underlying set. This data structure includes the fingerprints of the chaff pseudonyms used to sign chaff CAMs and chaff DENMs. When receiving a CAM or a DENM, an OBU could efficiently validate the attached pseudonym against the corresponding CF; if the membership test is positive, the CAM or the DENM is discarded; otherwise, the signature will be verified.

Chaff CAMs are to be disseminated until a vehicle reaches another mix-zone or the end of the trip duration. When a relaying vehicle intends to stop disseminating chaff CAMs, e.g., entering another mix-zone, it queries the PCA, signed under the private key of the chaff pseudonym, to remove that chaff pseudonym from the corresponding CF. Further dissemination of chaff CAMs using such a chaff pseudonym is considered a misbehavior and it can be identified by a misbehavior detection system, e.g. [104], that triggers the revocation. The CFs are frequently updated by the PCAs and pushed to the corresponding RSUs.

An RSU operating a mix-zone cannot filter out chaff pseudonyms, originated from other mix-zones; the PCA prepares a distinct set of chaff pseudonyms for each RSU, operating a mix-zone. As a result, an RSU cannot distinguish between a real pseudonym and a chaff one of another RSU. However, a vehicle might encounter other relaying vehicles with chaff pseudonyms obtained from other mix-zones. For example, when a vehicle is crossing mix-zone A and moving towards mix-zone B in Fig. 2, it might encounter chaff pseudonyms originated from mix-zone B. Thus, it needs to request and obtain the CF corresponding to mix-zone B. The vehicle could directly interact with the PCA and request to obtain the CFs, corresponding to the nearby mix-zones. The PCA needs to identify the physical location[3], e.g., [79], [97], [98], of requesting vehicles; in fact, requesting vehicles should be physically ‘‘close’’ to a mix-zone to obtain the corresponding CF for. Otherwise, an external adversary could request to obtain all CFs, thus filtering out all chaff pseudonyms exiting the mix-zones.

In what follows, we describe the credential acquisition protocols in Sec. IV-B. We present updated mix-zone advertisement with chaff pseudonym acquisition protocols in Sec. IV-C. Next, we present CF dissemination in Sec. IV-D. Table I provides a description of the functions and notion used.

B. Credentials Acquisition

A vehicle first requests an anonymous ticket [20], [21] from its LTCA, using it to interact with the desired PCA to obtain pseudonyms. Upon reception of a valid ticket, it generates Elliptic Curve Digital Signature Algorithm (ECDSA) public/private key pairs [1], [2] and sends the request to the PCA [20], [21]. Having received a request, the PCA verifies the ticket signed by the LTCA (assuming trust is established

[3]Physical identification of vehicles is also a key requirement in the original mix-zone scheme [34], [35], [37]; this prevents an adversary from remotely requesting the symmetric keys of the mix-zones.

TABLE I
NOTATION USED IN THE PROTOCOLS.

$Cert_{rsu}$	Long-term certificate of an RSU
CP	Chaff Pseudonym
CF^i	Cuckoo Filter corresponding to PCA^i
$E_k(msg)$, $D_k(msg)$	Encryption and decryption of msg using key k
(K_v^i, k_v^i)	Pseudonymous public/private key pairs
L_v^i	Length of vehicle i
(LK_v, Lk_v)	Long-term public/private key pairs
$(msg)_{\sigma_v}$	A signed message with the vehicle's private key
$(P_v^i)_{pca}$, P_v^i	A pseudonym signed by the PCA
Pos_{cmix}, R_{cmix}	The center and radius of a mix-zone
$Req_{SK}/Req_{CP}/Req_{CF}$	Requesting SK, CP, CF
$Sign(Lk, msg)$	Signing a message with the private key (Lk)
SK_{cmix}^i	Symmetric Session Key inside mix-zone i
t	Current timestamp
$Verify(LK, msg)$	Verifying a message with the public key

between the two). Then, the PCA initiates a proof-of-possession protocol to verify the ownership of the corresponding private keys by the vehicle. Finally, the PCA issues the pseudonyms and delivers the response. In order to achieve full unlinkability, each pseudonym should be obtained with a single ticket. For detailed security protocols and comprehensive performance evaluation, we refer interested readers to [8], [20], [105], [75].

C. Cryptographic Mix-zone Participation

Fig. 3 shows the mix-zone advertisement and chaff pseudonym acquisition protocols. An RSU periodically broadcasts the center of a mix-zone, Pos_{cmix}, its radius R_{cmix}, and timestamps, signed with the RSU private key; the RSU attaches it LTC as well (step 3.1, i.e., step 1 in Fig. 3). To join a mix-zone, the approaching vehicle first verifies the RSU LTC and then the mix-zone information, by validating the signature on the message (step 3.2). Each vehicle needs to obtain the mix-zone symmetric session key (SK_{cmix}^i), one chaff pseudonym (CP), and the current CF. It prepares a request and includes the vehicle length (L_v^i) and current timestamp (t); it signs the request with the private key corresponding to its currently valid pseudonym, and sends it to the RSU; it attaches its pseudonyms to facilitate message validation (step 3.3). Upon receipt of the request (step 3.4), the RSU verifies the signature (step 3.5) and delivers the response, first signed by the RSU and then encrypted by the vehicle's pseudonymous public key (step 3.6). The response includes the mix-zone symmetric key (SK_{cmix}^i), a chaff pseudonym (CP), current CFs signed by the PCAs ($(CF^i)_{\sigma_{Lk_{pca_i}}}$), and the timestamp. The RSU also provides the length of another vehicle (L_v^j), which needs to be emulated, i.e., disseminating chaff CAMs for it. Finally, the vehicle decrypts the response using its private key corresponding to the currently valid pseudonym (step 3.7) and it verifies the signature of the message using the public key of the RSU (step 3.8).

D. Cuckoo Filter (CF) Acquisition

The PCAs operating in a domain construct the CFs by inserting 'enough' chaff pseudonyms. The total number of required chaff pseudonyms depends on various factors, e.g., traffic conditions and desired level of privacy protection, further

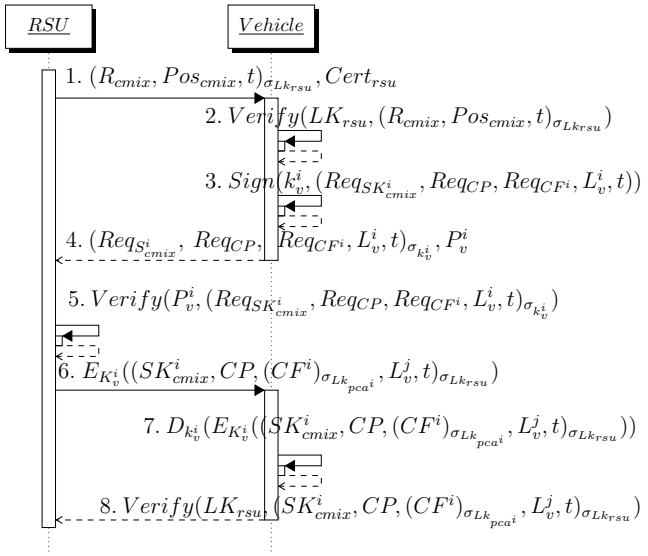


Fig. 3. Mix-zones advertisement & chaff pseudonym acquisition protocols.

evaluated in Sec. VI. Each PCA pushes a (signed) distinct CF to each RSU operating a mix-zone. RSUs provides CFs upon request (see step 6 in Fig. 3). In case of being outside of an RSU communication range, each vehicle broadcasts a signed query to its neighbors to fetch the latest CFs, e.g., similarly to [106]. Upon receiving an authentic query for the missing CFs, each vehicle searches its local repository and randomly chooses one of the requested CFs and broadcasts it. The signed CF is encrypted with the (pseudonymous) public key of the requesting vehicle. Upon reception, it decrypts the content using the private key of the currently valid pseudonym, it validates the signature of the CF (signed by the PCA), and stores them locally (evaluated in Sec. VI-E).

Each vehicle could also directly request the PCA to obtain the CFs corresponding to the nearby mix-zones. Thus, it can filter out all chaff pseudonyms it might encounter throughout its trajectory. Upon identification of the physical location, e.g., [79], [97], [98], of the requesting vehicle, the PCA provides the CFs corresponding to the nearby mix-zones. The vehicle-PCA communication is over mutually authenticated Transport Layer Security (TLS) [107] tunnels (or Datagram TLS (DTLS) [108]) leveraging the PCA's LTC and the vehicle's currently valid pseudonym. Still, a vehicle might receive chaff CAMs while it has not yet received the corresponding CF to discard them; however, it is equipped with other sensing systems, e.g., Radar and Lidar, to detect such chaff vehicles. In this case, it conducts an online pseudonym validation, e.g., Online Certificate Status Protocol (OCSP) [8], [109], to check the validity status of the pseudonym. Evaluating the impact of introducing decoy traffic on the operation of safety applications in various traffic conditions remains as one of our future work.

E. Chaff Pseudonym Resolution

In case of a suspicious action, a report is sent to the RA; the RA queries the PCA to identify the corresponding RSU, provided the chaff pseudonym. It then sends a request to the RSU to identify the pseudonymous identity, used to request

TABLE II
INFORMATION HELD BY HONEST-BUT-CURIOUS SYSTEM ENTITIES.

Honest-but-curious (colluding) Entities	Information Leaked	Security and Privacy Implications
$LTCA$	———	An LTCA infers no information during pseudonym changes since all the communication in a mix-zone is encrypted.
PCA^i	CP_{PCA^i}	A PCA can filter out the chaff pseudonyms it issued, but it cannot link any two pseudonyms upon pseudonym change or a pseudonym to a chaff one.
RSU^i	$CP_{PCA^i}^{RSU^j}, P^i, L_v^i$	An RSU knows a *distinct* set of chaff pseudonyms and the length of requesting vehicle. It can link a chaff pseudonym to the pseudonym of a requesting vehicle.
$LTCA$, PCA_H	CP_{PCA_H}	They can infer all chaff pseudonyms, issued by PCAs, operating in a domain.
$LTCA$, RSU_H	CP_{PCA_H}, P^i, L_v^i	They can infer all chaff pseudonyms, the length of vehicles and their pseudonyms.
PCA_H, RSU_H	$CP_{PCA_H}, P_{PCA_H}, L_v^i$	They can infer all chaff pseudonyms issued by all the PCAs and they can link a pseudonym to a chaff pseudonym. However, they cannot link two successive pseudonyms as they are issued fully unlinkable.
$LTCA$, PCA_H RSU_H	$CP_{PCA_H}, P_{PCA_H}, id_H, L_v^i$	Colluding LTCA, PCAs, and RSUs can link all successive pseudonyms to their corresponding real identities.

TABLE III
NOTATION USED IN SECURITY & PRIVACY ANALYSIS.

$LTCA^i$	$LTCA^i$ operating in a domain
PCA^i	PCA^i operating in a domain
PCA_H	A set of PCAs operating in a domain
RSU^i	RSU^i operating in a domain
RSU_H	A set of RSUs operating in a domain
id_H	Actual identities of the vehicles in a domain
P^i	A pseudonym issued by PCA^i
P_H	Pseudonyms issued by the PCAs in a domain
$CP_{PCA^i}^{RSU^j}$	Chaff pseudonyms issued by PCA^i for RSU^j
CP_{PCA_H}	Chaff pseudonyms issued by a set of PCAs in a domain

the chaff pseudonym. Having identified the pseudonym, the RA proceeds with the resolution process [8], i.e., interacting first with the PCA and then the LTCA. Due to lack of space, we do not present the detailed protocol description and we refer interested readers to our earlier work [8].

V. SECURITY AND PRIVACY ANALYSIS

We discuss how our scheme satisfies the security and privacy requirements, as well as operational requirements defined in Sec. III-C. For a detailed analysis of the information held by the VPKI entities during pseudonym provisioning, we refer to [8], [77]. Here, we only consider the privacy-sensitive information that can be inferred by system entities during mix-zone operations and pseudonym changes. Table II represents the privacy implications when honest-but-curious system entities collude, based on the notation summarized in Table III. We do not include RCA and RA in our analysis; the former only authorizes the operation of other entities [73], e.g., the LTCA and the PCA, and the latter is involved in the process of resolution. Moreover, we do not consider the disclosure of information if vehicles collude with other system entities. The effect of colluding a set of vehicles, crossing a mix-zone operated by RSU^i, would be equivalent to consider RSU^i colludes with other system entities.

All the V2X communication in a mix-zone is encrypted and hidden from an external observer. Upon a pseudonym change in a mix-zone, an external adversary, observing the encrypted communication cannot distinguish among vehicles sets towards correlating their corresponding pseudonyms (R1). A single entity cannot fully de-anonymize a user, link two successive pseudonyms, or link a chaff pseudonym to a pseudonymous identifier of a given vehicle. An LTCA or a PCA can infer no information to harm user privacy during changing pseudonyms since all communication inside a mix-zone is encrypted. An external adversary observing the communication could distinguish among pseudonym and chaff pseudonym sets based on the timing information [8]. To eliminate any distinction, the PCA issues pseudonyms and chaff pseudonyms with fully overlapping lifetimes, thus, timing information cannot harm user privacy. Moreover, the VPKI system issues fully unlinkable pseudonyms for all vehicles, thus, even if two pseudonyms are obtained by the same requester, they cannot be linked since each is requested using a distinct ticket [8], [20], [21]. LTCA cannot differentiate between a chaff pseudonym and a real one. A PCA can only differentiate chaff pseudonyms that it issued; in other words, it cannot distinguish a chaff pseudonym, issued by another PCA, from a real one. Moreover, a PCA cannot infer any information towards correlating a chaff pseudonym and an actual pseudonym: the RSU randomly assigns one chaff pseudonym to a relaying vehicle.

An honest-but-curious RSU learns the length of a requesting vehicle during mix-zone symmetric key acquisition process. However, this does not reveal additional information since the length is already included in the CAMs, frequently disseminated by the vehicle; thus, unlike the chaff-based CMIX [42] that requires vehicles provide their intended trajectory path to the RSUs, our scheme does not provide additional information (in comparison with the CMIX scheme [37]) to the RSUs. An RSU operating a mix-zone cannot filter out chaff pseudonyms originated from other mix-zones; this diminishes the probability of linking two successive pseudonyms belonging to the same vehicle; however, an RSU can filter out chaff pseudonyms that it provides towards linking successive pseudonyms upon pseudonym changes in the mix-zone. We quantitatively evaluated the successful linkability in the presence of honest-but-curious RSUs in Sec. VI-F. Collusion by PCA_A and PCA_B results in filtering out chaff pseudonyms they issued; but, they cannot observe the encrypted communication. Collusion

by RSU_H and PCA_H enable them to decrypt the encrypted communication and filter out all chaff pseudonyms. A collusion of the LTCA, PCA_H, and RSU_H enable them to link all pseudonyms issued in a given domain with their real identities. As a result, they can link any pseudonym to its prior or successive pseudonyms.

Issuing chaff pseudonyms, constructing and disseminating the CFs, and validating chaff pseudonym requests are all efficient processes (see Sec. VI-E). Each RSU, responsible for constructing a mix-zone, disseminates required information to the vehicles approaching the mix-zone, e.g., symmetric session key, mix-zone geometries, and CFs. This information is (signed by the RSU and) encrypted using the public key of a vehicle, approaching the mix-zone. All vehicle-RSU interactions are mutually authenticated using the currently valid vehicle's pseudonym and we leverage RSUs and car-to-car epidemic distribution to disseminate the CFs (R2). Non-cooperative vehicles could ignore changing their pseudonyms in order to degrade the anonymity set size of the mix-zone. However, as it is shown in Sec. VI-F, such behavior does not degrade the user privacy protection. Vehicles could also repeatedly request to obtain multiple chaff pseudonyms from the RSUs, monopolizing a substantial portion of the chaff pseudonyms (constructed by the PCA and pushed to the RSUs); however, each vehicle is equipped with an HSM which guarantees all outgoing signatures are signed under the private key of a single valid pseudonym at any time. In case of deviating from the system security policy, suspicious activities or (high-rate) spurious requests are sent to the RA to initiate a process to (possibly) resolve a pseudonym, thus identifying the long-term identity of a misbehaving vehicle, i.e., the pseudonym owner, and thus, their credentials will be revoked (R3).

The efficiency of the system stems from efficient CF construction of chaff pseudonyms (minimal overhead on the PCA side) and very fast validation (membership check) of chaff pseudonyms from a CF (minimal overhead on the vehicle side) (R4). Our scheme does not introduce extra computation overhead on the RSU side (in comparison with the CMIX scheme [37]) during mix-zone advertisement and symmetric key distribution. We allocate a small fraction of bandwidth for CFs distribution, which is sufficient to timely distribute CFs to all legitimate vehicles approaching a mix-zone (see Sec. VI). Our scheme introduces communication overhead to disseminate decoy traffic to enhance user privacy. In order to balance communication overhead and user privacy protection, our scheme also provides fine-grained adaptive mechanism to adjust the amount of decoy traffic in various situations, i.e., less decoy traffic during the rush-hours or more decoy traffic in sparse traffic conditions. Given a data rate of several Mbit/sec for modern IEEE 802.11p interfaces [110], dissemination of decoy traffic does not pose a significant communication overhead. Sec. VI-E provides a detailed quantitative analysis of our scheme on computation and communication overhead: disseminating decoy traffic for all vehicles introduces resealable computation and communication overhead.

Each CF is signed by the corresponding PCA which generated the chaff pseudonyms. Upon receiving a request from a vehicle, an RSU encrypts the CF (along with symmetric key and a chaff pseudonym) with the public key of the pseudonymous certificate of requesting vehicle. Thus, an eavesdropper cannot identify the chaff pseudonyms to filter out the decoy traffic. Moreover, an adversary cannot infer the number of active chaff pseudonyms from the size of a CF: each PCA overfills the CF with extra chaff pseudonyms; thus the size of a CF remains constant. This results in hiding the number of active chaff pseudonyms from an eavesdropper as well as diminishing the need to frequently update the CF and re-broadcast the updated fingerprint. There is a trade-off: the higher the number of chaff pseudonyms is, the larger the CF size becomes, thus the less frequent CF updates and broadcast are. Obtaining a large CF (e.g., valid for a day) could enable an adversary to filter out all chaff pseudonyms during that period. Thus, the more frequent CF updates, the lower the vulnerability window becomes.

Changing pseudonyms require changing addresses across a vehicle's protocol stack (i.e., MAC and IP addresses) to prevent their old and new pseudonyms from being (trivially) linked based on these interfaces [81], [82], [111], [112], [113]. For the CMIX scheme [37], each vehicle changes its addresses when changing pseudonym, i.e., once every pseudonym lifetime (τ_P); however, by leveraging our scheme, each vehicle should change its IP and MAC addresses twice every beacon interval (γ_v), e.g., 20 times per second if γ_v is 0.1s. To facilitate a fast handover, a vehicle could have potentially multiple virtual IP and MAC addresses at the same time, e.g., [114]. Thus, the relaying vehicles (and the RSUs), responsible for disseminating decoy traffic, would broadcast their actual CAMs and the chaff ones under distinct addresses. This eliminates (i) the trivial linking between the old and the new pseudonyms by the eavesdroppers, (ii) the ability of an adversary to filter out chaff CAMs based on the same interface identifier, and (iii) the need to change IP and MAC addresses twice every beacon interval.

Although communications inside the mix-zones are cryptographically protected, the physical properties of wireless radio signals, e.g., Received Signal Strength Indication (RSSI), time of arrival, Doppler shift, etc., could be used by an adversary to localize and identify propagation path from a transmitter, e.g., [91]. Tracking an object using such properties, e.g., [96], raises privacy concerns as such interfaces are uniquely associated with a single vehicle. Beyond that, by leveraging our scheme to disseminate decoy traffic, an adversary could filter out chaff CAMs from the actual ones since both are originating from the same transmitter, e.g., based on the Doppler shift and RSSI [115], [116], or by identifying the source Network Interface Card (NIC) of an IEEE 802.11 frame [91]. Based on our adversarial model, an adversary cannot differentiate decoy traffic from the actual ones using the properties of physical layer device identification. Leveraging such techniques to identify vehicles based on the signal's device-of-origin and track them accordingly requires a stronger adversary with more sophisticated resources to conduct such attacks; this requires a detailed investigation and remains as our future work.

Any CMIX scheme requires the VPKI system issuing pseudonyms with overlapping intervals. This facilitates transition to a new pseudonym at any time, e.g., when encountering a mix-zone. All vehicles registered in the system are provided with HSMs, ensuring that private keys never leave the HSMs,

Algorithm 1 Syntactic and Semantic Linking Attacks

```
1:  procedure LinkingSuccessivePseudonymsAlgorithm( )
2:      Fetch eavesdropped beacon and road layout information
3:      Classify eavesdropped beacons based on vehicle length
4:      Create a list with the first & last seen beacons for each identifier
5:      Filter out trivially linked pseudonyms (not changing their pseudonyms)
6:      MaxTravTime ← Maximum time to traverse a mix-zone
7:      MinTravTime ← Minimum time to traverse a mix-zone
8:      for Each B^i in BEACON_SET do
9:          B_f^i is the first seen message for beacon B^i
10:         B_l^i is the last seen message for beacon B^i
11:         for Each B_f^{i+1} in BEACON_SET do
12:             B_l^i and B_f^{i+1} are not correlated
13:             diff ← time difference between B_l^i and B_f^{i+1}
14:             if diff ≥ MinTravTime && diff ≤ MaxTravTime then
15:                 if pseudo-id for B_l^i and B_f^{i+1} not seen together then
16:                     if exists a road path from B_l^i to B_f^{i+1} then
17:                         if B_f^{i+1} direction is from an exit point then
18:                             B_l^i and B_f^{i+1} are correlated
19:                             break
20:                         end if
21:                     end if
22:                 end if
23:             end if
24:         end for
25:     end for
26: end procedure
```

thus mitigating Sybil attacks [117]. Note that chaff pseudonyms and their corresponding private keys are not required to be inside the HSMs; they can be stored in the OBUs. Thus, even if a vehicle is provided with multiple chaff pseudonyms, it cannot perform Sybil-based misbehavior since such chaff pseudonyms will be ignored by other vehicles and they cannot be used for any specific application.

VI. Quantitative Analysis

In order to evaluate the performance of our scheme, we need a tracking algorithm to conduct semantic and syntactic linking attacks. We do not constrain the choice and design of the tracking algorithm and we do not dwell on its performance. There are other tracking algorithms in the literature, e.g., [11], [14], [18], [42]. For example, the tracking algorithm in [42] was based on an exposure metric leveraging a vehicle's route length utilizing a pseudonym and the number of mix-zones traversed during a trip. However, in our work, we utilize a more sophisticated tracking algorithm by leveraging information in the CAMs in order to link two successive pseudonyms, thus tracking a vehicle. Due to the lack of a solid basis to compare the strength of the algorithms, we invented our own tracking algorithm, which is orthogonal to the defense mechanism.

A. Tracking Algorithm

We introduce a tracking algorithm towards conducting *syntactic* and *semantic* linking attacks. An adversary might observe an isolated pseudonym change, and associate the old and new pseudonymous identifiers through syntactic linking. Alternatively, an adversary could leverage physical constraints of the road layout, and CAMs or DENMs payload, e.g., location, velocity, and time, to predict a vehicle's trajectory towards linking messages semantically. The goal of an adversary is to link two successive pseudonyms upon pseudonym change within a mix-zone. Algorithm 1 shows our tracking algorithm: it first fetches eavesdropped beacon information and the road layout information (step 1.2, i.e., step 2 in Algorithm 1). It then classifies beacons based on the length of the vehicles (step 1.3). Next, it selects the first and the last observed beacons corresponding to each pseudonymous identifier (step 1.4). It then removes the beacons that enter and exit the mix-zone with the same pseudonymous identifiers, i.e., filtering out trivially linked pseudonyms (step 1.5). The minimum and maximum time to traverse a mix-zone is calculated based on the mix-zone geometry and vehicle speed limits (steps 1.6 – 1.7). The algorithm aims at linking the last observed beacon, in fact, the one seen before entering the mix-zone, to one of the messages exiting the mix-zone. Two pseudonyms are deemed correlated (i.e., belonging to the same vehicle) if (i) the time difference between the two observed beacons is within the minimum and maximum time to traverse the mix-zone, (ii) the two pseudonyms have not been seen together (i.e., syntactic linking [13]), (iii) there exists a road path from the last seen beacon (B_l^i) to the first seen beacon (B_f^{i+1}) [14], and (iv) the direction of the first seen beacon (B_f^{i+1}) is from one of the exit points of the mix-zone (steps 1.8 – 1.25).

TABLE IV
Simulation parameters for the experiments.

Parameters	Value	Parameters	Value
Beacon TX interval (γ_v)	0.2s, 0.5s, 1s	Snapshot interval	1s
Carrier frequency	5.89 GHz	Number of RSUs	100
TX power	20mW	RSUs transmission range	600 meters
Physical layer bit-rate	18Mbps	Number of Mix-zones	25
Sensitivity	-89dBm	Mix-zone transmission range	100 meters
Thermal noise	-110dBm	MxZ Advertisement interval (γ_{mz})	0.5s, 1s
Area size	15KM × 15KM	Number of eavesdroppers	25
Average trip duration	692.81s	Eavesdropping range	250 meters
Number of trips	287,939	Non-cooperative vehicles	0%-50%
Number of vehicles	138,259	CF distribution bandwidth ($\mathbb{B}$)	50 KB/sec
Duration of simulation	24 hours	CF TX interval	1s
Rush-hour periods	7-10, 12-14, 17-20	Fraction of honest-but-curious RSUs	0%-100%

B. Experimental Setup

We use OMNET++, the PREXT project [19], [118], and the Veins framework [119] to simulate a large-scale scenario using SUMO [120] with a realistic mobility trace, the LuST dataset [39]. V2X communication is over IEEE 802.11p [110]. For CF dissemination, we assume there is one PCA, even though the extension of our implementation with multiple PCAs is straightforward. RSUs broadcast a CF data structure, constructed and signed by a PCA. For CF operations (insertion and membership test), we used PYBLOOM [121]. To effectively place the RSUs and mix-zones, we sorted the intersections with the highest numbers of vehicles passing by. We then placed 100 RSUs and selected 25 to be mix-zones based on these *highly-visited* intersections with non-overlapping radio ranges. This aims at maximizing the chance for the vehicles to cross at least one mix-zone during their trajectory. We configured the transmission range of RSUs and mix-zones to be 600 and 100 meters, respectively. Near each mix-zone, we placed an eavesdropper with receiving antennas (250 meters range) capturing all broadcasted beacons. But, it cannot observe the communication within a cryptographically protected mix-zone. Vehicles are provided with pseudonyms

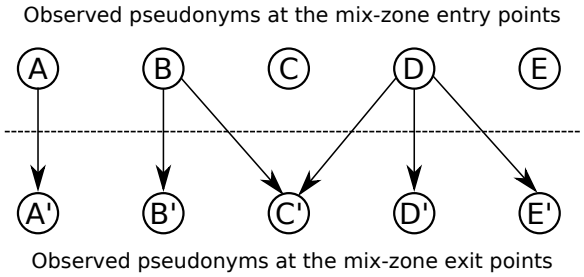


Fig. 4. Pseudonym transition from the adversary's viewpoint before entering and after exiting a mix-zone (without decoy traffic); the ground truth: $[A, A']$, $[B, B']$, $[C, C']$, $[D, D']$, $[E, E']$. Each arrow shows potential linkability between two pseudonyms based on Algorithm 1.

with overlapping intervals, compatible with the proposals of standardization bodies, i.e., IEEE 1609.2 WG [1] and ETSI [2]. They enter a cryptographically protected mix-zone and change their pseudonyms inside it. RSUs randomly assign a percentage of vehicles to be *relaying* ones to broadcast decoy traffic. Table IV shows the simulation parameters.

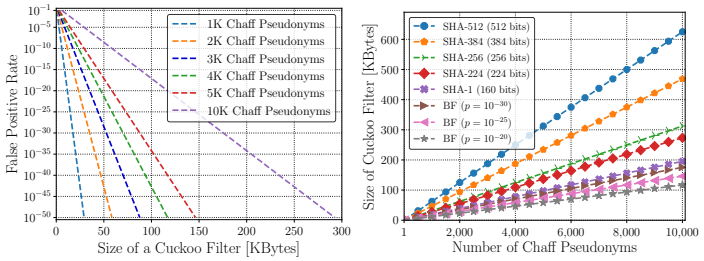
C. Metrics

To evaluate the performance of our scheme, we measure *end-to-end delay* to obtain CFs of chaff pseudonyms, i.e., from the time a vehicle approaches a mix-zone until it successfully downloads them. The maximum allocated bandwidth for CFs distribution, i.e., system parameter $\mathbb{B}$, is chosen to be much smaller than C, the bandwidth the data link support. We choose a small amount of bandwidth (50 KB/s) in order not to interfere with safety-critical operations. We also evaluate additional computation and communication overhead, imposed by our scheme, on the overall VC system components.

For privacy evaluation[4], we consider *average successful linkability*, the ratio of correctly linking two successive pseudonyms, A and B, belonging to the same vehicle (by leveraging Algorithm 1). We also consider *average tracking duration*, i.e., average traversed distance by any single vehicle that the adversary (eavesdropper deployed across multiple locations) can cumulatively track. This implies the cumulative successive correct linking of pseudonyms (and thus CAMs) across multiple CMIXs (including the trajectories from one CMIX to the next one).

Fig. 4 exemplifies a snapshot of pseudonym transitions from an adversary's viewpoint before and after a mix-zone. There are 5 pseudonym transitions happened in the mix-zone. The eavesdropper leverages algorithm 1 towards linking successive pseudonyms. Each arrow shows potential linkability between two pseudonyms based on algorithm 1.The linkability success rate is calculated as the probability of linking two pseudonyms belonging to the same vehicle. With respect to the ground truth, the probability of success rate for linking pseudonym A, Pr_A, is 1, Pr_B is $\frac{1}{2}$, Pr_C and Pr_E are zero, and Pr_D is $\frac{1}{3}$. Thus, the success rate for this example is $\frac{1+\frac{1}{2}+\frac{1}{3}}{5} = 0.36$.

[4]There are different metrics for quantifying location privacy, e.g., anonymity set size, distortion [122], and exposure degree [123]. Here, we focus on a fundamental metric to quantify location privacy. The selection of an optimal metric for quantifying location privacy and a full-blown comparison by leveraging various metrics warrant a separate investigation.



(a) (b)

Fig. 5. (a) The size of a CF as a factor of false positive rate. (b) The size of a CF as a factor of chaff pseudonyms numbers.

D. Summary of Results

Our scheme incurs low communication and computation overhead: the size of a CF with 1K chaff pseudonyms ($\rho = 10^{-25}$) is 14.63 KB, which is sufficient to disseminate decoy traffic[5] for 50% of vehicles by all the mix-zones for an hour. Moreover, such information can be timely disseminated across a region: $F_x(t = 6\,ms) = 0.99$ (Fig. 5 and Fig. 6(a)). Further, the additional computation overhead for a vehicle to validate a chaff pseudonym by performing a CF membership test with 1K chaff pseudonyms ($\rho = 10^{-25}$) is 3.68e-4 ms, which is highly efficient and scalable even with modest Nexcom OBUs [6] (Fig. 6(b) and Table V).

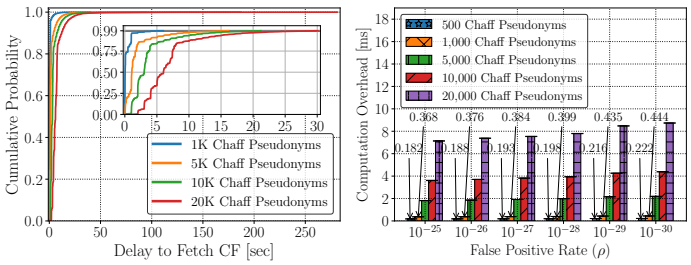
We compare our scheme with the CMIX [37], namely the *baseline* scheme, and the chaff-based CMIX [42]. Enhancing user privacy, i.e., preventing linking two successive pseudonyms by disseminating decoy traffic (for all vehicles) incurs low communication and commutation overhead: in comparison with the baseline scheme [37], the average communication overhead by RSUs increases from 0.26 KB/sec to 0.88 KB/sec, the average computation overhead for an RSU increases from 0.6 ms to 0.64 ms, and the average computation overhead on the vehicle side increases from 2.05 ms to 14 ms (Fig. 8). However, even with the modest computing resources, this extra computation overhead is reasonably low (Fig. 8).

Our scheme outperforms prior works [37], [42]: for the baseline scheme [37], an eavesdropper could successfully link ≈68% of pseudonyms after vehicles change their pseudonyms in a mix-zone. For chaff-based CMIX [42] with 50% decoy traffic, the average probability of linking pseudonyms is ≈50%. In contrast, by leveraging our scheme with decoy traffic for 50% of the vehicles, the average probability of linking pseudonyms is ≈19% (Fig. 7, Fig. 9, Fig. 10, Fig. 11, Fig. 12, and Fig. 13). Even in the presence of non-cooperative vehicles, not changing their pseudonyms while crossing the mix-zones, the average successful tracking is reasonably low (Fig. 14).

E. Performance Evaluation

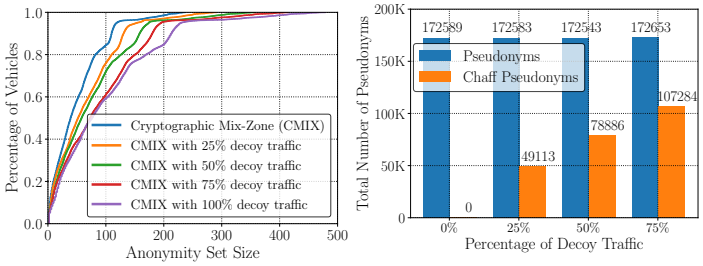
Representing chaff pseudonyms in a space-efficient CF trades off communication overhead for a false positive rate (p) [101]. Fig. 5(a) shows that the CF size linearly increases as the false positive rate decreases. For example, for 1000 chaff pseudonyms with $p = 10^{-30}$ (with the optimal number of hash functions), the CF size is 17.55 KB. This eliminates

[5]The percentage of decoy traffic indicates the percentage of (relaying) vehicles disseminating chaff CAMs.



(a) Communication Latency (b) Computation Latency

Fig. 6. (a) End-to-end delay to obtain CF of chaff pseudonyms to vehicles approaching mix-zones ($\rho = 10^{-30}$, $\mathbb{B} = 50KB/s$, $\gamma_v = 0.5s$, $\gamma_{rsu} = 1s$). (b) Computation overhead to validate chaff pseudonyms.



(a) (b)

Fig. 7. (a) CDF of anonymity set size for CMIX and CMIX with decoy traffic, observed by the eavesdroppers. (b) Total number of disseminated pseudonyms and chaff pseudonyms, derived from the vehicles (γ_v=*0.5s*).

TABLE V
LATENCY FOR VALIDATION OF CHAFF PSEUDONYMS USING A CF, EXECUTED ON A NEXCOM OBU, AVERAGED OVER 5000 RUNS.

chaff pseudonyms	false positive	CF size	delay	check/sec.
500	p=10^{-25}	7.31 KBytes	0.182 ms	2740
1000	p=10^{-25}	14.63 KBytes	0.368 ms	2719
5000	p=10^{-25}	73.13 KBytes	1.814 ms	2755
10000	p=10^{-25}	146.26 KBytes	3.611 ms	2769
20000	p=10^{-25}	292.51 KBytes	7.135 ms	2803
500	p=10^{-30}	8.78 KBytes	0.222 ms	2254
1000	p=10^{-30}	17.55 KBytes	0.444 ms	2254
5000	p=10^{-30}	87.75 KBytes	2.191 ms	2282
10000	p=10^{-30}	175.51 KBytes	4.387 ms	2279
20000	p=10^{-30}	351.02 KBytes	8.735 ms	2289

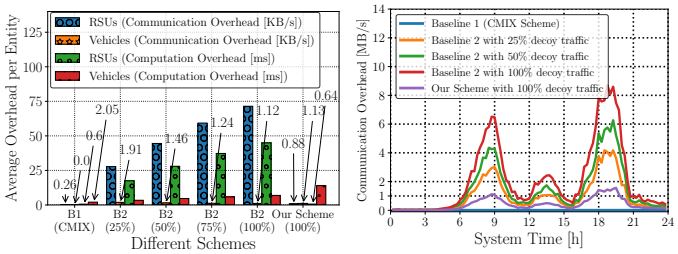
the need to validate chaff CAMs, thus enabling the correct functionality of safety applications. Note that there could be multiple CFs from different PCAs and the chaff pseudonyms are pro-actively integrated into CFs while they are updated over time, i.e., removing the expired ones and adding new ones. Given a data rate of several Mbit/sec for modern IEEE 802.11p interfaces [110], dissemination of CF updates do not pose a significant communication overhead. For example, the average number of chaff pseudonyms, per hour, for all the mix-zones to disseminate 25%, 50%, 75%, and 100% of decoy traffic is 688, 1140, 1567, and 1929, respectively.

The PCA can concatenate hash values for chaff pseudonyms. Fig. 5(b) compares our CF-based chaff pseudonyms fingerprint size with the five approved hash algorithms [124]: SHA-1, SHA-224, SHA-256, SHA-384 and SHA-512, each producing hash digest size of 160, 224, 256, 384 and 512 bits, respectively. For instance, by employing SHA-256 (32 bytes output size) as the pseudonym serial number, the size of a fingerprint for 5,000 chaff pseudonyms becomes 156 KB; whereas employing our scheme results in 73.13 KB ($p = 10^{-25}$) or 87.75 KB for the extremely low false positive rate ($p = 10^{-30}$). Alternatively, one can utilize truncated hash digests; however, truncated message digests must be carefully used: if the message digest length is too small, computation of pre-image, second pre-image or collisions becomes feasible [125]. All in all, truncation will not guarantee the expected security strength of a hash digest [124]. For a detailed investigation of different types of attacks on CFs (or BFs), we refer readers to [77], [125].

Fig. 6(a) shows the Cumulative Distribution Function (CDF) of end-to-end latencies to obtain the CFs with different number of chaff pseudonyms. We consider *end-to-end latency metric*, i.e., from the time a vehicle approaches a mix-zone until it successfully downloads all 'pieces' of a CF. Vehicle beacon frequency is $\gamma_v = 0.5s$ and RSUs beacon frequency is $\gamma_{rsu} = 1s$. The maximum allocated bandwidth to disseminate the CFs is $\mathbb{B} = 50$ KB/sec. In general, the higher the number of chaff pseudonyms, the larger CF size, thus the higher the latency to download CFs. For example, with 1000 chaff pseudonyms in a CF ($\rho = 10^{-30}$), 99% of the vehicles approaching a mix-zone received CF in less than 5s: $F_x(t = 5s) = 0.99$. Fig. 6(b) compares the computation delays for validating chaff pseudonyms in a CF with different number of inserted chaff pseudonyms. We performed our experiments on the Nexcom OBU boxes [6] (Dual-core 1.66 GHz, 1GB memory). For example, the average latency to perform a 1000 membership check for a CF with 1000 chaff pseudonyms ($\rho = 10^{-25}$) is ≈0.368 ms, i.e., the average latency to validate one chaff pseudonym is 3.68e-4 ms. Table V shows the latency for validating chaff pseudonyms with different false positive rates. For example, the latency to validate 10K chaff pseudonyms with 10K items in a CF with $p = 10^{-25}$ is ≈3.611 ms, i.e., 3.611e-4 ms to validate one chaff pseudonym. This is the extra overhead to filter out a chaff pseudonym, if seen, and discard all upcoming CAMs, signed under the private key of such a pseudonym. This shows that our scheme incurs minimal overhead on the vehicle side to filter out chaff pseudonyms from the real ones. We do not include latency for the PCA to insert chaff pseudonyms into a CF. Conducting such efficient operations on a PCA, with stronger computational resources, does not impose significant overhead. We refer to [77] on the evaluation of latency to insert items into a CF.

F. Performance Comparison

Fig. 7(a) shows the CDF of anonymity set size for the baseline and our scheme: dissemination of decoy traffic would increase the anonymity set size of the vehicles inside the mix-zones. This would diminish the power of an adversary to successfully track vehicles upon exiting the mix-zones. This obviously trades off communication overhead for a higher privacy protection level (see Fig. 8(b)). Fig. 7(b) shows the total number of pseudonyms and chaff pseudonyms for the baseline (0% of decoy traffic) and our scheme: the higher the percentage of decoy traffic is, the higher the number of chaff pseudonyms needed. For example, with 50% of vehicles chosen to be relaying nodes, ≈78K chaff pseudonyms needed for 24



(a) Overhead Comparison (b) Communication Comparison

Fig. 8. Comparison among CMIX (B1) [37], chaff-based CMIX (B2) [42], and our scheme: 1K chaff pseudonyms in a CF with $\rho = 10^{-25}$; beacon frequency: $\gamma_{mz} = 0.5$, $\gamma_v = 0.2$. (a) Computation and communication overhead. (b) Communication overhead, averaged every 300s.

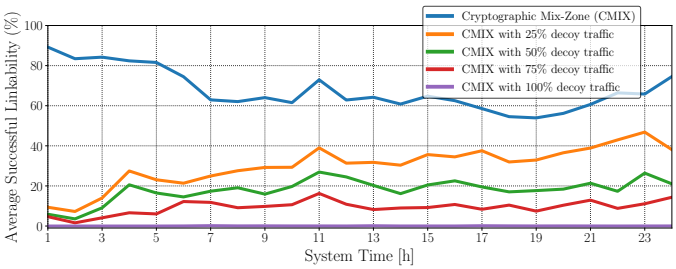


Fig. 9. Average successful linkability comparison with the CMIX scheme [37] through conducting syntactic and semantic linking attacks.

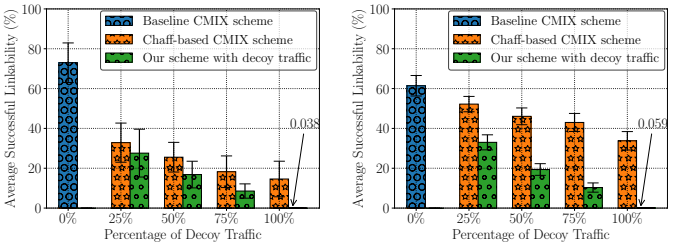
hours (for all 25 mix-zones). This is helpful to disseminate the decoy traffic via RSUs and the relaying vehicles. From our experimental results, the average number of chaff pseudonyms, per hour, required to disseminate 50% decoy traffic for each mix-zone is 52. Thus, a PCA (assuming there is only one) needs to construct a distinct CF with 52 chaff pseudonyms for each mix-zone.

Fig. 8(a) compares the computation and communication overhead for the CMIX [37], chaff-based CMIX [42], and our scheme. For our experiments, we assumed that RSUs are configured with *n1-standard-1* on the Google Cloud Platform (GCP) [126]. With this setup, the signature generation latency for ECDSA 256 key size is ≈0.3 ms and verification latency is ≈0.4 ms. Vehicles are provided with Nexcom boxes [6]: the signature generation latency is ≈3 ms and the verification latency is ≈3.5 ms. In our experiments, the size of a pseudonym (and a chaff pseudonym) is 140 bytes and the size of a CAM is 350 bytes [1], [15], [17], [127]. For the CMIX scheme [37], the computation and communication overhead is minimal: the average communication overhead for an RSU is 0.26 KB/sec and the average computation overhead is 0.6 ms. The communication overhead on the vehicles is zero while the average computation overhead is 2.05 ms. For the chaff-based CMIX [42], vehicles could request to obtain pre-generated CAMs from an RSU, operating a mix-zone, for the remaining trip duration. Our scheme outperforms the chaff-based CMIX [42]: by disseminating 100% decoy traffic for our scheme, the average communication overhead for an RSU is 0.88 KB/sec while for the chaff-based CMIX, the overhead is ≈72 KB/sec. The average computation overhead for an RSU with our scheme is 0.64 ms while this is 45 ms for the chaff-based CMIX scheme. Leveraging our scheme incurs higher computation overhead on the vehicle side in comparison with the chaff-based CMIX scheme [42] due to generation of chaff CAMs: the computation overhead for our scheme is 14 ms while this is 7 ms for the chaff-based CMIX scheme. However, even with the modest Nexcom box [6] computing resources, this extra computation overhead remains reasonably low.

Fig. 8(b) shows the total communication overhead for the CMIX scheme [37], chaff-based CMIX scheme [42], and our scheme. Mix-zones advertisement beaconing frequency is $\gamma_{mz} = 0.5$ and vehicles broadcast CAMs at the frequency of $\gamma_v = 0.2$. The size of a CF with 1K chaff pseudonyms ($\rho = 10^{-25}$) is 14.63 KB. The communication overhead is averaged every 300 seconds. As the figure shows, the total communication overhead for the CMIX scheme [37] is minimal, i.e., 6.146 KB/sec. However, the communication overhead for the chaff-based CMIX [42] scheme reaches ≈8 MB/sec when broadcasting chaff CAMs for all the vehicles. This is mainly due to the pre-generation of chaff CAMs by the RSUs. More precisely, RSUs pre-generate chaff CAMs and delivers the relaying vehicles. Thus, it has significant communication overhead during chaff CAMs acquisition process. In contrast, our scheme imposes reasonable communication overhead to the system even with the dissemination of decoy traffic for all the vehicles: the total communication overhead during the rush-hours reaches ≈1-1.5 MB/sec. This is due to the fact the each relaying vehicle would only receive a chaff pseudonym (along with the private key) from an RSU; thus, it has minimal communication overhead during chaff pseudonym acquisition.

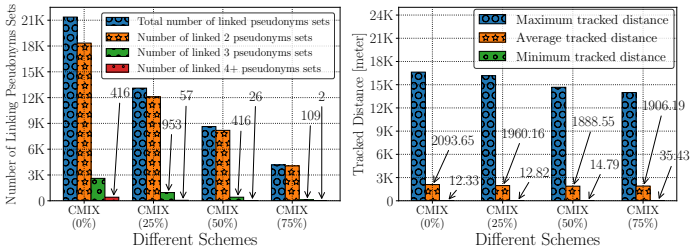
Based on the ground truth (included in the simulation results) and leveraging our novel tracking algorithm, we compute the *average successful linkability metric* towards linking pseudonyms before and after a cryptographically protected mix-zone. Fig. 9 shows the average pseudonym linkability by the eavesdroppers for a full-day realistic mobility pattern in the city of Luxembourg [39]. As we can see, the tracking algorithm could link pseudonyms for the CMIX scheme with high probability success rate during the non-rush hours period (until system time 6). The probability of linking two successive pseudonyms decreases when the traffic density increases; but still, it can successfully link the pseudonyms with ≈63% success rate at system time 7. By introducing decoy traffic for a fraction of vehicles, one can reduce the linkability: with 50% of vehicles to be the relaying vehicles, broadcasting decoy traffic, the probability of linking drops from ≈63% to ≈17% at system time 7. More so, one can eliminate (syntactic and semantic) pseudonym linking attacks by disseminating decoy traffic for all vehicles.

If the number of vehicles in a mix-zone is less than a predefined (system parameter) threshold, the RSU generates decoy traffic for all those vehicles. This stems from the results of tracking algorithm: if there are few vehicles inside a mix-zone, an adversary could easily track all those vehicles. In our simulation, we defined this threshold to be two, i.e., if there are one or two vehicles in a mix-zone, the RSU disseminates decoy traffic for all vehicles. This is also visible in Fig. 9: during very sparse traffic conditions (at system time 1), the average successful tracking is ≈7%-9%. Intuitively, the rate of decoy



(a) During Non-rush Hours (b) During Rush Hours

Fig. 10. Average successful linkability comparison with the CMIX [37] and the chaff-based CMIX [42] schemes.



(a) (b)

Fig. 11. (a) Linking pseudonym sets for the baseline and our scheme. (b) Successful tracked distance for the baseline and our scheme.

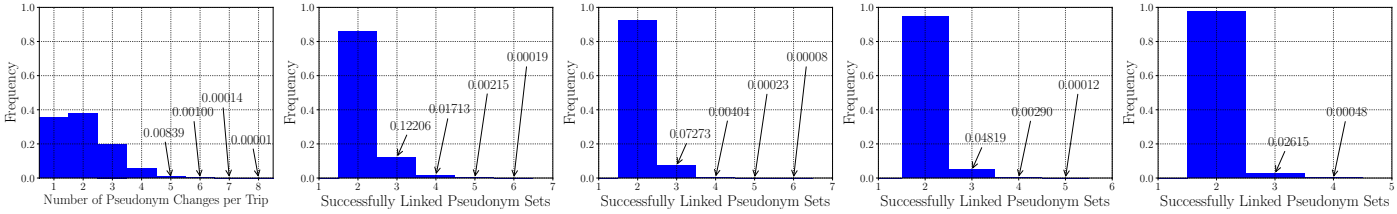
traffic should be inversely proportional to the traffic density, i.e., the higher the number of vehicles inside a mix-zone, the lower the probability of linking becomes, thus the less the number of chaff vehicles needed. Still, one needs to disseminate decoy traffic during the rush-hour periods: the probability of linking two successive pseudonyms during rush-hours, e.g., 7:00-10:00, is ≈ 62%. This trades off pseudonyms unlinkability for (communication and computation overhead) cost, which is important for balancing the effects of chaff messages on communication overhead in dense traffic scenarios.

Fig. 10 compares the average successful linkability for the CMIX scheme [37], chaff-based CMIX scheme [42], and our scheme. The average successful linkability for the CMIX scheme [37] during non-rush hours (Fig. 10(a)) is ≈73%; during rush-hours (Fig. 10(b)), it is ≈62%. With the chaff-based CMIX scheme [42], an adversary could filter out chaff CAMs from the real ones since the chaff CAMs are pre-generated by the RSUs (without considering the vehicles mobility pattern): if the distance of two CAMs, signed under two distinct pseudonyms, is less than the length of a vehicle, a chaff vehicle would stand out. The higher the percentage of decoy traffic, the higher the probability of filtering out chaff CAMs, thus the higher the average successful linkability. In contrast, with our schemes, vehicles disseminate chaff CAMs according to the traffic conditions. For the chaff-based CMIX scheme [42] with 50% decoy traffic, the average successful linkability, during the rush hours (Fig. 10(b)), is ≈46% while with the same set up for our scheme, the average successful linkability, during the rush hours, is ≈19%.

Fig. 11(a) considers the *average successful linkability metric* and compares the number of successfully linked pseudonyms sets for the baseline and our scheme. We refer to a successfully linked pseudonyms set as the number of pseudonyms, linked by the eavesdroppers, corresponding to the same vehicle. The figure shows the number of linked two-pseudonyms sets, three-pseudonyms sets, and four(+)-pseudonyms sets. For the baseline scheme, the total number of linked pseudonyms sets is 21367, i.e., 21367 sets of pseudonyms, each corresponding to the same vehicle, were successfully linked by the eavesdroppers. The total number of vehicles with two-pseudonyms sets linked is 18343, and the total number of vehicles with three-pseudonyms sets is 2608. Our scheme reduces the number of linked pseudonyms sets: the higher the percentage of decoy traffic is, the lower the number of linked pseudonyms sets becomes. With 75% of decoy traffic, the total number of linked pseudonyms sets is 4168, the total number of vehicles, linked with two-pseudonyms sets is 4057, and the total number of vehicles, linked with three-pseudonyms sets is 109. In Fig. 11(b), we consider the *tracking duration metric*, i.e., the total distance that was successfully tracked by the eavesdroppers. The average tracked distance diminishes by increasing the percentage of decoy traffic. These numbers were calculated based on the total number of linked pseudonyms sets and the distances observed by the eavesdroppers. More precisely, if eavesdroppers could link multiple-pseudonyms set corresponding to the same vehicle, then they could accumulate the total distance observed for all the CAMs, signed under the linked pseudonyms sets. For example, the average tracked distance for the baseline scheme [37] is 2093 meters; with 50% of vehicles disseminating decoy traffic, the average tracked distance becomes 1960 meters.

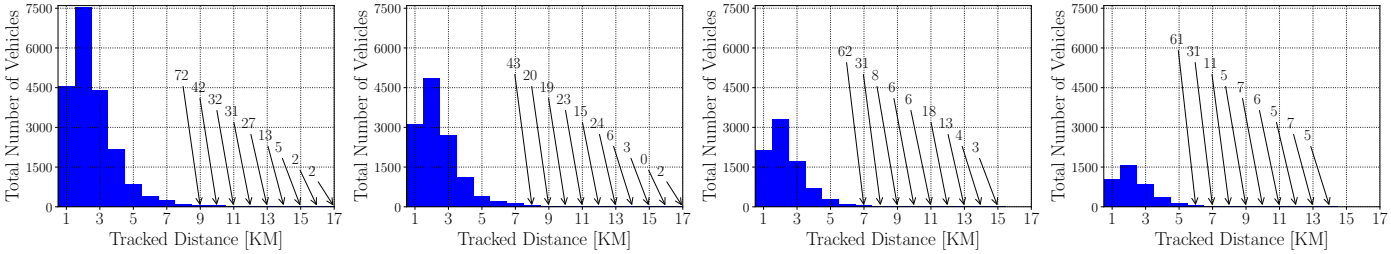
Fig. 12(a) shows the histogram of the number of pseudonym changes per trip. Vehicles only change their pseudonyms when they cross a mix-zone during their trip duration: 36% of the vehicles changed their pseudonyms only once, 38% changed twice, and 20% of them changed three times. There are also few vehicles changing their pseudonyms more than five times. The more mix-zones vehicles encounter, the higher the frequency of changing pseudonyms becomes; this would result in the higher number of unlinkable segments for any journey, thus enhancing user privacy protection. Fig. 12(b) - Fig. 12(e) show the histogram of successfully linked pseudonyms sets by the eavesdroppers for the baseline and our scheme. With the baseline scheme (Fig. 12(b)), the eavesdroppers could link 86% of two-pseudonyms sets while there are successfully linked sets with three-, four-, and five-pseudonyms. By disseminating decoy traffic, the percentage of linking pseudonyms sets decreases: with 75% of decoy traffic (Fig. 12(e)), the eavesdroppers link 97% of two-pseudonyms sets while there are very few three- or four-pseudonyms set, linked by the eavesdroppers (and no five- or six-pseudonyms sets). The higher the percentage of decoy traffic, the lower the probability of linking pseudonyms by the eavesdroppers, thus the smaller the number of linked pseudonyms corresponding to the same vehicle. This results in a smaller percentage of the trips which could be linked by the eavesdroppers to harm user privacy.

Fig. 13 shows the histogram of the number of vehicles, tracked by the eavesdroppers, based on the linked pseudonyms sets. With the baseline scheme (Fig. 13(a)), the eavesdroppers could link 4536 vehicles for 1 KM, 7532 vehicles for 2 KMs, and 4409 vehicles for 3 KMs. In contrast, by introducing decoy



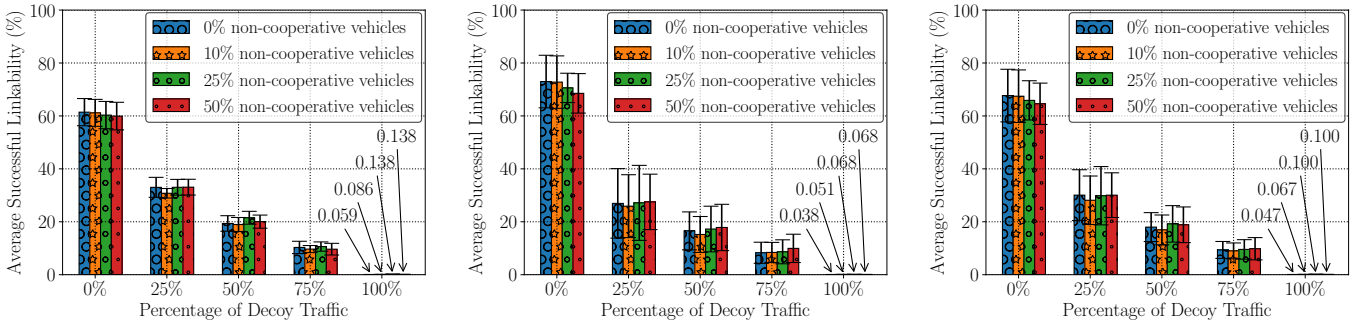
(a) Pseudonyms changes (b) CMIX: 0% decoy traffic (c) CMIX: 25% decoy traffic (d) CMIX: 50% decoy traffic (e) CMIX: 75% decoy traffic

Fig. 12. (a) Histogram of pseudonyms changes. (b) Histogram of successfully linked pseudonym sets for (b) the baseline scheme (CMIX), (c) our scheme (CMIX with 25% decoy traffic), (d) our scheme (CMIX with 50% decoy traffic), (e) our scheme (CMIX with 75% decoy traffic).



(a) CMIX: 0% decoy traffic (b) CMIX: 25% decoy traffic (c) CMIX: 50% decoy traffic (d) CMIX: 75% decoy traffic

Fig. 13. Histogram of tracked distances by eavesdroppers based on the linked pseudonyms sets for the baseline scheme (CMIX) and our scheme.



(a) During Rush Hours (b) During Non-rush Hours (c) During 24 Hours

Fig. 14. Average successful linkability in the presence of non-cooperative vehicles, not changing their pseudonyms while crossing the mix-zones.

traffic for vehicles exiting the mix-zones, the total number of vehicles, tracked by the eavesdroppers, drastically decreases: with 75% of decoy traffic (Fig. 13(d)), the eavesdroppers could only link 1044 vehicles for 1 KM, 1576 vehicles for 2 KMs, and 837 vehicles for 3 KMs. Note that by disseminating 100% decoy traffic, the probability of linking two successive pseudonyms by the eavesdroppers is very low, thus such tracking becomes ineffective (see Fig. 9 and Fig. 14).

Fig. 14 shows the average success rates in the presence of non-cooperative vehicles that try to diminish the anonymity set size of a mix-zone. Such vehicles exit the mix-zone without changing their pseudonyms; also, if chosen to be relaying vehicles, they do not disseminate decoy traffic. The tracking algorithm (step 4 in Algorithm 1) filters out these trivially linked pseudonyms, i.e., CAMs of vehicles that enter and exit the mix-zone with the same pseudonym. Fig. 14(a) shows the average successful tracking during the rush hours. The average successful tracking in the presence of non-cooperative vehicles for the CMIX scheme slightly decreases: the eavesdroppers filter out transcript of pseudonymously authenticated messages with the same pseudonym. Thus, non-cooperative vehicles, not changing their pseudonyms, do not help eavesdroppers link successive pseudonyms with higher percentage of successful tracking. During the non-rush hour periods (Fig. 14(b)), the average successful tracking for the CMIX scheme is higher than the one during the rush-hour periods: due to lower number of vehicles in a mix-zone, the probability of linking increases; still, non-cooperative vehicles that do not change their pseudonyms, when crossing a mix-zone, do not highly affect the anonymity set size. Fig. 14(c) shows the average successful tracking for the entire intervals: eavesdroppers could successfully link 68% of successive pseudonyms before and after pseudonym changes in the mix-zones.

The average successful tracking for our scheme is not considerably affected in the presence of non-cooperative vehicles thanks to dissemination of decoy traffic. Note that selection of non-cooperative vehicles is independent of selection of relaying vehicles, i.e., in each scenario, different sets of vehicles are selected to be non-cooperative. Thus, a direct comparison of the scenarios with different percentage of non-cooperative vehicles is not straightforward. In order to mitigate the effect of non-cooperative vehicles, an RSU could monitor the behavior of

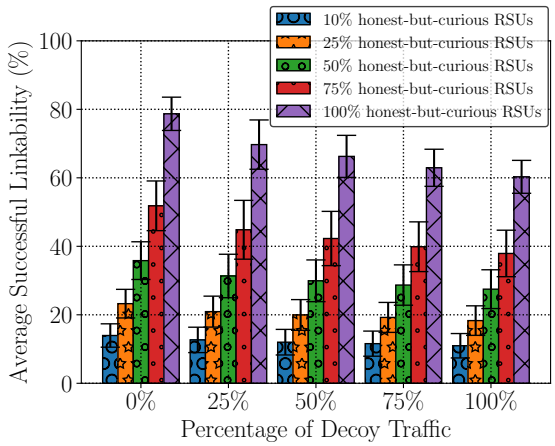


Fig. 15. Average successful linkability in the presence of a fraction of honest-but-curious RSUs, operating the mix-zones.

vehicles when entering and exiting the mix-zone; if a substantial fraction of vehicles exit the mix-zone without changing their pseudonyms, the RSU can increase the percentage of decoy traffic. Further investigation is one of our future work.

Fig. 15 shows the average successful linkability among pseudonyms sets by a fraction of honest-but-curious RSUs. Such entities have broader communication coverage and they can observe the communication inside the encrypted area. However, each RSU only knows a distinct set of CF, provided by the PCA and it cannot filter out chaff pseudonyms originated from other mix-zones. For the baseline scheme, the honest-but-curious RSUs could link the successive pseudonyms with higher probability in comparison with our scheme. For example, for the baseline scheme with 50% of RSUs to be honest-but-curious, the average successful pseudonym linkability is ≈36%. However, by introducing 100% decoy traffic, such linkability drops to ≈27%. Note that introducing chaff CAMs does not fully diminish the pseudonyms linkability in the presence of honest-but-curious RSUs. That requires introducing chaff CAMs combined with other techniques, e.g., simultaneously changing pseudonyms by all the vehicles inside a mix-zone, to fully diminish the syntactic and semantic linking attacks. This requires further investigation and remains as our future work.

VII. Conclusion and Future Work

We proposed a novel scheme to protect user privacy regardless of the geometry of the mix-zones, mobility patterns, vehicle density, and arrival rates. Our system enhances user privacy protection at the cost of low computation and communication overhead while it ensures that the operation of the safety applications remains unaffected by the dissemination of decoy traffic. Our results show that cooperative dissemination of decoy traffic, by relaying vehicles exiting a mix-zone, can significantly diminish syntactic and semantic pseudonym linking attacks. Moreover, our experiments show that the deployment of mix-zones can be cost-effective. As future work, we plan to expand our adversarial model and investigate the resiliency of our scheme against a fraction of malicious vehicles or compromised RSUs that covertly send the CMIX symmetric key or the CFs to other (internal or external) adversaries. Moreover, we plan to investigate the effect of mix-zone transmission range on the overall communication and computation overhead of the VC system. Further, we intend to improve our tracking algorithm towards tracking vehicles based on predicting vehicles trajectories using Kalman Filter and physical properties of the wireless radio signals. Moreover, we plan to investigate various metrics for quantifying location privacy and conduct a full-blown comparison of our scheme by leveraging different metrics. Finally, we intend to evaluate the impact of decoy traffic on the operation of safety applications in various traffic conditions.

Acknowledgements

Work supported by the Swedish Foundation for Strategic Research (SSF) SURPRISE project and the KAW Academy Fellowship Trustworthy IoT project.

References

[1] IEEE-1609.2, ''IEEE Standard for Wireless Access in Vehicular Environments - Security Services for Applications and Management Messages,'' Mar. 2016.

[2] ETSI, ''Intelligent Transport Systems (ITS); Security; ITS Communications Security Architecture and Security Management,'' ETSI TS 102-940, Nov. 2016.

[3] ''Commission Delegated Regulation (EU) of 13.3.2019: Supplementing Directive 2010/40/EU of the European Parliament and of the Council with Regard to the Deployment and Operational Use of Cooperative Intelligent Transport Systems,'' https://ec.europa.eu/transport/sites/transport/files/legislation/c20191789.pdf, European Commission, Tech. Rep., Mar. 2019.

[4] Car-to-Car Communication Consortium (C2C-CC), ''PKI Memo,'' https://www.car-2-car.org/, Feb. 2011.

[5] P. Papadimitratos, L. Buttyan, T. Holczer, E. Schoch, J. Freudiger, M. Raya, Z. Ma, F. Kargl, A. Kung, and J.-P. Hubaux, ''Secure Vehicular Communication Systems: Design and Architecture,'' *IEEE Comm. Mag.*, vol. 46, no. 11, pp. 100--109, Nov. 2008.

[6] M. Feiri, F. Kargl, A. Giannetsos, S. Gisdakis, H. Jin, M. Khodaei, and M. Sall, ''PREparing SEcuRe VEhicle-to-X Communication Systems,'' PRESERVE-Project, Deliverable 3.2, FOT Trial 2 Results, Tech. Rep., July 2015.

[7] W. Whyte, A. Weimerskirch, V. Kumar, and T. Hehn, ''A Security Credential Management System for V2V Communications,'' in *IEEE Vehicular Networking Conference (VNC)*, Boston, MA, Dec. 2013, pp. 1--8.

[8] M. Khodaei, H. Jin, and P. Papadimitratos, ''SECMACE: Scalable and Robust Identity and Credential Management Infrastructure in Vehicular Communication Systems,'' *IEEE Transactions on Intelligent Transportation Systems (ITS)*, vol. 19, no. 5, pp. 1430--1444, May 2018.

[9] P. Papadimitratos, V. Gligor, and J.-P. Hubaux, ''Securing Vehicular Communications-Assumptions, Requirements, and Principles,'' in *ESCAR*, Berlin, Germany, Nov. 2006, pp. 5--14.

[10] M. Gerlach and F. Guttler, ''Privacy in VANETs using Changing Pseudonyms - Ideal and Real,'' in *IEEE VTC*, Dublin, Ireland, Apr. 2007, pp. 2521--2525.

[11] L. Buttyán, T. Holczer, A. Weimerskirch, and W. Whyte, ''SLOW: A Practical Pseudonym Changing Scheme for Location Privacy in VANETs,'' in *IEEE Vehicular Networking Conference (VNC)*, Tokyo, Japan, Oct. 2009, pp. 1--8.

[12] J. Tijink, ''Position Paper Regarding Personal Data Protection Aspects in C-ITS CAR 2 CAR Communication Consortium,'' https://www.car-2-car.org/fileadmin/documents/General_Documents/C2CCC_TR_2045_Position_Paper_Personal_Data_Protection_R101.pdf, CAR 2 CAR Communication Consortium, Tech. Rep., Apr. 2017.

[13] M. Khodaei, H. Noroozi, and P. Papadimitratos, ''POSTER: Privacy Preservation through Uniformity,'' in *ACM WiSec*, Stockholm, Sweden, Jun. 2018, pp. 279--280.

[14] B. Wiedersheim, Z. Ma, F. Kargl, and P. Papadimitratos, ''Privacy in Inter-vehicular Networks: Why Simple Pseudonym Change is not Enough,'' in *WONS*, KG, Slovenia, Feb. 2010, pp. 176--183.

[15] ETSI-EN-302-637-2-V1.3.2, ''Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part 2: Specification of Cooperative Awareness Basic Service,'' ETSI EN 302 637-2 V1.3.2, Nov. 2014.

[16] M. Ullmann, T. Strubbe, and C. Wieschebrink, ''Technical Limitations, and Privacy Shortcomings of the Vehicle-to-Vehicle Communication,'' in *Proceedings of the IARIA VEHICULAR Conference*, Barcelona, Spain, Nov. 2016.

[17] S. Bai, ''US-EU V2V V2I Message Set Standards Collaboration,'' https://docbox.etsi.org/workshop/2014/201402_ITSWORKSHOP/S02_ITS_SomeBitsFromtheWorld/HONDA_BAI.pdf, Feb. 2013.

[18] K. Emara, W. Woerndl, and J. Schlichter, ''Vehicle Tracking using Vehicular Network Beacons,'' in *IEEE WoWMoM*, Madrid, Spain, Jun. 2013, pp. 1--6.

[19] K. Emara, ''Poster: PREXT: Privacy Extension for Veins VANET Simulator,'' in *IEEE Vehicular Networking Conference (VNC)*, Columbus, OH, USA, Dec 2016, pp. 1--2.

[20] M. Khodaei, H. Jin, and P. Papadimitratos, ''Towards Deploying a Scalable & Robust Vehicular Identity and Credential Management Infrastructure,'' in *IEEE Vehicular Networking Conference (VNC)*, Paderborn, Germany, Dec. 2014, pp. 33--40.

[21] M. Khodaei and P. Papadimitratos, ''Evaluating On-demand Pseudonym Acquisition Policies in Vehicular Communication Systems,'' in *Proceedings of the First International Workshop on Internet of Vehicles and Vehicles of Internet (IoV/VoI)*, Paderborn, Germany, Jul. 2016, pp. 7--12.

[22] M. Gruteser and X. Liu, ''Protecting Privacy in Continuous Location-Tracking Applications,'' *IEEE Security & Privacy*, no. 2, pp. 28--34, Mar. 2004.

[23] P. Golle and K. Partridge, ''On the Anonymity of Home/Work Location Pairs,'' in *Pervasive computing*. Springer, Berlin, Heidelberg, 2009, vol. 5538, pp. 390--397.

[24] C. Y. Ma, D. K. Yau, N. K. Yip, and N. S. Rao, ''Privacy Vulnerability of Published Anonymous Mobility Traces,'' *IEEE/ACM TON*, vol. 21, no. 3, pp. 720--733, Jun. 2013.

[25] H. Jin, M. Khodaei, and P. Papadimitratos, ''Security and Privacy in Vehicular Social Networks,'' in *Vehicular Social Networks*. Taylor & Francis Group, Mar. 2016.

[26] L. Huang, K. Matsuura, H. Yamane, and K. Sezaki, ''Enhancing Wireless Location Privacy using Silent Period,'' in *IEEE WCNC*, New Orleans, LA, USA, Mar. 2005, pp. 1187--1192.

[27] K. Sampigethaya, M. Li, L. Huang, and R. Poovendran, ''AMOEBA: Robust location privacy scheme for vanet,'' *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 25, no. 8, pp. 1569--1589, Oct. 2007.

[28] L. Buttyán, T. Holczer, and I. Vajda, ''On the Effectiveness of Changing Pseudonyms to Provide Location Privacy in VANETs,'' in *European Workshop on Security in Ad-hoc and Sensor Networks*, Berlin, Heidelberg., Jul. 2007, pp. 129--141.

[29] K. Emara, W. Woerndl, and J. Schlichter, ''CAPS: Context-Aware Privacy Scheme for VANET Safety Applications,'' in *ACM WiSec*, New York, NY, USA, Jun. 2015, pp. 1--12.

[30] L. Huang, H. Yamane, K. Matsuura, and K. Sezaki, ''Silent Cascade: Enhancing Location Privacy without Communication QoS Degradation,'' in *International Conference on Security in Pervasive Computing*, York, UK, Apr. 2006, pp. 165--180.

[31] A. Wasef and X. Shen, ''REP: Location Privacy for VANETs Using Random Encryption Periods,'' *Mobile Networks and Applications*, vol. 15, no. 1, pp. 172--185, Feb. 2010.

[32] S. Lefevre, J. Petit, R. Bajcsy, C. Laugier, and F. Kargl, ''Impact of V2X Privacy Strategies on Intersection Collision Avoidance Systems,'' in *IEEE Vehicular Networking Conference (VNC)*, Boston, MA, Dec. 2013.

[33] G. P. Corser, A. Arenas, and H. Fu, ''Effect on Vehicle Safety of Nonexistent or Silenced Basic Safety Messages,'' in *ICNC*, Kauai, HI, USA, Feb. 2016, pp. 1--5.

[34] A. R. Beresford and F. Stajano, ''Location Privacy in Pervasive Computing,'' *IEEE Pervasive computing*, vol. 2, no. 1, pp. 46--55, Jan. 2003.

[35] ------, ''Mix zones: User Privacy in Location-Aware Services,'' in *IEEE Annual Conference on Pervasive Computing and Communications Workshops*, Orlando, FL, USA, Mar. 2004, pp. 127--131.

[36] J. Freudiger, R. Shokri, and J.-P. Hubaux, ''On the Optimal Placement of Mix Zones,'' *Springer, Berlin, Heidelberg - Privacy Enhancing Technologies (PETS)*, pp. 216--234, Aug. 2009.

[37] J. Freudiger, M. Raya, M. Félegyházi, P. Papadimitratos, and J.-P. Hubaux, ''Mix-zones for Location Privacy in Vehicular Networks,'' in *Win-ITS*, Vancouver, BC, Canada, Aug. 2007.

[38] B. Palanisamy and L. Liu, ''Attack-Resilient Mix-zones over Road Networks: Architecture and Algorithms,'' *IEEE Transactions on Mobile Computing (TMC)*, vol. 14, no. 3, pp. 495--508, 2015.

[39] L. Codeca, R. Frank, and T. Engel, ''Luxembourg SUMO Traffic (LuST) Scenario: 24 Hours of Mobility for Vehicular Networking Research,'' in *IEEE Vehicular Networking Conference (VNC)*, Kyoto, Japan, Dec. 2015, pp. 1--8.

[40] J. Krumm, ''Inference Attacks on Location Tracks,'' in *International Conference on Pervasive Computing*, Toronto, Canada, May 2007, pp. 127--143.

[41] A. Tomandl, F. Scheuer, and H. Federrath, ''Simulation-based Evaluation of Techniques for Privacy Protection in VANETs,'' in *IEEE WiMob*, Barcelona, Spain, Oct. 2012, pp. 165--172.

[42] C. Vaas, M. Khodaei, P. Papadimitratos, and I. Martinovic, ''Nowhere to hide? Mix-Zones for Private Pseudonym Change using Chaff Vehicles,'' in *IEEE Vehicular Networking Conference (VNC)*, Taipei, Taiwan, Dec. 2018, pp. 1--8.

[43] C. Chen, D. E. Asoni, A. Perrig, D. Barrera, G. Danezis, and C. Troncoso, ''TARANET: Traffic-Analysis Resistant Anonymity at the Network Layer,'' in *IEEE European Symposium on Security and Privacy (EuroS&P)*, London, UK, Apr. 2018, pp. 137--152.

[44] ETSI, ''Intelligent Transport Systems (ITS); Security; Security Header and Certificate Formats,'' ETSI TS 103-097, Jun. 2015.

[45] ------, ''Intelligent Transport Systems (ITS); Security; Stage 3 Mapping for IEEE 1609.2,'' ETSI TS 102-867, Jun. 2012.

[46] J. Bellatti, A. Brunner, J. Lewis, P. Annadata, W. Eltarjaman, R. Dewri, and R. Thurimella, ''Driving Habits Data: Location Privacy Implications and Solutions,'' in *IEEE S&P*, vol. 38, no. 1, Jan. 2017, pp. 12--20.

[47] S. Narain, T. D. Vo-Huu, K. Block, and G. Noubir, ''Inferring User Routes and Locations Using Zero-Permission Mobile Sensors,'' in *IEEE S&P*, San Jose, CA, USA, May 2016, pp. 397--413.

[48] X. Gao, B. Firner, S. Sugrim, V. Kaiser-Pendergrast, Y. Yang, and J. Lindqvist, ''Elastic Pathing: Your Speed is Enough to Track You,'' in *ACM UbiComp*, Seattle, Washington, Sep. 2014, pp. 975--986.

[49] Q. Technologies, ''Leading the World to 5G: Cellular Vehicle-to-Everything (C-V2X) Technologies,'' https://www.qualcomm.com/media/documents/files/cellular-vehicle-to-everything-c-v2x-technologies.pdf, Jun. 2016.

[50] G. Greenwald, ''NSA Prism Program Taps in to User Data of Apple, Google and Others,'' https://www.theguardian.com/world/2013/jun/06/us-tech-giants-nsa-data, Jun. 2013.

[51] S. Era and B. Preneel, ''Cryptography and Information Security in the Post-Snowden Era,'' p. 1, May 2015.

[52] C. Cornelius, A. Kapadia, D. Kotz, D. Peebles, M. Shin, and N. Triandopoulos, ''Anonysense: Privacy-aware People-centric Sensing,'' in *Proceedings of the 6th international conference on Mobile systems, applications, and services (MobiSys)*, Breckenridge, CO, USA, June 2008, pp. 211--224.

[53] R. Yu, J. Kang, X. Huang, S. Xie, Y. Zhang, and S. Gjessing, ''MixGroup: Accumulative Pseudonym Exchanging for Location Privacy Enhancement in Vehicular Social Networks,'' *IEEE Transactions on Dependable and Secure Computing*, vol. 13, no. 1, pp. 93--105, Jan. 2016.

[54] M. Gruteser and D. Grunwald, ''Anonymous Usage of Location-based Services Through Spatial and Temporal Cloaking,'' in *ACM MobiSys*, San Francisco, USA, May 2003, pp. 31--42.

[55] H. Liu, X. Li, H. Li, J. Ma, and X. Ma, ''Spatiotemporal Correlation-Aware Dummy-Based Privacy Protection Scheme for Location-based Services,'' in *IEEE Conference on Computer Communications (INFOCOM)*, Atlanta, GA, USA, May 2017, pp. 1--9.

[56] ''ETSI TS 102 637-2, Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part 2: Specification of Co-operative Awareness Basic Service,'' Mar. 2011.

[57] D. Boneh, X. Boyen, and H. Shacham, ''Short Group Signatures,'' in *Advances in Cryptology CRYPTO*. Springer, 2004.

[58] G. Calandriello, P. Papadimitratos, J.-P. Hubaux, and A. Lioy, ''Efficient and Robust Pseudonymous Authentication in VANET,'' in *ACM VANET*, New York, USA, Sep. 2007, pp. 19--28.

[59] P. Papadimitratos, G. Calandriello, J.-P. Hubaux, and A. Lioy, ''Impact of Vehicular Communications Security on Transportation Safety,'' in *IEEE INFOCOM Mobile Networking for Vehicular Environments (MOVE) Workshop (IEEE MOVE)*, Phoenix, AZ, USA, Apr. 2008, pp. 1--6.

[60] G. Calandriello, P. Papadimitratos, J.-P. Hubaux, and A. Lioy, ''On the Performance of Secure Vehicular Communication Systems,'' *IEEE Transactions on Dependable and Secure Computing (TDSC)*, vol. 8, no. 6, pp. 898--912, Nov. 2011.

[61] M. Khodaei, A. Messing, and P. Papadimitratos, ''RHyTHM: A Randomized Hybrid Scheme To Hide in the Mobile Crowd,'' in *IEEE Vehicular Networking Conference (VNC)*, Torino, Italy, Nov. 2017, pp. 155--158.

[62] K. Sampigethaya, L. Huang, M. Li, R. Poovendran, K. Matsuura, and K. Sezaki, ''CARAVAN: Providing location privacy for VANET,'' in *Embedded Security in Cars*, Cologne, Germany, Nov. 2005.

[63] A. Boualouache, S.-M. Senouci, and S. Moussaoui, ''PRIVANET: An Efficient Pseudonym Changing and Management Framework for Vehicular Ad-Hoc Networks,'' *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 8, pp. 3209--3218, Aug. 2020.

[64] D. Förster, H. Löhr, A. Grätz, J. Petit, and F. Kargl, ''An Evaluation of Pseudonym Changes for Vehicular Networks in Large-Scale, Realistic Traffic Scenarios,'' *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 10, pp. 3400--3405, Dec. 2017.

[65] N. Ravi, C. M. Krishna, and I. Koren, ''Enhancing Vehicular Anonymity in ITS: A New Scheme for Mix Zones and their Placement,'' *IEEE Transactions on Vehicular Technology*, vol. 68, no. 11, pp. 10 372--10 381, Aug. 2019.

[66] M. Khodaei and P. Papadimitratos, ''Poster: Mix-Zones Everywhere: A Dynamic Cooperative Location Privacy Protection Scheme,'' in *IEEE Vehicular Networking Conference (VNC)*, Taipei, Taiwan, Dec. 2018, pp. 1--2.

[67] L. Zhang, ''OTIBAAGKA: A New Security Tool for Cryptographic Mix-zone Establishment in Vehicular Ad Hoc Networks,'' *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 12, pp. 2998--3010, Dec. 2017.

[68] B. Palanisamy and L. Liu, ''MobiMix: Protecting Location Privacy with Mix-zones Over Road Networks,'' in *IEEE 27th International Conference on Data Engineering*, Hannover, Germany, Apr. 2011, pp. 494--505.

[69] N. Guo, L. Ma, and T. Gao, ''Independent Mix Zone for Location Privacy in Vehicular Networks,'' *IEEE Access*, vol. 6, pp. 16 842--16 850, Apr. 2018.

[70] M. Li, K. Sampigethaya, L. Huang, and R. Poovendran, ''Swing & Swap: User-centric Approaches Towards Maximizing Location Privacy,'' in *ACM WPES*, Alexandria, Virginia, USA, Oct. 2006, pp. 19--28.

[71] P. K. Singh, S. N. Gowtham, S. Tamilselvan, and S. Nandi, ''CPESP: Cooperative Pseudonym Exchange and Scheme Permutation to Preserve Location Privacy in VANETs,'' *Vehicular Communications*, vol. 20, p. 100183, Dec. 2019.

[72] ETSI TR 102 731, ''Intelligent Transport Systems (ITS); Security; Security Services and Architecture,'' Sep. 2009.

[73] M. Khodaei and P. Papadimitratos, ''The Key to Intelligent Transportation: Identity and Credential Management in Vehicular Communication Systems,'' *IEEE Vehicular Technology Magazine*, vol. 10, no. 4, pp. 63--69, Dec. 2015.

[74] J. Sermersheim, ''Lightweight Directory Access Protocol (LDAP),'' RFC 4511, Tech. Rep., Jun. 2006.

[75] M. Khodaei, H. Noroozi, and P. Papadimitratos, ''Scaling Pseudonymous Authentication for Large Mobile Systems,'' in *ACM WiSec*, Miami, FL, USA, May 2019, pp. 174--184.

[76] M. Khodaei and P. Papadimitratos, ''Efficient, Scalable, and Resilient Vehicle-Centric Certificate Revocation List Distribution in VANETs,'' in *ACM WiSec*, Stockholm, Sweden, Jun. 2018, pp. 172--183.

[77] ------, ''Scalable & Resilient Vehicle-Centric Certificate Revocation List Distribution in Vehicular Communication Systems,'' *IEEE Transactions on Mobile Computing (TMC)*, Mar. 2020.

[78] P. Papadimitratos and A. Jovanovic, ''GNSS-based Positioning: Attacks and Countermeasures,'' in *IEEE MILCOM*, San Diego, CA, USA, Nov. 2008.

[79] P. Papadimitratos, M. Poturalski, P. Schaller, P. Lafourcade, D. Basin, S. Capkun, and J.-P. Hubaux, ''Secure Neighborhood Discovery: A Fundamental Element for Mobile Ad Hoc Networking,'' *IEEE Communications Magazine*, vol. 46, no. 2, pp. 132--139, Feb. 2008.

[80] M. Fiore, C. E. Casetti, C.-F. Chiasserini, and P. Papadimitratos, ''Discovery and Verification of Neighbor Positions in Mobile Ad Hoc Networks,'' *IEEE Transactions on Mobile Computing (TMC)*, vol. 12, no. 2, pp. 289--303, Feb. 2013.

[81] P. Papadimitratos, L. Buttyan, J.-P. Hubaux, F. Kargl, A. Kung, and M. Raya, ''Architecture for Secure and Private Vehicular Communications,'' in *IEEE ITST*, Sophia Antipolis, Jun. 2007, pp. 1--6.

[82] E. Fonseca, A. Festag, R. Baldessari, and R. L. Aguiar, ''Support of Anonymity in VANETs - Putting Pseudonymity into Practice,'' in *IEEE Wireless Communications and Networking Conference*, Kowloon, China, Mar. 2007, pp. 3400--3405.

[83] J. Petit, F. Schaub, M. Feiri, and F. Kargl, ''Pseudonym Schemes in Vehicular Networks: A Survey,'' *IEEE Communications Surveys and Tutorials*, vol. 17, no. 1, pp. 228--255, Aug. 2014.

[84] D. Eckhoff and C. Sommer, ''Marrying Safety with Privacy: A Holistic Solution for Location Privacy in VANETs,'' in *IEEE Vehicular Networking Conference (VNC)*, Columbus, OH, USA, Dec. 2016, pp. 1--8.

[85] E. Schoch, F. Kargl, T. Leinmüller, S. Schlott, and P. Papadimitratos, ''Impact of Pseudonym Changes on Geographic Routing in VANETs,'' *European Workshop on Security and Privacy in Ad hoc and Sensor Networks (ESAS)*, pp. 43--57, Sep. 2006.

[86] A. Festag, P. Papadimitratos, and T. Tielert, ''Design and Performance of Secure Geocast for Vehicular Communication,'' *IEEE Transactions on Vehicular Technology (TVT)*, vol. 59, no. 5, pp. 2456--2471, Jun. 2010.

[87] M. Humbert, M. H. Manshaei, J. Freudiger, and J.-P. Hubaux, ''On the Optimal Placement of Mix Zones: A Game-theoretic Approach,'' in *ACM Conference on Computer and Communications Security (CCS)*, Chicago, IL, USA, Nov. 2009, pp. 324--337.

[88] P. Golle, D. Greene, and J. Staddon, ''Detecting and Correcting Malicious Data in VANETs,'' in *Proceedings of the 1st ACM international workshop on Vehicular ad hoc networks*, Philadelphia, PA, USA, Oct. 2004, pp. 29--37.

[89] B. Xiao, B. Yu, and C. Gao, ''Detection and Localization of Sybil Nodes in VANETs,'' in *Proceedings of the 2006 workshop on Dependability issues in wireless ad hoc networks and sensor networks*, Los Angeles, CA, USA, Sep. 2006, pp. 1--8.

[90] K. B. Rasmussen and S. Capkun, ''Implications of Radio Fingerprinting on the Security of Sensor Networks,'' in *IEEE International Conference on Security and Privacy in Communications Networks and the Workshops-SecureComm*, Nice, France, Jun. 2007, pp. 331--340.

[91] V. Brik, S. Banerjee, M. Gruteser, and S. Oh, ''Wireless Device Identification with Radiometric Signatures,'' in *Proceedings of the ACM international conference on Mobile computing and networking*, San Francisco, California, USA, Sep. 2008, pp. 116--127.

[92] B. Danev, D. Zanetti, and S. Capkun, ''On Physical-Layer Identification of Wireless Devices,'' *ACM Computing Surveys (CSUR)*, vol. 45, no. 1, pp. 1--29, Dec. 2012.

[93] L. Xiao, L. Greenstein, N. Mandayam, and W. Trappe, ''Fingerprints in the Ether: Using the Physical Layer for Wireless Authentication,'' in *IEEE International Conference on Communications*, Glasgow, UK, Aug. 2007, pp. 4646--4651.

[94] L. Xiao, L. J. Greenstein, N. B. Mandayam, and W. Trappe, ''Using the Physical Layer for Wireless Authentication in Time-variant Channels,'' *IEEE Transactions on Wireless Communications*, vol. 7, no. 7, pp. 2571--2579, Jul. 2008.

[95] A. Al-Momani, R. W. van der Heijden, F. Kargl, and C. Waldschmidt, ''Exploiting Propagation Effects for Authentication and Misbehavior Detection in VANETs,'' in *IEEE Vehicular Networking Conference (VNC)*, Columbus, OH, USA, Dec. 2016, pp. 1--4.

[96] C. Vaas, M. Roeschlin, P. Papadimitratos, and I. Martinovic, ''Poster: Tracking Vehicles Through Encrypted Mix-Zones Using Physical Layer Properties,'' in *IEEE Vehicular Networking Conference (VNC)*, Taipei, Taiwan, Dec. 2018, pp. 1--2.

[97] P. Papadimitratos and A. Jovanovic, ''Protection and Fundamental Vulnerability of GNSS,'' in *IEEE IWSSC*, Toulouse, France, Oct. 2008, pp. 167--171.

[98] S. Brands and D. Chaum, ''Distance-Bounding Protocols,'' in *Workshop on the Theory and Application of Cryptographic Techniques*, Berlin, Heidelberg, May 1993, pp. 344--359.

[99] S. Narain, A. Ranganathan, and G. Noubir, ''Security of GPS/INS Based On-road Location Tracking Systems,'' in *IEEE Symposium on Security and Privacy (SP)*, San Francisco, CA, USA, May 2019, pp. 587--601.

[100] K. C. Zeng, S. Liu, Y. Shu, D. Wang, H. Li, Y. Dou, G. Wang, and Y. Yang, ''All Your GPS Are Belong to Us: Towards Stealthy Manipulation of Road Navigation Systems,'' in *27th USENIX Security Symposium*, Baltimore, MD, USA, Aug. 2018, pp. 1527--1544.

[101] B. Fan, D. G. Andersen, M. Kaminsky, and M. D. Mitzenmacher, ''Cuckoo Filter: Practically Better Than Bloom,'' in *ACM CoNEXT*, Sydney, Australia, Dec. 2014, pp. 75--88.

[102] B. H. Bloom, ''Space/Time Trade-offs in Hash Coding with Allowable Errors,'' *Communications of the ACM*, vol. 13, no. 7, pp. 422--426, Jul. 1970.

[103] M. Mitzenmacher, ''Compressed Bloom Filters,'' *IEEE transactions on networking*, vol. 10, no. 5, pp. 604--612, Dec. 2002.

[104] N. Bißmeyer, ''Misbehavior Detection & Attacker Identification in Vehicular Adhoc Networks,'' Ph.D. dissertation, TU, Darmstadt, Dec. 2014.

[105] H. Noroozi, M. Khodaei, and P. Papadimitratos, ''DEMO: VPKIaaS: A Highly-Available and Dynamically-Scalable Vehicular Public-Key Infrastructure,'' in *ACM WiSec*, Stockholm, Sweden, Jun. 2018, pp. 302--304.

[106] S. Das, A. Nandan, and G. Pau, ''SPAWN: A Swarming Protocol for Vehicular Ad-Hoc Wireless Networks,'' in *ACM workshop on VANET*, Philadelphia, PA, USA, Oct. 2004, pp. 93--94.

[107] T. Dierks, ''The Transport Layer Security Protocol,'' Aug. 2008.

[108] E. Rescorla and N. Modadugu, ''Datagram Transport Layer Security V.1.2,'' Jan. 2012.

[109] M. Myers, R. Ankney, A. Malpani, S. Galperin, and C. Adams, ''X. 509 Internet Public Key Infrastructure Online Certificate Status Protocol-OCSP,'' RFC 2560, Tech. Rep., Jun. 1999.

[110] ''IEEE Standard for Wireless Access in Vehicular Environments (WAVE) - Networking Services,'' *IEEE Vehicular Technology Society*, Jan. 2016.

[111] R. M. Hinden and S. E. Deering, ''Internet Protocol Version 6 (IPv6) Addressing Architecture,'' *RFC 3513*, Apr. 2003.

[112] D. Eastlake, J. Schiller, and S. Crocker, ''Randomness Requirements for Security,'' *RFC 4086*, Jun. 2005.

[113] T. Narten, R. Draves, and S. Krishnan, ''Privacy Extensions for Stateless Address Autoconfiguration in IPv6,'' *RFC 3041*, Sep. 2007.

[114] K. Ishibashi, A. Okubo, T. Sakakura, and M. Kuroda, ''A Proposal of Fast Vertical Handover by Virtual MAC Address Scheme on Mobile Ethernet,'' in *The 13th IEEE Workshop on Local and Metropolitan Area Networks (LANMAN)*, Mill Valley, CA, USA, Oct. 2004, pp. 145--149.

[115] Y. Yao, B. Xiao, G. Wu, X. Liu, Z. Yu, K. Zhang, and X. Zhou, ''Voiceprint: A Novel Sybil Attack Detection Method Based on RSSI for VANETs,'' in *IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, Denver, CO, USA, Jun. 2017, pp. 591--602.

[116] S. So, J. Petit, and D. Starobinski, ''Physical Layer Plausibility Checks for Misbehavior Detection in V2X Networks,'' in *Proceedings of the 12th Conference on Security and Privacy in Wireless and Mobile Networks*, Miami, Florida, USA, May 2019, pp. 84--93.

[117] J. R. Douceur, ''The Sybil Attack,'' in *ACM Peer-to-peer Systems*, London, UK, Mar. 2002, pp. 251--260.

[118] PREXT: Privacy Extension for Veins VANET Simulator, https://github.com/karim-emara/PREXT, Apr. 2017.

[119] The Open Source Vehicular Network Simulation Framework, https://veins.car2x.org/, Jul. 2019.

[120] M. Behrisch, L. Bieker, J. Erdmann, and D. Krajzewicz, ''SUMO - Simulation of Urban MObility: An Overview,'' in *The 3rd International Conference on Advances in System Simulation*, Barcelona, Spain, Oct. 2011.

[121] ''Scalable Bloom Filter Implemented in Python,'' https://github.com/jaybaird/python-bloomfilter/, Mar. 2013.

[122] R. Shokri, J. Freudiger, M. Jadliwala, and J.-P. Hubaux, ''A Distortion-Based Metric for Location Privacy,'' in *Proceedings of the 8th ACM workshop on Privacy in the electronic society*, Chicago, Illinois, USA, Nov. 2009, pp. 21--30.

[123] H. Jin and P. Papadimitratos, ''Resilient Privacy Protection for Location-Based Services Through Decentralization,'' *ACM Transactions on Privacy and Security (ACM TOPS)*, vol. 22, no. 4, pp. 21:1--36, Sep. 2019.

[124] Q. Dang, *Recommendation for Applications using Approved Hash Algorithms*. US Department of Commerce, National Institute of Standards and Technology, Aug. 2015.

[125] T. Gerbet, A. Kumar, and C. Lauradoux, ''The Power of Evil Choices in Bloom Filters,'' in *IEEE/IFIP International Conference on Dependable Systems and Networks (IFIP DSN)*, Rio de Janeiro, Brazil, Jun. 2015, pp. 101--112.

[126] ''Google Cloud Platform,'' https://cloud.google.com/, Feb. 2020.

[127] J. Kenney and B. Gallagher, ''Harmful Interference to DSRC Systems,'' https://mentor.ieee.org/802.11/dcn/13/11-13-1309-00-0reg-harmful-interference-to-dsrc-systems.pptx, Nov. 2013.

Mohammad Khodaei earned his Ph.D degree from KTH Royal Institute of Technology, Stockholm, Sweden, in 2020. He is currently a postdoctoral researcher at the Networked Systems Security Group, KTH, under the supervision of Prof. Panos Papadimitratos. His research interests include security and privacy in smart cities, the Internet of Things, distributed systems, and cloud computing.

Panos Papadimitratos earned his Ph.D. degree from Cornell University, Ithaca, NY. At KTH, Stockholm, Sweden, he leads the Networked Systems Security lab, and he is a member of the steering committee of the Security Link center. He has delivered numerous invited talks, keynotes, panel addresses, and tutorials in flagship conferences. He serves or served as: Associate Editor of the IEEE TMC, the ACM/IEEE ToN and the IET IFS journals. He is a member of the PETS Editorial and Advisory Boards, and the ACM WiSec and CANS conference steering committees. He was a program chair for the ACM WiSec'16, TRUST'16 and CANS'18 conferences; a general chair for ACM WISec'18, PETS'19, and IEEE EuroS&P'19. He is a Fellow of the Young Academy of Europe, a Knut and Alice Wallenberg Academy Fellow, and an IEEE Fellow. His group web-page is: https://www.eecs.kth.se/nss.