Privacy-enhancing Interleaved Pseudonym Distribution for Vehicular Communication Systems

Keyao Huang, Hongyu Jin, and Panos Papadimitratos
Networked Systems Security Group
KTH Royal Institute of Technology
Stockholm, Sweden
{*keyao, hongyuj, papadim*}@kth.se

Abstract*—Vehicular Communication Systems (VCSs) enhance safety by enabling communication among vehicles and road-side infrastructure. To provide secure communication and privacy, a Vehicular Public-Key Infrastructure (VPKI) distributes long-term and temporary credentials (pseudonyms) to validate Cooperative Awareness Messages (CAMs) and other communication by registered vehicles. However, improperly designed pseudonym-based privacy techniques are still vulnerable to pseudonym linking based on their issuer, i.e., a specific Pseudonymous Certification Authority (PCA) that signed the pseudonyms in one batch, in the presence of PCAs. As each PCA digitally signs pseudonyms, the presence of distinct PCAs makes linking pseudonyms easier. We propose two strategies, Interleaved Pseudonym Distribution (IPD) and Random Interleaved Pseudonym Distribution (RIPD), to mitigate linking. Multi-PCA collaboration allows distributing interleaved pseudonyms (issued by different PCAs), reducing linkability. To assess our scheme privacy enhancement, a PCA_{id}-based pseudonym linking algorithm is developed utilizing a Kalman filter and road information to track vehicles. Simulation results show that IPD provides a moderate improvement in pseudonym confusion, while RIPD significantly disrupts PCA_{id}-based linking, enhancing vehicle privacy. The proof-of-concept implementation of our scheme shows the efficiency of our scheme, compared to the baseline that a single PCA issues all pseudonyms within one request.

***Index Terms*—Pseudonymous authentication, VANET, Sybil-resilient**

I. Introduction

Vehicular Communication Systems (VCSs) facilitate Vehicle-to-Everything (V2X) communication, including Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communication, to increase traffic efficiency and road safety. Safety critical Vehicular Communication (VC) security is based on public key cryptography, notably with the help of a so-called Vehicular Public-Key Infrastructure (VPKI) [1]. To enhance privacy along with strong security, the VPKI separates duties, distributing digital certificates via a Long-Term Certification Authority (LTCA) and one or more Pseudonymous Certification Authorities (PCAs) [2]. It allows each vehicle to obtain a Long-Term Certificate (LTC), a long-term credential used to request short-term anonymized credentials, pseudonyms (i.e., pseudonymous certificates) to authenticate V2V/V2I messages; including notably V2V Cooperative Awareness Messages (CAMs) and Decentralized Environmental Notification Messages (DENMs), thus making vehicle tracking harder. Vehicles use anonymized tickets obtained from the LTCA to acquire a set of pseudonyms from PCAs. The LTCA cannot link vehicle pseudonyms to their real identity, while the PCA is not able to link tickets to the real identities of vehicles.

One LTCA and multiple PCAs in a region/city are deployed to distribute certificates for vehicles [3]. The existence of multiple PCAs in a region eliminates the risk that PCA is attacked (e.g., Denial of Service (DoS)) or compromised. If there is only one PCA and it is compromised, the attacker has control of or denies the whole pseudonym issuance process. Moreover, it will be a bottleneck for pseudonym issuance. A cloud-based instantiation of the VPKI creates multiple cloned instances of PCAs to automatically load balance pseudonym requests [4]. Typically, these PCA replicas are still essentially the same PCA.

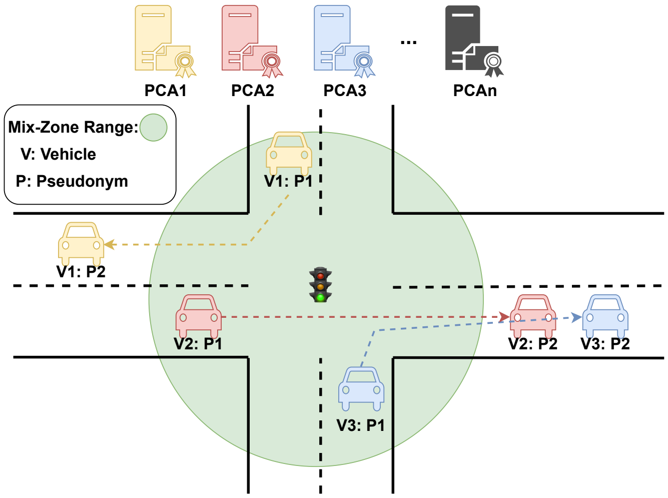

Fig. 1: Pseudonyms Linking based on PCA_{id}.

Deploying multiple PCAs is natural, e.g., multiple companies/organizations could show interest. However, this would enable pseudonym linking and distinction based on the issuing PCA identity (PCA_{id}), allowing an attacker to more easily distinguish and link pseudonyms. Fig. 1 illustrates an example of linking based on PCA_{id} even in the presence of a Cryptographic Mix Zone (CMIX) [5]. Vehicles change pseudonyms in the mix-zone, keeping the V2V messages confidential by encrypting them with the CMIX shared symmetric key. Three vehicles (V_1, V_2, V_3) acquired pseudonyms from different PCAs (PCA_1, PCA_2, PCA_3) from the VPKI. Each vehicle changes its pseudonym within the CMIX. While simultaneous pseudonym changes within the CMIX introduce ambiguity, the tracker can still identify different vehicles based on the PCA_{id} of the pseudonyms before entering and after exiting the mix-zone.

Contributions: To address the issue, we propose a VPKI pseudonym issuance approach interleaving pseudonyms from multiple PCAs: they issue pseudonyms jointly in response to a vehicle pseudonym request. Instead of having the vehicle using a set of pseudonyms with the same PCA_{id}, it now uses pseudonym with alternating PCA_{id}. With vehicles having such differing PCA_{id} patterns, returning to Fig. 1, switching to a pseudonym issued from a different PCA in the mixzone, linking pseudonyms merely based on PCA_{id} is mitigated. It is important to note however that, current CMIX pseudonym-changing strategies [5]–[7] are incompatible with ticket-based VPKI schemes that manage pseudonym lifetime so that only one pseudonym per vehicle is valid at any point in time. Our schemes addresses this by having at most two simultaneously valid pseudonyms, enabling multiple pseudonym change strategies in the VCS. Furthermore, we develop a pseudonym-linking algorithm based on PCA_{id} to evaluate our schemes. This algorithm allows the adversary to link pseudonyms in deployments with two pseudonym change strategies (changing strictly based on pseudonym lifetime and changing in CMIXs) used simultaneously.

In the rest of the paper, related work is presented in Section II. We describe the system and adversarial model, and research objectives in Section III, and present our proposed solution in Section IV. Privacy evaluation and performance evaluation are presented in Sections V and VI, respectively, before conclusions (Section VII).

II. Related Work

In VeSPA [8], the vehicle registers and obtains an encrypted ticket from the LTCA, then uses it to obtain a set of pseudonyms from the PCA to replenish its pseudonym pool. SECMACE [3] inherited conditional pseudonym disclosure by the Resolution Authority (RA) for accountability proposed in VeSPA but improved the scheme for multi-domain scenarios, including cross-domain verification. SECMACE+ [4] extended SECMACE by implementing it in a cloud environment, with many PCA instances that share the same private key created for load balancing. The pseudonym issuance overhead is distributed to multiple PCA instances, with the cloud-based PCA creating extra instances under heavy load of pseudonym issuance requests. However, relying on a single PCA, or multiple cloned instances with the same private key, may not be the natural deployment scenario, especially at VCS mature. More so, multiple PCAs could be a way to mitigate the risk of PCA compromise.

PREXT [9] is a privacy evaluation framework that assesses pseudonym-switching strategies using metrics such as anonymity set, entropy, and traceability. It employs the Nearest-Neighbor Probabilistic Data Association (NNPDA) algorithm, which utilizes a Kalman Filter (KF) for vehicle tracking [10]. However, NNPDA relies solely on location data from CAMs to link pseudonyms and does not consider the influence of PCA-signed pseudonyms on vehicle privacy. Our linking algorithm adapts PREXT by integrating with PCA_{id}-based linking for the adversary.

III. System and Adversarial Models

System model: We assume that each domain, for example, a city/region, has one LTCA that registers vehicles and issues LTCs and tickets, and multiple PCAs. Communication between VPKI entities is end-to-end protected by Transport Layer Security (TLS) channels, ensuring data confidentiality, authenticity, and integrity. Certification Authorities (CAs) in the system comply with system protocols. Each vehicle is equipped with an On-Board Unit (OBU) for message communication and a Hardware Security Module (HSM) to prevent private key extraction. Additionally, only one pseudonym can be active at each time to prevent Sybil-based attacks.

Adversary model: We assume a global external eavesdropper (tracker), capable of eavesdropping all CAMs transmitted in the VCS by deploying eavesdropping devices. The tracker also possesses detailed knowledge of city layout, the road network, and the lane connectivity. The eavesdropping devices collect CAMs in real time and transmit them to the tracker, which analyses the collected CAMs after a period of time (e.g., one hour) and mounts a pseudonym linking attack to infer the journeys of different vehicles. The primary purpose of the tracker is to link pseudonyms and track vehicles despite pseudonym changes, by leveraging information from CAMs, including location, speed, and the PCA_{id}.

IV. Our Scheme

A. Overview

The system consists of multiple PCAs, each managed by different operators. Each independent PCA possesses a distinct ECDSA public-private key pair. The vehicle first obtains a ticket from the LTCA by authenticating itself with LTC, then it requests pseudonyms with the ticket. The interactions with the VPKI are achieved through an Access Point (AP), a Base Station (BS), or a Roadside Unit (RSU). These requests are distributed to different PCAs for load-balancing. To enhance privacy and mitigate PCA_{id}-based pseudonym linking, we propose Interleaved Pseudonym Distribution (IPD) and its improved version, Random Interleaved Pseudonym Distribution (RIPD). These schemes introduce multi-PCA collaboration to make linking of pseudonyms based on PCA_{id} harder.

Interleaved Pseudonym Distribution (IPD): IPD allows a vehicle to acquire a batch of pseudonyms from n target PCAs. A vehicle chooses target PCAs to request pseudonyms from, to which guarantees that consecutive pseudonyms have different PCA_{id}, increasing the difficulty of linking pseudonyms by breaking PCA_{id} continuity. Fig. 2 illustrates the IPD. In this example, the vehicle requests 20 pseudonyms over a time period Γ, and the lifetime of each pseudonym (τ) overlaps with

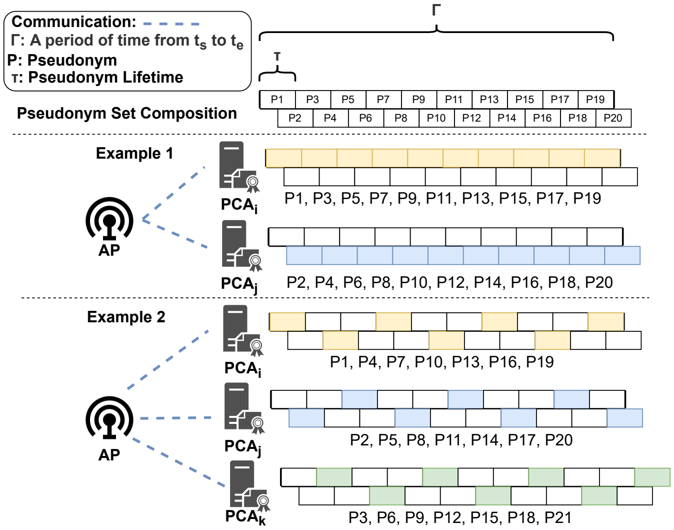
Communication:
Γ: A period of time from t_s to t_e
P: Pseudonym
τ: Pseudonym Lifetime
Γ
τ
Pseudonym Set Composition
P1 P3 P5 P7 P9 P11 P13 P15 P17 P19
P2 P4 P6 P8 P10 P12 P14 P16 P18 P20
Example 1
AP
PCA_i
P1, P3, P5, P7, P9, P11, P13, P15, P17, P19
PCA_j
P2, P4, P6, P8, P10, P12, P14, P16, P18, P20
Example 2
AP
PCA_i
P1, P4, P7, P10, P13, P16, P19
PCA_j
P2, P5, P8, P11, P14, P17, P20
PCA_k
P3, P6, P9, P12, P15, P18, P21

Fig. 2: IPD with two and three (Example 1 and 2) target PCAs.

that of the adjacent pseudonym, ensuring flexible pseudonym changing at any time, compared to strictly non-overlapping pseudonyms. Example 1 shows that two target PCAs (PCA_i and PCA_j) are chosen to issue interleaved pseudonyms, while the vehicle obtains interleaved pseudonyms from three target PCAs (PCA_i, PCA_j and PCA_k) in Example 2.

Compared to state-of-the-art ticket-based VPKI schemes, IPD allows vehicles to obtain pseudonyms from multiple target PCAs in a strictly alternating order, to disrupt the tracker's ability to link a set of pseudonyms based on fixed PCA_{id}. However, the predictability of PCA_{id} sequences becomes an obvious limitation. Since the pseudonyms have a fixed alternating PCA_{id}, the tracker can infer the PCA_{id} of the follow-up pseudonyms based on the current PCA_{id}. For instance, assuming that one vehicle alternates between PCA_i and PCA_j, the tracker can predict that the next pseudonym of the said vehicle is issued by PCA_i observing the current pseudonym is issued by PCA_j.

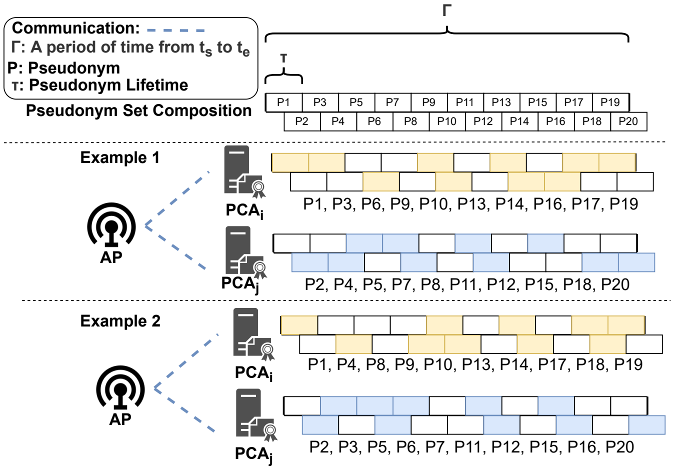
Communication:
Γ: A period of time from t_s to t_e
P: Pseudonym
τ: Pseudonym Lifetime
Γ
τ
Pseudonym Set Composition
P1 P3 P5 P7 P9 P11 P13 P15 P17 P19
P2 P4 P6 P8 P10 P12 P14 P16 P18 P20
Example 1
AP
PCA_i
P1, P3, P6, P9, P10, P13, P14, P16, P17, P19
PCA_j
P2, P4, P5, P7, P8, P11, P12, P15, P18, P20
Example 2
AP
PCA_i
P1, P4, P8, P9, P10, P13, P14, P17, P18, P19
PCA_j
P2, P3, P5, P6, P7, P11, P12, P15, P16, P20

Fig. 3: RIPD with two target PCAs

Random Interleaved Pseudonym Distribution (RIPD): RIPD extends IPD by introducing a randomized order for pseudonym acquisition, disrupting the predictability of PCA_{id} sequences. As illustrated in Fig. 3, a vehicle generates two shuffled subsets, each containing ten elements, and sends them to two target PCAs. Each PCA then issues pseudonyms according to its assigned subset, resulting in non-sequential pseudonym lifetimes. By eliminating the strict alternation of PCA_{id}, RIPD prevents the tracker from reliably predicting the next pseudonym PCA_{id}, even when a vehicle acquires pseudonyms from only two PCAs. This increased uncertainty hinders the tracker linking pseudonyms accurately. In addition, compared to [11], which requires the vehicle to swap the private key of the pseudonym with the nearby vehicle to obscure the tracking, RIPD achieves the same goal without any key exchange, thereby eliminating the associated security risks.

B. Protocols

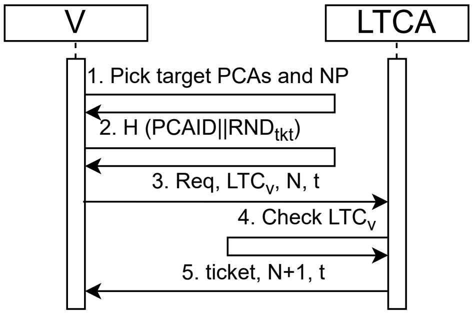

Fig. 4: Ticket Acquisition Process

Ticket Acquisition (Fig. 4): After obtaining the LTC, the vehicle, V, needs to select the number of pseudonyms (NP) as well as the n PCAs it trusts and decides to obtain pseudonyms from, as target PCAs ($NP \bmod n = 0$). Subsequently, the vehicle hashes the target PCA_{id} with a random number (Rnd_{tkt}) to hide the target PCA information from the LTCA. The Req sent by the vehicle to the LTCA is: $Req = \{Id_{req}, NP, PCAIDList, t_s, t_e\}$, in which $PCAIDList = \{H(PCAID_i||Rnd_{tkt}), ..., H(PCAID_n||Rnd_{tkt})\}$. After verifying LTC_v, the LTCA generates the ticket ($ticket = \{NP, PCAIDList, CK_{n-tkt}, t_s, t_e, t_{expire}\}$), in which $CK_{n-tkt} = H(LTC_v||t_s||t_e||Rnd_{CK_{n-tkt}})$. The ticket is signed and returned to the vehicle.

Pseudonym Acquisition (Protocol 1 and Protocol 2): The vehicle generates public-private key pairs based on the predetermined number of pseudonyms (NP) and target PCAs (n), embedding them into its pseudonym request(s), then sending to the respective target PCAs. If the RIPD scheme is used, the vehicle shuffles the pseudonym sequence subsets; otherwise, it follows a sequential interleaving pattern. The pseudonym request includes the set of public keys $\{\alpha_1, ..., \alpha_n\}$ and each the subset of the pseudonym sequence indices $\{\beta_1, ..., \beta_n\}$.

Upon receiving the request, each target PCA verifies whether it is included in the target PCAs, checks the requested

Protocol 1 Vehicle Generates Pseudonym Request

```
$m \leftarrow NP/n$
for $i \in (1, n)$ do
    $(K_v^1, \cdots, K_v^m) \leftarrow GenKeyPairs(m)$
    $\alpha_i \leftarrow \{(K_v^1, \ldots, K_v^m)_{\sigma_{k_v^1}}, \ldots, (K_v^1, \ldots, K_v^m)_{\sigma_{k_v^m}}\}$
end for
if Scheme = IPD then
    $(\beta_1, ..., \beta_n) \leftarrow GenSeq(NP, n)$
else if Scheme = RIPD then
    $(\beta_1, ..., \beta_n) \leftarrow GenRndSeq(NP, n)$
end if
return $(Rnd_{tkt}, (n - tkt)_{\sigma_{LTCA}}, (\alpha_1, ..., \alpha_n), (\beta_1, ..., \beta_n), N, t_{now})$
```

pseudonym count, and validates the public key by verifying the signatures of the public key sets. After that, it assigns the start and end time (t_s and t_e) to each pseudonym based on its position in the sequence set (β_i) and the lifetime τ_p. A hash chain enhances security by facilitating efficient pseudonym revocation and preventing a compromised RA from repeatedly resolving pseudonyms from the same vehicle [3]. Finally, the PCA signs and returns the set of pseudonyms to the vehicle.

Protocol 2 Pseudonyms Issuance by PCA

```
for $i \in (1, n)$ do
    $H(PCAID_{this}||Rnd_{tkt}) \overset{?}{=} H(PCAID_i||Rnd_{tkt})$
end for
$m \leftarrow size(\alpha_i)$
if $m \neq (NP/n)$ then
    Break
end if
$NP, t_s, t_e, CK_{tkt} \leftarrow Extract(ticket)$
for $i = 1$ to $m$ do
    $Verify(K_v^i, (K_v^1, \ldots, K_v^m)_{\sigma_{k_v^i}})$
end for
for $i \in \beta_i$ do
    $t_s^i \leftarrow t_s + ((i - 1) \times \tau_p)/2$
    $t_e^i \leftarrow t_s^i + \tau_p$
    $CK_{P_v^i} \leftarrow H(CK_{tkt}||K_v^i||t_s^i||t_e^i||H'(H^i(Rnd_p)))$
    if $i$ is the first element in $\beta_i$ then
        $SN^i \leftarrow H(CK_{P_v^i}||H'(H^i(Rnd_p)))$
    else
        $SN^i \leftarrow H(SN^{i-1}||H'(H^i(Rnd_p)))$
    end if
    $\xi \leftarrow (SN^i, K_v^i, CK_{P_v^i}, t_s^i, t_e^i)$
    $(P_v^i)_{\sigma_{this\_PCA}} \leftarrow Sign(Lk_{this\_PCA}, \xi)$
end for
return $(\{(P_v^1)_{\sigma_{PCA_{this}}}, \ldots, (P_v^n)_{\sigma_{PCA_{this}}})\}, Rnd_p, N + 1, t_{now})$
```

Certificates Resolution: The process of resolving the pseudonym and the ticket can align with the protocols described in [4], allowing the resolution of the actual, long-term identity of a vehicle (LTC_v).

Certificates Revocation (Protocol 3): The RA sends the ticket with its signature to the LTCA, which adds the LTC bound on the ticket to the Certificate Revocation List (CRL) and then sends an acknowledgment (ACK) to the RA, whcih then generates a pseudonym revocation request that includes the resolved suspected pseudonym (P_v^i) and the ticket ($n-tkt$), used by the PCA to trace issued pseudonyms. The RA sends the request to the PCA that issued the specific pseudonym ($PCA_{P_v^i}$). Based on $n - tkt$, the PCA searches the local database for all pseudonyms generated by that ticket and adds them to the CRL. The PCA then extracts Rnd_{tkt} from the suspicious pseudonym (P_v^i) and returns it to the RA, which needs to locate other target PCAs with Rnd_{tkt}. The RA locates other target PCAs by looping through the IDs of all PCAs within the domain, performing a hash operation, comparing the result with the target PCAs in the ticket, and subsequently sending the pseudonym revocation request to those target PCAs, which in turn perform the aforementioned operations to revoke the pseudonyms they issued.

Protocol 3 LTC and Pseudonym Revocation

```
$RA : \xi \leftarrow (n - tkt, N, t_{now})$
$RA \rightarrow LTCA : (\xi)_{\sigma_{RA}}$
$LTCA : (LTC_v) \leftarrow Resolve(n - tkt)$
$LTCA : AddToCRL(LTC_v)$
$LTCA : \delta \leftarrow (ACK, N + 1, t_{now})$
$LTCA \rightarrow RA : (\delta)_{\sigma_{LTCA}}$
$RA : \xi \leftarrow (n - tkt, P_v^i, N, t_{now})$
$RA \rightarrow PCA_{P_v^i} : (\xi)_{\sigma_{RA}}$
$PCA_{P_v^i} : (P_v^i, \ldots, P_v^n) \leftarrow SearchPsnyms(n - tkt)$
$PCA_{P_v^i} : AddToCRL((P_v^i, \ldots, P_v^n))$
$PCA_{P_v^i} : Rnd_{tkt} \leftarrow Extract(P_v^i)$
$PCA_{P_v^i} : \delta \leftarrow (Rnd_{tkt}, N + 1, t_{now})$
$PCA_{P_v^i} \rightarrow RA : (\delta)_{\sigma_{PCA_{P_v^i}}}$
$RA : \{$
for $PCA_x$ in $PCA_{in\_domain}$ do
    Find all target PCAs
end for
for $PCA_x$ in (target PCAs) do
    if $PCA_x \neq PCA_{P_v^i}$ then
        $\gamma \leftarrow (n - tkt, N, t_{now})$
        $RA \rightarrow PCA_x : \gamma_{\sigma_{RA}}$
    end if
end for$\}$
$PCA_x : \{$
$(P_v^i, \ldots, P_v^n) \leftarrow SearchPsnyms(n - tkt)$
$AddToCRL((P_v^i, \ldots, P_v^n))$
$\delta \leftarrow (ACK, N + 1, t_{now})$
$PCA_x \rightarrow RA : (\delta)_{\sigma_{PCA_x}}\ \}$
```

V. Privacy Evaluation

The experiment aims to evaluate whether the proposed schemes improve privacy protection compared to the existing Sybil-resilient ticket-based scheme. We define two metrics (Tracking Coverage and Pseudonym Confusion Degree) to evaluate the performance, which is used to link vehicle pseudonyms across the Baseline, IPD, and RIPD schemes. In the baseline scheme, a single PCA issues all pseudonyms within one pseudonym request.

A. Tracking Algorithm

The scenario considers two pseudonym-switching situations: triggered by pseudonym expiration outside CMIX and changing within the CMIX. The pseudonym linking algorithm for the former is based on KF-based location prediction and the linking algorithm for the latter is based on road topology. The algorithm operates in three modes, corresponding to the baseline, IPD, and RIPD schemes: **Mode 1 (Baseline):** Vehicles maintain a consistent PCA_{id}, so the tracker only considers new pseudonyms with $PCAID_{new} = PCAID_{old}$. **Mode 2 (IPD):** PCA_{id}s alternate sequentially ($PCAID_i, PCAID_j$). When a pseudonym with $PCAID_i$ disappears, the tracker searches for its successor with $PCAID_j$. **Mode 3 (RIPD):** PCA_{id}s randomly alternate. Therefore, the tracker maintains a record of all associated PCA_{id}s, and it has to consider all newly appeared pseudonyms with both $PCAID_i$ and $PCAID_j$ as candidates.

For pseudonym changes outside CMIXs, KF predicts the target's next position Pos_{predic} and velocity $speed_{predic}$. The tracker evaluates differences Pos_{diff} and $speed_{diff}$ between predicted and newly appeared pseudonyms. A score-based system confirms a match when the score exceeds a threshold and is the highest among candidates. For pseudonym changes within CMIXs, the tracker utilizes lane information $(LaneID_1, LaneID_2)$ to determine whether there is a valid path between them [7]. A distance threshold ensures that both pseudonyms originate near the same CMIX, improving the tracking accuracy.

B. Metrics

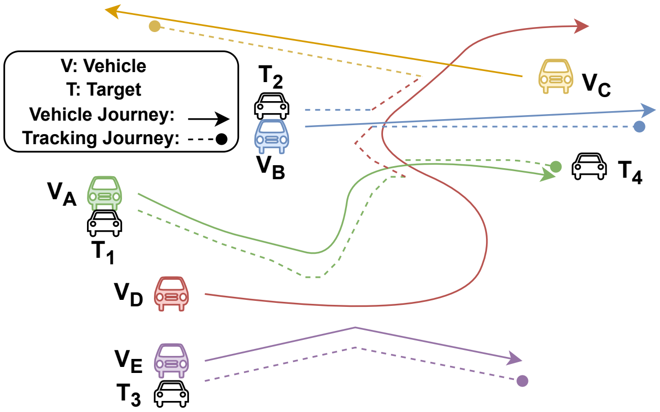

Fig. 5: Vehicle Journey and Tracking Journey

Tracking Coverage: This metric is used to measure the effectiveness of tracking vehicles; the better the tracking, the less protective the scheme is of the vehicle's privacy. To illustrate this, in Fig. 5, Target 1 initially tracks $Vehicle_A$, but due to pseudonym changes in CMIX, it mistakenly links to $Vehicle_D$ and later to $Vehicle_B$. As a result, Target 1 contains partial journeys of three different vehicles. When Vehicle A reappears with new pseudonym, the tracker starts a new tracking target (Target 4), failing to link it to Target 1. This discontinuity enhances privacy by preventing full trip reconstruction.

Privacy depends on the extent the tracker monitors vehicle trips. Fully tracking all trips of a vehicle (for example, Target 3 completely tracks $Vehicle_E$'s trips in Fig. 5), driver behavior patterns would emerge.

To quantify tracking effectiveness, we define **Correct Pseudonyms (CP)** as the pseudonyms of $Vehicle_i$ correctly linked by its **Main Target** (the target covering the longest segment of its journey). **Total Pseudonyms (TP)** represents all pseudonyms used by the vehicle. **Total Pseudonyms in the Main Target (TPMT)** refers to all pseudonyms recorded in the main target, including potential mis-linked pseudonyms belonging to other vehicles. Therefore, *Tracking Coverage* is defined as: $Coverage_i = \left(\frac{CP_i}{TP_i}\right) \times \left(\frac{CP_i}{TPMT_i}\right)$.

TABLE I: Parameters for Privacy Experiment

Parameter	Value
Beacon Interval	1s
TX Power	20mW
Sensitivity	-89dBm
Carrier Frequency	5.89 GHz
Physical Layer Bit-rate	18Mbps
Area Size	14KM x 12KM
Number of Eavesdroppers	2,247
Number of CMIXs	100
CMIX Coverage Range	50m
CMIX Advertisement Interval	3s
Duration of Simulation	1 hour (16:00-17:00)
Number of PCAs	3,5,7,11,15
Number of Target PCAs	2
Percentage of Vehicles	40%
Number of Vehicles	6321

$\frac{CP_i}{TP_i}$ measures how much of the journey of Vehicle i is continuously tracked, while $\frac{CP_i}{TPMT_i}$ evaluates the proportion of correctly assigned pseudonyms in the target. These two submetrics both contribute to the confusion of the tracking performance. Therefore, multiplication is used to indicate accuracy. We calculate the average value of tracking coverage for all vehicles as the final metric.

Pseudonym Confusion Degree (PCD): This metric is defined to assess the effectiveness of the tracker in identifying vehicle pseudonym changes and determining the success rate of pseudonym linking: $PCD = 1 - (\frac{OPCE}{TPCE} \times \frac{SL}{OPCE})$.

Pseudonym Change Event (PCE): A PCE represents a pseudonym change by a vehicle. **Total PCE (TPCE)** presents the total number of pseudonym change events that occurred during the experiment, while **Observed PCE (OPCE)** represents the number of PCEs detected by the tracker. The PCE is considered unobserved if a newly appeared pseudonym is mistakenly assigned to a new vehicle. **Successful Link (SL)** presents the number of correct pseudonym linking.

The fraction $\frac{OPCE}{TPCE}$ quantifies how many pseudonym changes the tracker detects, while $\frac{SL}{OPCE}$ measures its precision in linking detected events. Their product represents the tracker's effectiveness, with values closer to 1 indicating better tracking accuracy. PCD, ranging from 0 to 1, reflects the system's confusion: Higher values imply greater uncertainty in tracking pseudonym changes.

C. Experimental Setup

The experiments were conducted using OMNeT++ with the Veins framework [12] for large-scale traffic simulations. The Luxembourg SUMO Traffic (LuST) dataset [13] provided real-world traffic flow data. To simulate tracking challenges, 2,247 eavesdropping devices were placed at intersections, with 100 intersections randomly selected for CMIX deployment (each with a 50 m range). Vehicles and eavesdroppers had a communication range of 300 m. The simulation ran from 16:00 to 17:00. The proposed IPD and RIPD schemes allowed vehicles to select multiple target PCAs. Based on the same simulation, 40% of the vehicles were extracted for further

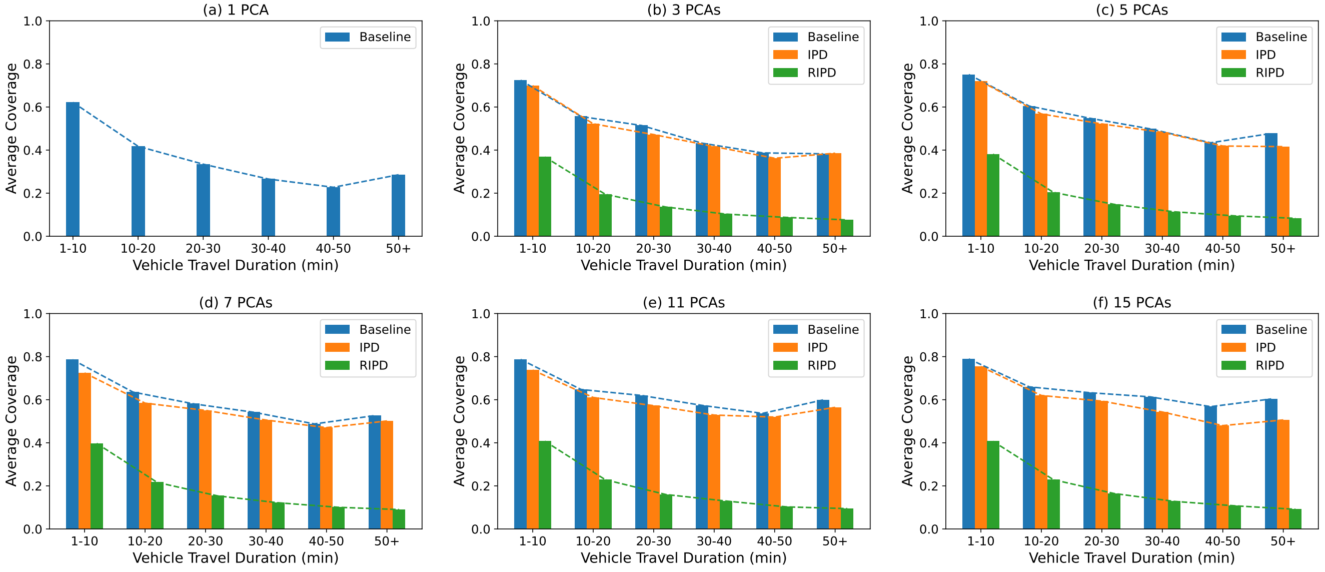

Fig. 6: Average coverage under Baseline, IPD, and RIPD schemes across 1, 3, 5, 7, 11, and 15 PCA configurations.

analysis, reflecting a realistic scenario where not all vehicles participate. Table I outlines all parameters.

D. Results

Tracking Coverage: Fig. 6 shows the average coverage for various groups of vehicle travel durations in different scenarios. The horizontal axis of each subplot is grouped on the basis of the duration of the vehicle's travel. For instance, the group "10-20" indicates that the vehicle used 10-20 pseudonyms during the experiment. The lifetime of pseudonyms is 60 seconds, so this group of vehicles corresponds to a travel time of 10-20 minutes in the simulation. The results indicate that IPD exhibits a slight reduction in terms of average coverage across all groups compared to the Baseline scheme, with a decrease of approximately 5%. RIPD demonstrates a substantial reduction in average coverage in all groups, approximately 30% lower than both Baseline and IPD. In particular, the coverage value decreases significantly as the duration of vehicle travel increases, which reflects the reduced tracking performance of the tracker. For vehicles traveling more than 50 minutes, the coverage drops to around 9%, indicating that higher privacy protection is offered by RIPD.

Pseudonym Confusion Degree (PCD): Fig. 7 illustrates the Pseudonym Confusion Degree (PCD) for the Baseline, IPD, and RIPD schemes across varying quantities of accessible PCAs. In all circumstances, the order of privacy protection performance remains unchanged: RIPD surpasses IPD and Baseline, indicating that RIPD generates the greatest degree of pseudonym confusion, hence complicating the tracker's ability to associate pseudonyms and sustain continuous vehicle tracking. The findings demonstrate that IPD enhances PCD by around 0.05 relative to the Baseline, but RIPD elevates it by approximately 0.4 compared to IPD and 0.45 compared to the Baseline. This underscores the RIPD exceptional capacity to thwart tracking. It is interesting that PCD decreases as the number of PCAs grows, regardless of the scheme utilized. For instance, Baseline PCD declines from 0.1307 (3 PCAs) to 0.0960 (15 PCAs), while IPD reduces from 0.1937 to 0.1614, and RIPD diminishes from 0.6231 to 0.5780. This drop arises as an increased number of PCA_{id}s diminishes the tracker's search space, resulting in fewer erroneous connections. In summary, RIPD retains a significant advantage in safeguarding vehicle privacy across all PCA setups.

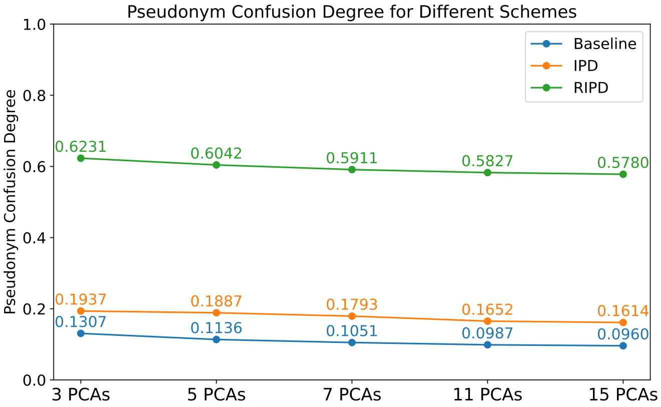

Fig. 7: Pseudonym Confusion Degree of Three Schemes

Overall Average Coverage: Fig. 8 shows the overall average coverage of the Baseline, IPD, and RIPD schemes across varying quantities of PCAs. The single-PCA case is presented, wherein all pseudonyms possess the identical PCA_{id}, rendering PCA_{id}-based linking attack ineffectual. Due to the necessity of several PCAs for IPD and RIPD, these methods are inapplicable in this instance, and only the Baseline scheme is evaluated. As the number of PCAs goes

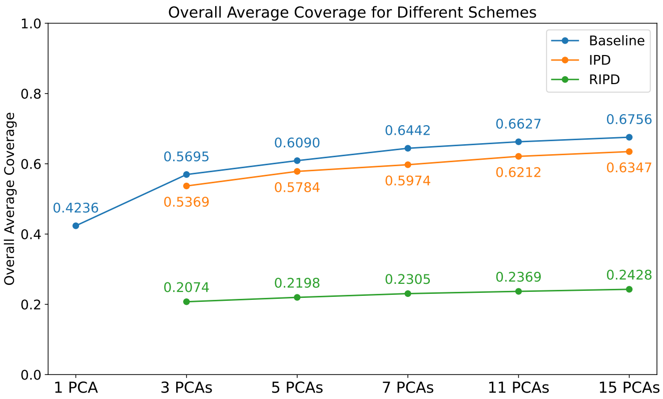

Fig. 8: Overall Average Coverage of Three Schemes

TABLE II: Delay for Obtaining 20 Pseudonyms.

Scheme	Mean	Median	SD
Baseline	0.62s	0.59s	0.16s
IPD	0.43s	0.46s	0.15s
RIPD	0.61s	0.57s	0.14s

from 1 to 3, the Baseline scheme average coverage climbs from 0.4236 to 0.5695, thus a reduction in vehicle privacy. The existence of several PCAs generates additional PCA_{id}s, hence enhancing the tracker's capacity to associate pseudonyms and monitor vehicles. A similar pattern is noted for IPD, which outperforms Baseline marginally while still demonstrating an increase in average coverage with additional PCAs, rising from 0.5369 (3 PCAs) to 0.6347 (15 PCAs). Nevertheless, RIPD has inferior average coverage across all cases. Despite a minor increase in coverage from 0.2074 (3 PCAs) to 0.2428 (15 PCAs), the increase is far lower than that for the other two schemes, thanks to the RIPD randomized PCA_{id} selection that makes forecasting the PCA_{id} of the next pseudonym harder. The uniformity of the RIPD performance across different PCA setups demonstrates its resilience to PCA_{id}-based linking attacks, offering superior privacy protection compared to Baseline and IPD.

VI. Overhead Evaluation

Our scheme inherits the ticket issuance process from earlier VPKI design [3], [4]. The experiment measures the time delay between the generation of pseudonym requests and pseudonym issuance by the PCA(s). The experiment implements the core pseudonym issuance process for the Baseline, IPD, and RIPD schemes using Python, including key generation, pseudonym request, and pseudonym response. Two Virtual Machines (VMs) are assigned dual-core 2.3 GHz CPUs and 4GB RAM to emulate PCAs. In the Baseline scheme setting, a single VM was utilized, as vehicles acquired pseudonyms from a single PCA. Two VMs settings emulate IPD and RIPD. The experiment evaluates the delay to obtain 20 pseudonyms for each scheme, averaged over 10 experiments.

Table II average the delay measured for the three schemes, baseline, IPD, and RIPD, in acquiring 20 pseudonyms. The findings indicate that the performance of the proposed IPD and RIPD are comparable to the baseline scheme. The average latencies for the three systems are 0.62s, 0.43s, and 0.61s, respectively. IPD and RIPD do not degrade the efficiency of pseudonym issuance while improving user privacy.

VII. Conclusion

We propose IPD and RIPD, scalable frameworks that enable multiple PCAs to collaborate in pseudonym distribution to mitigate pseudonym linking based on PCA_{id}. Compared to IPD issuing pseudonyms by chosen PCAs in a strictly alternating order, RIPD further enhances privacy by introducing randomness, allowing vehicles to obtain pseudonyms in an unpredictable order from multiple PCAs. This disrupts the tracker ability to anticipate the next pseudonym's PCA_{id}, increasing tracking uncertainty. Experimental results show that RIPD significantly raises pseudonym confusion and reduces the tracking coverage compared to the existing Sybil-resilient ticket-based VPKI schemes. Additionally, efficiency tests show that RIPD enhances privacy protection without sacrificing pseudonym issuance performance (latency).

Acknowledgment

Work supported by the Swedish Research Council (VR).

References

[1] T. Leinmüller, L. Buttyan, J.-P. Hubaux, F. Kargl, R. Kroh, P. Papadimitratos, M. Raya, and E. Schoch, "Sevecom-secure vehicle communication," in *IST mobile and wireless communication summit*, 2006.

[2] S. Khan, F. Luo, Z. Zhang, M. A. Rahim, M. Ahmad, and K. Wu, "Survey on issues and recent advances in vehicular public-key infrastructure (VPKI)," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 3, pp. 1574–1601, 2022.

[3] 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*, vol. 19, no. 5, pp. 1430–1444, 2018.

[4] M. Khodaei, H. Noroozi, and P. Papadimitratos, "SECMACE+: Upscaling Pseudonymous Authentication for Large Mobile Systems," *IEEE Transactions on Cloud Computing*, vol. 11, no. 3, pp. 3009–3026, 2023.

[5] J. Freudiger, M. Raya, M. Félegyházi, and P. Papadimitratos, "Mix-zones for Location Privacy in Vehicular Networks," in *ACM WiN-ITS*, Vancouver, British Columbia, Canada, August 2007.

[6] 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)*, 2018.

[7] M. Khodaei and P. Papadimitratos, "Cooperative location privacy in vehicular networks: Why simple mix zones are not enough," *IEEE Internet of Things Journal*, vol. 8, no. 10, pp. 7985–8004, 2020.

[8] N. Alexiou, M. Laganà, S. Gisdakis, M. Khodaei, and P. Papadimitratos, "Vespa: Vehicular security and privacy-preserving architecture," in *ACM workshop on Hot topics on wireless network security and privacy*, 2013.

[9] K. Emara, "Poster: Prext: Privacy extension for veins vanet simulator," in *IEEE Vehicular Networking Conference (VNC)*, 2016.

[10] K. Emara, W. Woerndl, and J. Schlichter, "Vehicle tracking using vehicular network beacons," in *IEEE WoWMoM*, 2013.

[11] A. P. Mdee, M. T. R. Khan, J. Seo, and D. Kim, "Security compliant and cooperative pseudonyms swapping for location privacy preservation in vanets," *IEEE TVT*, vol. 72, no. 8, pp. 10 710–10 723, 2023.

[12] C. Sommer, R. German, and F. Dressler, "Bidirectionally Coupled Network and Road Traffic Simulation for Improved IVC Analysis," *IEEE Transactions on Mobile Computing*, vol. 10, no. 1, pp. 3–15, 2011.

[13] 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*, Dec. 2015.