2019-01-0487 Published 02 Apr 2019

Evaluating Trajectory Privacy in Autonomous Vehicular Communications

Abdelnasser Banihani Oakland University

Abdulrahman Zaiter and George P. Corser Saginaw Valley State University

Huirong Fu and Abdulrahman Alzahrani Oakland University

Citation: Banihani, A., Zaiter, A., Corser, G.P., Fu, H. et al., "Evaluating Trajectory Privacy in Autonomous Vehicular Communications," SAE Technical Paper 2019-01-0487, 2019, doi:10.4271/2019-01-0487.

Abstract

Autonomous vehicles might one day be able to implement privacy preserving driving patterns which humans may find too difficult to implement. In order to measure the difference between location privacy achieved by humans versus location privacy achieved by autonomous vehicles, this paper measures privacy as trajectory anonymity, as opposed to single location privacy or continuous privacy. This paper evaluates how trajectory privacy for randomized driving patterns could be twice as effective for autonomous vehicles using diverted paths compared to Google Map API generated shortest paths. The result shows vehicles mobility patterns could impact trajectory and location privacy. Moreover, the results show that the proposed metric outperforms both K-anonymity and KDT-anonymity.

Introduction

Privacy as a citizen's right has confounded consumers and legislators, more and more as technology seems to chip away at personal anonymity. Now, in the dawning of the age of Autonomous Vehicles (AVs), engineers and business people may soon have to grapple with privacy as a quantifiable, marketable feature of automotive products. It is possible that one day automotive manufacturers may have to compete not only on price and gas mileage, but also on the degree to which vehicles protect users' information. In the case of vehicles, a valuable piece of information is location. Laws have been proposed to protect drivers' *location privacy*. But, how location privacy is measured? More importantly, how should we measure location privacy in AVs? If two manufacturers claim to protect drivers' anonymity, can there be any way for consumers to determine a measurable difference between their products? Let's call this the *autonomous vehicle location privacy* (AVLP) measurement problem.

Location Based Services (LBSs) rely on users' or vehicles' locations to provide the requested services. However, adversaries can exploit the LBSs, such as navigation systems, to locate a vehicle either by pinpointing the location of a vehicle in real time or by establishing its position at a given time in the past. Thus, without privacy protocols in place, the vehicles' location privacy could be jeopardized.

We assume AVs use the LBS for the same reasons humans do today. Therefore, AVs may be equally vulnerable to privacy attacks. Internet of Things (IoT) and Vehicular ad Hoc Networks (VANETs) may also rely on the LBS to provide a variety of services, including navigation and traffic management, with increasing speed, frequency and precision, and without human intervention. Wireless eavesdroppers, malicious the LBS administrators, or hackers could track specific vehicles by connecting vehicles trajectories to a specific vehicle. Car manufacturers might be concerned about the misuse of their customers' data as such a breach that might hurt a company's reputation.

In addition, there may be legal reasons to implement location privacy. European Union (EU) mandates car manufacturers preserve privacy. Whereas, the US Legislators have recognized the problem of location privacy and expected to follow EU by mandating location privacy. The US Legislators proposed national level legislation in the United States to address digital location privacy threats includes the Location Privacy Act and the Geolocation Privacy and Surveillance Act. Privacy attacks could be perpetrated by a spouse, paparazzi, and other stalkers to track and exploit vehicles' locations.

Location privacy measurement is complex. It is necessary to consider the protocols involved. If the LBS uses Dedicated Short Range Communications (DSRC), then it is possible to implement Pseudo-Identifiers (Pseudo-IDs) for each car. Different approaches have been proposed to measure location privacy, such as snapshot (point in time) privacy or continuous location privacy [1, 2]. Sometimes, distance between anonymized vehicles is a factor [2]. Sometimes, the time or duration that a vehicle is able to maintain its anonymity is considered. In almost all cases, Pseudo-IDs and mix zones are used to achieve K-anonymity. That is, the attacker may know that a car is active, but the attacker does not know which car is the target car among *k* cars in an anonymity set.

AVs might use preplanned paths to travel from a source to a destination, or a central the LBS might compute shortest paths for AVs. These paths may be more predictable than paths chosen by human drivers [4, 6]. However, AVs could implement diversion tactics, redirecting in mix zones, tactics which humans might find too difficult to execute. In order to measure the difference, a clear definition and metric for trajectory privacy is needed. The main proposal of this paper is to break vehicle trajectories into multiple sub-trajectories by implementing Pseudo-IDs and mix zones, then evaluate privacy protections with the new metrics.

Briefly, this paper makes the following contributions to the ongoing field of location privacy:

- Proposes a new definition for trajectory privacy.
- Presents trajectory-anonymity metric to evaluate trajectories privacy, which precisely evaluates trajectory privacy and shows existing metrics insufficiency.
- Derives formulas to calculate theoretical values of *trajectory-anonymity*.
- Evaluates trajectory-anonymity with existing metrics including K-anonymity and KDT-anonymity.

The remainder of the paper is organized as follows. The Background section states the assumptions on which this work is built and specifies the protocol stack, system model, and threat model under investigation. It also provides an overview of Autonomous Vehicles. Then, the Location Privacy Protection Techniques section demonstrates different techniques to preserve trajectory privacy. Next, the Proposed Metric section illustrates the difference between shortest path and diverted route trajectory privacy under different conditions. After that, the Performance Analysis section shows graphs of privacy levels achieved. Finally, the Conclusion and Future Works section concludes this paper and provides future directions.

Background

This work discusses the trajectory privacy problem by showing the case where VANETs and AVs work together to achieve the desired goals of transportations in the future. A brief overview of AVs and VANETs is discussed below. All vehicles in VANETs are able to communicate with each other by Vehicle-to-Vehicle (V2V) communication. Another kind of communication is called vehicle-to-Infrastructure (V2I), where vehicles communicate with other entities to access Internet applications like infotainment or the LBSs. To achieve that, vehicles send messages to an intermediate node that is responsible to convey messages between vehicles and other entities, called Road Side Unit (RSU).

A. Protocol Stack

This paper assumes that the communications between vehicles are done by using DSRC. Fig. 1 illustrates the DSRC protocol stack, the FCC dedicates a 75 MHz spectrum in the 5.9 GHz band for wireless communication between vehicles. IEEE and SAE have established standards, DSRC/WAVE, to achieve interoperability between devices communicating in this spectrum. The protocol stack features two distinct protocol sets. IPv6/TCP/UDP typically would be used in V2I communication. WAVE Short Message Protocol (WSMP), would typically be used in V2V communication, such as safety applications. This paper assumes the use of WSMP using security protocols as defined in IEEE 1609.2 and message types defined in SAE J2735. Message sets in SAE J2735 are used to implement privacy protocols. The Basic Safety Message (BSM) is defined in SAE J2735.

B. System Model

VANETs and AVs rely on accurate Global Positioning Systems (GPS) for navigation and for sending safety messages to other vehicles. In the future, vehicles will communicate with application servers to utilize a variety of LBS such as transportation management system (TMS) via RSU by using V2I communication as shown in Fig. 2.

C. Autonomous Vehicles (AVs)

AV is defined as the ability of a vehicle to sense the surrounding and navigate without any human assistance. Therefore, the AVs guide themselves by using path planning algorithms, which generate a path between a source and a destination. AVs' navigation consists of three stages: localization, path planning, and vehicle control [3]. In this paper, only the impacts of path planning are considered. Path planning stage starts when a user enters a destination's address to acquire a path. Then, the map service contacts the On-Board Unit

FIGURE 1 System model: includes vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications. Roadside units (RSUs) relay queries to application servers such as transportation management systems (TMSs).

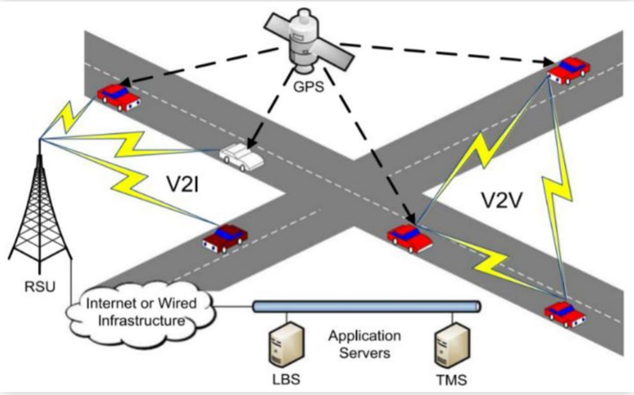
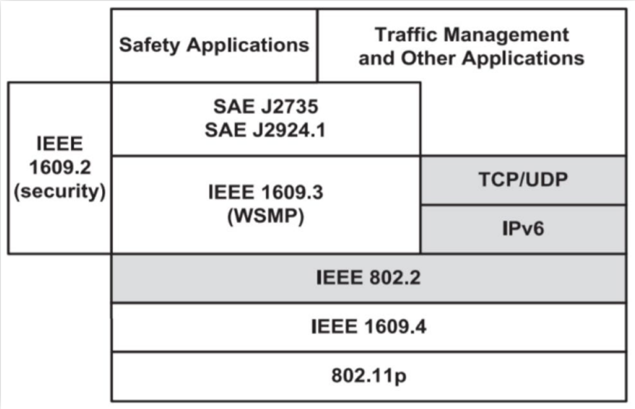


FIGURE 2 DSRC protocol stack: includes traditional protocols such as IPv6 and TCP/UDP, but also new protocols such as WAVE Short Message Protocol (WSMP) and SAE J2735.



(OBU), which is responsible for sending and receiving all messages, to obtain an attainable path and other information. After that, the OBU requests the TMS to return the best path based on some factors, such as traffic condition.

This work assumes all AVs use shortest path to travel between a source and a destination. The AVs are illustrated as an example of the impacts of trajectories and pre-planned trips on location privacy. The AVs are used to clarify the need of trajectory privacy metrics to accurately evaluate vehicles trajectory privacy.

D. Threat Model

Leakage in trajectory privacy could possibly lead to many privacy concerns. For instance, as stated in [2], two surveys have reported that 55% of the LBS users considered location privacy loss when using a service. Furthermore, more than 50% of US residents have privacy concerns when using social networks. In particular, trajectory privacy could be exploited to reduce the anonymity possibility by utilizing other external factors to reveal location privacy. Hence, trajectory privacy could be applied to track vehicle's locations, which might be used in different harmful scenarios.

This work utilizes the threat model illustrated in [1] to demonstrate trajectory privacy techniques and privacy concerns, which assumes a Global Passive Adversary (GPA). Moreover, different AVs riding approaches could come to the existence, such as ridesharing. This approach might be vulnerable against attacks, such as T-PAAD [6], by considering other factors including credit card usage. However, this work only considers a single person case (one owner) to prove the possible impacts of the path planning techniques on trajectory privacy.

AVs may need to communicate with the LBS to find the best path based on the road condition, such as traffic. Hence, the LBS is utilized to obtain the TMS services, such as traffic controller that guides AVs based on the road condition. Thus, the US Department of Transportation (US DoT) is expected to mandate the use of the TMS for different reasons, such as economic reasons, distributed traffic, and roads durability. Therefore, the GPA could have the following means; road cameras, traffic lights, the LBS data, road maps and its topology, and expected mobility profiles such as autonomous vehicles mobility algorithms, which the TMS might have access to all of them. In addition, the TMS could easily find homeowners' names and specific drivers addresses and license plate numbers.

The GPA actions are assumed to be passive. The GPA would eavesdrop and monitor transmitted data without attempting to change or alter it. Moreover, the GPA, such as the TMS manager, could observe data over a very wide region and the actions may be long term since the GPA could eavesdrop for a long period of time. The goal of the GPA is to determine a specific vehicle trajectory, location at specific time, or track a specific vehicle in real time. By analyzing the mobility patterns of vehicles and the expected paths provided by the TMS, the GPA can reveal a specific vehicle's location or trajectory.

Location Privacy Protection Techniques

This section discusses the location privacy protection techniques that are mentioned in [2]. Four techniques are discussed to find the feasibility and the efficiency for each technique when it is applied to VANETs and AVs trajectories to preserve location privacy. In addition, this section shows how those techniques neglect AVs special characteristics possible impacts on trajectory privacy such as path planning.

A. Spatial Cloaking

This technique aims to blur users' locations in the cloaked region by guaranteeing that each cloaked region has other users to be confused with the targeted user. Thus, the user location is indistinguishable from other users in that region [8]. The spatial cloaking problem occurs when the LBS does not determine a precise location for the user to provide the requested service. The cloaking technique is useful since such service does not need a precise location to provide the service. On the other hand, AVs need precise locations to function properly and more importantly for safety reasons where location accuracy is crucial. Finally, this technique needs an intermediate node or an anonymizer to provide cloaked regions that guarantee the existence of other vehicles within the cloaked region.

B. Dummies Trajectories

In this technique fake trajectories are generated by mobile users or vehicles to be sent together with the real trajectories. The goal of generating fake trajectories is to guarantee the indistinguishability between dummies and the targeted users' real trajectories [9]. The dummies challenge the adversary to distinguish the fake trajectories from the real trajectories. Dummies produce a huge computation and communication overhead to the system. Unlike spatial cloaking, dummies do not need any intermediate anonymizer node. Hence, fake trajectories should simulate real user trajectories and patterns to become indistinguishable. Otherwise, it can be excluded by the adversary. Moreover, considering fake trajectories as real confuses the TMS. As a result, counting fake trajectories affects the whole system functionality and quality of service.

C. Mix-Zone

It is defined as a specific geographical area where vehicles enter and silence their wireless communications for a period of time in order to change their Pseudo-IDs. A vehicle's Pseudo-ID is changed frequently aiming to keep the vehicle's trajectories un-linkable to a specific vehicle [10]. Different techniques have been proposed to find the best mix-zones' locations to be deployed in a fixed location, such as intersections or on-the-fly locations, where locations are decided by vehicles based on their mobility and availability. On-the-fly mix-zone provides a better location privacy compared to fixed mix-zones. Mix-zones aim to divide a trajectory into sub-trajectories that are difficult to be connected to each other. Each Pseudo-ID represents a sub-trajectory. Thus, the sub-trajectory is established when a vehicle enters a mix-zone to obtain a brand new Pseudo-ID that is un-linkable to the old one.

D. Path Confusion with Mobility Prediction and Data Caching

This Technique works by avoiding contacting the LBS server for all the requested messages [11]. An LBS server sends messages to the requested users based on the predicted future locations. Thus, when the user needs to request a location service, it might be already sent to the user even before requesting the service. Therefore, the LBS server does not have a complete path or trajectory because some queries do not reach the LBS and users will find it cached in their devices. This technique might be helpful to protect continuous location privacy. On the other hand, the impacts of caching do not affect the trajectory privacy unless a whole sub-trajectory or Pseudo-ID has been cached and the Pseudo-ID's requests do not reach the LBS server.

Platooning is a technique that is expected to be widely used in autonomous vehicles. This technique might be utilized in the future to work as a path confusion technique to achieve trajectory privacy. To the best of our knowledge, platooning is not yet considered to preserve vehicles location privacy and more studies and investigations need to be done in order to benefit from this technique. Moreover, platooning could be utilized to reduce the number of queries send by each vehicle to the server by relying on a platoon leader to send LBS queries for all platoon members. Each vehicle in the platoon has a unique path and needs to use the TMS services, which introduces new concerns for platooning to be adopted in protecting vehicles' location privacy [12, 13].

Measuring Location Privacy

This section introduces metrics that are utilized to evaluate location privacy to be compared with trajectory privacy metrics. A microscopic location privacy, as illustrated in [7], is the anonymity set size at a specific point and specific time. That means, the number of indistinguishable entities that could be confused among each other at specific point and time to be calculated along the whole path between a source to a destination. For vehicular privacy, trajectory K-anonymity, entropy of trajectory K-anonymity, tracking probability, and KDT-anonymity are widely used to evaluate vehicular FPL problem. This work adopts the definition for the mentioned metrics as in [2] as follows:

1. *K-Anonymity*: Which is a technique that has been widely used in literature to evaluate privacy by counting the numbers of entities to be confused with the targeted entity.

$$AS_i = \{j \mid j \in ID, \exists T_j \ s.t. \ p(i,j) \neq 0\} \tag{1}$$

where AS_i represents the anonymity set of a targeted LBS user i, j represents the collections of all LBS users including i, ID represents the LBS users' identifications set, T_j represents the set of j trajectories whose trajectories are indistinguishable from the i trajectory T_i, and p represents the probability of T_i within T_j.

2. *Entropy of K-Anonymity* (H_i): which is defined as the level of uncertainty in the correlations between trajectory T_i and trajectories T_j.

$$H_i = -\sum_{j \in AS_i} p(i,j) \times \log_2\left(p(i,j)\right) \tag{2}$$

3. *Tracking Probability* (Pt_i): which is defined as the probability that the anonymity set size of a vehicle's trajectory $|ASi|$ is equal to one.

$$Pt_i = \Pr\left(\left\|As_i\right\|\right) = 1 \tag{3}$$

4. *KDT-Anonymity:* This metric combines K-anonymity, Distance and time to evaluate continuous location privacy. The metric uses traditional K-anonymity, average distance time between entities over period of time and average anonymity duration over period of time [5].

As discussed in [5], a need to computationally find an accurate metric to evaluate location privacy is inevitable. The above four metrics are calculated at specific time or by taking the average of anonymity over a period of time. The aforementioned metrics lack of defining a precise trajectory anonymity metric. Based on the existing metrics, if a vehicle has an anonymity set size of five for a trajectory time period and the vehicle is linked at the end point of that trajectory by comparing new trajectory with old trajectories, the anonymity set size would be five. However, the anonymity set size is one. Trajectories affect each other and could lead to a privacy leakage. Thus, a precise evaluation for trajectory privacy cannot be neglected with all new technologies and features come along with evolution of the new communication and computational devices characteristics.

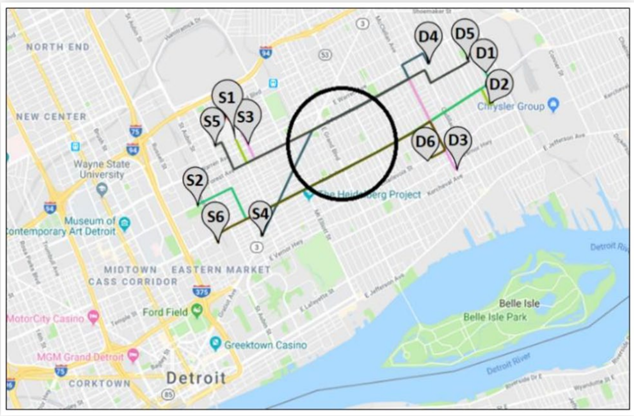
Proposed Trajectory Anonymity

Existing anonymity metrics for location privacy or continuous location privacy do not consider the entire trajectory of a vehicle. For instance, location privacy is measured for vehicle anonymity in a specific point at a specific time. Whereas, continuous location privacy is evaluated based on the anonymity average of contiguous series of time intervals. However, a trajectory, which is a sequence of geospatial points to construct the path between a source and a destination, raises new privacy concerns. For example, some trajectories are not compatible with each other. To illustrate that, Fig. 3 shows six vehicles' starting and ending points with one mix-zone. This scenario is created by using Google Map API to ensure that each vehicle has the shortest path between its source and destination. Based on existing anonymity metrics, the anonymity set size equals to 6 for each vehicle. In fact, using the trajectory-anonymity metric 4, vehicles could have different anonymity set sizes. Hence, vehicles at S2 and S4 would have an anonymity set size of 3. Therefore, new evaluation techniques are needed to accurately measure trajectory privacy.

A. Definitions of Privacy

- *Privacy:* The degree to which an entity cannot be linked to its identity.
- *Location Privacy:* The degree to which a spatial characteristic of an entity cannot be linked to its identity.
- *Continuous Location Privacy:* The degree to which, over a contiguous series of time intervals, an entity cannot be linked to its identity.
- *Trajectory Privacy:* The degree to which a path defined by a set of spatiotemporal characteristics of an entity cannot be linked to its identity.

FIGURE 3 Mix Zone. Google Map API creates sources and destinations randomly and finds the shortest path between them. The circle represents a mix zone where vehicles stop communicating with LBS.



Map data © 2019 Google.

B. Performance Metrics

First, to evaluate trajectory privacy, *trajectory anonymity set size*, denoted by $|As_T^i|$, is considered as a performance metric, which is the average number of entities confused with the targeted entity during a trip between a source and a destination for a specific vehicle (the average of all Pseudo-IDs' anonymity).

$$|As_T^i| = \frac{\sum_{m=1}^{M} N_m^i}{M} \tag{4}$$

where M represents the number of Pseudo-IDs used during a trip for a specific vehicle, i, which is equal to two for each vehicle in Fig. 3. m represents a specific Pseudo-ID of a specific vehicle, i. N represents the number of entities confused with a specific Pseudo-ID, m, or the number of sub-trajectories confused with each other at a specific time interval. For example, Fig. 3 shows two sub-trajectories for each vehicle, which are the trajectory from the source to the mix-zone and from the mix-zone to the destination.

Furthermore, using Equation 4, if N_m^i of a specific Pseudo-ID equals to one, that Pseudo-ID is linked to the previous Pseudo-ID. Therefore, N_m^i should be greater than one in order to preserve the vehicle's trajectory privacy. Using same equation, if N_m^1 and N_m^4 equal to one, their corresponding Pseudo-IDs are linked to each other, which means that N_m^2 and N_m^3 equal to one, as well. Hence, $|As_T^i|$ can be defined as the unlinkability between new and old Pseudo-IDs, where the goal of the mix-zone is to divide trajectories into unlinkable sub-trajectories.

Second, *trajectory probability*, denoted by P_{Tri}, calculates the probability of a trajectory that is used by the targeted vehicle, i, to be confused with other possible trajectories. Equation 5 can be utilized to find all possible trajectories between a source and a destination within a trip as follows:

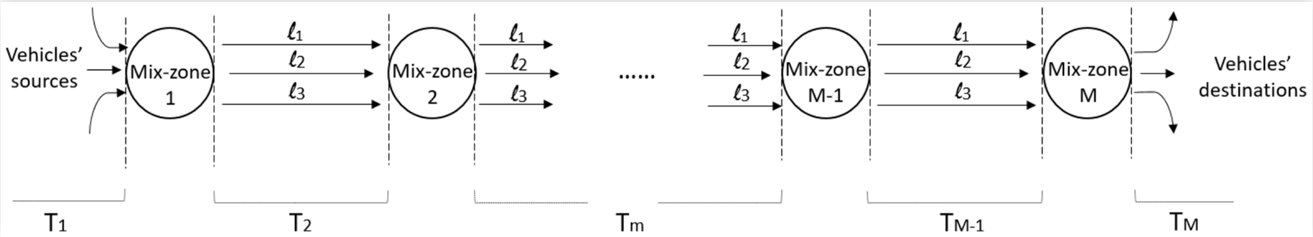
$$P_{Tr_i} = \frac{1}{\prod_{m=1}^{M} N_m^i} \tag{5}$$

Thus, this equation is used to find all possible trajectories that are confused with the targeted vehicle's trajectory. Therefore, the anonymity set size is the compatibility between the generated sub-trajectories.

Moreover, all Pseudo-IDs is considered to be beneficial if their trajectories are included in one of the possible trajectories in Equation 5. Accordingly, in a single mix-zone scenario as shown in Fig. 3, if shortest path technique is applied, S4 cannot appear in D1, D3, or D6 end-points when compared to S4 trajectory before entering the mix-zone. Same for S2, which cannot appear in D1, D4, or D5. In other words, by applying shortest path technique, the trajectory to reach D1, D4 and D5 must be different.

In multiple mix-zones scenario, Fig. 4 shows three vehicles entering Mix-zone 1 at time interval T1 and traveling

FIGURE 4 Multiple mix-zones representation model, where vehicles travel through multiple mix-zones and change their Pseudo-IDs until reaching their destinations.



through multiple mix-zones to reach their destinations. In this scenario, each pseudo-ID has an anonymity set size of three at each time interval. Consequently, at each time interval, every vehicle has three different possible paths (denoted as l1, l2, and l3 at any time interval) to reach the next mix-zone. Applying Equation 5, all possible trajectories from a source to a destination of a specific vehicle can be calculated. Hence, every Pseudo-ID from the anonymity set calculated by Equation 4 must use at least one of the possible trajectories calculated by Equation 5. Otherwise, the Pseudo-ID will be excluded from the anonymity set.

Performance Evaluation Analysis

A. Simulation Setup

We utilize Google Map API to generate paths between two randomly chosen source and destination as shown in Table 1. The work uses real roads on maps to generate the shortest paths between all sources and destinations on the map. Thus, if we have five vehicles, the simulation generates random sources and destinations for the five vehicles. Then, it finds the shortest path for each vehicle to reach the targeted destination. All sources and destinations rely on 3000 m × 3000 m in Detroit, MI area.

The simulation also deployed a mix-zone in the middle of the targeted area. All vehicles start from the source nodes with the same speed and enter the mix-zone before any of them leaving the mix-zone. Vehicles change their Pseudo-IDs that are used before entering the mix-zone to new ones to be used after leaving the mix-zone. A comparison between new Pseudo-ID and old Pseudo-ID is conducted to exclude some trajectories based on the shortest path compatibility. The shortest path compatibility is calculated by comparing the last point of the new Pseudo-ID possible trajectories with the first point of the old Pseudo-ID trajectory. Then, it finds the possible paths that fit the new Pseudo-IDs trajectories. Otherwise, the path is excluded from the list.

TABLE 1 Simulation Setup.

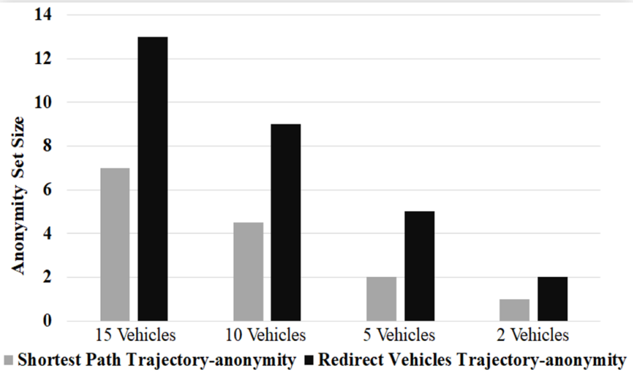
Area	3000 m * 3000 m
Location	Detroit area
Mobility pattern	Shortest path
Source	Random point within the area
Destination	Random point within the area
Vehicle Speed	Same speed for all vehicles.
Number of Vehicles	2, 5, 10, 15

Another comparison between K-anonymity, KDT-anonymity, and Trajectory-anonymity is conducted to illustrate the differences among them compared to our proposed metric. Finally, redirected vehicles paths are added to vehicles mobility to evaluate the impacts of adding some vehicles that do not use shortest paths to confuse the adversary. Thus, redirected vehicles paths are chosen randomly by making 20% of the vehicles taking any paths but shortest paths to reach destination. Then, the privacy evaluation of adding those redirected paths to the simulation is studied to evaluate the impacts of AVs mobility pattern on the overall trajectory privacy by using the proposed technique to show the differences in anonymity sets size after applying trajectory privacy.

B. Discussion and Results

Fig. 5 shows the result of trajectory privacy when the shortest path is applied for all vehicles and after redirecting some vehicles to use longer paths to confuse adversaries. 2, 5, 10, and 15 vehicles are used in the simulation to show the comparison. In addition, both shortest path and redirected paths are evaluated by using the proposed technique to show the differences in anonymity sets size after applying trajectory privacy.

FIGURE 5 Shortest path vs redirected vehicles impact on trajectory privacy.

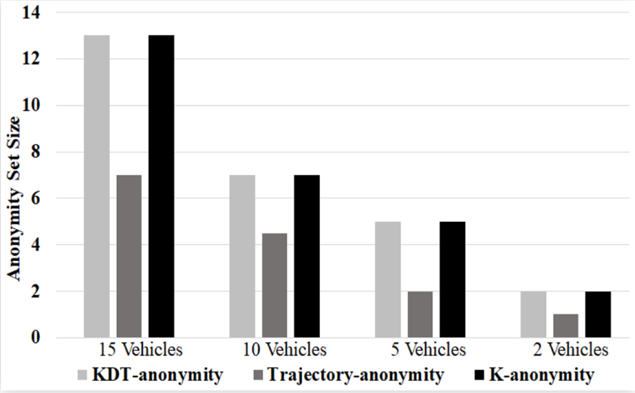


The result shows that when vehicles are redirected to use different paths, it becomes more complicated for the adversary to find the targeted vehicle trajectory based on shortest path compatibility. Since some vehicles uses non-shortest paths, that makes all trajectories valid and possible to be used by any other vehicle. This work demonstrates the indistinguishability based on the shortest path compatibility as defined early in this paper. However, applying entropy for redirected paths measure privacy based on the probability for each vehicle to use one of the paths in the targeted area. Thus, having 20% redirected vehicles mean that a probability might be different based on the number of vehicles that are chosen to use longer paths to confuse adversaries. In other words, 10%, 20%…etc. have different probabilities and impacts on the trajectory privacy by taking other factors into account.

Fig. 6 shows a comparison between K-anonymity, KDT-anonymity and the proposed metric. Same as in Fig. 5, the simulation used 2, 5, 10, and 15 vehicles to compare the aforementioned metrics. The results show almost the same values for K-anonymity and KDT-anonymity due to the fact that these metrics are only evaluated on one mix-zone and the average time and distance do not affect the final result to differentiate between them. However, comparing the proposed metric results to K-anonymity and KDT-anonymity indicates a huge difference that cannot be overlooked. For instance, the result when two vehicles are used shows that anonymity sets size are two based on K-anonymity and KDT-anonymity, which means vehicles are indistinguishable.

However, based on the proposed metric the anonymity set size is equal to one, which means both vehicles linked to their old Pseudo-IDs. Likewise, when having five vehicles, the result shows that the anonymity set size is five based on K-anonymity and KDT-anonymity although the result of the proposed metric shows anonymity set size of two. Consequently, adversary has 50% chance to find the targeted vehicle's location or to use other factors to exclude on vehicle such as home address. On the other hand, based on K-anonymity and KDT-anonymity the adversary has a chance of 20% to find the targeted vehicle and deanonymize it. If a trajectory is excluded by other factors, the trajectory remains anonymized.

FIGURE 6 Comparison between Trajectory-anonymity, K-anonymity and KDT-anonymity to find the anonymity set size for vehicles within an area.



In addition, having more vehicles, such as 10 or 15 vehicles, indicates that applying the proposed metric reduces the anonymity set size almost by half of the anonymity set size returned after applying K-anonymity and KDT-anonymity. Accordingly, many of the trajectories that categorized as indistinguishable when applying K-anonymity and KDT-anonymity are successfully excluded from the anonymity set size when applying the proposed metric. In other words, the trajectory impacts are neglected when applying K-anonymity and KDT-anonymity.

Conclusions and Future Works

This work presents the impacts of trajectories on the evaluation of location privacy and how neglecting trajectories correlation impacts lead to imprecise anonymity set size. Imprecise anonymity set size means over rated privacy levels or in some cases deanonymize vehicles that are categorized previously as protected. Furthermore, AVs mobility might negatively affect privacy in the future when vehicles need to use predefined path planning techniques. In contrast, AVs could change the mobility model to have longer trip time to achieve better privacy. As a future work, more simulation factors would be considered, such as multiple mix-zones impacts and trip distance, which may lead to more reduction in the anonymity set size and more deanonymized vehicles based on trajectories comparisons.

References

1. Shokri, R., Freudiger, J., and Hubaux, J.-P., "A Unified Framework for Location Privacy," No. EPFL-REPORT-148708, 2010.
2. Chow, C.-Y. and Mokbel, M.F., "Trajectory Privacy in Location-Based Services and Data Publication," *ACM Sigkdd Explorations Newsletter* 13(1):19-29, 2011.
3. Levinson, J., Askeland J., Becker J., Dolson J. et al., "Towards Fully Autonomous Driving: Systems and Algorithms," in *Intelligent Vehicles Symposium (IV), 2011 IEEE*, 2011, 163-168, IEEE.
4. Xue, A.Y., Zhang, R., Zheng, Y., Xie, X. et al., "Destination Prediction by Sub-Trajectory Synthesis and Privacy Protection against Such Prediction," in *Data Engineering (ICDE), 2013 IEEE 29th International Conference on*, 2013, 254-265, IEEE.
5. Corser, G.P., Fu, H., and Banihani, A., "Evaluating Location Privacy in Vehicular Communications and Applications," *IEEE Transactions on Intelligent Transportation Systems* 17(9):2658-2667, 2016.
6. Banihani, A., Alzahrani, A., Alharthi, R., Fu, H. et al., "T-PAAD: Trajectory Privacy Attack on Autonomous Driving," in *2018 IEEE Conference on Communications and Network Security (CNS)*, 2018, 1-2, IEEE.

7. Weerasinghe, H., Fu, H., and Leng, S., "Anonymous Service Access for Vehicular Ad Hoc Networks," in *Information Assurance and Security (IAS), 2010 Sixth International Conference on*, 2010, 173-178, IEEE.
8. You, T.-H., Peng, W.-C., and Lee, W.-C., "Protecting Moving Trajectories with Dummies," in *Mobile Data Management, 2007 International Conference on*, 2007, 278-282, IEEE.
9. Mokbel, M.F., Chow, C.-Y., and Aref, W.G., "The New Casper: Query Processing for Location Services without Compromising Privacy," in *Proceedings of the 32nd International Conference on Very large Data Bases*, 2006, 763-774, VLDB Endowment.
10. Palanisamy, B. and Liu, L., "Mobimix: Protecting Location Privacy with Mix-Zones over Road Networks," in *Data Engineering (ICDE), 2011 IEEE 27th International Conference on*, 2011, 494-505, IEEE.
11. Meyerowitz, J. and Choudhury, R.R., "Hiding Stars with Fireworks: Location Privacy through Camouflage," in *Proceedings of the 15th Annual International Conference on Mobile Computing and Networking*, 2009, 345-356, ACM.
12. Dadras, S. and Winstead, C., "Cybersecurity of Autonomous Vehicle Platooning," 2017.
13. Dadras, S., Gerdes, R.M., and Sharma, R., "Vehicular Platooning in an Adversarial Environment," in *Proceedings of the 10th ACM Symposium on Information, Computer and Communications Security*, 2015, 167-178, ACM.

Contact Information

Abdelnasser Banihani
abanihani@oakland.edu

Positions and opinions advanced in this work are those of the author(s) and not necessarily those of SAE International. Responsibility for the content of the work lies solely with the author(s).

ISSN 0148-7191