T-PAAD: Trajectory Privacy Attack on Autonomous Driving

Abdelnasser Banihani*, Abdulrahman Alzahrani*, Raed Alharthi*, Huirong Fu*, George P. Corser†
* Oakland University, Michigan, USA
Email: {abanihani, aalzahrani, rsalharthi, fu}@oakland.edu
† Saginaw Valley State University, Michigan, USA
Email: gpcorser@svsu.edu

***Abstract*—Despite extensive research efforts over the last decade, security and privacy concerns on self-driving cars and other autonomous vehicles are still prevalent. This paper presents a new possible attack, *Trajectory Privacy Attack on Autonomous Driving* (T-PAAD), which aims to deanonymize trajectories by exploiting the relation between autonomous vehicles path planning algorithms and trajectories.**

I. Introduction

In the future, *location based service* (LBS) may collect *frequent precise location* (FPL) data from autonomous vehicles (AVs). The more frequent and precise the location data, the more difficult for AVs to protect location privacy. This has been called the FPL problem [1]. AVs guide themselves using path planning algorithms, which generate a path between a source and a destination. The AVs' navigation consists of three stages: localization, path planning, and vehicle control [2].

This paper suggests that, utilizing path planning algorithms impacts privacy preserving techniques negatively. Existing trajectory privacy techniques, as summarized in [3], fail to defend against T-PAAD attack, which relies on the impacts of path planning algorithms. This problem is important to pave the road toward predicting threat models and preserving location privacy, for the inevitable adoption of AVs. This paper presents a new attack for trajectory location privacy in AVs called T-PAAD, where an adversary de-anonymizes trajectories using existing path planning techniques.

II. Background

A. System Model

This section discusses the system model of AVs navigation. As shown in Figure 1. It consists of four main entities:

User Interface: A user enters a destination to acquire a path. Afterward, the map service contacts the *on-board unit* (OBU) to obtain an attainable path and other information.

Autonomous Vehicle (AV): The OBU is utilized to exchange information with the *roadside units* (RSUs). The AV contains four major layers as follows. First, LBS application layer is responsible for submitted queries to LBS server via trusted RSUs. Second, path planning layer, which is responsible for finding an itinerary to the requested destination, then declares the path to the LBS application layer. Third, privacy aware layer, which applies techniques to preserve users' privacy via a centralized or decentralized trusted third party approach.

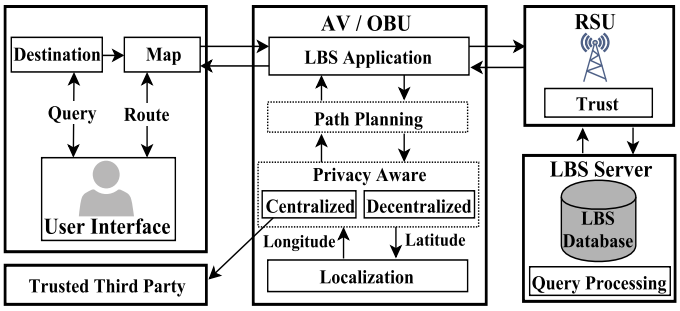

Fig. 1: System model

Finally, localization layer returns longitude and latitude to be used by the AVs applications.

Roadside Units (RSUs): RSUs are intermediary communication nodes between AVs and LBS server.

Location Based Service (LBS) Server: LBS servers process and store queries requested by AVs.

B. Terms Definitions

To simplify the illustration of T-PAAD scenarios, a few terms used in this paper are defined as follows:

Mix Zone: A specific geographical area where vehicles enter and silence their wireless communications for a period of time in order to change their *PseudoIDs (Ps)*. A vehicle's PseudoID changes frequently aiming to keep the vehicle's trajectories unlinkable to a specific vehicle.

Trajectory (T_r): A sequence of geospatial points to construct a path between a source and a destination.

Trajectory k-Anonymity: It is defined as unlinkability between new and old pseudoIDs where the goal of mix zone is to divide trajectories into unlinkable sub trajectories (PseudoIDs' trajectories T_r(P)). Moreover, a trajectory is defined to be anonymized (k) if: (1) there must be a point that all anonymized vehicles are located inside the targeted mix zone; (2) At least, two trajectories (k = *2*) must be interchangeably indistinguishable based on the path planning techniques used by the AVs. As shown in Figure 2, $T_r(P_A^1)$ is indistinguishable from $T_r(P_B^1)$ after leaving *mix zone 1*.

Entropy of Trajectory k-Anonymity: It represents the level of uncertainty in the correlations between trajectory Tr_i and a set of trajectories Tr_j in a certain mix zone.

978-1-5386-4586-4/18/$31.00 ©2018 IEEE

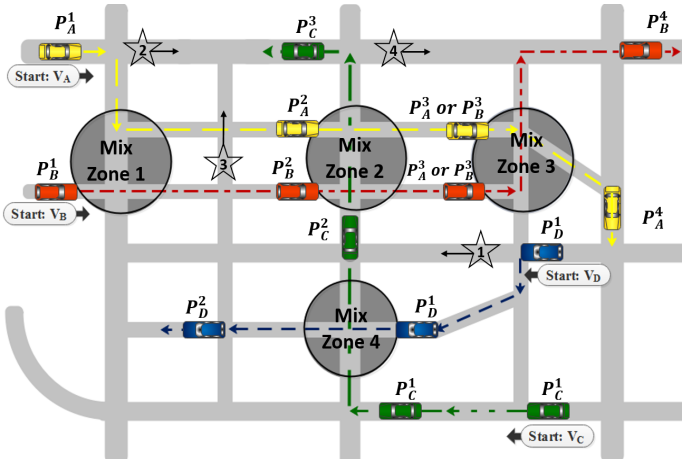

Fig. 2: T-PAAD: Mapping old PseudoIDs to new PseudoIDs

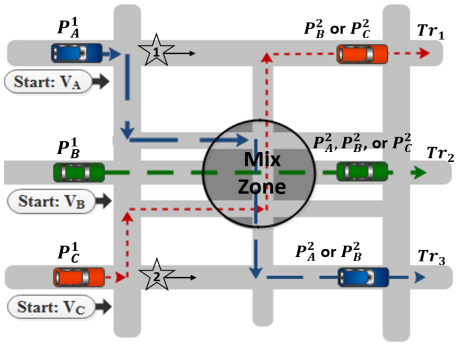

Fig. 3: T-PAAD: Decreasing anonymity set

III. T-PAAD Model

This paper assumes a *global passive adversary* (GPA), which would only eavesdrop and would not alter the data being transmitted. Consider a hypothetical LBS, a *traffic management system* (TMS) which (1) collects the positions of vehicles, (2) estimates likely vehicle traffic congestion, and (3) determines optimal vehicle paths in real time.

The adversary's objective is to find a specific vehicle, attempting to link it to only one of the trajectories and decrease the anonymity set at mix zones in order to deanonymize trajectories. Based on AVs' path planning techniques, the adversary assigns a probability for each vehicle to be mapped to one of the existing trajectories in the anonymity set. Therefore, the adversary determines which trajectory fits path planning algorithms by checking the correlation between old and new sub trajectories.

Figure 2 shows a road map with four vehicles (V_A, V_B, V_C, *and* V_D) and multiple mix zones. Each vehicle V_X has one trajectory and a sequence of changeable PseudoIDs (P_X^1, P_X^2, ..., P_X^n). For instance, the first PseudoID for vehicle V_A is represented as P_A^1, where the first PseudoID trajectory is represented as $T_r(P_A^1)$. Below are three scenarios on how T-PAAD exposes vehicle trajectory privacy.

Short Term Scenario: As shown in Figure 2, vehicles V_C and V_D with PseudoIDs P_C^1 and P_D^1 respectively, start from different points toward *mix zone 4*. Once they enter the mix zone 4, they change their PseudoIDs to P_C^2 and P_D^2, respectively. After leaving the mix zone 4, both $T_r(P_C^2)$ and $T_r(P_D^2)$ should be indistinguishable. However, adversaries in this scenario can distinguish vehicles' trajectories since there is no correlation between P_D^1 and P_C^2. In other words, the shortest path for V_D to reach V_C position is a different path, as the symbol (1) represents.

Long Term Scenario: V_A and V_B with PseudoIDs P_A^1 and P_B^1 respectively, start from different points and travel through multiple mix zones. In *mix zone 1*, both vehicles change their PseudoIDs to P_A^2, P_B^2 respectively. After leaving the mix zone 1, both $T_r(P_A^2)$ and $T_r(P_B^2)$ are still indistinguishable. In *mix zone 2*, V_A and V_B meet with V_C, where all vehicles change their PseudoIDs to P_A^3, P_B^3, and P_C^3, respectively. However, adversaries in this scenario can distinguish $T_r(P_C^3)$ since there is no correlation between the PseudoIDs P_A^1, P_B^1, and P_C^3, as the symbols (2) and (3) represent the shortest paths for V_A and V_B. Further, V_A and V_B enter *mix zone 3* and change their PseudoIDs to P_A^4 and P_B^4 respectively. The adversaries can distinguish $T_r(P_A^4)$ and $T_r(P_B^4)$ since there is no correlation between the PseudoIDs P_A^1 and P_B^4, as the symbols (2) and (4) represent the shortest paths for V_A and V_B.

Decreasing Anonymity Set Scenario: As mentioned above in *mix zone 2* where $T_r(P_C^3)$ is linked to $T_r(P_C^2)$, the anonymity set of all vehicles in *mix zone 2* is decreased by 1. Figure 3 shows another case where three vehicles V_A, V_B and V_C start with three PseudoIDs P_A^1, P_B^1, and P_C^1, respectively, and enter the mix zone. After leaving the mix zone, these PseudoIDs have been changed to P_A^2, P_B^2, and P_C^2, creating new PseudoIDs trajectories $T_r(P_A^2)$, $T_r(P_B^2)$, and $T_r(P_C^2)$, respectively. No path planning algorithms would ever compute P_A^1 to follow Tr_1. Thus, that scenario can be eliminated, reducing the anonymity set from $k = 3$ to $k = 2$ for vehicle V_A, likewise for vehicle V_C. Vehicle V_B could follow any of the three paths. Hence, its anonymity set size remains $k = 3$.

Acknowledgment

This research work is partially supported by the National Science Foundation under Grants CNS-1460897 and DGE-1623713. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

[1] G. P. Corser, H. Fu, and A. Banihani, "Evaluating location privacy in vehicular communications and applications," *IEEE Transactions on intelligent transportation systems*, vol. 17, no. 9, pp. 2658–2667, 2016.

[2] J. Levinson, J. Askeland, J. Becker, J. Dolson, D. Held, S. Kammel, J. Z. Kolter, D. Langer, O. Pink, V. Pratt *et al.*, "Towards fully autonomous driving: Systems and algorithms," in *Intelligent Vehicles Symposium (IV), 2011 IEEE*. IEEE, 2011, pp. 163–168.

[3] C.-Y. Chow and M. F. Mokbel, "Trajectory privacy in location-based services and data publication," *ACM Sigkdd Explorations Newsletter*, vol. 13, no. 1, pp. 19–29, 2011.