Toward an Open Source Location Privacy Evaluation Framework for Vehicular Networks

David Eckhoff*, Mykola Protsenko† and Reinhard German*
* Computer Networks and Communication Systems, Dept. of Computer Science, University of Erlangen, Germany
† IT Security Infrastructures, Dept. of Computer Science, University of Erlangen, Germany
{david.eckhoff,mykola.protsenko,reinhard.german}@fau.de

I. Introduction

Inter-Vehicle Communication (IVC), i.e., the ad hoc information exchange between vehicles, is believed to help increase traffic safety, improve traffic efficiency, and also increase the comfort of drivers. Often simulation is used to evaluate these applications. The fact that this becomes a challenging task in the context of privacy is one of the reasons that proper privacy protection is often disregarded [1]. The concept of privacy is very abstract, meaning it is difficult to quantify the level of privacy protection and find easy to understand metrics [2]. Also, in order to compare two different privacy protection mechanisms, the exact settings of the simulation have to be known. In vehicular network simulation, and more so in privacy simulation, results heavily depend on the used parameters and due to space limitations they are often not in the article presenting a privacy protection scheme [3]. To make matters worse, there is still no publicly available simulation environment, requiring each researcher to implement their own evaluation tools.

In this paper we discuss the outline and present building blocks for a comprehensive open source location privacy evaluation framework to enable researchers to reproducibly assess the effectiveness of a given privacy protection algorithm. We also present a proof of concept evaluation. By extending the well established Veins simulation framework [4] that couples the traffic simulator SUMO and the network simulator OMNeT++ we allow for an easy setup and integration with existing simulation scenarios or already implemented protocols. We hope that our framework lowers the complexity of privacy evaluation and thereby makes certain protection measures more likely to be considered in future vehicular networks.

II. Requirements

1) Attacker model: The effectiveness of a privacy protection mechanism may differ greatly depending on the deployed attacker model. Our model supports both the widely employed omniscient attacker and an attacker who actually relies on the successful reception of broadcast messages over the wireless channel. The latter can be designed to have set up multiple (connected) access points, making tracking highly dependable on the area coverage and even on packet loss.

2) Metrics: Díaz et. al discussed the crucial requirement for easy-to-understand metrics when evaluating privacy algorithms [2]. Currently the most used metrics include the anonymity set size (i.e., the number of vehicles a targeted vehicle could possibly be), the entropy of that set accounting for different probabilities, and the tracking time (i.e., the time it was possible to track the vehicle).

For the sake of comparability, a privacy evaluation framework must support these metrics. However, using these metrics to compare one mechanism to another can be incomprehensible and misleading. We therefore introduce a light-weight metric, the tracking probability, that is, the probability of a vehicle being successfully tracked throughout the entire network.

3) Scenarios: Finding standard scenarios is still an ongoing challenge in the simulation of vehicular network applications and protocols. This also applies to the simulation of privacy protection. For easy comparison our simulator therefore already includes simple scenarios such as crossings, freeways, roundabouts, and small city segments. Using consistent and public scenarios increases the credibility and reproducibility of simulation results.

III. Vehicle Tracking

Cars in vehicular networks periodically emit unencrypted broadcast messages containing position information, speed, heading, and a (volatile) source address. If an attacker is able to exploit this information to track a vehicle through the network the privacy of that driver would be undoubtedly violated. The quality of a location privacy protection mechanism is tightly linked to the feasibility of such an attack. It is therefore crucial to base the assessment of these mechanisms on realistic scenarios in terms of mobility and network transmissions. The Veins framework already provides this, making it a perfect candidate to be the basis for a privacy simulation framework.

Our simulator uses the broadcast transmissions of the vehicles in the scenario as input. Using the information obtained, a given attacker then deploys a tracking algorithm to track the vehicles' movement. Each track T_i consists of different way points, called *Observations*, each observation representing one received broadcast message.

Assuming an attacker overheard messages $O_1, O_2, ..., O_n$ in the last time interval, the problem of tracking is to find a valid continuation of a track T_i using an observation O_j. To efficiently find the correct successor of track T_i, we deploy a gating algorithm that only selects observations that can possibly be the true successor. This selection can be done based on the limits of vehicular mobility (maximum possible

(a) Fully covered crossing scenario

(b) Freeway scenario with blind spot in between attacker access points

Figure 1. Example of simulation scenarios possible to investigate using our privacy evaluation framework

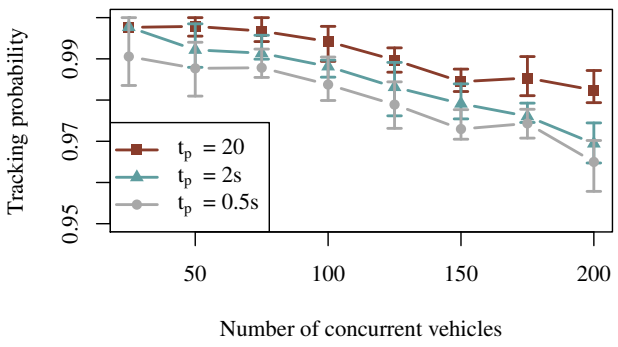

Figure 2. Impact of address changing on tracking in the crossing scenario

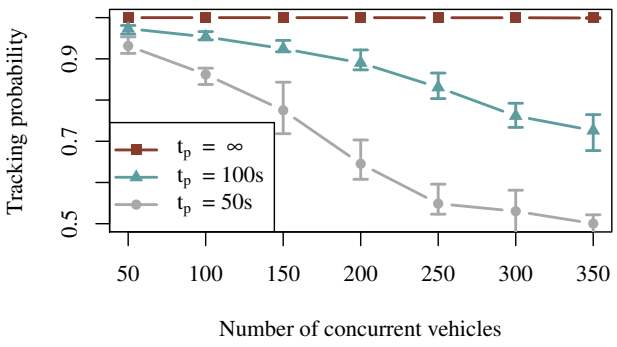

Figure 3. Impact of address changing on tracking in the freeway scenario

distances, etc.). Once a set of feasible observations has been found, each tuple (T_i,O_j) is assigned a value that represents the likelihood of T_i being continued by O_j. In state-of-the-art tracking systems, this value is found by first extrapolating the expected next position of T_i and then computing the statistical distance between this point and the observations [5].

The most favorable overall assignment solution is the one that minimizes the sum of statistical distances. This problem can be easily solved by converting it to the well known maximum weighted matching problem [6]. However, the best matching does not necessarily have to be the correct one; it is therefore desirable to store the best K solutions as hypotheses. Using a multi hypothesis tracking algorithm, a wrong assignment is then likely to result in a higher overall statistical distance in the next time step. A hypothesis is discarded when its probability (computed using the JPDA method [5]) falls below a certain threshold.

IV. Evaluation

To show the versatility of our simulator we evaluated address changing strategies (at a broadcast frequency of 1 Hz) in two different scenarios shown in Fig. 1. While the crossing scenario was fully covered by the attacker, the freeway had a blind spot between two deployed attacker access points. The investigated strategies impose a maximum validity time t_p until a vehicle must change its source address. Vehicles draw a random number $r \in [0, t_p]$ and change their source address after r seconds ($t_p = \infty$ meaning addresses were never changed).

It was almost impossible to confuse the attacker in the Crossing scenario (Fig. 2); the main reason for a failed tracking here was packet loss. This indicates the effectiveness and correctness of the deployed tracking algorithm, and also the importance to account for properties of the wireless channel.

The blind spot in the freeway scenario (Fig. 3) had a considerable impact on the tracking probability. When vehicles changed their pseudonym while not being overheard by an attacker, there was a great chance that the attacker would lose their track due to a wrong assignment.

V. Conclusion

We presented the outline and building blocks of an open-source location privacy evaluation framework. We showed how we addressed necessary requirements and presented first results. Future work will include the deployment of even more effective filter algorithms and the preparation and release of the source code under the GNU GPL.

References

[1] D. Eckhoff and C. Sommer, "Driving for Big Data? Privacy Concerns in Vehicular Networking," *IEEE Security and Privacy*, vol. 12, no. 1, pp. 77–80, February 2014.

[2] C. Díaz, "Anonymity Metrics Revisited," in *Dagstuhl Seminar on Anonymous Communication and its Applications*, Dagstuhl, Germany, October 2005.

[3] S. Kurkowski, T. Camp, and M. Colagrosso, "MANET Simulation Studies: The Incredibles," *ACM SIGMOBILE Mobile Computing and Communications Review (MC2R)*, vol. 9, no. 4, pp. 50–61, October 2005.

[4] 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, Jan. 2011.

[5] S. Blackrnan and R. Popoli, *Design and Analysis of Modern Tracking Systems*. Artech House Boston, 1999.

[6] J. Edmonds, "Maximum Matching and a Polyhedron with 0, l-vertices," *J. Res. Bur. Stand*, vol. 69B, no. 1-2, pp. 125–130, January 1965.