Protecting Vehicle Privacy using Dummy Events

Nahed Alnahash, George Corser, Huirong Fu
Department of Computer Science
Oakland University
Rochester, MI 48309

Ye Zhu
Department of Electrical and Computer Engineering
Cleveland State University
Cleveland, OH 44115

Email: { nalnahas, gpcorser, fu }@oakland.edu, zhuye06@gmail.com

Abstract

VANETs (Vehicular Ad Hoc Networks) can improve safety and enable a wide range of internet services, such as location based services. The problem is VANETs can also be used for tracking vehicles, raising privacy concerns. To confuse potential eavesdroppers some researchers have approached the problem by having network nodes transmit dummy data along with true data. Few such proposals consider vehicular applications where network nodes, vehicles, often travel along predictable paths, roadways. In this paper our approach is to examine one generic dummy-based scheme which we applied in a vehicular context. The results we found were that, with modifications, the scheme may be useful in certain vehicular situations.

1. Introduction

This paper considers the application level *vehicle location tracking problem*. How can vehicles conceal their locations from a location based service, LBS, an internet service provider which requires vehicle position as user input?

Spatial cloaking is a well researched solution to this problem but spatial cloaking requires a trusted third party, TTP. We consider only cases when spatial cloaking and TTP are undesirable, for example, at times when there is only one vehicle (or very few vehicles) using the LBS.

Vehicular ad-hoc networks, VANETs, are self-configuring communication architectures which enable vehicles in motion to intercommunicate as nodes in computer networks. VANETs present distinctive location privacy challenges for motorists because vehicles tend to travel along predictable routes, roadways. If a motorist uses LBS while driving, an LBS administrator may be able to monitor the motorist's position. Vehicles may attempt to confuse LBS by sending false/dummy location data. However, LBS might identify a sequence of locations as a set, or trajectory. Trajectories consisting of multiple dummy events must follow a pattern similar to a vehicle path or the dummy trajectory will be detectable as fake by LBS that uses vehicle pattern analysis. Further, each dummy location must correspond to a real location on a roadway otherwise the dummy location could be detectable as fake if LBS uses map deanonymization, cross-referencing the dummy location to a real road map on Mapquest or Google Maps.

The vehicle location tracking problem is important and has received attention from both legislators and researchers. Legislation has been introduced at the national level, including the Location Privacy Act and the Geolocation Privacy and Surveillance Act. This legislation has

been difficult to pass perhaps because its technical implications are unclear. Technical solutions are proposed frequently but counterproposals defeating the solutions are proposed almost equally frequently. If no technical solution emerges to enable drivers to protect the location privacy of their vehicles' from LBS, then surveillance may become unpreventable; therefore perhaps inevitably it may become socially acceptable. LBS might even sell this unavoidably public data as additional services. Employers might monitor an employee's car parked at a competitor's office (revealing an employee's job interview) or at specialized medical facilities (revealing an employee's health condition). It is not difficult to construct further privacy breaches arising from vehicle surveillance by spouses and ex-spouses, or paparazzi and other stalkers.

The location privacy challenge from a technical standpoint is large-scale and complicated in VANETs. Equipment supporting wireless/Wi-Fi networks is already being installed in new vehicles. Industry representatives estimate that 90% of vehicles will be Wi-Fi-connected within the decade [1]. LBS usage continues to grow rapidly [2] and is expected to expand to VANET platforms [3]. Standards governing VANETs [4] have outlined sophisticated encryption schemes to enable privacy, but researchers continue to find privacy vulnerabilities inherent in VANET protocols and vehicle mobility patterns.

Our primary contribution is demonstrating the technique in [5] in vehicle scenarios and evaluating some of the implications. The rest of this paper is organized as follows. Section 2 describes related work done in [5] upon which our results are founded. Section 3 describes our demonstration of the original model and our modifications for vehicular situations. Section 4 presents simulation results assuming Manhattan-style roadways and compares similar results using no roadways. Section 5 concludes the paper.

2. Related Work

The foundation for our research is presented in [5], which attempts to minimize the number of dummies required for given levels of short term disclosure (*SD*), long-term disclosure (*LD*) and distance deviation (*dst*). [5] evaluates human-like trajectories of mobile users, illustrated in Fig. 1 (a). Our work evaluates vehicle-like trajectories as illustrated in Fig. 1 (b).

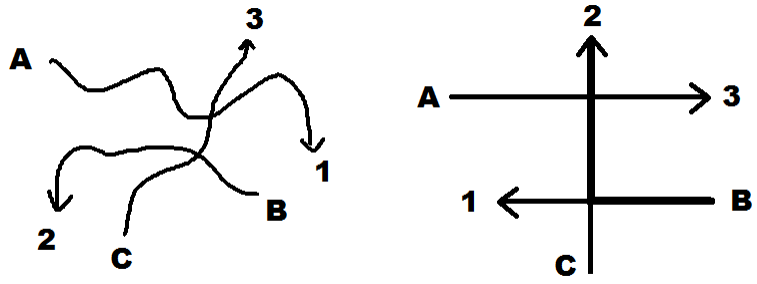

Fig. 1: (a) Human-like trajectories. (b) Vehicle-like trajectories.

Short term disclosure (*SD*) is a measure of the probability of an eavesdropper successfully identifying any particular true location given a set of true and dummy locations over a presumably short time. If there are m time slots and D_i is the set of true and dummy locations at time slot i, where $| D_i |$ is the size of D_i, then

$$SD = \frac{1}{m} \sum_{i=1}^{m} \frac{1}{|Di|} \quad (1)$$

Long term disclosure (*LD*) is a measure of the probability of an eavesdropper successfully identifying a true trajectory given a set of true and dummy trajectories. The more trajectories overlap, the lower the probability of detection. If there are n total trajectories and k trajectories that overlap, then there are $n - k$ trajectories that do not overlap. If T_k is the number of possible trajectories amongst the overlapping trajectories, then

$$LD = 1 / (T_k + (n - k)) \quad (2)$$

Distance deviation (*dst*) is the average of distance between trajectories of dummies and the true user. To define dst_i as the distance deviation of user i, let PL^j_i be the location of user i at the jth time slot and let L^j_{dk} be the location of the kth dummy at the jth time slot. The function *dist()* express the distance between the true user location and the dummy locations. Then

$$dst_i = \frac{1}{m} * \frac{1}{n} * \sum_{k=1}^{n} \sum_{j=1}^{m} dist(PL^j_{i,} L^j_{dk}) \quad (3)$$

[5] proposes two dummy generation methods, random pattern scheme and rotation pattern scheme. The *random pattern scheme* would arbitrarily choose a starting point, ending point and points in between for dummy trajectories. The *rotation pattern scheme* would ensure overlap, extending the random pattern scheme by also arbitrarily choosing an intersection point and rotation angle for dummy trajectories. Recall: The more trajectories overlap, the lower the *LD*.

3. Demonstration

To illustrate the difference between human-like trajectories and vehicle-like trajectories consider Fig. 2. While humans may roam freely and move relatively slowly, vehicles tend to move in more predictable patterns much more quickly. Fig. 2 (a) shows how dummies **d1** and **d2** may be undetectable as dummies because their movement patterns are human-like. Fig. 2 (b) shows how dummies **d1** and **d2** are detectable as fakes in vehicular contexts. Dummy **d1** is detectable because it does not follow a vehicle-like pattern. Dummy **d2** is detectable because, while it moves in a vehicle-like pattern, some of the positions are not on roadways, or it is easy to detect that it is impossible to move from one position to another by using a known roadway. Numerical data for the true user and dummies in Fig. 2 (b) are presented in Table 1.

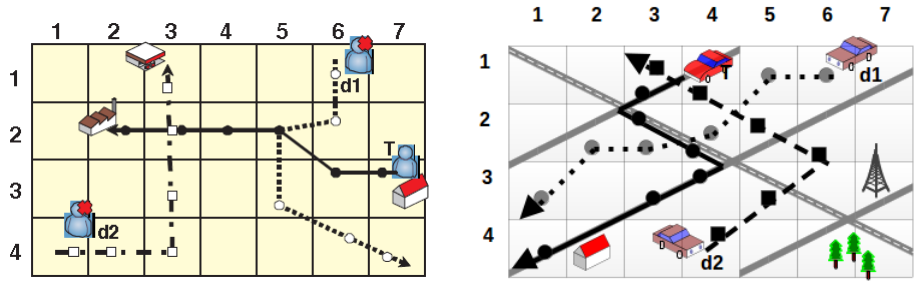

Fig. 2: Random patterns for human-like trajectories, (a) from [5], and (b) viewed over roadways.

Table 1: Privacy measurement of random pattern scheme dummy trajectories from Fig. 2 (b)

Time slot, i (m = 6)	1	2	3	4	5	6
True vehicle location	(3,1)	(3,2)	(4,2)	(4,3)	(3,3)	(1,4)
Dummy 1 location	(6,1)	(5,1)	(4,2)	(3,2)	(2,2)	(1,3)
Dummy 2 location	(4,4)	(3,3)	(6,2)	(5,2)	(4,1)	(3,1)
$\|D_i\|$, number of unique locations	3	3	2	3	3	3
Distance (*dist()*)	3.1	2.2	1.0	1.4	1.8	2.3

From equations (1), (2) and (3) we can compute *SD*, *LD* and *dst* for this user. See equations (4), (5) and (6) below.

$$SD = (1/6) * (1/3 + 1/3 + 1/2 + 1/3 + 1/3 + 1/3) = 0.3611 \quad (4)$$

$$LD = 1 / (5 + (3 - 2)) = 0.1667 \quad (5)$$

$$dst = (1/6) * (3.1 + 2.2 + 1.0 + 1.4 + 1.8 + 2.3) = 2.0 \quad (6)$$

The scenario above illustrates the random pattern scheme. To illustrate the implications of the rotation pattern scheme consider Fig. 3. For human-like movement rotation angle may be chosen arbitrarily. For vehicle-like movement, because of often perpendicular roadways, it may be more advantageous to constrain dummy trajectories to rotations in increments of 90 degrees.

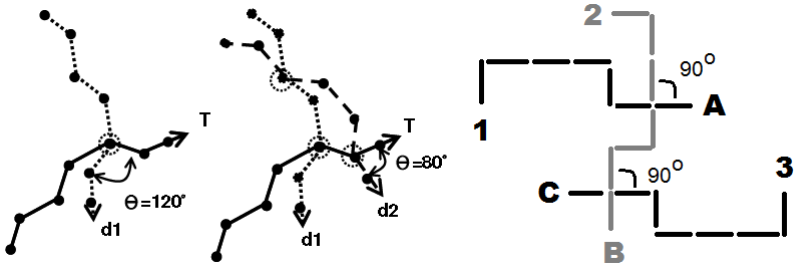

Fig. 3: Rotation patterns for (a) human-like trajectories, from [5], and (b) vehicle-like trajectories.

Restricting the rotation angle offers both advantages and disadvantages. Constraints may be considered disadvantageous because vehicles have fewer potential paths to choose from as they move, which degrades *SD*. However, fewer potential positions for vehicles implies more potential overlap, which may improve *LD*. Consider the illustration in Fig. 4 and compare with Fig. 2 (b). The former shows trajectories more overlapping positions. Since vehicles frequently transmit precise positions it is possible for a vehicle to construct realistic dummy trajectories using real or realistic data from other vehicles on the road.

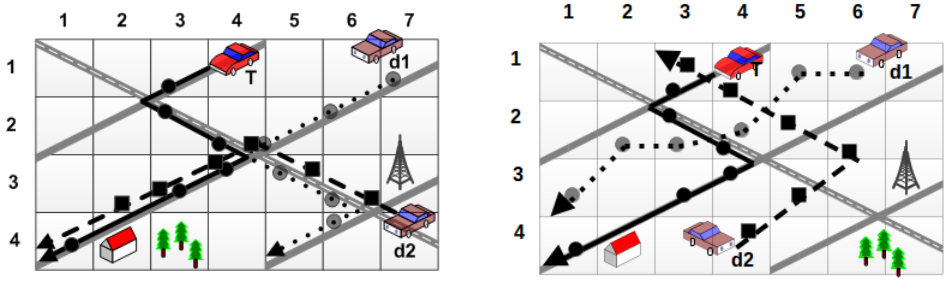

Fig. 4: Modified rotation patterns for (a) vehicle-like trajectories and (b) human-like trajectories over roadways, from Fig. 2 (b).

Table 2: Privacy measurement of random pattern scheme dummy trajectories from Fig. 4 (a)

Time slot, *i* (m = 6)	1	2	3	4	5	6
True vehicle location	(3,1)	(3,2)	(4,2)	(4,3)	(3,3)	(1,4)
Dummy 1 location	(7,1)	(6,2)	(4,2)	(5,3)	(6,3)	(6,4)
Dummy 2 location	(6,3)	(5,3)	(4,2)	(4,3)	(3,3)	(2,3)
$\|D_i\|$, number of unique locations	3	3	1	2	2	3
Distance (*dist()*)	3.8	2.6	0.0	1.0	3.0	3.6

From equations (1), (2) and (3) we can compute *SD*, *LD* and *dst* for this user. See equations (7), (8) and (9) below.

$$SD = (1/6) * (1/3 + 1/3 + 1/1 + 1/2 + 1/2 + 1/3) = 0.5278 \quad (7)$$

$$LD = 1 / (9 + (3 - 3)) = 0.1111 \quad (8)$$

$$dst = (1/6) * (3.8 + 2.6 + 0.0 + 1.0 + 3.0 + 3.6) = 2.0 \quad (9)$$

4. Simulation

We simulated scenarios similar to the illustration above, except we used 20 time slots and 5 to 25 dummies on a grid of 50x50 squares. We computed *SD*, *LD* and *dst* for each scenario for each of two conditions, one where roadways were restricted to exist only in squares which had one dimension evenly divisible by 10, the other with no such restriction. We ran each scenario nine times and recorded the run with the median number of trajectory intersections.

Table 3: Simulation data for scenarios (a) with road restrictions, and (b) without road restrictions

(a) Roads restricted to every 10 squares in grid					
dummies	25	20	15	10	5
rr = 10	10	10	10	10	10
SD	0.04052	0.05031	0.06547	0.09464	0.17
LD	0.01282	0.01538	0.02272	0.03703	0.1
intersects	26	22	14	8	2
dst	25.793	22.7686	26.5048	25.8764	20.6967

(b) No road restrictions (rr)					
dummies	25	20	15	10	5
rr = 0	0	0	0	0	0
SD	0.03853	0.04761	0.06270	0.09136	0.16666
LD	0.03571	0.04761	0.05555	0.07692	0.16666
intersects	1	0	1	1	0
dst	23.8404	24.5140	25.6914	23.8121	31.6829

The charts in Fig. 5 show how restricting locations to realistic road paths reduces the chance of long term disclosure, *LD* (Fig. 5, right), while maintaining minimal effect on short term disclosure, *SD* (Fig. 5, left).

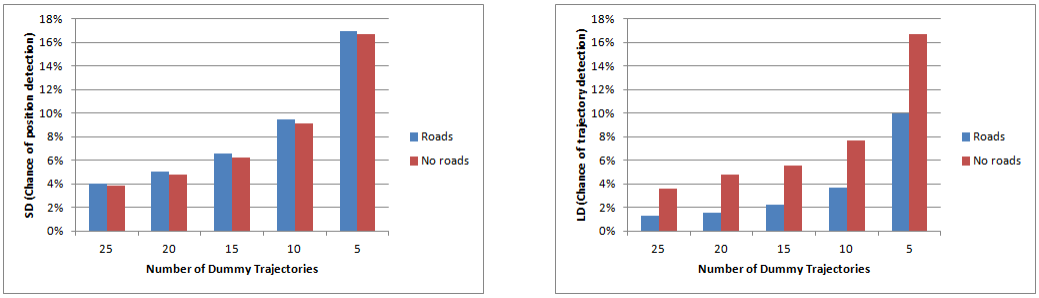

Fig. 5: For varying numbers of dummy trajectories, calculations for (a) SD and (b) LD

5. Conclusion

Dummy locations which are not realistic in vehicular context may be detectable as fakes because location coordinates can be cross-referenced and validated using maps. Our solution overcomes this problem by restricting dummy vehicles to roadways. This reduces the number of locations and therefore increases the number of overlapping trajectories which improves LD. The method in [5] modified by a 90 degree rotation, instead of random rotation, is more realistic in vehicular context and better protects privacy because the chance of long term disclosure is reduced.

Two key areas of future work include evaluation of realistic distance deviation and frequency of LBS requests. The foundation paper in [5] and our paper present purely abstract quantities. In vehicular settings we can and should estimate the privacy protection using realistic distances which will likely depend on the precision of GPS devices used in on-board VANET components. Continuous precise location tracking also remains a problem even more challenging than the general vehicle location tracking problem. If LBS receives dozens, hundreds or even thousands of requests per hour the privacy of a vehicle may become more difficult to protect. Further, dummy trajectories may become so numerous that the resulting congestion on the LBS server due to unnecessary database queries may render the technique impractical.

6. References

[1] Bush, I. (2013, Feb 25). GM, AT&T readying in-vehicle wi-fi. http://philadelphia.cbslocal.com/2013/02/25/gm-att-readying-in-vehicle-wi-fi/

[2] Johnson, L. (2012, Oct 31). Location-based services to bring in $4b revenue in 2012: study. http://www.mobilemarketer.com/cms/news/research/14115.htmlhttp://www.mobilemarketer.com/cms/news/research/14115.html

[3] Koslowski, T. (2012, Jan 3). Your connected vehicle is arriving. http://www.technologyreview.com/news/426523/your-connected-vehicle-is-arriving/

[4] IEEE Standard for Wireless Access in Vehicular Environments Security Services for Applications and Management Messages," IEEE Std 1609.2-2013 (Revision of IEEE Std 1609.2-2006) , vol., no., pp.1,289, April 26 2013, doi: 10.1109/IEEESTD.2013.6509896

[5] You, T. H., Peng, W. C., & Lee, W. C. (2007, May). Protecting moving trajectories with dummies. In Mobile Data Management, 2007 International Conference on (pp. 278-282). IEEE.