

基于车辆协作的混淆路径轨迹隐私保护机制

赵子文, 叶阿勇, 金俊林, 孟玲玉

(福建师范大学数学与信息学院, 福建 福州 350007)

摘要: 在车联网中, 车辆通过和第三方共享位置信息获得基于位置的服务, 这可能会导致车辆轨迹隐私泄露。针对该问题, 提出基于车辆协作的混淆路径轨迹保护机制。首先, 车辆轨迹熵达到自定义的轨迹保护阈值后, 减少车辆混淆, 解决了路径混淆开销大的问题。然后, 设计路径混淆算法来增加车辆在路口路径混淆的机会, 提高了车辆轨迹保护程度。最后, 仿真实验从轨迹熵和轨迹跟踪成功率验证了该方法的有效性和高效性。

关键词: 车联网; 轨迹保护; 混淆路径; 轨迹熵; 基于位置的服务

中图分类号: TP393

文献标识码: A

doi: 10.11959/j.issn.2096-109x.2019038

Trajectory privacy protection mechanism of obfuscating paths based on vehicles cooperation

ZHAO Ziwen, YE Ayong, JIN Junlin, MENG Lingyu

College of Mathematics and Informatics, Fujian Normal University, Fuzhou 350007, China

Abstract: Vehicles sent location information to third parties to obtain location-based services in the Internet of vehicles, which may lead to vehicle trajectory privacy leakage. To address the privacy problem, trajectory privacy protection mechanism of obfuscating paths based on vehicles cooperation was proposed. Firstly, after the vehicle trajectory entropy reaches the custom trajectory protection threshold, the overhead of paths confusion saves by reducing the number of vehicle paths confusion. Then, paths confusion algorithm was designed to increase chances of vehicle paths confusion at intersections, which can improve the degree of vehicle trajectory protection. Finally, the simulation experiment verified the validity and efficiency of the method from the trajectory entropy and trajectory tracking success rate.

Key words: Internet of vehicles, trajectory protection, obfuscating paths, trajectory entropy, location-based service

收稿日期: 2019-01-03; 修回日期: 2019-03-18

通信作者: 叶阿勇, yay@fjnu.edu.cn

基金项目: 国家自然科学基金资助项目 (No.61771140, No.6187208, No.61872090); 福建省自然科学基金资助项目 (No.2018J01780)

Foundation Items: The National Natural Science Foundation of China (No.61771140, No.6187208, No.61872090), The Natural Science Foundation of Fujian Province (No.2018J01780)

论文引用格式: 赵子文, 叶阿勇, 金俊林, 等. 基于车辆协作的混淆路径轨迹隐私保护机制[J]. 网络与信息安全学报, 2019, 5(4): 52-62.

ZHAO Z L, YE A Y, JIN J L, et al. Trajectory privacy protection mechanism of obfuscating paths based on vehicles cooperation[J]. Chinese Journal of Network and Information Security, 2019, 5(4): 52-62.

1 引言

随着移动互联网的迅速发展,车联网已经成为汽车产业的必然趋势。它是由行驶中的车辆建立的分布式、自组织的通信网络。通过在车辆装置车载通信单元并在道路周围布设路边基础设施,使车辆之间以及车辆和路边基础设施可以安全快捷交换信息,车联网作为未来智能交通系统的重要组成部分,旨在达到改善车主驾驶体验,提高交通效率的目的。例如,当乘客处在出租车司机服务区域内时,出租车司机可以被提醒,或车辆在驾驶的过程中,使用电子地图,规划行驶路线,获得良好的驾驶体验。

虽然车联网为行驶的车辆有效改善了道路安全 and 提高了驾驶者的驾乘环境,但同时也带来了诸多问题隐患,如轨迹隐私、位置服务提供商(LBSP, location-based service provider)或可访问用户轨迹数据的第三方组织通过车辆连续不断的位置更新,可以跟踪某个车辆的行驶轨迹,分析车辆驾驶者的驾驶习惯,如常去的地方、工作地点、家庭住址等,进而揭示出用户的身份信息,给用户的隐私带来极大的危害。

针对上述车联网轨迹隐私泄露问题,学术界涌现了大量的解决方案,大致可以划分为3类:基于混合区的轨迹隐私保护方法、基于噪声数据的轨迹隐私保护方法和基于路径混淆的轨迹隐私保护方法。其中,基于路径混淆的轨迹隐私保护方法的基本思想为:通过地理位置邻近的车辆合作生成若干个虚假查询点,以至车辆轨迹出现交叉来混淆车辆在不同时间段的匿名轨迹之间的关联性,使攻击者无法分辨出车辆的完整轨迹,从而实现轨迹隐私保护,其具有简单易用,计算开销小,灵活,高效的优点。到目前为止,最典型的车辆混淆路径轨迹隐私保护方法是 Lim^[1]提出的相互混淆路径(MOP, mutually obfuscating

paths)。如图1所示,车辆A、B在 t_0 时刻到达路口时,如果车辆A计算出车辆B的虚假轨迹和车辆A的真实轨迹在 $t_0+\alpha$ 时刻相交,车辆A给车辆B发送一条请求路径混淆消息,车辆B收到请求消息时,验证车辆B的真实轨迹和车辆A的虚假轨迹能否在 $t_0+\beta$ 时刻相交,如果可以,则车辆B向车辆A发送确认信息。

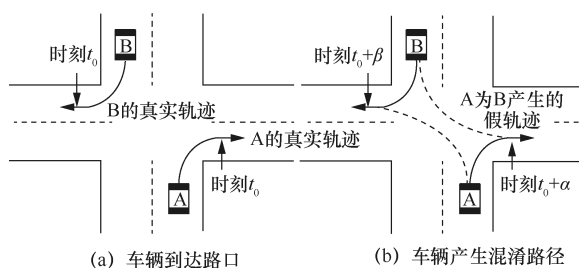


图1 相互混淆路径 MOP

此后,车辆A、B在请求基于位置的服务时各自向LBSP一次建立两个TCP连接,分别发送包含自身车辆真实位置、配对车辆虚假位置的两个查询。当自身车辆真实轨迹和配对车辆虚假轨迹重合时,目标车辆只发送包含自身位置的查询,但此方案存在以下2个缺点。

1) 在每个路口一部车辆只能与周边另外一部车辆相互混淆产生一条虚假轨迹,浪费多部车辆相互协作制造多条虚假轨迹的机会,而且车辆的行驶路径具有社会属性,即80%车辆会经过20%的路口,造成此方案不适用于车辆稀疏的农村地区。

2) 车辆需要在每个路口执行相互混淆路径,增加了车辆执行相互混淆路径的计算、存储、通信等开销,忽视了系统开销和隐私保护平衡问题。

为了解决上述方案缺点,本文提出了一种基于车辆协作的混淆路径轨迹隐私保护机制,本文的主要贡献如下。

1) 定义了用于量化车辆轨迹隐私保护度的轨迹熵,车辆轨迹熵达到自定义轨迹隐私保护阈值后,减少车辆混淆,解决了车辆路径混淆系统

开销和隐私保护的平衡问题。

2) 提出了车辆协作的路径混淆轨迹保护算法,增加了车辆在路口路径混淆机会,提高了车辆轨迹隐私保护度,使本方案在不同车辆密度道路环境具有高效的轨迹隐私保护效果。

2 相关工作

2.1 基于混合区的轨迹隐私保护方法

基于混合区的轨迹隐私保护方法基本思想是用户通过交换使用多个无映射关系的假名代替真实身份请求 LBS,将轨迹进行垂直分割,割裂不同时间间隔轨迹关联关系,从而达到保护位置隐私的目的。

Beresford^[2]首次提出了 Mix-zone 概念并将 Mix-zone 方案应用于位置隐私保护,其基本思想是将地图划分为应用区域、混淆区域,当用户处于应用区域时可正常使用 LBS,但是当用户处于混淆区域时,用户应停止使用 LBS,并和区域内成员交换假名,当用户离开混合区域时,用户开始使用新的假名请求 LBS。Freudiger^[3]为了解决攻击者对混合区域进行监听的问题,提出了适用于车辆自组织网络的 mix-zone 方案,该方案结合了 mix-network 的思想,提出了加密混合区(CMIX, cryptographic mix-zones)方案。针对 mix-zone 方法容易遭受时间推理攻击和转移概率推理攻击的问题,Palanisamy^[4]和 Palanisamy^[5]提出了 MobiMix 利用多边形及其组合构造混合区域,提高移动用户轨迹保护程度。攻击者可以凭借大数据环境下的辅助信息轻松地推测出一个 mix-zone 新假名和旧假名的对应关系,因此一个 mix-zone 并不能为基于位置服务的用户提供充分的轨迹隐私保护。只有在现实环境中同时构建合适数量的 mix-zone 才能高效保护 LBS 用户轨迹。又因在 mix-zone 中用户必须停止使用基于位置的服务,为保证用户的服务质量将限制部署的

mix-zone 数量。针对这一矛盾问题,Freudiger^[6]提出了一种基于组合优化多混合区部署方法。LIU^[7]考虑了路网和车流量限制,提出了一种基于整数规划多混合区部署方法,力图在部署有限个混合区的前提下,尽可能地提高 LBS 用户的轨迹保护程度。

2.2 基于噪声数据的轨迹隐私保护方法

基于噪声数据的轨迹隐私保护方法^[8-9]的思想是将 LBS 用户的真实位置发送 LBSP 同时,按照合理的策略生成若干个包含虚假位置的伪查询,并发送给 LBSP 作为 LBS 用户真实查询的“掩护”,使攻击者无法区分出 LBS 用户的真实位置。文献[10-12]针对连续查询中的伪查询添加机制进行了探索。Kido^[10]根据上一时刻 LBS 用户位置,按随机速度和方向生成位置点,并作为虚假位置进行发布,但这种方案生成的虚假位置往往与真实的移动特征不符合,并且这些虚假位置本身可能是现实环境上不可达的位置点。针对这一问题,Suzuki^[11]在生成虚假位置时考虑了速度、路网等限制条件。Kato^[12]认为 LBS 用户不会一直连续移动,所以 Kato 在生成噪声位置时会让 LBS 用户根据周围环境随机地产生若干个停顿点来防止攻击者辨别出噪声位置点。然而在现实世界中,LBS 用户移动容易遇到各种突然事件,如何处理这些突然事件,以生成更“真实”的噪声轨迹数据仍是一个巨大的挑战。

2.3 基于路径混淆的轨迹隐私保护方法

基于路径混淆的轨迹保护的基本思想是车辆在发送 LBS 请求前通过与物理位置邻近车辆协作,使在真实查询中上添加若干个虚假查询让用户行驶路径出现交叉,割裂不同时间段路径的映射关系,从而使攻击者无法分辨出目标车辆的完成轨迹。

Cui^[13]提出一种基于虚拟位置和路径混淆的车辆位置隐私保护方案,其方案是车辆在发送

LBS 请求时通过监听周边车辆的信标信息来挑选符合要求的邻近车辆作为车辆的“掩护”，被掩护车辆收到掩护车辆的位置信标消息后，保持消息中除被掩护车辆位置外的其他内容相同，向 LBSP 分别发送包含掩护车辆位置和被掩护车辆位置的两个不同请求，当 LBSP 回复请求后，被掩护车辆根据自己的真实位置丢弃对应掩护车辆的响应消息，从而得到属于基于自己位置的响应消息。但此方案并不能为 LBS 用户提供有效的轨迹隐私保障。例如，在一个只有两部车辆 A、B 的路口，两部车辆呈十字形行驶，一部车辆 A 在这个路口由于某种原因未选择周边的车辆作为影子而直接直行，另一部车辆 B 选择车辆 A 作为影子来制造假查询，两部车辆一直连续地查询基于位置的服务，然而攻击者轻而易举地识别车辆 B 生成的无效虚假轨迹。

Lim^[1]设计了一种有效的分布式车辆相互混淆方案，在方案中车辆依靠自身强大的计算和存储能力在路口选择速度相近的合作车辆，在使用基于位置的服务时，通过同时发送真实车辆真实位置查询和合作车辆虚假位置查询使轨迹相互混淆。此方案具有兼顾服务质量与用户隐私的优点，但仍存在没有充分利用路口多部车辆相互混淆的机会，且存在计算、存储、通信等开销过大的问题。

3 预备知识

3.1 基本概念

轨迹熵是排除虚假轨迹所需要知识的度量，表示 LBSP 或可访问用户轨迹数据的第三方组织对车辆真实轨迹的不确定程度；也用于量化车辆轨迹隐私保护度，车辆的轨迹熵越高，攻击者推断出其真实轨迹的概率越低，车辆的隐私保护性越好。

假设随机变量 X 表示一部车辆在路口真实轨迹、虚假轨迹情况，随机变量 X 元素个数依赖于

路口路段的数量，随机变量 X 对应的概率分布为

$$P = \{p_1, p_2, p_3, \dots, p_h\} \quad (1)$$

其中，在不考虑每条路段拥挤状况是否相同、车辆到达目的地最短路径等复杂情况下， P 服从均匀分布 U ，车辆在一个路口的轨迹熵 E_n 可以表示为

$$E_n = -\sum_{j=1}^h p_j \cdot \lg p_j \quad (2)$$

车辆经过 M 个路口在 Q 个路口执行路径混淆，车辆真实轨迹和虚假轨迹数目总数 S 最少为

$$S = \sum_{i=1}^Q 2^{E_i} - Q + 1 \quad (3)$$

车辆总轨迹熵 E_S 表示为

$$E_s = \lg S \quad (4)$$

3.2 威胁模型

本文假设 LBS 服务器是半可信的，即 LBS 服务器会严格按照协议提供用户所需要的服务，但 LBS 服务器可能通过关联分析的方式将用户的匿名轨迹与车辆关联起来，进而挖掘用户轨迹中蕴含的隐私信息。文中的攻击限定为 LBSP 或其他可访问用户轨迹数据的第三方组织。LBSP 为用户提供服务后可得到这些用户的匿名轨迹信息 $C(u_1) = \langle l_1, t_1 \rangle, \dots, \langle l_m, t_m \rangle$ ，其中 (l_i, t_i) 是序列 $C(u_1)$ 中的一个元素，表示用户 u_1 在时刻 t_i 处于位置 l_i 。如图 2 所示，用户 1 和用户 2 在使用 Google 地图导航时，LBSP 可得到用户 1 和用户 2 两个用户的匿名轨迹。

从用户 1 轨迹中 LBSP 或者其他可访问用户轨迹数据的第三方可以推断出车主本人或家属可能患有肿瘤，从用户 2 轨迹中可以推测出用户 2 可能信仰基督教。因此，基于位置的服务在给人们带来巨大便利的同时，也泄露了用户的位置信息、个性习惯、健康状况、政治倾向、社会地位等敏感信息，给用户构成巨大的威胁。

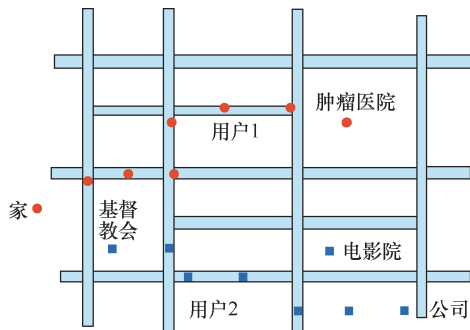


图 2 轨迹隐私推测

4 多路径混淆的车辆轨迹隐私保护

4.1 系统架构

本文提出的基于车辆协作的分布式轨迹保护机制，其系统架构如图 3 所示，主要由车载单元 (OBUS, on-board units)、路边单元 (RSU, road side unit)、互联网服务提供商 (ISP, Internet service provider)、基于位置服务提供商 (LBSP, location-based services provider) 4 部分组成。其各个组成部分的功能职责如下。

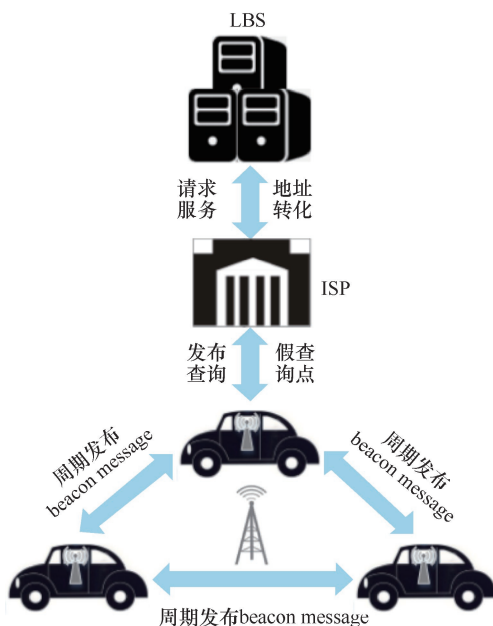


图 3 车辆自组织网络架构

1) 车载单元

车载单元是装载在移动车辆中处理单元，包括无线通信和数据分析等模块。车载单元是车辆

中最基础的组成部分，可以将车载单元看作无线通信系统中的移动终端。通过车载单元，车辆可以实时获取周边环境信息并且与周边车辆进行通信。

2) 路边单元

在车联网中，路边单元是部署在路边的基础设施，车联网除了能实现短距离通信外，还可以通过路边单元连接移动网络获得更加丰富的增值服务。

3) 互联网服务提供商

车辆通过移动通信网络访问 LBS，移动通信网络通过 NAT 技术为车辆提供了匿名保障，保护了车辆的身份隐私。

4) 基于位置服务提供商

LBSP 能提供各种类型基于位置的信息服务。例如，电子地图服务提供商可为用户提供实时交通导航，还可为用户提供实时路况信息，并规划最优线路。

车联网中移动通信分为车-车 (V2V, vehicle to vehicle) 和车与路边基础设施 (V2I, vehicle to infrastructure) 两个部分。在车-车通信中，车辆通过周期广播包括速度、位置、方向、加速度等信标信息给邻近车辆，从而为车辆协作提供了保障。在车-路边基础设施通信中，车辆根据自己的位置向邻近的 RSU 请求服务，RSU 起到一个中转站的作用，将消息转发给服务提供商。

4.2 总体方案

针对集中式单点瓶颈、可扩展性差的问题，本文在不改变 LBS 服务提供商已有的服务模式基础上，利用车辆自身的计算能力和存储能力，提出基于车辆协作的轨迹隐私保护机制。其基本思想是用于量化目标车辆轨迹隐私保护度的轨迹熵 E_S 未达到车辆自定义的隐私保护阈值 Z 时，目标车辆在路口根据周围车辆行驶情况建立混淆查询组 G ，预测并广播自身在 θ 时间将到达的位置 (X_θ, Y_θ) ，邻近车辆通过专用短程通信技术监听目标车辆的信标信息，在 θ 时间内能到达 (X_θ, Y_θ) 的邻

近车辆申请加入混淆查询组 G , 目标车辆按照在满足车辆轨迹隐私需求基础上, 尽可能降低车辆通信和计算代价的原则生成若干个虚假查询, 以至车辆轨迹出现交叉点来混淆车辆在不同时间段的匿名轨迹之间的关联性, 使攻击者无法分辨出车辆的完整轨迹, 实现轨迹隐私保护。

4.3 多路径混淆生成算法

本节将详细描述基于车辆协作的混淆路径轨迹隐私保护机制。如图4所示, 假设有 A 、 B 、 C 、 D 4 部车辆在 t_0 时刻到达路口, 4 部车辆各自执行混淆路径轨迹隐私保护机制。以车辆 D 为代表具体说明。车辆 D 在路口如果其轨迹熵小于个性化隐私需求, 即 $E_s < Z$, 车辆 D 则建立混淆查询组, 在建立之前, 车辆 D 计算出满足式(5)的混淆组成员数量 n , 若车辆轨迹熵 E_s 远小于个性化轨迹保护阈值 Z , 车辆根据周围车辆行驶情况, 按照在这个路口轨迹熵最大的原则来确定混淆组成员数量 n 。

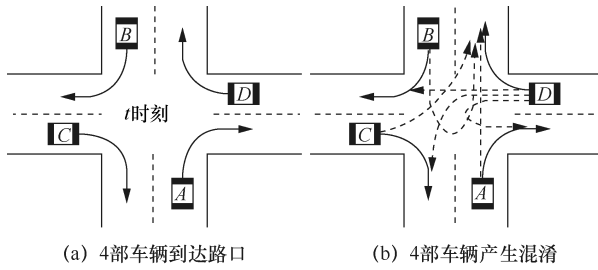


图4 车辆协作多路径混淆

$$E_s + E_n \geq Z \quad (5)$$

之后, 车辆 D 在路口建立一个 $n=3$ 的混淆组, 并广播建立组的四元组消息 $G=(W, V, S_U, T)$ 。其中 W 表示组号, V 表示车辆速度, S_U 表示加速度, T 表示时间。

临近车辆 A 、 B 、 C 根据式(6)计算其速度是否与车辆 D 的速度匹配。

$$|V_l - V| < \beta V \quad (6)$$

其中, V_l 表示申请加入组的车辆速度, V 表示建立组的车辆速度。 β 是预先限定的速度变化量。

如果临近车辆速度与车辆 D 匹配、方向与车辆 D 不同, 周围车辆向车辆 D 发送申请加入消息 $I=(W, ID, L, V, S_U)$, W 表示组号, ID 表示车辆标识符, L 表示车辆目前所在的位置, V 表示车辆的速度, S_U 表示车辆的加速度。

车辆 D 收到周围车辆 A 、 C 、 B 发送的消息后, 按照最早申请原则选择 3 辆汽车作为组成员, 并按照式(7)分别计算出车辆 A 、 C 、 B 的虚假轨迹与车辆 D 真实轨迹的相交时间段 $(T_{\min}, T_{\theta})$ 。

$$\frac{(X_N^U, Y_N^U) \rightarrow (X_{\theta}^D, Y_{\theta}^D)}{(1 \pm \beta)V_U} < \theta \quad (7)$$

其中, (X_N^U, Y_N^U) 表示组成员现在所在位置, $(X_{\theta}^D, Y_{\theta}^D)$ 表示车辆 D 在 θ 时间段后将到达的位置, $(X_N^U, Y_N^U) \rightarrow (X_{\theta}^D, Y_{\theta}^D)$ 表示组成员目前所在位置到 $(X_{\theta}^D, Y_{\theta}^D)$ 之间距离。 $(1 \pm \beta)V_U$ 表示组成员的速度。

车辆 D 计算出车辆 A 、车辆 B 和车辆 C 轨迹相交时间段的交集 $(T_{\min}^{se}, T_{\theta}^{se})$, 即

$$(T_{\min}^{se}, T_{\theta}^{se}) = (T_{\min}^A, T_{\theta}^A) \cap (T_{\min}^B, T_{\theta}^B) \cap (T_{\min}^C, T_{\theta}^C) \quad (8)$$

车辆 D 在 $(T_{\min}^{se}, T_{\theta}^{se})$ 内计算出距离车辆 A 、 B 和 C 距离之和最小的点, 即

$$D_i = (X_N^A, Y_N^A) \rightarrow (X_i, Y_i) + (X_N^B, Y_N^B) \rightarrow (X_i, Y_i) + (X_N^C, Y_N^C) \rightarrow (X_i, Y_i) \quad (9)$$

$$D_{op} = (X_N^A, Y_N^A) \rightarrow (X_{op}, Y_{op}) + (X_N^B, Y_N^B) \rightarrow (X_{op}, Y_{op}) + (X_N^C, Y_N^C) \rightarrow (X_{op}, Y_{op}) \quad (10)$$

$$D_i \geq D_{op} \quad (11)$$

其中, (X_i, Y_i) 表示车辆 D 在 $(T_{\min}^{se}, T_{\theta}^{se})$ 内任意可能到达的位置点, $(X_N^A, Y_N^A) \rightarrow (X_i, Y_i)$ 表示车辆 A 目前所在位置 (X_N^A, Y_N^A) 到 (X_i, Y_i) 位置点之间的距离。 (X_{op}, Y_{op}) 表示车辆 A 、 B 、 C 的虚假轨迹与车辆 D 真实行驶轨迹相交一点时距离之和最短的位置点, D_{op} 表示车辆 A 、 B 、 C 从目前所在位置 (X_N^U, Y_N^U) 到最优位置点 (X_{op}, Y_{op}) 的距离。

位置点 (X_{op}, Y_{op}) 作为车辆 A 、 B 、 C 虚假轨迹

与车辆 D 真实轨迹最优相交位置点。当车辆 A 、 B 、 C 、 D 以速度 V_A 、 V_B 、 V_C 、 V_D 行驶 T_{OP} 时间段后同时到达 (X_{OP}, Y_{OP}) 。

$$V_A = \frac{(X_N^A, X_N^A) \rightarrow (X_{OP}, Y_{OP})}{T_{OP}} \quad (12)$$

$$V_B = \frac{(X_N^B, X_N^B) \rightarrow (X_{OP}, Y_{OP})}{T_{OP}} \quad (13)$$

$$V_C = \frac{(X_N^C, X_N^C) \rightarrow (X_{OP}, Y_{OP})}{T_{OP}} \quad (14)$$

车辆 D 在为组内成员生成混淆查询点关键在于相邻两次虚假查询点之间的距离比例等于 $V_A:V_B:V_C:V_D$ 。更进一步地说, 假设在 t_1 时刻车辆 D 在 $(X_{t_1}^D, Y_{t_1}^D)$ 位置发布查询请求, 在 t_0 时刻发布查询时, 车辆 D 位于 $(X_{t_0}^D, Y_{t_0}^D)$, 车辆 D 在为车辆 A 生成假查询点, 应按照式(15)算出车辆 A 在 t_1 时刻假查询点位置 $(X_{t_1}^A, Y_{t_1}^A)$, 构造混淆轨迹, 车辆 D 为组内的其他成员生成假查询点, 构造混淆轨迹的方法如同为车辆 A 构造虚假轨迹。

$$\frac{(X_{t_0}^D, Y_{t_0}^D) \rightarrow (X_{t_1}^D, Y_{t_1}^D)}{(X_{t_0}^A, Y_{t_0}^A) \rightarrow (X_{t_1}^A, Y_{t_1}^A)} = \frac{V_D}{V_A} \quad (15)$$

综上所述, 本文在算法 1 中给出了具体的多路径混淆算法伪代码, 对应的详细算法流程如图 5 所示。

算法 1 多路径混淆生成算法

输入 轨迹隐私保护阈值 Z

输出 假查询位置序列 L_{array}

- 1) 计算车辆的轨迹熵 E_s
- 2) if ($E_s \geq Z$)
- 3) return(“车辆自身真实查询点”)
- 4) else{
- 5) 根据式(5)计算混淆组成员数量 n
- 6) 组领导车辆创建混淆组并广播组消息
- 7) 周围车辆申请加入混淆组
- 8) 组领导车辆计算组成员虚假轨迹和组领

导车辆真实轨迹相交时间段 $(T_{min}^e, T_{\theta}^e)$

9) 组领导车辆根据 $(T_{min}^e, T_{\theta}^e)$ 计算出组成员虚假轨迹与组领导车辆真实轨迹的相交的最近位置点 (X_{OP}, Y_{OP})

10) 组领导车辆计算出从目前位置到 (X_{OP}, Y_{OP}) 这段路程中组成员的速度比

11) 组领导车辆将根据速度比计算出假查询点位置 (x, y) 并存储 (x, y) 到 L_{array} 。

12) 返回 L_{array}

13) }

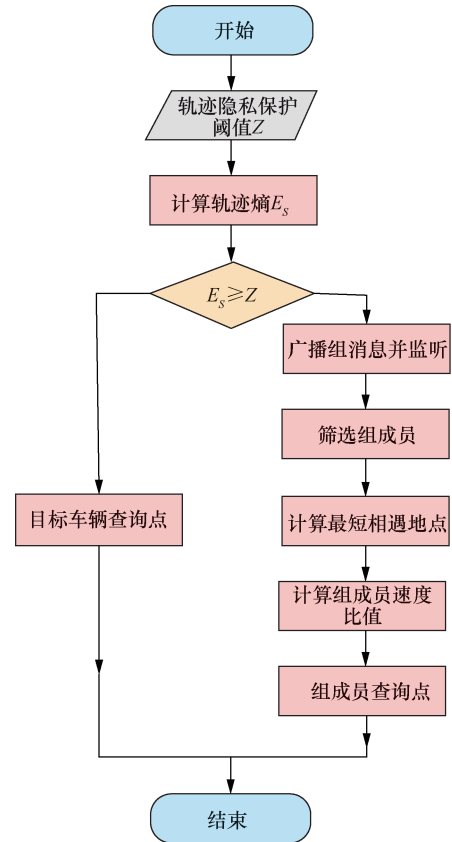


图 5 算法流程

4.4 安全性分析

车辆采用本文所提出的方案生成假轨迹时, 申请加入周围车辆所创建的组生成的假轨迹可定义为入度 r , 车辆自身创建的组生成的假轨迹可定义为出度 s , 则车辆的轨迹隐私保护程度由 $\min(r, s)$ 决定 ($\min(r, s)$ 是车辆出度、入度的最小

值)。更具体地说,如图 6(a)所示,此时车辆 B 出度为 2 (轨迹 1、轨迹 3)、入度为 2 (轨迹 2、轨迹 4),车辆 B 的 $\min(r, s)$ 值为 2,攻击者猜测出车辆 B 真实轨迹的概率为 $P = \frac{1}{3}$ 。如果车辆 B 考虑到生成假轨迹的成本问题,不愿意为其他车辆产生假轨迹,如图 6(b)所示,此时车辆 B 的出度为 1、入度为 2,则车辆 B 的 $\min(r, s)$ 为 1,车辆 B 的真实轨迹被攻击者识别的概率为 $P = \frac{1}{2}$ 。因此车辆在路口必须在车辆自身创建的组生成不少于入度的虚假轨迹,才能保证车辆个性化轨迹隐私需求。从根本上避免了路口自私车辆为了节省通信成本和计算成本不为周边车辆产生虚假轨迹的问题。

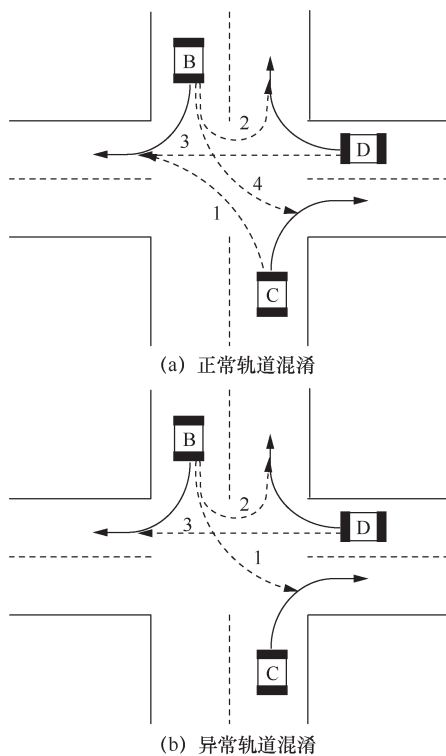


图6 轨迹安全分析

在生成虚假轨迹的过程中,车辆通过速度计算选择速度相近的组加入,从而可以防止攻击者通过速度推断出车辆的真实轨迹。另外,一个组的虚假轨迹中查询点个数相同,查询频

率相同,并且虚假轨迹相交于一点,使攻击者无法区分组的创建者,从而无法识别车辆的真实轨迹。

因此,本文基于假轨迹的轨迹保护机制是一种高效的轨迹隐私保护方案。

5 实验及仿真

5.1 基本设置

为了验证本方案的有效性和高效性,本文通过交通仿真工具 sumo 与网络仿真工具 NS-3 联合实施实验仿真。使用交通仿真工具 sumo 在山东省济南市以北纬 $36^{\circ}40'$ 东经 117° 为中心生成 $6\text{ km} \times 6\text{ km}$ 道路拓扑结构图,分别收集了车辆数目 $n=30$ 、 $n=60$ 、 $n=120$ 和 $n=240$ 在 $6\text{ km} \times 6\text{ km}$ 道路拓扑结构图上行驶 10 min 的车辆运动轨迹文件。在网络仿真 NS-3 节点中部署了基于位置服务的车载应用程序和车辆通信协议,并设置车辆每隔 0.3 s 发送更新位置请求,节点将按照交通仿真工具 sumo 生成的轨迹进行移动来完成实验仿真。实验环境为 lenovo m415, i5 处理器, 8 GB 内存, 操作系统为 fedora 12。

实验涉及的参数设置如下。1) Z 为车辆自定义轨迹隐私保护阈值与轨迹熵值息息相关,设置为 (3 bit, 8.5 bit) 范围内随机值。2) λ 为车辆选择时速度差量系数,默认 $\lambda=0.2$ 。

5.2 实验结果

本节比较了所提方案和 Lim 提出的 MOP 方案在车辆数目 $n=30$ 、 $n=60$ 、 $n=120$ 和 $n=240$ 时车辆行驶 10 min 轨迹熵的整体分布情况,如图 7 所示,从图中可以清晰地观察到本文方案车辆轨迹熵明显分布在 MOP 方案的轨迹熵的上部。这表明了本方案在车辆稀疏的农村、乡镇地区达到的轨迹隐私保护程度大致相当,但在车辆稠密的城市地区本方案的执行效果显著优于 MOP 方案。

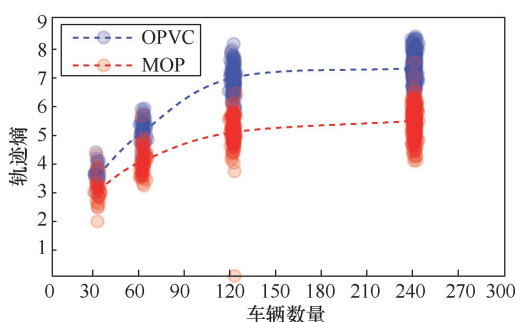


图7 轨迹熵整体分布

在图8中给出了本文方案在不同车辆密度的情况下车辆 10 min 内的轨迹熵变化, 同时也在图9中给出了MOP方案车辆在10 min内的轨迹熵随时间的变化情况作为参考。从图中可以看到, 车辆密度 $n=30$ 时, 本文方案的轨迹熵和MOP方案的轨迹熵大致相同, 这是由于分布在 $6\text{ km} \times 6\text{ km}$ 区域内的车辆大多数的情况下是两部车辆同时到达路口, 只有少数情况下会有超过3部车辆同时出现在交叉口。然而, 车辆密度 $n=60$ 时, 本文的轨迹熵值在10 min达到5 bit, 由轨迹熵的定义可知, 轨迹熵为 X bit 代表车辆从起点行驶到目前所在位置的过程中至少生成了 $2X-Q+1$ 条虚假轨迹。更加具体地说, 攻击者跟踪轨迹熵值为5 bit 汽车的行驶轨迹时至少需要猜测32条轨迹, 这比MOP方案相同车辆密度下多了16条虚假轨迹, 这充分说明了本文方案比MOP方案在保护车辆轨迹隐私做得更好。从图中可以观察到车辆密度为 $n=120$ 和 $n=240$ 的情况下更加证实这一点, 在车辆密度 $n=120$ 和 $n=240$ 时, 轨迹熵分别在10 min达到6.9 bit、7.23 bit 比MOP方案相同车密度下多1.9 bit、1.78 bit。另外, 在4 min之前本文方案轨迹熵的增长速率快, 这表明本方案能够在很短的时间内达到良好的轨迹隐私保护效果, 由于本文方案考虑了部分车辆轨迹熵达到了自己的轨迹隐私保护程度, 在短时间内停止执行生成假轨迹, 在4 min之后增长速率减缓一些, 这既可以节省车辆的计算、电量等资源, 又可以保证车辆的轨迹熵值满足个性化的需要。

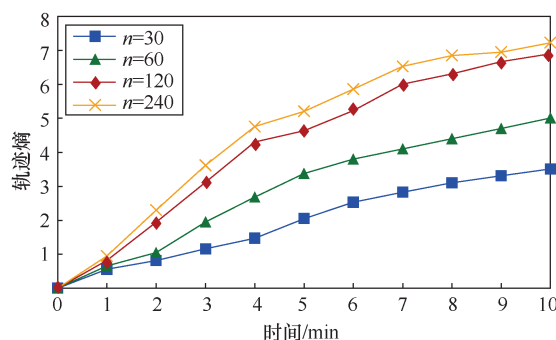


图8 OPVC方案轨迹熵变化

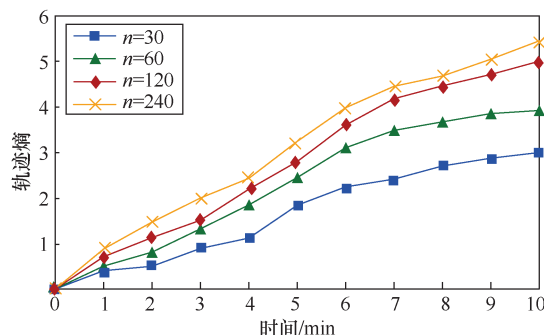


图9 MOP方案轨迹熵变化

图10和图11从轨迹跟踪成功率的角度分别给出了本文方案和MOP方案在不同车辆密度下的车辆轨迹隐私保护效果。从图中可以观察到两种方案在1 min内的轨迹跟踪成功率大致相同, 这是由于车辆在 $6\text{ km} \times 6\text{ km}$ 区域内的起点分散, 在1 min内多部车辆相遇可行性很小, 这也对应着现实世界车辆起始位置分散情景。

由于本文方案在未达到车辆个性化轨迹隐私保护程度时, 尽可能在路口多生成不同方向的虚假轨迹, 而MOP方案在每个路口只生成一条虚假轨迹, 故而在1 min到4 min本方案轨迹跟踪率比MOP方案轨迹跟踪率下降迅速, 这正好对应车辆在1 min到4 min轨迹熵的变化趋势。从图10车辆密度在(60, 240)区间的轨迹跟踪率曲线不难看出, 本文方案在5 min轨迹跟踪率下降到0.1, 比MOP方案早1 min, 这充分说明了本文方案具有高效率。同时, 在10 min即使车密度 $n=30$, 轨迹跟踪率下降到了0.0877, 这进一步表明本文方案在车辆轨迹隐私保护方面具有良好的效果。

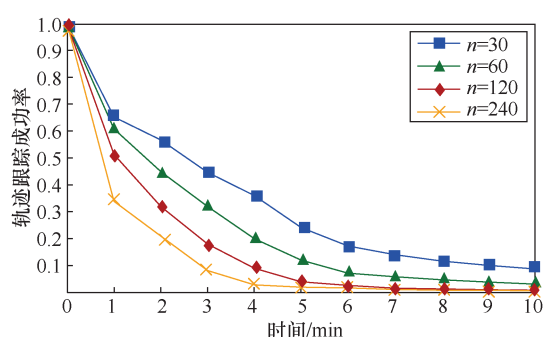


图10 OPVC 方案轨迹成功跟踪率

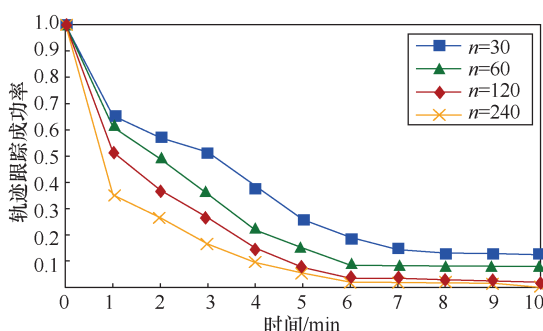


图11 MOP 方案轨迹跟踪成功率

最后,对本文方案和现存在的比较流行的轨迹隐私保护方案做对比分析,比较结果如表1所示。

表1 各方案对比

方法	轨迹跟踪率	位置精度	盲区
OPVC	3%	高	无
Mix-zone	6.7%	高	有
Mix-group	3%	高	无
Silent period	10%	高	有

在基于混合区域的方案中,一般选择车辆比较密集的路口,这样会造成缺乏灵活性,同时在混合区中不能实时更新位置,造成了盲区。在基于静默时间的轨迹隐私保护方案中,由于在持续的一段时间进行假名交换,不能进行实时位置更新,这也造成了盲区。基于混合组的轨迹隐私保护是一种比较先进的假名交换和组签名的相结合的轨迹隐私保护方法,但算法复杂,方案的执行效率低。综合以上对比分析,本文提出的方案相比现存在的大多数轨迹隐私保护方案而言,具有

良好的轨迹隐私保护效果,位置精确,满足个性化轨迹保护需求,无盲区的优势。

6 结束语

车联网作为物联网在交通系统的智能系统,在缓解交通拥挤和提高驾驶体验方面有着巨大的贡献,但安全隐私问题是车联网系统迫切需要解决的问题,如果不能实现安全隐私保护,车联网是无法被广泛普及的。本文针对车辆通过和第三方共享位置信息以获得基于位置的服务时面临轨迹泄露的问题,针对该问题,本文从两个角度展开研究,首先研究如何利用车辆的存储能力和计算能力,在不依赖于第三方设备的提前下,保护车辆的轨迹隐私;其次由于车辆的计算能力和存储能力仍然相对有限,本文将研究兼顾系统开销与用户隐私的平衡;最后提出一种用户轨迹度量方法,定量计算车辆的轨迹隐私保护程度。仿真对比实验分析表明了本文方案的有效性和高效性。

参考文献:

- [1] LIM J, YU H, KIM K, et al. Preserving location privacy of connected vehicles with highly accurate location updates[J]. IEEE Communications Letters, 2017, 21(3):540-543.
- [2] BERESFORD A, STAJANO F. Location privacy in pervasive computing[J]. IEEE Pervasive Computing, 2003, 2(1):46-55.
- [3] FREUDIGER J, RAYA M, FELEGYHAZI M, et al. Mix-zones for location privacy in vehicular networks[C]// Win-ITS 07.2007.
- [4] PALANISAMY B, LIU L. MobiMix: protecting location privacy with mix-zones over road networks[C]// IEEE 27th International Conference on Data Engineering(ICDE 2011).2011: 494-505.
- [5] PALANISAMY B, LIU L. Attack-resilient mix-zones over road networks: architecture and algorithms[J]. IEEE Transactions on Mobile Computing, 2015,14(3):495-508.
- [6] FREUDIGER J, SHOKRI R, HUBAUX J P. On the optimal placement of mix zones[C]// The 9th International Symposium on Privacy Enhancing Technologies. Springer Berlin Heidelberg,2009.
- [7] LIU X, ZHAO H, PAN M, et al. Traffic-aware multiple mix zone placement for protecting location privacy[J]. Proceedings of the IEEE Infocom, 2012, 131(5):972-980.
- [8] NIU B, LI Q, ZHU X, et al. Achieving k -anonymity in privacy-aware location-based services[C]// IEEE Infocom. 2014: 754-762.
- [9] NIU B, LI Q, ZHU X, et al. Enhancing privacy through caching in location-based services[C]// IEEE Conference on Computer Com-

munications. 2015:1017-1025.

- [10] KIDO H, YANAGISAWA Y, SATOH T. Protection of location privacy using dummies for location-based services[C]// International Conference on Data Engineering. Tokyo: IEEE Computer Society, 2005:1248.
- [11] SUZUKI A, IWATA M, ARASE Y, et al. A user location anonymization method for location based services in a real environment[C]//Sigspatial International Conference on Advances in Geographic Information Systems. 2010:398-401.
- [12] KATO R, IWATA M, HARA T, et al. A dummy-based anonymization method based on user trajectory with pauses[C]// International Conference on Advances in Geographic Information Systems. 2012:249-258.
- [13] CUI J, WEN J, HAN S, et al. Efficient privacy-preserving scheme for real-time location data in vehicular ad-hoc network[J]. IEEE Internet of Things Journal, 2018, 5(10): 3491-3498.



叶阿勇（1977-），男，福建漳州人，博士，福建师范大学教授，主要研究方向为信息隐私与安全。



金俊林（1994-）男，江苏盐城人，福建师范大学硕士生，主要研究方向为手机端数据隐私保护。

[作者简介]



赵子文（1992-）男，山东枣庄人，福建师范大学硕士生，主要研究方向为车联网数据隐私与安全。



孟玲玉（1994-），女，黑龙江安达人，福建师范大学硕士生，主要研究方向为位置隐私保护。