轨迹隐私保护研究综述

许志凯，张宏莉，余翔湛

(哈尔滨工业大学 计算机科学与技术学院，哈尔滨 150001)

摘 要： 随着智能终端的普及和无线通信技术的发展，基于位置的服务已渗入到人们的日常生活当中。这些服务在给人们的日常生活带来便利的同时，也带来隐私泄漏的风险。针对轨迹数据的推理攻击不仅可分析出目标用户的家庭住址、工作地点等敏感位置信息，甚至可推测出用户的生活习惯、健康状态、宗教信仰等隐私信息。轨迹隐私能否得到妥善保护已成为制约移动互联网发展的瓶颈问题。本文对已有的轨迹隐私保护方法进行了分类描述，并分析已有工作的优缺点，最后指明未来的研究方向。

关键词： 轨迹隐私；隐私保护；位置隐私；网络安全

中图分类号： TP391.41 **文献标志码：** A **文章编号：** 2095-2163(2017)01-0125-03

Survey on trajectory privacy protection techniques

XU Zhikai, ZHANG Hongli, YU Xiangzhan

(*School of Computer Science and Technology, Harbin Institute of Technology, Harbin* 150001, *China*)

Abstract: With the rapid development of GPS-enabled mobile devices and wireless communication technology, location-based services (LBS) have become an essential part of daily life. However, with trajectory information, an adversary can easily infer several facets of users' lifestyles, such as living habit, health conditions, exercise hobbies, and religious belief, beyond just the locations. The potential abuse of trajectory information by unauthorized entities is evolving into a serious concern in mobile internet. The paper analyzes the existing trajectory privacy protection techniques, and puts forward the future research works.

Keywords: trajectory privacy; privacy protection; location privacy; network security

0 引 言

随着智能终端的普及和无线通讯技术的发展，基于位置的服务(Location-based Service，LBS)已渗入到人们的日常生活当中。然而，许多基于位置的服务，如电子地图、运动计步、移动广告，需用户实时提交自己的位置信息。这些服务可为人们的生活带来巨大的便利。以电子地图服务提供商Google地图、百度地图为例，这些应用不仅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。然而，这些服务也带来隐私泄漏的风险，在使用这类服务时，LBS用户需实时地将自己的位置信息提交给LBS服务器，但这些轨迹数据往往含有丰富的时空信息。针对轨迹数据的推理攻击不仅可得出用户在什么时间去过什么位置，还可分析出目标用户的家庭住址、工作地点等敏感位置信息，甚至可推测出用户的生活习惯、健康状态、宗教信仰等隐私信息。因此，轨迹隐私保护受到用户及研究者的广泛关注。

针对上述问题，本文介绍基于位置的服务，在此基础上分别综述位置隐私保护技术的主要研究现状及存在的问题，同时，根据目前研究的不足指出未来可能的研究方向。

基金项目：国家重点基础研究发展计划("973"计划)(2011CB302605，2013CB329602)；国家自然科学基金(61202457，61402149)。

作者简介：许志凯(1988-)，男，博士研究生，主要研究方向：隐私保护；张宏莉(1973-)，女，博士，教授，博士生导师，主要研究方向：网络安全、网络测量、网络计算等；余翔湛(1973-)，男，研究员，博士生导师，主要研究方向：网络安全、网络测量、并行计算等。

收稿日期：2016-05-26

1 基于位置的服务

图1表示了基于位置服务的一般架构，该架构包含3个实体：

1) 为LBS用户提供定位服务的导航定位基础设施，主要包括GPS卫星、无线网络基站、WIFI等。

2) 持有移动智能终端的LBS用户(本文的研究中将LBS用户与移动智能终端可视为同一主体)。移动智能终端可通过硬件(如GPS芯片)和软件(如基站信号定位、WIFI指纹定位)技术确定该LBS用户所在地理位置，并通过无线信号与LBS服务器进行通信。

3) 为LBS用户指定基于位置服务的服务提供商，如百度地图、Google地图、大众点评等。

连续型LBS服务指的是用户需实时提交的自己的位置信息才能获取到相应服务的LBS服务，这类服务主要包括智能导航服务、无人驾驶汽车、基于位置的新闻(广告)推送、运动计步及某些社交类APP(如定位附近与我兴趣相同的人)等。以智能导航服务为例，一次典型的连续型LBS服务如图1所示。具体可做如下阐释：

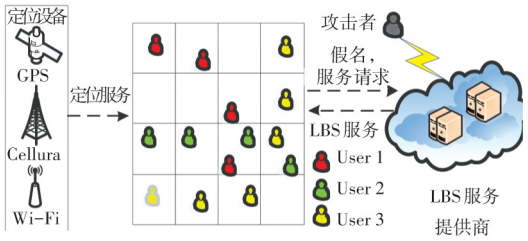


图 1 基于位置服务的架构

Fig. 1 LBS system model

1) 用户通过定位服务如 GPS 等获取自身位置。

2) 用户将自身位置及所需的服务请求发送给 LBS 服务提供商。

3) LBS 服务提供商根据用户当前的位置及道路拥塞情况为用户规划行驶线路，并提供导航服务。

4) 重复过程 1) 至 3) 直到满足用户的服务要求，如完成一次从地点 A 到地点 B 的导航服务。

2 威胁模型

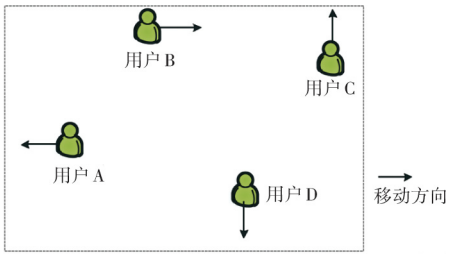
在轨迹隐私保护的研究中，研究者一般认为 GPS 等定位设备是可信的，即用户可获取其位置坐标的过程是安全的，而 LBS 服务提供商是不可信的，即 LBS 服务提供商可能会利用用户的轨迹信息挖掘用户的隐私信息。这是因为用户在将位置数据提交给 LBS 提供商后没有能力验证服务提供商是否可信，其次可信的服务提供商也可能会被恶意第三方攻击，导致用户位置隐私的泄漏。

3 轨迹隐私保护技术

近些年轨迹隐私保护受到研究者的广泛关注。根据用户的查询请求在到达 LBS 服务器之前变换方式的不同，当前的轨迹隐私保护方法可分为以下 3 类，分别是：基于 k 匿名泛化的方法、基于噪声数据的方法、基于动态假名的方法。在此，将针对各类方法给出如下研究阐释与分析。

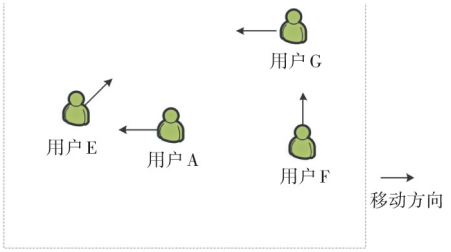
3.1 基于 k 匿名泛化的轨迹隐私保护方法

k 匿名泛化法[1]是一种经典的位置隐私保护技术。其基本思想是：在发送服务请求时，以一块空间区域代替用户的精准位置，即通过降低用户位置的精度的方式满足用户的隐私需求。部分研究者[2-3]将 k 匿名泛化应用到轨迹隐私保护中。一个直观的思路是将连续的 2 次查询视为 2 个独立的 LBS 服务请求，即分别为这 2 次查询构建匿名空间。但这种方法易受基于用户移动速度的推理攻击。针对上述问题，文献[2]基于用户移动速度构建匿名区域，以保证 2 个连续提交的匿名区域在速度上可达。然而，这一方法并不能充分保证用户的位置隐私。如图 2 所示，设 A 正在沿道路行驶，并在实时地查询周边的东北菜馆。图 a) 是用户 A 在时刻 t_i 生成的匿名空间，图 b) 是用户 A 在时刻 t_{i+1} 生成的匿名空间。由于用户的移动方向与速度并不会完全相同，2 块匿名空间中只包含用户 A 这一共同用户，即使 2 块匿名空间在速度上是完全可达的，攻击者也可通过比对 2 块匿名空间中的用户轻易地推测出用户 A 正在查询东北菜馆的信息。



a) t_i 时刻的匿名空间

a) Anonymous space at t_i moment



b) t_{i+1} 时刻的匿名空间

b) Anonymous space at t_{i+1} moment

图 2 基于 k 匿名泛化的位置隐私保护

Fig. 2 Spatial Cloaking for anonymous LBSs

综上问题论述开展研究，文献[3]认为在连续查询中，用户所提交的匿名区域应包含 k 个相同的用户。基于上述思想，有针对性地提出一种基于贪心的匿名区域构建算法，但通过该方法构造的匿名空间的面积会随着查询次数的上升而带来线性增加，从而形成严重的通信和计算开销。轨迹隐私保护需建立用户服务可用性的基础上，如何在保护用户轨迹隐私的同时，尽可能地减小匿名空间的面积，提高基于服务的可用性，是该类方法未来的研究重点。

3.2 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[4-5]的实现思想是：在将用户的真实位置发送给 LBS 提供商的同时，以一定策略生成若干虚假位置(dummy)并发送给 LBS 服务提供商作为用户真实位置的“掩护”，使攻击者无法分辨出用户真实位置。文献[6-8]针对连续查询中的噪声数据添加机制进行了研究。文献[6]根据上一时刻的位置，按随机速度和方向进行移动，并将获得的随机的位置点作为虚假位(dummy)进行发布，但这种方法生成的虚假位置点往往与用户真实的移动特征不符，且这些虚假位置点本身可能是一些实际上不可能到达的位置。针对这一问题，文献[7]在生成虚假位置点时加入了移动速度、路网等约束条件。文献[8]认为移动用户不会始终连续性移动，因此其在生成噪声数据点时会让移动对象根据周边的环境随机地产生一些停顿，以防止攻击者识别出噪声数据。然而，现实社会中，用户行驶过程易遇到各种意外事件，如何处理这些意外事件，以生成更“真实”的噪声轨迹数据仍是一个巨大的挑战。此外，上述方法对攻击者的背景知识假设较为保守，当攻击者获得了从用户的日常行为中提取的背景知识时，即使用户的生成的噪声轨迹无法模拟出真实用户的移动轨迹，攻击者也可辨别出用户的真实轨迹。

3.3　基于动态假名的轨迹隐私保护方法

基于假名的轨迹隐私保护方法的基本思想为：用户在发送基于位置的服务请求是以一个假名(pseudonym) 来代替用户的真实身份。然而，研究者发现长时间使用同一假名并不能有效地保护用户的隐私，这是因为在路网中移动用户是公开可见的，一次偶然的隐私的泄漏就可能会导致用户整个移动轨迹的泄漏。因此，研究者[3]提出基于 mix-zone 的动态假名技术。

当前大部分研究工作围绕如何构建单个 mix-zone 区域展开。文献[10]提出的 MobiMix 即是其中的代表性方法，对应基本思想为针对现实中可能被攻击者利用的背景知识，如移动速度，转移概率等，利用不规则的多边形及其组合建立混淆区域，使其适应当前场景，提高隐私保护程度。然而，在现实中只有同时部署多个混淆区域才能有效地保护用户的轨迹隐私，但在 mix-zone 内用户必须停止使用基于位置的服务，部署过多的 mix-zone 可能会严重影响服务质量。在此问题基础上，文献[11]提出一种基于组合优化的混淆区域部署方案。文献[12]进一步地考虑路网和交通流量的限制，提出一种基于整数规划的混淆区域部署方案，可在部署有限个混淆的区域的同时，尽可能提高用户的隐私保护水平。单个 mix-zone 的设计目前已经有较为成熟的方案，如何在城市环境下协调部署多个 mix-zone 则是该方向未来的研究重点。

4　结束语

随着移动互联网的发展，轨迹隐私保护受到研究者的广泛关注。本文对近些年来该方向已有的研究成果进行了回顾，对比和分析了已有的方法和技术，并指出仍然存在的问题和可能的研究方向。整体来说，目前轨迹隐私保护的研究仍然处于起步阶段，仍有许多关键性问题尚未解决。

参考文献：

[1] SWEENEY L. k-anonymity: a model for protecting privacy [J]. International Journal on Uncertainty, Fuzziness and Knowledge-based Systems, 2002, 10(5): 557-570.

[2] GHINITA G, DAMIANI M L, SILVESTRI C, et al. Preventing velocity-based linkage attacks in location-aware applications [C]//Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. Washington: ACM, 2009: 246-255.

[3] WANG Y, XU D, HE X, et al. L2P2: Location-aware location privacy protection for location-based services [C]//Proceedings of IEEE INFOCOM. Orlando, Florida: IEEE, 2012: 1996-2004.

[4] NIU B, LI Q, ZHU X, et al. Achieving k-anonymity in privacy-aware location-based services [C]//Proc. of IEEE INFOCOM. Toronto: IEEE, 2014: 754-762.

[5] NIU B, LI Q, ZHU X, et al. Enhancing privacy through caching in location-based services [C]//Proc. of IEEE INFOCOM. Hong Kong: IEEE, 2015: 1017-1025.

[6] KIDO H, YANAGISAWA Y, SATOH T. Protection of location privacy using dummies for location-based services [C]//Proc. of the 21st Int'l Conf. on Data Engineering. Tokyo: IEEE, 2005: 1248.

[7] SUZUKI A, IWATA M, ARASE Y, et al. A user location anonymization method for location based services in a real environment [C]//Proc. of the 18th ACM SIGSPATIAL Int'l Symp. on Advances in Geographic Information Systems. Sana Jose: ACM, 2010: 398-401.

[8] KATO R, IWATA M, HARA T, et al. A dummy-based anonymization method based on user trajectory with pauses [C]//Proc. of the 20th ACM SIGSPATIAL Int'l Conf. on Advances in Geographic Information Systems. Redondo: ACM, 2012: 249-258.

[9] FREUDIGERJ, RAYA M, FLEGYHAZI M, et al. Mix-zones for location privacy in vehicular networks [C]//WiN-ITS. Vancouver, British Columbia, Canada: ACM, 2007: 1-7.

[10] ALANISAMY B, LIU L. Attack-resilient mix-zones over road networks: architecture and algorithms [J]. IEEE Transactions on Mobile Computing, 2015, 14(3): 495-508.

[11] FREUDIGER J, SHOKRI R, HUBAUX J P. On the optimal placement of mix zones [C]//Proc. of the 9th International Symposium on Privacy Enhancing Technologies (PETS'09). Seattle: IEEE, 2009: 216-234.

[12] LIU X, ZHAO H, PAN M, et al. Traffic-aware multiple mix zone placement for protecting location privacy [C]//Proceedings of the IEEE INFOCOM 2012. Orlando, Florida: IEEE, 2012: 972-980.

(上接第 124 页)

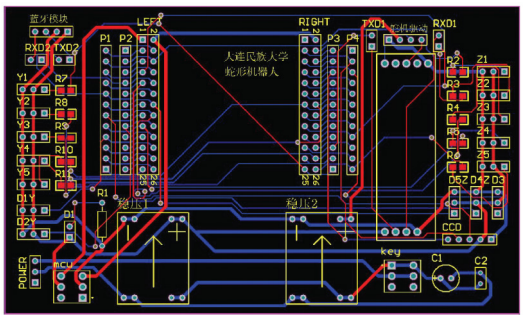


图 9　PCB 图

Fig. 9　The PCB diagram

3　结束语

蛇形机构的选择直接决定了其运动的功能。未来蛇形机器人结构设计和连接方式的开发将遵循通用性、经济性和鲁棒性的原则，且其结构设计也将沿着可重构性方向获得延拓升级发展。本文研制开发的蛇形机器人基本是用舵机串联而成，设计紧凑，动作连贯舒展，并可完成各种目标需求动作。

参考文献：

[1] GRAY J. The mechanism of locomotion in snakes [J]. Journal of Experimental Biology, 1946, 23(23): 101-124.

[2] 李斌. 蛇形机器人的研究及在灾难救援中的应用 [J]. 机器人技术与应用, 2003(3): 22-26.

[3] 任志敏. 基于 AVR 单片机的舵机驱动电路研究 [J]. 自动化技术与应用, 2008, 27(6): 85-87.

[4] 董晓坡，王绪本. 救援机器人的发展及其在灾害救援中的应用 [J]. 防灾减灾工程学报, 2007, 27(1): 112-117.

[5] 陈丽，王越超，李斌. 蛇形机器人研究现况与进展 [J]. 机器人, 2002, 24(6): 559-563.