doi:10.11959/j.issn.1000-436x.2016279

位置隐私保护技术研究进展

万盛[1]，李凤华[1,2]，牛犇[2]，孙哲[2]，李晖[1]

(1. 西安电子科技大学综合业务网理论及关键技术国家重点实验室，陕西 西安 710071；
2. 中国科学院信息工程研究所信息安全国家重点实验室，北京 100195)

摘 要：日趋流行的基于位置服务(LBS, location-based service)在为人们日常生活带来便利的同时也严重威胁到用户隐私。位置隐私保护技术逐渐成为研究热点，并涌现出大批研究成果。首先介绍位置隐私保护背景知识，包括位置服务应用场景、位置服务体系框架、隐私保护目标和系统构架；接着讨论 LBS 中的攻击者模型和隐私保护度量指标；然后对 4 种基于泛化和模糊的 LBS 隐私保护技术进行深入分析和总结；最后给出了未来 LBS 隐私保护技术潜在的研究方向。

关键词：位置服务；隐私保护；位置隐私；隐私度量；攻击者模型

中图分类号：TN929 **文献标识码**：A

Research progress on location privacy-preserving techniques

WAN Sheng[1], LI Feng-hua[1,2], NIU Ben[2], SUN Zhe[2], LI Hui[1]

(1. State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an 710071, China;
2. State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100195, China)

Abstract: While providing plenty of convenience for users in daily life, the increasingly popular location-based service(LBS) posed a serious threat to users' privacy. The research about privacy-preserving techniques for LBS is becoming a hot spot, and there are a large number of research results. First, background information of privacy protection for LBS was introduced, including application scenarios of LBS, the LBS framework, objects of privacy protection and system architectures of privacy protection. Second, adversary models and metrics for privacy protection in LBS was discussed. Third, four types of privacy-preserving techniques based on generalization and obfuscation for LBS were analyzed and summarized thoroughly. Finally, the potential research directions for privacy-preserving techniques for LBS in the future were shown.

Key words: location-based service, privacy protection, location privacy, privacy metrics, adversary model

1 引言

随着移动设备的普及和定位技术的发展，多样化的基于位置服务应用给人们的日常生活提供了便利，如寻找附近的餐厅或者银行、搜索从住所到公司的最短路线等。然而，LBS 为用户提供便利的同时也带来了隐私泄露的威胁[1]。通过收集用户发送的服务请求中的敏感信息，如位置或者兴趣点(POI, point of interest)等，服务提供商能够获取并推断出用户的更多隐私信息：1) 通过获取用户持续更

收稿日期：2016-08-12；**修回日期**：2016-10-10
通信作者：牛犇，niuben@iie.ac.cn
基金项目：国家自然科学基金—广东联合基金资助项目(No.U1401251)；国家高技术研究发展计划（“863”计划）基金资助项目(No.2015AA016007)；国家自然科学基金—青年科学基金资助项目(No.61502489)；国家“核高基”科技重大专项基金资助项目（No.2015ZX01029101）

Foundation Items: The National Natural Science Foundation of China—Guangdong Provincial People's Government of the Joint Natural Science Fund Projects (No.U1401251), The National High Technology Research and Development Program of China (863 Program)(No.2015AA016007), The National Natural Science Youth Science Foundation of China (No.61502489), The National Science and Technology Major Project of China (No.2015ZX01029101)

新的位置信息，可分析其出行规律[2]，预测未来所处位置；2) 利用某些时间段内用户位置信息的统计数据，可推断出用户家庭地址和单位[3]；3) 结合地图等背景知识，可推断出用户的健康状况、生活习惯及宗教信仰等[4]。这些隐私信息的泄露可能对用户造成难以估计的损失。因此，如何保证用户隐私信息已成为此类服务中亟待解决的问题。

由于隐私信息能够带来巨大的商业价值，现有研究普遍认为服务提供商不可信或为恶意攻击者[5]，隐私泄露威胁主要源于服务提供商或其工作人员。即使服务提供商可信，恶意第三方也可以通过各种手段获取服务器所掌握的部分甚至全部用户信息，如窃听无线信道或攻破并控制服务器。相对而言，用户设备端在不考虑病毒或木马入侵的情况下可信。LBS 隐私泄露问题大致分为 2 类：位置隐私(location privacy)泄露和查询隐私(query privacy)泄露。位置隐私指与用户位置信息相关的敏感信息，包括实时位置，可以帮助攻击者准确定位用户；用户访问过的位置，统计这些位置可以帮助攻击者了解该用户的基本信息（家庭地址和工作单位）；由位置信息推断出的其他敏感信息。查询隐私指与查询内容相关的敏感信息，查询内容包括用户爱好和需求等信息。

针对上述 2 类隐私泄露问题，近年来，学者们提出了大量解决方案，取得了一定进展。例如，通过加密技术保护隐私信息、使用假名代替真实用户名、在位置信息中加入噪声，以及发布模糊区域信息取代具体位置坐标等，这些技术在一定程度上降低了攻击者识别用户隐私信息的能力，但仍不同程度地存在不足。

本文关注 LBS 隐私保护最常用的 4 类技术，分别对每一类进行深入分析。首先介绍 LBS 隐私保护背景知识，包括位置服务应用场景、体系框架、隐私保护目标和系统构架，接着阐述攻击者模型和隐私保护度量指标，然后深入讨论和分析空间隐匿(spatial cloaking)技术、位置偏移和模糊(perturbation and obfuscation)技术、伪造虚假位置(dummy)技术和群组协作(collaborative group)技术，最后总结现有技术并展望未来发展。

2　LBS 隐私保护背景知识

2.1　位置服务应用场景

目前，位置服务在人们的日常生活中已经得到了广泛的应用，本文关注常用的 5 类应用场景：基于位置的 POI 检索服务、基于位置的精确导航服务、基于位置的社交网络服务、基于位置的运动检测服务和基于位置的广告推送服务。

1) 基于位置的 POI 检索服务

Yelp、大众点评和百度糯米等都提供基于位置的 POI 检索服务，用户使用该类软件可搜索任何感兴趣的位置信息，如附近的餐厅、酒店、医院和银行等公共基础设施，除了具体的地理位置之外，该类软件还会向用户简要介绍 POI 的其他信息，如网友点评等，以提升用户体验。在这类服务场景中，由于用户只是偶尔地提出位置服务请求，这导致服务提供商获取的信息通常是单个孤立的请求。目前，大量的 LBS 隐私保护方案都是基于这类场景。

2) 基于位置的精确导航服务

Autonavi Navigation、百度地图和滴滴出行等都提供基于位置的精确导航服务，用户使用这类软件时，持续地发送实时位置信息，服务提供商通过计算，向用户返回实时路况信息，并推荐最佳路线完成导航。在这类服务场景中，服务提供商获取的信息是用户的精确位置和运动轨迹。

3) 基于位置的社交网络服务

Foursquare、Facebook 和微信等都提供基于位置的社交服务，包括近邻检测服务和签到(check-in)服务。近邻检测服务是指用户向服务提供商上传位置信息，服务提供商根据 2 个用户的物理位置判断是否为近邻，这样便于用户发现新的朋友，与附近的朋友见面。签到服务是指用户在某个语意(semantic)位置签到，服务提供商返回该位置附近的信息，并将该位置信息通知其朋友。通过该类软件，用户可了解朋友在做什么，从中发现共同爱好。在这类服务场景中，服务提供商和朋友获取的信息是单个孤立位置点。

4) 基于位置的运动检测服务

UltimatePedometer、Nike+和咕咚运动等都提供基于位置的运动检测服务，用户使用该类软件(同时佩戴智能手环)可统计其每天运动的步数、距离和路径等信息，为分析用户的运动情况提供依据，引导用户健康生活。在这类服务场景中，服务提供商获取的信息是用户的运动轨迹。

5) 基于位置的广告推送服务

ShopAlerts 和 Getyowza 等都提供基于位置的广

告推送服务，当用户位于某些商店附近时，该类软件向用户推送商店的广告信息，并赠送折扣券。这样既能帮助商店推销商品，又方便用户选择合适的商品。在这类服务场景中，服务提供商获取的信息是单个孤立的位置点。

上述 5 类基于位置服务的应用场景已经得到了全面的应用和推广，为用户日常生活带来了极大便利。除此以外，还有一些位置服务应用场景，如历史位置验证[6,7]、基于位置的游戏[8]等。

2.2 位置服务体系框架

位置服务相关的应用场景大都对应如图 1 所示的体系框架，该体系框架包括 4 个主要实体：移动设备、定位系统、通信网络和服务提供商。移动设备指的是人们常用的智能手机或者平板笔记本；定位系统可以使用 GPS 定位、Wi-Fi 接入点定位或基站定位[9]；通信网络可以是 3G 或 4G 网络。其服务过程如下：首先，用户利用移动设备向服务提供商提出服务请求，请求中包含的位置信息由定位系统提供；然后，服务请求通过通信网络发送到服务提供商；最后，服务提供商根据用户的位置和请求内容进行相应的处理，并将结果通过通信网络返回用户设备。

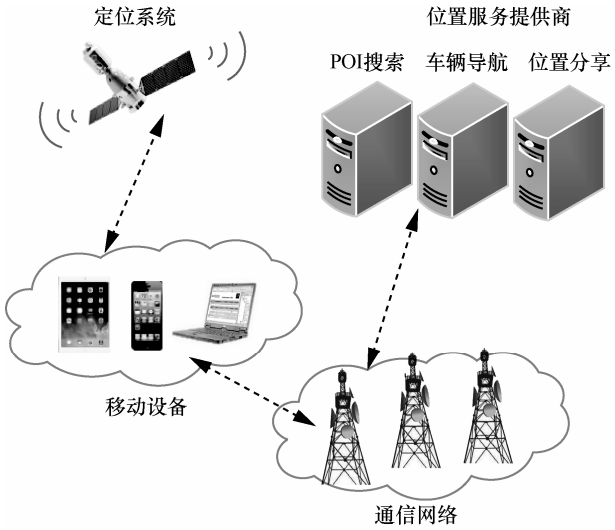


图 1 位置服务体系框架

2.3 隐私保护目标

现有基于位置的隐私保护方案所保护的用户属性主要包括：用户 ID、位置信息、时间信息以及 POI，用四元组表示为〈用户 ID，位置信息，时间信息，POI〉。在设计方案之初，需明确隐私保护目标，通常为 4 个属性的任意组合。

1) 用户 ID

用户 ID 是指用户姓名、身份证号或其他能够唯一确定用户的属性组。如在导航应用中，用户提供实时位置信息的同时可使用假名以隐藏用户 ID，但保护效果有限，原因在于用户会周期性往返于住所和公司，通过这 2 个信息，可推测出用户 ID[10]的概率较大。

2) 位置信息

位置信息指用户所处的位置坐标，通常不同应用对位置信息有不同精确度需求。例如导航应用需要用户相对精确的位置，而天气预报应用只需要用户所在的城市即可。

3) 时间信息

时间信息是指用户处于某个位置时所对应的时刻或时间范围。用户在某些时间段内很可能处于同一位置，比如夜晚一般位于家中，此类时间段可辅助攻击者推断用户家庭地址相关的隐私信息。此外，用户使用导航应用时，若能够获取时间信息，则可计算出用户的行驶速度，从而推测用户的出行工具等。

4) POI

POI 指用户查询的服务内容，包括查询附近餐厅、酒店、购物中心、医院和银行等。POI 的暴露很可能导致用户隐私泄露。例如，某用户经常查询附近的酒吧，则可推断出其可能有酗酒习惯；某用户经常查询周围的医院，则可推断出其可能患有某种慢性疾病。

2.4 隐私保护系统构架

根据隐私保护机制(LPPM, location privacy protection mechanism)主体的不同，隐私保护系统构架可以分为 2 类。第一类依赖可信第三方(TTP, trusted third party)的系统构架，如图 2 所示，由可信匿名服务器(trusted anonymity server)获取所有用户的位置信息，并负责执行 LPPM。第二类不依赖可信第三方的系统构架，如图 3 所示，由用户移动设备执行 LPPM，直接发送服务请求并接收查询结果。这 2 类系统构架各有特点，前者的优势在于 TTP 能够获取海量用户的位置信息，并且辅助用户过滤服务数据，其缺点在于 TTP 可能成为攻击者的目标，在实际使用中它的可信性难以保障。后者的优势在于其部署 LPPM 的简易性，方便用户根据其隐私保护需求调整隐私保护粒度，其缺点为 LPPM 的实现受移动设备性能的限制，且无法获取全局用户位置信息。

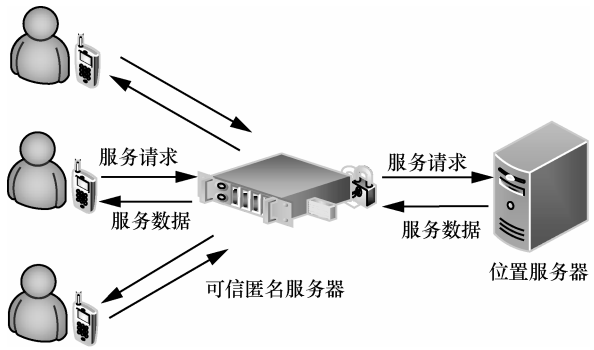


图 2　依赖可信第三方的系统构架

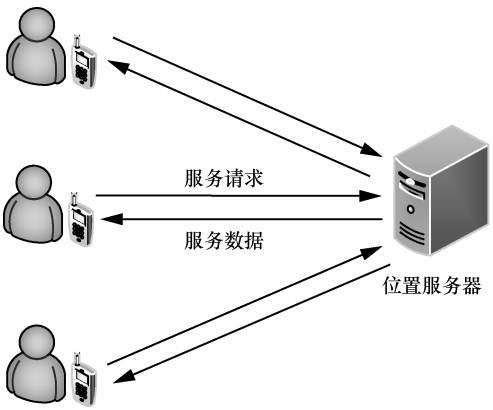


图 3　不依赖可信第三方的系统构架

3　攻击者模型

从时间维度上，攻击者模型可分为 2 类，一类攻击者能获取某一时刻用户的单次服务请求，称为单点攻击模型；另一类攻击者可获得用户在一段时间内的服务请求，甚至是整个运动过程中所有服务请求，称为多点攻击模型。从背景知识维度上，攻击者模型也分为 2 类，其中一类模型不考虑或假设攻击者不掌握任何背景知识（电话簿、统计数据等），只依靠单次或几次服务请求中的信息推断用户的隐私信息，称为无背景知识模型；另一类假设攻击者长期收集和利用各类相关的背景知识，称为背景知识模型。

3.1　无背景知识模型

在无背景知识模型下，攻击者的攻击方式主要包括位置同质攻击、区域中心攻击和位置分布攻击。

位置同质攻击[11]是指攻击者分析用户提交的隐匿区域内的多个位置，若距离非常接近，如位于同一个建筑内，虽然达到 *k*-匿名的要求，但由于隐匿空间太小，用户的位置隐私仍然无法得到很好的保护。

区域中心攻击[12]是指部分空间隐匿算法可能造成用户的真实位置以相对较高的概率出现在隐匿区域中心。如在某些群组协作隐私保护方案中，如图 4 所示，发起者向周围的协助用户发送请求，不断扩张其匿名集合，从而满足隐私保护要求。这种方式往往导致发起者位于隐匿区域的中心，攻击者可极大地缩小用户所在范围。

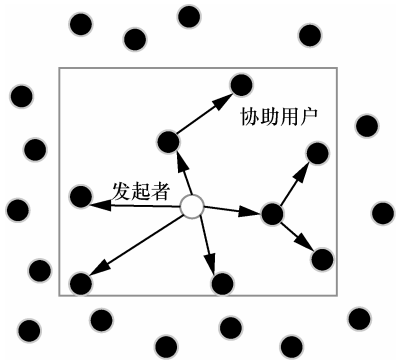


图 4　区域中心攻击

位置分布式攻击[13]主要利用用户发送位置分布不均匀的信息推测用户真实位置。例如，若攻击者观察到隐匿区域中只有一个用户位于人口稀少区域，而其他 *k*–1 个用户位于人口密集区域，那么位于人口稀少区域的用户很可能是该服务请求的发送者。如图 5 所示，用户 *A* 位于人口稀少区域，用户 *B*、*C*、*D* 和 *E* 位于人口密集区域，假设 *k*=4。当用户 *A* 发送查询请求时，由于其位于人口稀少区域，为了达到 *k*=4 的要求，其隐匿区域 R_1 必须扩展到人口密集区域，而其他 4 个用户无需扩大隐匿区域。由此，攻击者推断出用户 *A* 发送隐匿区域 R_1 的概率极高。

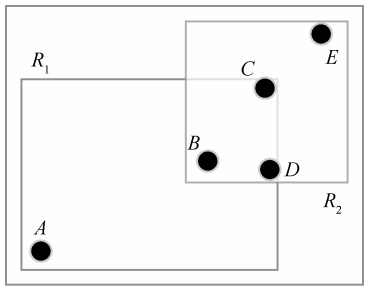


图 5　位置分布式攻击

3.2　背景知识模型

在背景知识模型下，攻击者的攻击方式主要包括同源攻击、概率分布攻击和位置相似性攻击。

同源攻击[14]是指攻击者了解用户会在同一位置多次发送服务请求并收集该信息，结合 LPPM，攻击者可计算出最有可能产生这一系列请求的用

户位置。如图 6 所示，攻击者了解到用户在某家咖啡店，该用户使用空间隐匿作为 LPPM。经过一段时间，攻击者收集到该用户发送的隐匿区域 R_1、R_2 和 R_3，求出 3 个隐匿区域的交集，即可将用户真实位置限定在极小的范围内。

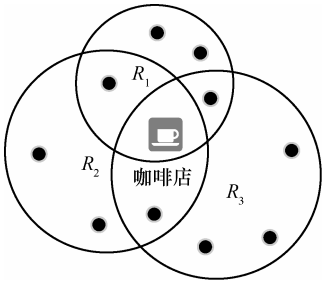


图 6 同源攻击

概率分布攻击[15]是指攻击者能够收集到用户的服务请求历史记录，由此计算出其查询概率分布。当用户发送查询请求时，若隐匿区域内概率分布不均匀，则攻击者可以推断出用户很可能位于概率较高的位置。

位置相似性攻击[16]是指攻击者分析隐匿区域内的语义信息，若该区域仅包含一种语义信息，比如医院或学校等，则攻击者可推断出用户的行为。如图 7 所示，X 表示用户 A 的真实位置，其隐私保护要求为 $k\geqslant 5$，S_1、S_2 和 S_3 分别表示医院、学校和银行 3 个区域。若隐匿空间 R 仅包含 S_1，虽然可以达到隐私保护要求，但攻击者可以推断出用户患有某种疾病，正在进行治疗。因此，用户应当将隐匿区域扩展到包含 S_2 和 S_3，以抵御该类攻击。

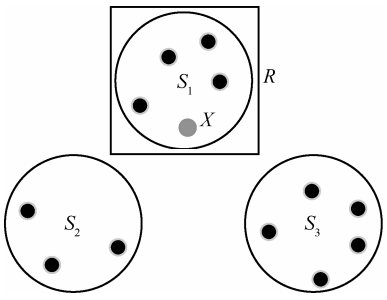


图 7 位置相似性攻击

3.3 多点攻击模型

在多点攻击模型下，攻击者的攻击方式主要包括位置依赖攻击、连续位置攻击和位置追踪攻击。

位置依赖攻击[17]主要根据用户 2 次连续提交的隐匿区域在时间上的相关性，结合前一隐匿区域和用户最大运动速度计算出最大移动边界(MMB, maximum movement boundary)，将该边界区域与后一隐匿区域求交集，可缩小隐匿区域大小，甚至直接定位用户位置。如图 8 所示，在 t_1 时刻，用户 A 需要查询附近的医院，于是向服务提供商发送包含 A、B 和 C 的隐匿区域 R_1。经过一段时间，在 t_2 时刻，用户 A 需要查询附近的加油站，于是向服务提供商发送包含 A、D 和 E 的隐匿区域 R_2。若攻击者能够获取隐匿区域 R_1 和用户 A 的最大运动速度，通过计算可获得用户 A 从 t_1 时刻到 t_2 时刻的最大移动边界。将最大移动边界与 R_2 求交集，可定位 A 的真实位置。

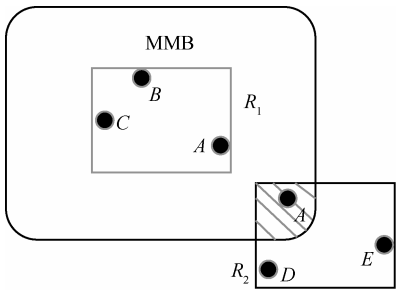


图 8 位置依赖攻击

连续位置攻击[18,19]是指运动中的用户连续发送自己的位置给服务提供商，即使位置信息是一个隐匿区域，攻击者可将连续的隐匿区域链接在一起，从而识别出查询发起者。如图 9 所示，假设存在 11 个用户 A~K，用户 A 的隐私保护要求为 k=5。在 t_1 时刻，用户 A 提出查询请求，空间隐匿算法产生一个包含 A、B、C、D 和 E 的隐匿区域，攻击者有 $\frac{1}{5}$ 的概率猜测出查询发送者。在 t_2 时刻，各用户的位置信息发生变化，攻击者观察到的隐匿区域包含 A、B、F、G 和 H，通过联系 t_1 时刻和 t_2 时刻的隐匿区域，攻击者可推断出只有 A 和 B 可能是该查询的发送者。类似地，攻击者再联系 t_3 时刻的隐匿区域，可以推测出用户 A 为该连续 LBS 查询的发送者。

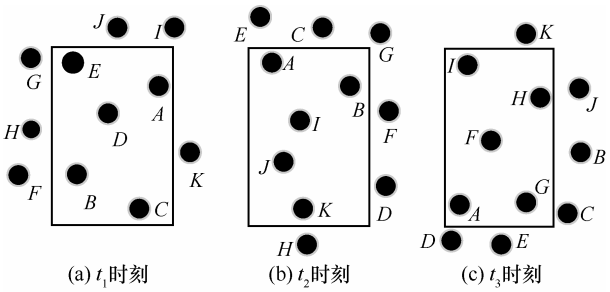


图 9 连续位置攻击

位置追踪攻击[20]主要利用用户使用导航应用时，实时更新其精确位置信息。通过收集和分析这些信息，攻击者可以获取用户的运动规律和行为习惯。若用户只采用随机变换假名的方法，没有合理使用混淆区(mix zone)等技术，则无法抵御这类攻击。

4　隐私保护度量指标

k-匿名和位置熵是目前被广泛使用的隐私保护度量指标，除此以外，一些学者结合各类应用场景提出了很多度量指标。

4.1　*k*-匿名

隐私保护度量指标中最为经典的是 *k*-匿名，它首先出现在数据库方向[21]，被定义为：对于准标识符属性，任意一条记录无法与其他至少 $k-1$ 条记录区分，准标识符属性是指家庭地址、生日和邮政编码等可以相结合以唯一确定某人身份的属性。在位置服务中，该度量指标可以总结为“向服务提供商发布的区域中，至少包含其他 $k-1$ 个不可区分的用户信息[22]”。

目前，多数隐私保护方案采用 *k*-匿名作为隐私保护度量指标。如部分基于匿名服务器的方案[20,22~26]由匿名服务器产生一个隐匿区域，包含查询用户与其他至少 $k-1$ 个用户，攻击者从这 k 个用户中识别出查询者的概率为 $\frac{1}{k}$；部分不依赖可信第三方的方案[27~31]，由移动设备产生至少 $k-1$ 个虚假位置，将用户真实位置混入其中一并发送给服务提供商，攻击者从中识别出真实位置的概率为 $\frac{1}{k}$。由于 *k*-匿名没有考虑隐匿区域内建筑的语义信息，若隐匿区域正处于某医院范围内，攻击者可推测出用户患有某种疾病。*l*-多样性[32,33]作为 *k*-匿名的一种扩展，要求隐匿区域内至少包含 l 种不同语义类型的小区域，可有效避免上述问题。另一种扩展是强 *k*-匿名[34]，要求用户在连续的查询请求中，组成隐匿区域的 k 个用户尽可能不变，可防止攻击者通过计算多个隐匿区域的交集来定位用户位置。

4.2　位置熵

k-匿名得到了广泛应用，但也存在明显缺陷。比如当隐匿区域中用户出现的概率不相等时，*k*-匿名不能准确反映隐私保护水平。因此，很多学者提出了位置熵的概念。位置熵起源于香农熵[35]，是一种度量不确定性的方法。目前，很多方案[35~40]基于位置熵设计隐私保护度量指标。Palanisamy 等[39]针对基于混淆区的方案提出成对熵(pairwise entropy)的概念。理想情况下，用户进入混淆区后会停留任意随机时间，且从任何一个出口离开的概率是相同的。若用户 i 离开混淆区时的假名为 i'，则用户 i' 的位置熵为

$$H(i')=-\sum_{j\in A}p_{i'\to j}\mathrm{lb}(p_{i'\to j})$$

其中，A 为混淆区内所有用户的集合，$p_{i'\to j}$ 为 i' 映射到用户 j 的概率，理想情况下，i' 映射为 A 中任意用户的概率相等。在实际中，混淆区一般设置在交叉路口，很多因素会影响其效果：1) 时间因素，先进入混淆区的用户先离开的概率大于后进入混淆区的用户；2) 路网环境，用户直行、拐弯和调头的概率各不相同；3) 道路状况，比如某个出口堵车，会提高攻击者推断出用户从某些出口驶出的概率。因此，实际使用时 i' 映射为 A 中任意用户的概率不相同。若少量映射关系发生的概率很高，而大部分映射关系发生的概率很低，那么即使位置熵很高，攻击者也能以较高概率推测出映射关系。为了解决该问题，在成对熵的概念中，假设混淆区只有 2 个用户 i 和 j，离开混淆区时的假名为 i' 和 j'，则成对熵 $H_{\mathrm{pair}}(i,j)$可以表示为

$$H_{\mathrm{pair}}(i,j)=-(p_{i'\to i}\mathrm{lb}p_{i'\to i}+p_{i'\to j}\mathrm{lb}p_{i'\to j})$$

若 i' 映射到用户 i 和 j 的概率 $p_{i'\to i}$ 和 $p_{i'\to j}$ 相等，则 $H_{\mathrm{pair}}(i,j)$等于 1，攻击者能够联系前后假名的可能性最低。但是应当同时考虑成对熵 $H_{\mathrm{pair}}(j,i)$，若 j' 映射到用户 i 和 j 的概率 $p_{j'\to i}$ 和 $p_{j'\to j}$ 有较大差距，攻击者联系前后假名的可能性依然很高。因此，用户 i 和 j 的成对熵应当是 $H_{\mathrm{pair}}(i,j)$和 $H_{\mathrm{pair}}(j,i)$中的较小者，有效的混淆区应当满足任意 2 对用户的成对熵都接近 1。针对群组协作的隐私保护方案，Niu 等[37]提出了单个用户和全局隐私度量标准。当用户直接向服务器提出服务请求时，隐私保护水平通过单个用户隐私度量标准来衡量。文献[37]中采用了伪造虚假位置的方法，地图被分为 $N\times N$ 个单元，从中选择 k 个位置，根据背景知识确定 k 个位置的查询概率 q_i，然后计算出每个位置成为真实位置的概率

$$p_i=\frac{q_i}{\sum_{j=1}^{k}q_i}$$

最后，计算出位置熵，即攻击者推断出真实位

置的不确定性，位置熵越大，隐私保护水平越高。当所有 p_i 相等时，位置熵最大，即 lb$k$。同时该文献中群组组员缓存并共享数据，减少了向服务提供商提出请求的次数，提高了整体隐私保护水平，从而提出了全局隐私度量标准。对于由缓存应答的服务请求，服务器无法获得任何信息，因此，攻击者推断出真实位置的不确定性为 lbN^2。全局隐私度量标准可表示为

$$\lambda=\frac{\sum_{q\in Q_{\text{server}}} H_q+\text{lb}N^2\left|Q_{\text{cache}}\right|}{\left|Q_{\text{server}}\right|+\left|Q_{\text{cache}}\right|}$$

其中，Q_{cache} 表示由缓存提供的服务次数，Q_{server} 表示由服务器提供的服务次数，H_q 表示位置熵。可以看出，应当从两方面提高整体隐私保护水平，一方面提高每一次查询的位置熵，另一方面提升缓存的命中率，使更多请求通过缓存应答。

4.3 其他度量方法

除了 k-匿名和位置熵，学者们也提出了许多新的度量方法，扭曲度和差分隐私是其中较为典型的 2 种度量方法。

4.3.1 扭曲度

Shokri 等[41]提出隐私保护度量指标应当满足 3 点要求：1) 从攻击者角度，考虑攻击者利用背景知识最小化出错概率；2) 考虑真实位置与攻击者推断位置间的差异性，多数情况下该差异性可表示为真实位置与推断位置间的距离；3) 适用于各类隐私保护技术。文献[41]中提出一种位置隐私度量框架，包括用户、轨迹、攻击者及 LPPM。由于攻击者的目的是将经过 LPPM 变换后的用户位置重新转化为用户的真实轨迹，因此，各类 LBS 隐私保护技术所提供的隐私保护水平取决于用户真实的轨迹与攻击者推断的轨迹之间的差异，即扭曲度，扭曲度越大，隐私保护水平越高。对于用户 u，在时刻 t，扭曲度可以表示为

$$\sum_{Y} D\left(whereis\left(u,t\right),loc\left(tail\left(Y\right)\right)\right)\pi\left(Y\right)$$

其中，*whereis*(u,t)表示 u 在 t 时刻的真实位置，Y 表示 u 在一段时间内可能发生的某条轨迹，*loc*(*tail*(Y))表示轨迹 Y 在 t 时刻对应的位置。由于函数 $\pi(\cdot)$表示攻击者通过背景知识推断出的 u 按照某条轨迹运动的概率，$D(\cdot)$表示 2 个位置的差异，且各类 LBS 隐私保护技术都适用，因此满足上述 3 点要求。Shokri 等[15]细化了文献[41]中的位置隐私度量框架：用户将真实位置通过 LPPM 进行调整和扭曲，变成某种形式的假位置，并将用户名换成假名。同时攻击者可以获取三方面的信息：1) LPPM；2) 服务提供商所获取的信息；3) 背景知识，如用户查询历史记录等。基于此，攻击者设法推断用户的真实位置。此外，文献[15]从攻击者的角度确切地阐述了准确性(accuracy)、确定性(certainty)及正确性(correctness)这 3 个指标，其中，确定性对应位置熵，正确性对应扭曲度。准确性：由于攻击者不可能了解用户的所有相关信息，故其推断出的用户位置分布律(Pr(x|o)，x 为用户位置的可能值，o 为提交的虚假位置)与理想情况下拥有所有信息时推断出的分布律是有差距的，差距越小，准确性越高。确定性：利用攻击者推断出的用户位置分布律，计算出位置熵来度量确定性，即

$$H(x)=-\sum_{x}\Pr(x\mid o)\text{lb}\frac{1}{\Pr(x\mid o)}$$

位置熵越高，确定性越低，攻击者就越不容易确定出用户真实位置。正确性：即使攻击者收集到足够多的背景知识、推断出的用户位置分布律很准确、而且从用户位置分布律中很容易确定出用户位置，即确定性很高，但正确性未必很高，如在某次查询中，用户真实位置在以往数据中从未出现过。通过计算用户真实位置 x_c 与攻击者猜测的所有可能值的期望距离来度量正确性，即

$$\sum_{x}\Pr(x\mid o)\left\|x-x_c\right\|$$

用户最关心的是攻击者能否推断出正确位置，以及攻击者推断结果与真实位置有多大差距，因此，正确性是度量隐私保护水平最重要的指标。

4.3.2 差分隐私

以上度量方式均存在一个明显缺陷，即在度量隐私保护水平时需要假设攻击者掌握的背景知识，而差分隐私的提出能很好地解决该问题。差分隐私[42](differential privacy)首先出现在数据发布隐私保护领域，其不考虑攻击者拥有多少背景知识，通过向查询或者在分析结果中添加适当噪音的方式来达到一定的隐私保护效果。近几年来，差分隐私也被引入到位置隐私保护领域，取得了一定进展。若按照严格的差分隐私要求，要求位置上的任何移动对于 LPPM 处理后的结果都没有影响，或者影响极小。然而，严格的差分隐私要求使服务提供商无法提供精准的服务信息。因此，

Andres 等[43]借鉴差分隐私思想，提出通过位置不可分辨性(geo-indistinguish ability)来表示用户的位置隐私保护需求。文献[43]指出向服务提供商发送隐匿区域的方法使攻击者能够准确地获取用户所处区域，因此，并未达到理想的隐私保护水平。用户期望隐私保护水平应当与距离相关，例如，攻击者从用户提供的位置信息中推断出用户位于半径为 1 km 的某区域的概率极低，而随着半径的扩大，概率也随之提高，如攻击者推断出用户位于某城市的概率几乎是百分之百。根据上述思想，文献[43]中给出了不可分辨性的 3 种形式化定义，第 3 种定义为

$$\frac{\Pr(S \mid x)}{\Pr(S \mid x')} \leqslant \mathrm{e}^{\varepsilon r}, \forall r > 0, \forall x, x' : d(x, x') \leqslant r$$

其中，r 表示 x 和 x'所处区域半径，函数 $d(\cdot)$表示 2 个位置的距离，该定义表明 2 个位置的距离越近，LPPM 产生相同结果 S 的概率应当接近，反之，LPPM 产生相同结果的概率相差可较大。不同于文献[38]要求一个圆形区域内的所有位置都满足位置不可分辨性，Xiao 等[44]定义了基于 δ-位置集的差分隐私，δ-位置集 X_t 选取用户出现概率较高的位置，在这些位置上应当满足差分隐私，使任意 t 时刻，攻击者不能根据用户所发送的位置 Z_t 区分出 X_t 中的任意 2 个位置，具体定义如下

$$\frac{\Pr(A(x_1) = Z_t)}{\Pr(A(x_2) = Z_t)} \leqslant \mathrm{e}^{\varepsilon}, \forall x_1, x_2 \in X_t$$

其中，$A(\cdot)$表示 LPPM，使用 δ-位置集的好处在于对于掌握道路环境和用户运动模型等精确背景知识的攻击者而言，将真实位置偏移到概率较低的位置意义不大，故基于 δ-位置集的差分隐私更符合实际要求。

5　LBS 隐私保护技术

常用 LBS 隐私保护方法主要包括基于策略的方法、基于密码学的方法和基于泛化和模糊的方法。

基于策略的方案[45,46]主要通过制定隐私保护规则、标准和详细规范来约束服务提供商，禁止位置信息被不正当的使用。虽然此类方案能够同时保护位置隐私和查询隐私，但缺陷明显：1) 服务提供商能否有效执行相关策略往往依赖于法律制度和社会舆论的压力；2) 没有度量隐私保护水平的标准。

基于密码学的方案[47~49]主要通过相关密码技术处理向服务提供商发送的请求并获取服务数据以保护用户隐私信息。此类方案无需匿名服务器，位置信息加密后发送给服务提供商，后者解密后执行相应查询，并将结果加密后返回，因此安全性很高，但有着较明显的缺陷：1) 服务器端需大量的代码修改以实现相关功能，且计算开销较大；2) 对于服务提供商而言，隐私信息可产生巨大的经济价值，故没有足够动机用高昂成本来保护用户隐私信息，因此，此类方案可行性较差。

以下主要介绍并详细分析基于泛化和模糊的 LBS 隐私保护技术，包括常用的 4 种技术：空间隐匿、位置偏移和模糊、伪造虚假位置和群组协作。

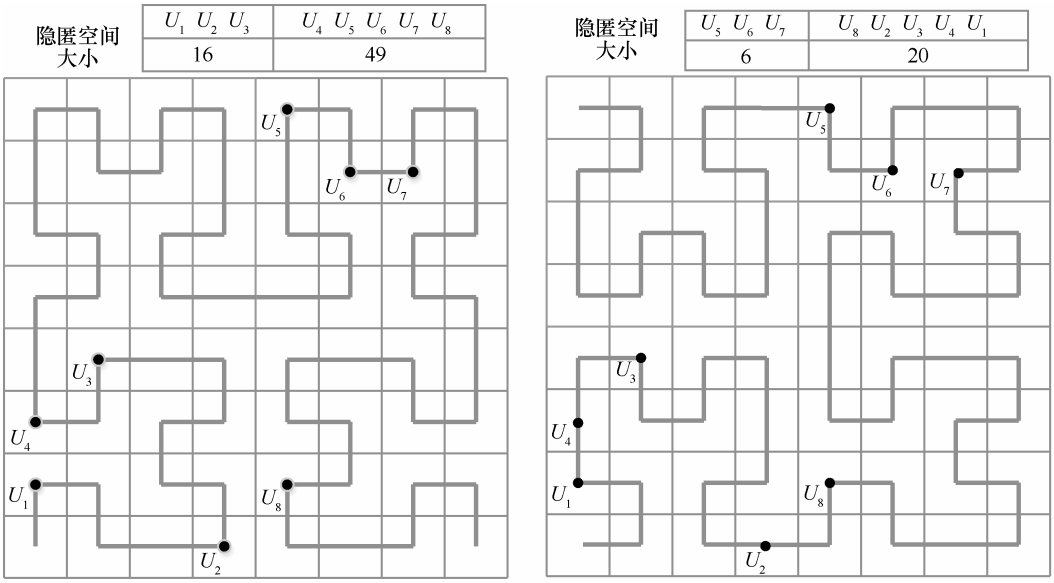
5.1　空间隐匿

空间隐匿技术通过为每个查询用户形成一个包含 k 个真实用户的隐匿区域，使服务提供商难以从该隐匿区域中确定用户的真实身份和准确位置。

Gruteser 等[20]将 k-匿名的概念引入到位置隐私领域，并首次提出了一种简单而有效的空间隐匿方案，其基本思想为：从一个足够大的区域开始，将该区域分割成 4 个小区域，计算真实用户所在区域的用户数，若用户数大于 k 则继续分割，直到用户数小于 k 为止。倒数第 2 次分割形成的区域就是隐匿区域。考虑到当用户处在人口密度较小地区时只有增大隐匿区域才可能覆盖 k 个用户，服务质量会明显降低。因此，文献[20]提出了改进方案，其基本思想为：用户设置隐匿区域面积，经若干次分割后，隐匿区域缩小到用户的设定值，观察该隐匿区域内出现的用户数量，当有 k–1 个用户访问过该区域时，将此隐匿区域和这段时间间隔发送给服务提供商。该方案主要问题是用户无法根据需求设置不同的 k 和隐匿区域。Chow 等[22]提出个性化 k-匿名方案，将整个区域层级划分，划分方法为每层对上层的每个单元进行四等分，从而形成自顶向下的金字塔型数据结构，金字塔中的每个单元都记录该区域的用户数量。同时，在一张散列表中记录每个用户所在的底层单元和隐私需求，用户可设置不同的 k 和隐匿区域大小。由于用户位置的持续变化，金字塔形数据结构和散列表均需实时更新。当用户需要位置服务时，从用户散列表条目中获取隐私需求和用户所在单元，若用户所在单元的用户数和区域面积都超过隐私要求，则将该单元发送给服务提供商，否则考察该单元水平方向的邻居单元和竖直方向的邻居单元，若用户所在单元与其中一个邻居单

元合并后满足用户隐私要求，则将合并后的区域发送给服务提供商，否则返回到该单元的父单元，并重复执行上述过程直到找到满足条件的区域。然而，上述 2 种空间隐匿方案均不能抵抗位置推断攻击，为此文献[50]提出利用希尔伯特曲线(Hilbert curve)构造隐匿区域。根据希尔伯特曲线对用户进行排序，若采用 *k*-匿名，则将连续 *k* 个用户分为一组，最后一组不超过 2*k*−1 个用户，由于同一组用户形成相同的隐匿区域，因此，该方案能够抵御位置推断攻击。Ngo 等[51]指出，由于空间维度的降低，该方案[50]容易产生较大的隐匿区域，由此给服务提供商带来大量的计算开销和网络通信开销。如图 10 所示，假设有 8 个用户，隐私保护需求为 3-匿名，由此将用户分为 2 组。图 10(a)中 2 个隐匿区域大小分别是 16 个单元和 49 个单元，平均值为 32.5 个单元。通过对希尔伯特曲线的翻转或旋转操作，可减少大部分用户的隐匿区域，例如图 10(b)中 2 个隐匿区域大小的平均值为 13 个单元。由此，若对所有可能的希尔伯特曲线进行筛选，可为每个用户找到其最小的隐匿区域，但该方式又不能抵御推断攻击。为了同时解决这 2 个问题，文献[51]提出了一种基于差分隐私的隐匿区域扰动方案，将隐匿区域最小的希尔伯特曲线版本扰动到所有可能的希尔伯特曲线版本。根据差分隐私的性质，相邻用户一般会得到较为接近的扰动结果以抵御推断攻击。另一方面，其噪声函数保证隐匿区域越小的希尔伯特曲线版本有较高概率作为扰动结果。

以上空间隐匿方案均采用了 *k*-匿名的度量标准。除了 *k*-匿名，位置熵的度量标准也被广泛应用于该类方案。Beresford 等[35]提出基于混淆区(mix zone)的隐私保护方案。该方案将区域分为 2 类：混淆区和应用区。当用户位于应用区时，匿名服务器可向服务提供商发布位置信息。而当用户进入混淆区时，不允许匿名服务器发布信息。首先所有用户都被赋予假名，当用户离开混淆区时更换新假名。如图 11 所示，假设某时刻假名为 *A* 和 *B* 的用户进入混淆区，此时立即停止向服务提供商发布位置信息。当 2 名用户离开时，匿名服务器为他们分配新的假名 *X* 和 *Y*。由于服务提供商无法找出前后 2 个假名间的联系，则可认为达到 2-匿名的效果。然而，若考虑用户的运动模式，比如用户直行的概率远大于折返的概率，那么这种情况很难达到 2-匿名的效果。因此，该场景中使用位置熵的度量方法衡量隐私保护水平更为准确。Palanisamy 等[39]将混淆区应用到路网环境下的车辆位置隐私保护，为了保证混淆区有较大的成对熵，必须使同时进入混淆区的用户在混淆区内停留的时间尽可能相同。由此，文献[39]提出了 3 种改进方案。方案 1 在普通混淆区基础上加入时间窗口，需要变换假名的车辆需满足进入混淆区的时间相近，从而提高成对熵，后 2 种改进都包含时间窗口功能。方案 2 考虑了道路状况，如堵车，攻击者通过观察车辆驶出时间，可推断出车辆从各个出口驶出的概率。针对该问题，方案 2 提出根据实际情况适当偏移矩形混淆区，保证从每个入



(a) 希尔伯特曲线原始状态　　(b) 希尔伯特曲线顺时针旋转90°

图 10　希尔伯特曲线变换

口到达交叉路口中心的时间相同。方案 3 考虑用户速度偏离平均速度，其从各个出口驶出的概率不同。针对该问题，方案 3 采用非矩形混淆区，即混淆区只包含交叉路口中心区和出口路段，使用户在混合区中停留的时间尽可能相同。Xu 等[23]扩展了位置熵的概念，提出利用区域人口密度来体现隐私保护需求。文献[23]指出隐私是用户的一种主观概念，用户无法判断其隐私保护需求应当对应多大的 k 值，因此，利用 k 值反映用户的隐私保护要求是不准确的。由此，文献[23]中提出由用户确定某个示例区域，该区域能够满足其隐私保护需求，由匿名服务器计算该区域的人口密度，并作为用户的隐私保护需求。在计算区域人口密度时，不能简单地统计该区域中的用户数，因为某些用户在该区域出现的次数要远高于其他用户，如用户 A 的工作单位位于该区域，则用户 A 提出服务请求的概率较大，因此，利用位置熵定义区域人口密度更为准确。匿名服务器通过收集历史数据计算出位置熵，包括经过该区域的用户数和每个用户出现的次数等，从而得到该区域人口密度。空间隐匿技术也存在缺点，如在人口稀少区域，为了达到隐私保护需求，服务质量将受到较大影响。此外，虽然依赖可信第三方的系统构架能够获取海量的用户位置信息，为实现空间隐匿技术提供保障。但是此类系统构架缺陷明显：1) 随着用户规模和应用软件的剧增，匿名服务器的负载较大，将成为位置服务生态系统的性能瓶颈；2) 单点失效问题严重，将极大地影响用户体验；3) 若 TTP 被攻击者控制，用户的位置信息和服务请求历史记录都会被泄露。不依赖可信第三方的系统构架可以避免上述问题，以下 3 类技术均针对不依赖可信第三方的系统构架。

图 11　混淆区

5.2　位置偏移和模糊

位置偏移和模糊技术[43,52~57]通过加入噪音降低位置精确度，如真实位置的小范围移动或者以某个区域代替真实位置，来达到保护用户隐私的效果。

Yiu 等[52]提出 SpaceTwist 方案，主要解决 k 近邻(k nearest neighbor)查找问题。首先，在真实位置附近挑选一个位置设为锚点(anchor)，然后以其为位置信息连续向服务器发送查询请求，查询范围不断扩大，最后计算返回结果与真实位置之间的距离，挑选 k 个最为邻近的结果。如图 12 所示，q 为用户真实位置，g 为锚点，算法以锚点为中心逐次找到满足服务请求的近邻 a、b、c，同时更新 m 和 n(m 为 g 到最新发现的近邻的距离，n 为 q 与最近近邻的距离，m 逐渐增大，n 逐渐减小)，直到满足 $m > n + distance(q, g)$ 时算法结束。锚点离真实位置越远，隐私保护水平越高，但相应的计算开销也越大。Ardagna 等[54]提出位置模糊可通过几种基本操作及其组合来完成。该方案考虑到现有定位技术存在一定误差，用户只能确定其真实位置的所在区域，该区域大小表示某种定位技术的最高精确度。由此引出相关性(relevance)的概念，即相对于某种定位技术最高精确度的损失，它可同时度量精确性和隐私保护水平，通常情况下，LPPM 产生的精确度损失越大，相关性就越低，同时隐私保护水平就越高，用户可以通过相关性准确描述其隐私保护需求。为达到用户指定的相关性要求，文献[54]提出 3 种位置模糊基本操作：扩大半径、缩小半径以及圆心偏移。通过以上 3 种操作的组合，能够以多种方式实现用户的相关性要求，灵活完成位置模糊处理。Perazzo 等[53]指出已有的位置偏移和模糊方案存在缺陷[43,54,55]，此类方案仅简单地偏移位置并扩大隐匿区域，没有考虑到攻击者可能了解偏移算法，或者统计大量数据得出偏移向量概率分布等信息。通过此类信息，攻击者可计算出用户在隐匿区域内分布的概率密度，若不为常数，即非均匀分布，则攻击者可推断出用户在隐匿区域的某一范围内出现概率较大，导致隐私保护水平下降。由此，文献[53]提出了一种偏移向量的概率分布函数，使用户在隐匿区域内分布的概率密度为常数，从而达到均匀分布的效果。Andres 等[43]借鉴差分隐私的思想，提出通过位置不可分辨性表示用户的隐私保护需求，并设计了一种满足该特性的噪音生成机制。由于拉普拉斯分布(Laplace distribution)能够满足位置不可分辨性，方案中将二维拉普拉斯分布转化为极坐标形式作为噪声函数，其中，夹角满足均匀分布，半径满足伽玛分布(Gamma distribution)。此外，由于坐标精度有限，需将该噪声函数产生的结果离散化，从而得到理想的位置偏移。文献[43]没有考虑位置偏移对于服务质量的影响，不能很好地满足实用性要求。Shokri 等[58]提出基于零和贝叶斯博弈

(zero sum Bayesian game)的策略，在最优化隐私保护水平的同时确保服务质量损失小于给定阈值。该策略假设攻击者已获取先验知识，用户和攻击者轮流进行博弈，从而最优化各自利益。其中，用户在确保服务质量损失小于给定阈值的情况下最大化隐私保护水平，而攻击者根据先验知识和虚假位置力求最小化隐私保护水平，形成博弈双方。但该方案也存在不足，当攻击者掌握细粒度先验知识时，比如在时间维度上，某用户早晨和晚上有完全不同的习惯，攻击者则可在这 2 个时段使用不同的先验知识来降低隐私保护水平。由于位置不可分辨性基于差分隐私，因此，位置不可分辨性具有不基于先验知识的特性，可弥补文献[58]中方案的不足。Bordenabe 等[59]结合文献[43]和文献[58]方案的优势，构建一种最优化服务质量机制，在满足位置不可分辨性的前提下，最小化服务质量损失。Xiao 等[44]关注用户位置的时间相关性，充分考虑动态场景下的隐私保护。如图 13 所示，用户从学校到咖啡馆的过程中，利用位置偏移技术，向服务提供商发送了 3 次位置信息。虽然 3 个时刻的位置信息都得到了很好的保护，但攻击者仍然可利用道路环境和用户运动模型，推断出用户位于咖啡馆。该方案中时间相关性可用马尔可夫链描述，其马尔可夫转换矩阵为

$$\boldsymbol{p}_t^- = \boldsymbol{p}_{t-1}^+ \boldsymbol{M}$$

其中，$\boldsymbol{p}_t^-$ 表示 t 时刻用户位置先验概率向量，$\boldsymbol{p}_{t-1}^+$ 表示 $t-1$ 时刻的用户位置后验概率向量，$\boldsymbol{M}$ 表示用户运动规律矩阵，即用户从一个位置到另一位置的概率分布矩阵，其中，用户位置先验概率和后验概率体现了每一次用户发送位置信息都会改变用户位置分布率。此外，针对掌握道路环境和用户运动模型等精确背景知识的攻击者，该方案定义了基于 δ-位置集的差分隐私，δ-位置集选取用户出现概率较高的位置，在这些位置上应当满足差分隐私。在任意时刻，如在 t 时刻，计算出用户位置先验概率向量 $\boldsymbol{p}_t^-$，若需要发送位置信息，则根据 $\boldsymbol{p}_t^-$ 构建 δ-位置集。然后利用基于 δ-位置集的差分隐私机制计算出偏移位置 Z_t，发送给服务提供商，同时利用 Z_t 更新后验概率向量 $\boldsymbol{p}_t^+$，用于之后计算 $t+1$ 时刻的先验概率向量 $\boldsymbol{p}_{t+1}^-$。

5.3 伪造虚假位置

由于位置偏移和模糊技术降低了位置精确度[60]，因此，返回的服务数据可能不准确。而伪造虚假位置技术[28,29,36,61]则不存在该问题，但通信开销相对较高。

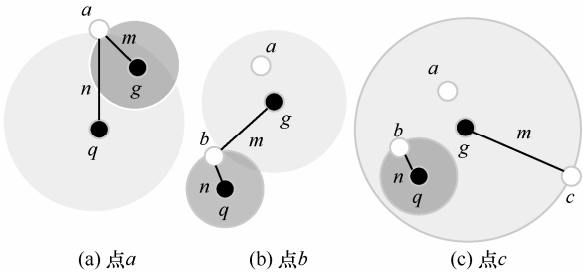


图 12 查询处理过程示例

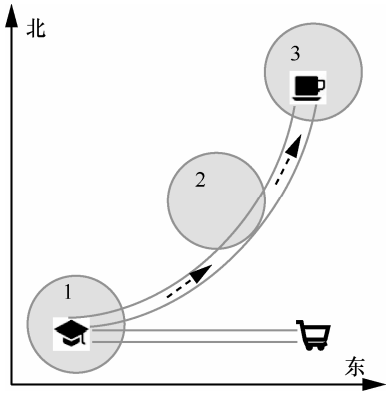


图 13 时间相关性引起的隐私泄露

Kido 等[28]通过产生多个虚假位置(dummy)，将用户真实位置混入其中，一起发送给服务提供商以实现隐私保护。服务提供商由于无法区分真实位置和 dummy，只能对每个所提交的位置提供所需服务。该方案中 dummy 的产生基于随机运动模型，没有考虑 dummy 的选取质量问题，而是把重点放在如何降低通信消耗上。应当如何选取 dummy 才能接近真实用户的运动模型是这类技术一直努力解决的问题。Lu 等[29]提出 2 种伪造虚假位置方案，称为 CirDummy 和 GridDummy。如图 14 所示，CirDummy 方案基于包含用户真实位置的虚拟圆产生 dummy，GirdDummy 方案基于包含用户真实位置的虚拟方格产生 dummy。该方案主要考虑两方面问题，一方面 dummy 分布不均匀时，零散的 dummy 很容易被攻击者排除掉，故该方案采用均匀产生 dummy 的方法(图 14(a)中所有的夹角 θ 都相同，图 14(b)中任意相邻 dummy 的间距相同)；另一方面考虑了 dummy 所形成的隐私保护区域应当超过某个阈值，以防止隐私保护区域太小直接泄露用户位置。这 2 种方案都没有考虑背景知识[62,63]，如部分 dummy 位于湖泊、海洋、高山和沼泽等区域时，攻击者可以轻易地排除这部分 dummy。Niu 等[36]提出

DLS 方案，以每个位置的查询概率(在历史查询记录中，所有用户在该位置上的查询次数除以整张地图总的查询次数)来表示背景知识，通过位置熵来度量隐私保护水平。

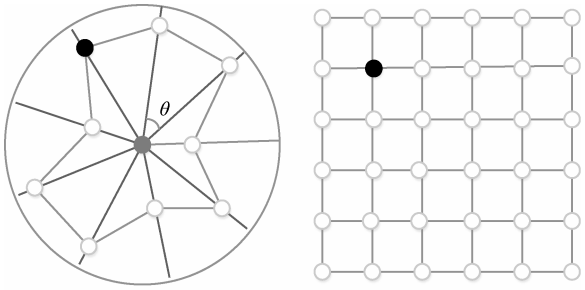


图 14 伪造虚假位置示例

如图 15 所示，用户随机产生 9 个 dummy，并与真实位置一起发送给服务提供商，理论上达到了 10-匿名的效果。然而，若服务提供商利用背景知识排除掉一部分 dummy，则实际上达不到 10-匿名的效果。因此，该场景中使用位置熵的度量方法衡量隐私保护水平更为准确，如图 15 中位置 1、2、3 的查询概率高于其他 7 个位置，则这次查询的位置熵为 lb3，而不是 lb10。理想情况下，当所有 dummy 的查询概率与真实位置相同时，熵值最大，隐私保护水平最高。为了获得较大的位置熵，文献[36]首先选取与真实位置查询概率最接近的 $4k$ 个备选 dummy，并随机挑选 m 组，每组包含真实位置和 $2k-1$ 个备选 dummy，然后通过计算每组的位置熵，找到熵值最大的一组，最后从熵值最大的组中选出 $k-1$ 个 dummy，选择的标准是尽可能形成较大的隐匿区域。在具体实现过程中，该方案采用了一种探索式的方法：定义输出集和备选集 2 个集合，输出集在初始时只包含真实位置，备选集包含 $2k-1$ 个备选 dummy。每轮从备选集中选出一个 dummy 移除，并将其加入到输出集，在 $k-1$ 轮后得到输出结果。每轮选择时，计算备选集中每一个 dummy 被选中的概率与它的权重(该 dummy 与输出集中所有位置的距离之积)成正比。通过该方法不仅达到了较高的熵值，还使 dummy 分布在较大的空间中。Pingley 等[61]提出如图 16 所示 DUMMY-Q 方案，通过伪造多个 POI 的方式保护查询隐私。具体地，首先根据历史查询次数，对空间进行四叉树形式的划分，保证划分后的每一个区域有大致相同的历史查询次数。然后根据用户的运动模型和当前位置预测出用户可能经过的区域集合，并根据规则建立 POI 池。最后基于 POI 池相应地伪造请求与真实请求一起发送给服务提供商，使攻击者难以分辨。Fawaz 等[4]提出一种位置隐私保护框架 LP-Guardian，该框架利用伪造虚假位置技术，在不修改位置服务 APP 的前提下，满足不同类型位置服务 APP 的隐私保护需求。对于用户正在使用的 APP 根据其制定的规则进行处理，结果分为 3 种情况：直接发送、不发送和隐匿处理后发送。若需隐匿处理，则根据 APP 类型选择不同的隐匿处理方式，从而满足不同的隐私保护需求：对于持续收集用户位置的 APP(如健康监测软件)，通过伪造某运动过程的起点和终点来达到隐私保护的目的，同时保持距离和速度与实际情况相一致；对于需要偶尔提供精确位置的 APP，该框架考虑到攻击者持续收集用户信息，并计算每个用户的运动模型。若某用户的运动模型与所有其他用户差别较大，则攻击者可识别该用户的身份。因此，当某用户的运动模型与理论模型偏离较大时，通过发送虚假位置来矫正该用户的运动模型；对于需要提供城市级别位置的 APP 和后台运行的 APP，只提供用户所在城市。LP-Guardian 权衡了隐私保护粒度、APP 可用性和服务质量，在实际应用中取得了较好的效果。Niu 等[37]提出群组协作方案中，当需要向服务提供商发送请求时，可利用伪造虚假位置技术来提高缓存(cache)命中率。服务提供商对每一个 dummy 都会提供位置服务，因此，将协作组中没有缓存记录的位置作为 dummy，可以达到提高缓存命中率的效果。文献[37]在保证隐私保护水平的同时考虑了影响缓存命中率的多个因素：1) 缓存位

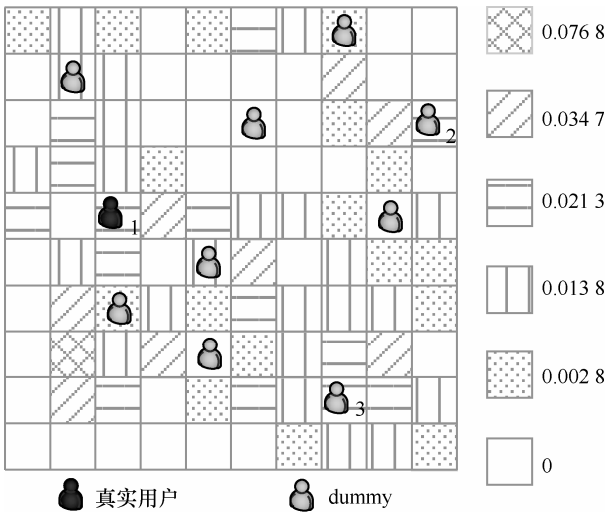


图 15　虚假位置筛选示例

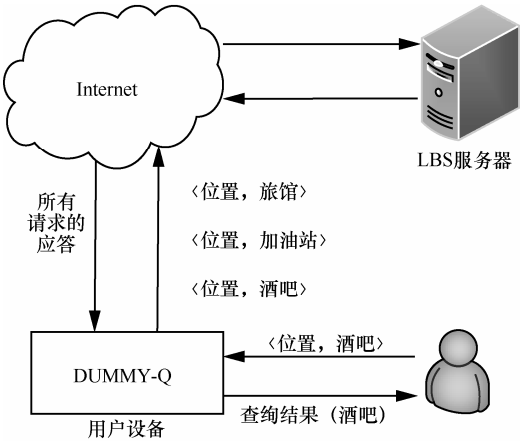


图 16 DUMMY-Q 处理过程示例

置的查询概率，应当尽量缓存查询概率高的位置；2) 缓存位置与用户真实位置的距离，应当尽量缓存真实位置的近邻位置；3) 缓存位置的新鲜度，应当尽量缓存接近缓存期限的位置。综合上述 3 个因素，可定量计算每组 dummy 对缓存做出的贡献。为了保证隐私保护水平，该方案首先挑选与真实位置查询概率最接近的 $4k$ 个 dummy 以确保较大的位置熵，然后从其中随机选取 $2k$ 个作为备选，最后选出 k–1 个对缓存贡献最大的 dummy。

5.4 群组协作

群组协作技术[27,30,31,38,64~66]通过用户自组织的 P2P 网络能够获取附近用户的位置信息。该技术可分为 2 类：陌生人之间的协作和基于信任关系的协作。

5.4.1 陌生人之间的协作

陌生人之间的协作通常使用蓝牙通信技术，用户之间的信息交互发生在 10 m 之内。

Manweiler 等[64]提出陌生人之间的协作方案：首先用户在相遇时产生随机对称密钥，此密钥作为此次相遇的凭证，并将密钥的散列值前缀发送给服务器。然后当用户需要发送消息时，用该密钥对消息加密，并将密钥的散列值前缀和密文一并发送给服务器。最后服务器通过匹配散列值前缀的方式，找到拥有相同前缀的用户，将密文发送给他们。该方案用 k-匿名来度量和设置用户的隐私保护水平，如果用户将密钥的散列值前缀长度缩短，那么服务器需要将消息发送给更多的用户(k 就越大)，隐私保护水平也就相应提高，但系统开销也随之增长。为了更好地降低系统开销，且便于用户对系统开销进行细粒度的控制，Niu 等[27]提出细粒度的隐私保护方案。首先，每个用户在本地缓冲寄存器中周期性地记录其历史位置，当与其他用户相遇时，随机分享部分历史数据。为了便于用户控制通信消耗，用户设置一个交换时间比，每隔一段时间开启位置数据交换模块，工作一定时间后关闭。接着，用希尔伯特曲线按照查询次数分布将地图划分为多个单元，查询概率越高的区域划分粒度越细，每个单元按顺序被赋予希尔伯特值，该过程可以离线执行。然后将所有单元(cell)划分为 k 段，若真实位置位于某一段的第 i 个单元，则选出其他段的第 i 个单元作为备选。最后，对于每一个备选单元，在本地缓冲寄存器中找到与其距离最近的位置，若距离足够近(在单元内部)，则用此位置替换备选位置，否则将备选位置做小范围的偏移。该方案利用希尔波特曲线降低了空间维度，从而减小了系统开销。此外，结合陌生人之间的协作技术能够收集真实用户位置信息的特点和伪造虚假位置技术计算和通信开销较小的特点，实现用户对系统开销的细粒度控制。

5.4.2 基于信任关系的协作

基于信任关系的协作通常使用蓝牙或 ad hoc 网络建立协作组，协作组成员之间相互信任，用以分享位置或其他信息。

Chow 等[31]提出 P2Pcloaking 方案，可同时满足 k-匿名和隐匿区域要求。如图 17 所示，用户向距离为 1 的邻近组员广播请求，每个邻近组员返回其 ID 和位置信息，若返回结果少于 k 个，则将距离设为 2 重新发送请求。当组员发现请求中的距离大于 1 时，不仅返回自身 ID 和位置信息，还需向其相邻组员广播请求，该请求中的距离是其收到请求中的距离减 1。当距离等于 1 时，组员只需返回自身 ID 和位置信息。用户通过不断增加距离，使返回信息的组员数量至少为 k–1。例如，图 17(a)中 m_8 为需要位置服务的用户，设 k 等于 5，先向距离为 1 的邻近组员广播请求，发现组员 m_5、m_7 和 m_9。由于没有达到 5-匿名的要求，因此将距离设为 2 重新发送请求，m_5、m_7 和 m_9 收到请求后，发现距离大于 1，则返回自身 ID 和位置信息的同时向其相邻组员广播请求，请求中的距离等于 1。如图 17(b)所示，组员 m_4、m_6、m_{10} 和 m_{15} 收到请求，由于距离等于 1，故只返回自身 ID 和位置信息。此时 m_8 收到了 7 个组员的应答，满足 5-匿名要求。接着用户从这些返回信息的组员中，挑选 k–1 个最近组员形成匿名集。如图 17(c)所示，用户选择 m_4、m_5、m_7 和 m_9 形成匿名集。最后若匿名集形成的隐匿区域小于用

户的最小隐匿区域要求，则向外扩张使其符合要求，如图 17(d)所示，扩大隐匿区域的边长。Niu 等[30]指出文献[31]存在的问题：1) 由于 ad hoc 网络是一种短距离通信方式，且具有广播特性，因此，该方案挑选的 k–1 个组员都在真实用户周围，真实用户位于隐匿区域中心的概率极高；2) 由于该方案采用逐跳挑选组员的方式，相对远离用户的组员很难被选中，因此方案所形成的隐匿区域较小。虽然 Chow 等在文献[12]中对算法进行了改进，首先调整了隐匿区域的中心，以用户位置和任意相邻组员的中点作为中心，其次根据用户要求适当扩大隐匿区域，但通过实验证明该修改方案仍然不能抵御文献[30]中提出的基于方差的攻击。针对上述问题，文献[30]中提出了如下解决方案：首先 Alice 向距离一跳的邻近组员广播请求，收到应答后随机选择 $f(1\leqslant f\leqslant k)$ 个组员，并将这些组员和用户一并作为备选集。接着，从备选集中随机选出一个组员 Bob，将备选集发送给 Bob，由 Bob 再次广播请求，收到应答后同样随机选择 f 个组员，并将这些组员加入到备选集中。重复上一步操作，直到备选集中组员的数量大于或等于 k。最后一个广播请求的组员 Frank 将备选集发送给 Alice，由 Alice 构建包含备选集中所有组员的隐匿区域。该方案中，f 是一个重要参数，首先对于 Bob 而言，Alice 的真实位置隐匿在 f 个位置中。其次，f 可以用来调节隐私保护水平和系统开销。若 f 取最大值 k，则该方案与文献[27]相同，可能只需一跳就能完成，系统开销很小，但无法抵御相关攻击。若 f 取最小值 1，则可能需要多跳完成该算法，这样便可形成较大的隐匿区域，但系统开销较大。目前，大量基于群组协作的位置隐私保护方案提出了缓存的思想[37,38,65]。用户将服务信息缓存一段时间，并向其他需要该信息的组员转发，从而减少用户向服务提供商提出请求的次数，以降低用户隐私泄露的可能性。Shokri 等[65]提出 MobiCrowd 方案，此方案对于服务器而言，需要做的改进是对每条服务信息进行数字签名，用户则可使用服务提供商的公钥来验证位置服务信息，以防止篡改位置服务信息或传播过期位置服务信息的行为。每个用户设备都维持一个缓冲寄存器，用以存储其获得的其他组员或者服务提供商发送的服务信息。每条服务信息中包含到期时间，在过期前数据一直存储在缓冲寄存器中。当用户在缓冲寄存器中没有找到满足服务请求的信息时，服务请求将通过 ad hoc 网络发送给其他组员，若某些组员缓存了该服务请求所对应的服务信息，则向该用户发送应答。一般会设置一个服务周期以降低每个用户的通信开销。当用户没有收到任何应答时，才向服务提供商提出服务请求。MobiCrowd 方案没有考虑当缓存数据无法满足用户需求时如何保护隐私。Niu 等[38]提出 EPclock 方案以解决上述问题。当用户需要位置服务时，先执行“本地搜索算法”，通过 ad hoc 网络从其他组员获得服务数据，若服务覆盖率超过阈值，则用户需求得到满足。否则，执行空间隐匿算法，逐跳将服务请求发送给某个组员，即虚拟请求者。虚拟请求者将其用户名加入服务请求后，继续转发服务请求给另一个组员，即发送者，由发送者向服务提供商发送该请求，请求范围覆盖用户所需区域，并将应答返回给用户。由于服务请求中，用户 ID 来源于虚拟请求者，POI 来源于用户，位置来源于发送者，因此攻击者难以获得任何参与者的隐私信息。

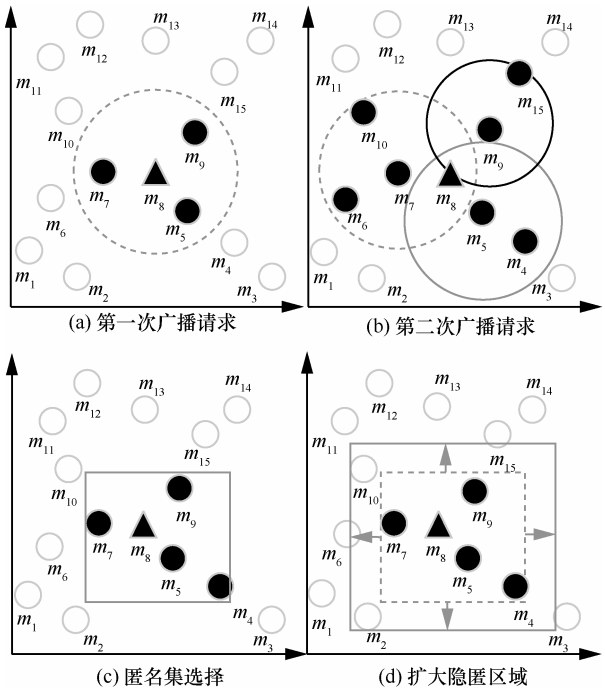


图 17　P2P 空间隐匿算法示例

6　现有技术对比分析

6.1　4 类技术优缺点分析

本文综述了位置服务中 2 种隐私保护系统构架下基于泛化和模糊的 4 类 LBS 隐私保护技术。

依赖可信第三方系统构架中，现有方案一般采

用空间隐匿技术将真实位置隐藏在包含其他 k–1 个真实用户的隐匿区域中。该技术既能保护位置隐私，又能保护查询隐私，一般使用 k-匿名和位置熵作为隐私保护度量标准。其优势在于：1) 能够同时满足用户的 k-匿名和隐匿区域要求；2) TTP 能够获取海量用户的位置信息；3) TTP 能够辅助用户过滤服务数据，降低用户端的计算和存储开销。缺点在于：1) 在人口稀少地区，该技术容易导致隐匿区域过大或服务延迟过高；2) 隐匿区域较大时，服务器端需消耗大量计算资源处理查询请求；3) 系统负载的增加和单点失效的发生，使 TTP 成为系统瓶颈。

不依赖可信第三方的系统构架中，现有方案主要集中于 3 类技术：位置偏移和模糊技术、伪造虚假位置技术和群组协作技术。

位置偏移和模糊技术在真实位置中加入噪声以降低位置精确度。该技术主要用于保护位置隐私，使用相关性和位置不可分辨性作为隐私保护度量标准。其优势在于：1) 计算和存储开销较小；2) 通过调整噪声函数的相关参数，可达到较高的隐私保护水平；3) 根据不同应用需求，用户可调节位置精确度损失。其缺点在于：1) 降低了位置精确度，可能难以满足用户服务需求；2) 在人口稀少地区，难以达到应有的隐私保护水平；3) 该技术主要针对位置隐私，不能保护查询隐私。

伪造虚假位置技术将真实位置和多个虚假位置一并发送给服务提供商。该技术既能保护位置隐私，又能保护查询隐私，一般使用 k-匿名和位置熵作为隐私保护度量标准。其优势在于：1) 不受人口密度限制，无论用户处于何位置都能满足用户的隐私保护要求；2) 由于向服务提供商发送的信息中包含真实位置，因此返回的服务数据较为精确；3) 便于用户随时调整隐私保护策略。其缺点在于：1) LPPM 的实现受移动设备性能的限制；2) 由于无法收集和利用真实的用户位置，攻击者可以通过各种背景知识排除虚假位置；3) 虚假位置会带来额外的通信消耗。

群组协作技术通过组员间分享位置信息，相互协作完成隐私保护工作。该技术既能保护位置隐私，又能保护查询隐私，一般使用 k-匿名和位置熵作为隐私保护度量标准。其优势在于：1) 可获取周围用户的真实位置，能够抵御基于背景知识的攻击；2) 通过缓存和共享服务信息减少查询次数，提高整个群组的隐私保护水平；3) 可结合其他几种技术提高隐私保护水平或者缓存命中率。其缺点在于：1) ad hoc 网络通信距离较短，无法形成较大隐匿区域；2) 若用户处于人口稀少区域，该方法可行性较差；3) 组员间分享和缓存相关信息增加了额外的通信和存储开销。

6.2 现有方案存在问题

通过分析现有工作，不难看出 LBS 隐私保护方案有以下几方面问题亟需解决。

1) 如表 1 所示，现有方案主要针对位置信息和用户 ID 这 2 个目标，将 POI 和时间信息作为保护目标的相对较少，尤其是对时间信息的保护尚未得到足够重视。对于掌握细粒度背景知识的攻击者，可利用多个位置的时间相关性来推测用户真实位置，对于时间信息的保护能够抵抗此类攻击。所以，在很多应用场景中，单一的隐私保护目标不能达到很好的隐私保护效果，只有综合考虑多个隐私保护目标才能设计出更有效的方案。

2) 现有方案大都没有定量权衡隐私保护水平和位置精确度。为提高隐私保护水平，很多技术牺牲了一定的位置精确度，如使用隐匿区域代替真实位置、在真实位置中加入噪声等。然而，若位置精确度太低，会导致服务数据不能满足用户需求，影响可用性，隐私保护失去意义。此外，不同应用对位置精确度的要求各异，如对于 POI 搜索应用，真实位置和虚假位置的距离应有严格限制；对于天气预报应用，真实位置和虚假位置的距离要求较低。因此，设计方案时应对隐私保护水平和位置精确度进行均衡分析。

3) 如表 2 所示，现有研究大多数都是从提高隐私保护水平角度出发，很少考虑系统开销问题。然而，由于用户设备端的系统资源有限，精度损失、通信开销、存储开销和计算开销等都会影响用户体验。如大量计算开销使移动设备的处理速度变慢，大量通信开销使用户额外费用增加，大量电量开销影响移动设备在户外的使用，最终阻碍位置服务的发展。因此，设计方案时应当充分考虑降低系统开销。

4) 现有研究大多针对如何设计和改进隐私保护方案，专门针对隐私保护度量指标的研究尚不完善，由此导致对于不同隐私保护方案的评判没有一个统一的标准。因此，应当研究如何制定统一的隐私度量指标使各类方案都能得到合理的评估。

表 1　现有位置隐私保护技术的目标

位置隐私保护技术	身份	位置	时间	POI
空间隐匿	是	是	否	否
位置偏移和模糊	否	是	否	否
伪造虚假位置	是	是	否	是
群组协作	是	是	否	否

表 2　现有位置隐私保护技术的系统开销

位置隐私保护技术	精度损失	通信开销	计算开销	存储开销
空间隐匿	中	高	高	高
位置偏移和模糊	高	低	中	中
伪造虚假位置	高	中	高	高
群组协作	中	高	高	中

7　结束语

LBS 已成为人们日常生活中不可或缺的一种信息服务方式，因此，如何在使用 LBS 过程中有效地防止隐私泄露成为隐私保护领域的研究热点。本文全面介绍了 LBS 隐私保护背景知识，讨论了现有攻击者模型和隐私度量指标，并深入分析和对比了 4 类基于泛化和模糊的隐私保护技术，指出了现有方案的优点和不足。未来面对云计算、大数据等新技术带来的机遇，现有的隐私保护技术面临着各种新的挑战，研究者适时提出隐私计算理论体系[67]。在未来位置隐私保护方面，亟待从以下几方面为用户隐私数据全生命周期保护提供理论与技术支撑。

1) 扩展背景知识，建立更为可靠的 LBS 隐私保护机制。通过挖掘和利用背景知识，攻击者可发动更为有效的攻击。现有方案缺乏对时间、用户运动和行为模型等背景知识的充分考虑，而攻击者可能会收集此类信息。因此，对背景知识的扩展研究，并将其应用到具体方案中是一个重要的研究方向。

2) 设计合适的 LBS 隐私保护度量标准。*k*-匿名和位置熵是最为广泛的 LBS 隐私保护度量标准，但很难准确表达用户多样化的隐私保护要求。差分隐私作为一种严格的、可证明的度量标准，在数据库隐私保护方向已经取得了很多成果，如拉普拉斯机制和指数机制。将差分隐私和相应隐私保护机制引入到位置隐私保护领域是一个重要的研究方向。

3) 平衡隐私保护和服务质量。很多方案提高隐私保护水平需以牺牲服务质量为代价。博弈论的方法能够较好地解决双方或多方的利益均衡问题，如零和贝叶斯博弈、纳什均衡等。因此，将博弈论的方法引入到 LBS 隐私保护技术中是一个重要的研究方向。

参考文献：

[1] 王璐, 孟小峰. 位置大数据隐私保护研究综述[J]. 软件学报, 2014, 25(4): 693-712.
WANG L, MENG X F. Location privacy preservation in big data era: a survey[J]. Journal of Software, 2014, 25(4): 693-712.

[2] HOH B, GRUTESER M. Protecting location privacy through path confusion[C]//1st IEEE International Conference on Security and Privacy for Emerging Areas in Communications Networks (SecureComm'05). Athens, Greece, 2005: 194-205.

[3] KRUMM J. Inference attacks on location tracks[C]//5th Springer International Conference on Pervasive Computing(PERVASIVE'07). Toronto, Canada, 2007: 127-143.

[4] FAWAZ K, SHIN K G. Location privacy protection for smartphone users[C]//21st ACM Conference on Computer and Communications Security(CCS'14). Scottsdale, United States, 2014: 239-250.

[5] SHIN K G, JU X, CHEN Z, et al. Privacy protection for users of location-based services[J]. IEEE Wireless Communications, 2012, 19(1): 30-39.

[6] ZHANG Y, TAN C C, XU F, et al. VProof: lightweight privacy-preserving vehicle location proofs[J]. IEEE Transactions on Vehicular Technology, 2015, 64(1): 378-385.

[7] WANG X, PANDE A, ZHU J, et al. STAMP: enabling privacy-preserving location proofs for mobile users[J]. IEEE Transactions on Networking, 2016, PP(99): 1-14.

[8] DICKINSON A, LOCHRIE M, EGGLESTONE P. UKKO: enriching persuasive location based games with environmental sensor data[C]//2th ACM Annual Symposium on Computer-Human Interaction in Play(CHI Play'15). London, United Kingdom, 2015: 493-498.

[9] 周傲英, 杨彬, 金澈清, 等. 基于位置的服务: 架构与进展[J]. 计算机学报, 2011, 34(7): 1155-1171.
ZHOU A Y, YANG B, JIN C Q, et al. Location-based services: architecture and progress[J]. Journal of Software, 2011, 34(7): 1155-1171.

[10] WERNKE M, SKVORTSOV P, DÜRR F, et al. A classification of location privacy attacks and approaches[J]. Personal and Ubiquitous Computing, 2014, 18(1): 163-175.

[11] MACHANAVAJJHALA A, KIFER D, GEHRKE J, et al. *L*-diversity: privacy beyond *k*-anonymity[J]. ACM Transactions on Knowledge Discovery from Data, 2007, 1(1): 3.

[12] CHOW C Y, MOKBEL M F, LIU X. Spatial cloaking for anonymous location-based services in mobile peer-to-peer environments[J]. GeoInformatica, 2011, 15(2): 351-380.

[13] MOKBEL M F. Privacy in location-based services: state-of the-art and research directions[C]//8th IEEE International Conference on Mobile Data Management (MDM'07). Mannheim, Germany, 2007: 228-228.

[14] THEODORAKOPOULOS G. The same-origin attack against location privacy[C]//14th ACM Workshop on Privacy in the Electronic Soci-

ety(WPES'15). Denver, United States, 2015: 49-53.

[15] SHOKRI R, THEODORAKOPOULOS G, BOUDEC J Y LE, et al. Quantifying location privacy[C]//32nd IEEE Symposium on Security and Privacy(S&P'11). Berkeley, United States, 2011: 247-262.

[16] LEE B, OH J, YU H, et al. Protecting location privacy using location semantics[C]//17th ACM international conference on Knowledge discovery and data mining(SIGKDD'11). San Diego, United States, 2011: 1289-1297.

[17] PAN X, XU J, MENG X. Protecting location privacy against location-dependent attacks in mobile services[J]. IEEE Transactions on Knowledge and Data Engineering, 2012, 24(8): 1506-1519.

[18] 刘海, 李兴华, 王二蒙, 等. 连续服务请求下基于假位置的用户隐私增强方法[J]. 通信学报, 2016, 37(7): 140-150.
LIU H, LI X H, WANG E M, et al. Privacy enhancing method for dummy based privacy protection with continuous location-based service queries[J]. Journal on Communications, 2016, 37(7): 140-150.

[19] 潘晓, 郝兴, 孟小峰. 基于位置服务中的连续查询隐私保护研究[J]. 计算机研究与发展, 2010, 47(1): 121-129.
PAN X, HAO X, MENG X F. Primacy prospering towards continuous query in location-based services[J]. Journal of Computer Research and Development, 2010, 47(1): 121-129.

[20] GRUTESER M, GRUNWALD D. Anonymous usage of location-based services through spatial and temporal cloaking[C]//1st ACM International Conference on Mobile Systems, Applications, and Services(MobiSys'03). San Francisco, United States, 2003: 31-42.

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

[22] CHOW C Y, MOKBEL M F, AREF W G. Casper*: query processing for location services without compromising privacy[J]. ACM Transactions on Database Systems, 2009, 34(4): 1-48.

[23] XU T, CAI Y. Feeling-based location privacy protection for location-based services[C]//16th ACM Conference on Computer and Communications Security(CCS'09). Chicago, United States, 2009: 348-357.

[24] GEDIK B, LIU L. Protecting location privacy with personalized *k*-anonymity: architecture and algorithms[J]. IEEE Transactions on Mobile Computing, 2008, 7(1): 1-18.

[25] GEDIK B, LIU L. Protecting location privacy with personalized *k*-anonymity: architecture and algorithms[J]. IEEE Transactions on Mobile Computing, 2008, 7(1): 1-18.

[26] 倪巍伟, 马中希, 陈萧. 面向路网隐私保护连续近邻查询的安全区域构建[J]. 计算机学报, 2016, 39(3): 628-642.
NI W W, MA Z X, CHEN X. Safe region scheme for privacy-preserving continuous recent neighbor query on road networks[J]. Journal of Computers, 2016, 39(3): 628-642.

[27] NIU B, LI Q H, ZHU X Y, et al. A fine-grained spatial cloaking scheme for privacy-aware users in location-based services[C]//23rd IEEE International Conference on Computer Communication and Networks (ICCCN'14). Shanghai, China, 2014: 1-8.

[28] KIDO H, YANGISAWA Y, SATOH T. An anonymous communication technique using dummies for location-based services[C]//2nd IEEE International Conference on Pervasive Services(ICPS'05). Santorini, Greece, 2005: 88-97.

[29] LU H, JENSEN C S, YIU M L. Pad: privacy-area aware, dummy-based location privacy in mobile services[C]//7th ACM International Workshop on Data Engineering for Wireless and Mobile Access(MobiDE'08). Vancouver, Canada, 2008: 16-23.

[30] NIU B, ZHU X Y, LI Q H, et al. A novel attack to spatial cloaking schemes in location-based services[J]. Future Generation Computer Systems, 2015, 49(1): 125-132.

[31] CHOW C Y, MOKBEL M F, LIU X. A peer-to-peer spatial cloaking algorithm for anonymous location-based services[C]//14th ACM International Symposium on Advances in Geographic information Systems(ACM-GIS'06). Arlington, USA, 2006: 171-178.

[32] CICEK A E, NERGIZ M E, SAYGIN Y. Ensuring location diversity in privacy-preserving spatio-temporal data publishing[J]. The VLDB Journal, 2014, 23(4): 609-625.

[33] ZHANG X, XIA Y, BAE H Y. A novel location privacy preservation method for moving object[J]. International Journal of Security and Its Applications, 2015, 9(2): 1-12.

[34] ZHANG C, HUANG Y. Cloaking locations for anonymous location based services: a hybrid approach[J]. GeoInformatica, 2009, 13(2): 159-182.

[35] BERESFORD A, STAJANO F. Location privacy in pervasive computing[J]. IEEE Pervasive Computing, 2003, 2(1): 46-55.

[36] NIU B, LI Q H, ZHU X Y, et al. Achieving *k*-anonymity in privacy-aware location-based services[C]//33th IEEE International Conference on Computer Communications(INFOCOM'14). Toronto, Canada, 2014: 754-762.

[37] NIU B, LI Q H, ZHU X Y, et al. Enhancing privacy through caching in location-based services[C]//34th IEEE International Conference on Computer Communications(INFOCOM'15). Hong Kong, China, 2015: 1017-1025.

[38] NIU B, ZHU X Y, LI W H, et al. EPcloak: an efficient and privacy-preserving spatial cloaking scheme for LBSs[C]//11th IEEE International Conference on Mobile Ad Hoc and Sensor Systems(MASS'14). Philadelphia, United States, 2014: 398-406.

[39] 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.

[40] 彭长根, 丁红发, 朱义杰, 等. 隐私保护的信息熵模型及其度量方法[J]. 软件学报, 2016, 27(8): 1891-1903.
PENG C G, DING H F, ZHU Y J, et al. Information entropy models and privacy metrics methods for privacy protection[J]. Journal of Software, 2016, 27(8): 1891-1903.

[41] SHOKRI R, FREUDIGER J, JADLIWALA M, et al. A distortion-based metric for location privacy[C]//8th ACM Workshop on Privacy in the Electronic Society(WPES'09). Chicago, United States, 2009: 21-30.

[42] DWORK C, LEI J. Differential privacy and robust statistics[C]//41st Annual ACM Symposium on Theory of Computing(STOC'09). Bethesda, United States, 2009: 371-380.

[43] ANDRES M E, BORDENABE N E, CHATZIKOKOLAKIS K, et al. Geo-indistinguishability: differential privacy for location-based systems[C]//20th ACM Conference on Computer and Communications Security(CCS'13). Berlin, Germany, 2013: 901-914.

[44] XIAO Y, XIONG L. Protecting locations with differential privacy under temporal correlations[C]//22nd ACM Conference on Computer and Communications Security(CCS'15). Denver, United States, 2015:

1298-1309.

[45] IETF. Geographic location/privacy (geopriv) [EB/OL]. http://datatracker.ietf. org/wg/geopriv/charter/

[46] World Wide Web Consortium(W3C). Platform for privacy preferences(P3P) project. [EB/OL]. http://www.w3. org/P3P

[47] PAULET R, KAOSAR M G, YI X, et al. Privacy-preserving and content-protecting location based queries[J]. IEEE Transactions on Knowledge and Data Engineering, 2014, 26(5): 1200-1210.

[48] GHINITA G, KALNIS P, KHOSHGOZARAN A, et al. Private queries in location based services: anonymizers are not necessary[C]//27th ACM Conference on Management of Data(SIGMOD'08). Vancouver, Canada, 2008: 121-132.

[49] YI X, PAULET R, BERTINO E, et al. Practical approximate *k* nearest neighbor queries with location and query privacy[J]. IEEE Transactions on Knowledge and Data Engineering, 2016, PP(99): 1-14.

[50] KALNIS P, GHINITA G, MOURATIDIS K, et al. Preventing location-based identity inference in anonymous spatial queries[J]. IEEE Transactions on Knowledge and Data Engineering, 2007, 19(12): 1719-1733.

[51] NGO H, KIM J. Location privacy via differential private perturbation of cloaking area[C]//28th IEEE Computer Security Foundations Symposium(CSF'15). Verona, Italy, 2015: 63-74.

[52] YIU M L, JENSEN C S, MØLLER J, et al. Design and analysis of a ranking approach to private location-based services[J]. ACM Transactions on Database Systems, 2011, 36(2): 1-42.

[53] PERAZZO P, DINI G. A uniformity-based approach to location privacy[J]. Computer Communications, 2015, 64(1): 21-32.

[54] ARDAGNA C A, CREMONINI M, DE CAPITANI DI VIMERCATI S, et al. An obfuscation-based approach for protecting location privacy[J]. IEEE Transactions on Dependable and Secure Computing, 2011, 8(1): 13-27.

[55] DÜRR F, SKVORTSOV P, ROTHERMEL K. Position sharing for location privacy in non-trusted systems[C]//9th IEEE International Conference on Pervasive Computing and Communications(PerCom'11). Seattle, United States, 2011: 189-196.

[56] PINGLEY A, YU W, ZHANG N, et al. Cap: a context-aware privacy protection system for location-based services[C]//29th IEEE International Conference on Distributed Computing Systems(ICDCS'09). Montreal, Canada, 2009: 49-57.

[57] 周长利, 马春光, 杨松涛. 基于敏感位置多样性的 LBS 位置隐私保护方法研究[J]. 通信学报, 2015, 36(4): 14-25.
ZHOU C L, MA C G, YANG S T. Research of LBS privacy preserving based on sensitive location diversity[J]. Journal on Communications, 2015, 36(4): 14-25.

[58] SHOKRI R, THEODORAKOPOULOS G, TRONCOSO C, et al. Protecting location privacy: optimal strategy against localization attacks[C]//19th ACM Conference on Computer and Communications Security(CCS'12). Raleigh, United States, 2012: 617-627.

[59] BORDENABE N E, CHATZIKOKOLAKIS K, PALAMIDESSI C. Optimal geo-indistinguishable mechanisms for location privacy[C]//21st ACM Conference on Computer and Communications Security(CCS'14). Scottsdale, United States, 2014: 251-262.

[60] 倪巍伟, 陈萧, 马中希. 支持偏好调控的路网隐私保护 *k* 近邻查询方法[J]. 计算机学报, 2015, 38(4): 884-896.
NI W W, CHENG X, MA Z X. Location privacy preserving *k*-nearest neighbor query method on road network in presence of user's preference[J]. Chinese Journal of Computers, 2015, 38(4): 884-896.

[61] PINGLEY A, ZHANG N, FU X, et al. Protection of query privacy for continuous location based services[C]//30th IEEE International Conference on Computer Communications (INFOCOM'11). Shanghai, China, 2011: 1710-1718.

[62] MA C Y T, YAU D K Y, YIP N K, et al. Privacy vulnerability of published anonymous mobility traces[J]. IEEE Transactions on Networking, 2013, 21(3): 720-733.

[63] LIU X, LIU K, GUO L, et al. A game-theoretic approach for achieving k-anonymity in location based services[C]//32nd IEEE International Conference on Computer Communications(INFOCOM'13). Turin, Italy, 2013: 2985-2993.

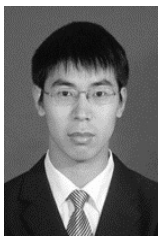
[64] MANWEILER J, SCUDELLARI R, COX L P. Smile: encounter-based trust for mobile social services[C]//16th ACM Conference on Computer and Communications Security(CCS'09). Chicago, United States, 2009: 246-255.

[65] SHOKRI R, THEODORAKOPOULOS G, PAPADIMITRATOS P, et al. Hiding in the mobile crowd: location privacy through collaboration[J]. IEEE Transactions on Dependable and Secure Computing, 2014, 11(3): 266-279.

[66] 黄毅, 霍峥, 孟小峰. CoPrivacy: 一种用户协作无匿名区域的位置隐私保护方法[J]. 计算机学报, 2011, 34(10): 1976-1985.
HUANG Y, HUO Z, MENG X F. CoPrivacy: a collaborative cocation privacy-preserving method without clocking region[J]. Chinese Journal of Computers, 2011, 34(10): 1976-1985.

[67] 李凤华, 李晖, 贾焰, 等. 隐私计算研究范畴及发展趋势[J]. 通信学报, 2016, 37(4): 1-11.
LI F H, LI H, JIA Y, et al. Privacy computing: concept, connotation and its research trend[J]. Journal on Communications, 2016, 37(4): 1-11.

作者简介：



万盛（1987-），男，江苏南通人，西安电子科技大学博士生，主要研究方向为网络安全与隐私保护。

李凤华（1966-），男，湖北浠水人，博士，中国科学院信息工程研究所副总工程师、研究员、博士生导师，主要研究方向为网络与系统安全、隐私计算、可信计算。

牛犇（1984-），男，陕西西安人，博士，中国科学院信息工程研究所助理研究员，主要研究方向为网络安全、隐私计算。

孙哲（1987-），男，安徽安庆人，中国科学院信息工程研究所博士生，主要研究方向为信息安全、隐私保护。

李晖（1968-），男，河南灵宝人，博士，西安电子科技大学教授、博士生导师，主要研究方向为密码学、无线网络安全、云计算安全、信息论与编码理论。