Computer Science and Application 计算机科学与应用, 2016, 6(6), 354-367
Published Online June 2016 in Hans. http://www.hanspub.org/journal/csa
http://dx.doi.org/10.12677/csa.2016.66044

Hans汉斯

A Survey on Location Privacy Preserving Techniques

Xiaodan Lu, Lefeng Zhang, Ping Xiong

School of Information Security Engineering, Zhongnan University of Econimocs and Law, Wuhan Hubei
Email: lxd0937@126.com

Received: Jun. 15th, 2016; accepted: Jun. 24th, 2016; published: Jun. 28th, 2016

Abstract

While location-based services (LBSs) have become increasingly popular and provided enormous benefits in daily life, location privacy of individuals has been confronted with serious concerns. To address the issue, a number of location privacy preserving techniques have been proposed during the last decade. This paper surveys the state of the art of location privacy preserving techniques. Firstly, we introduce the general framework of LBS system as well as the potential threats to LBS user. Then we group the location privacy preserving techniques into four categories including space cloaking, dummy-based method, private information retrieval, and differential privacy-based method. The general principle and representative techniques of each category are discussed in detail respectively and a comparison of the techniques is presented. Finally, we summarize some new research directions and make a conclusion.

Keywords

Location-Based Services, Location Privacy, Privacy Preserving, Differential Privacy

位置隐私保护技术研究综述

卢小丹，张乐峰，熊　平

中南财经政法大学信息与安全工程学院，湖北　武汉
Email: lxd0937@126.com

收稿日期：2016年6月15日；录用日期：2016年6月24日；发布日期：2016年6月28日

文章引用：卢小丹，张乐峰，熊平．位置隐私保护技术研究综述[J]．计算机科学与应用, 2016, 6(6): 354-367.
http://dx.doi.org/10.12677/csa.2016.66044

摘　要

随着位置信息服务的日益普及，位置信息中包含的个人隐私信息逐渐受到了人们的广泛关注。学术界近年来对位置隐私保护问题进行了深入研究并提出了一系列实现技术。本文对位置隐私保护技术的研究进展进行综述。首先介绍基于位置的服务系统的基本框架及其面临的风险，然后将位置隐私保护技术划分为四类，包括空间模糊化、虚拟对象技术、隐私信息检索和差分隐私保护技术，详细讨论它们的基本原理及有代表性的实现方法，并在此基础上进行性能上的分析和比较。最后归纳总结位置隐私保护进一步的研究方向。

关键词

基于位置的服务，位置隐私，隐私保护，差分隐私

1. 引言

随着智能手机和无线网络的广泛应用，移动应用近年来呈现出爆炸式的增长。据统计 2014 年全球移动应用的使用量增长了 76% [1]。至 2014 年底，谷歌 Play 拥有 143 万款应用，Apple Store 拥有 121 万款应用[2]，其中基于位置的服务(Location-based Service, LBS)备受用户青睐。以百度地图为例，从 2012 年成立至今，用户数从最初 7000 万增长到超 2 亿活跃用户[3]，它提供的定位、导航、查询等位置信息服务为现代生活带来了极大的便利。然而，用户提交的位置服务请求和自己的位置信息在一定程度上会对个人的隐私造成泄露风险[4] [5]。一方面，对某些用户而言，其位置信息本身就是隐私数据；另一方面，攻击者可以根据位置信息来推测用户的个人身份、工作性质、健康状况或者兴趣爱好等隐私信息[6]-[9]。

近年来，学术界对位置隐私保护问题进行了广泛的研究，并提出了一系列保护方法。从已有的位置隐私保护技术来看，可以将其分为四类，即空间模糊化、虚拟对象、隐私信息检索(Privacy Information Retrieval, PIR)和差分隐私保护。其中，空间模糊化技术是将用户的真实位置模糊成一个满足用户个性隐私需求的空间，并用模糊后的空间代替精确位置提交给位置信息服务器处理；虚拟对象技术将虚拟的对象与真实对象混合在一起作为位置服务请求发送者，使攻击者无法实现位置与用户的准确映射，这种方法的研究重点在于如何合理的选择虚拟对象；PIR 技术是基于不可信数据库提出的隐私保护技术，它能实现用户访问服务器的同时阻止服务器获知用户访问内容，提供了高水平的隐私保护，该技术最大的挑战在于设计一个好的检索算法来加快检索效率和降低存储空间；差分隐私是近年来提出的一种新的隐私保护定义，由于其独立于攻击者的背景知识，并提供了严格的、可证明的隐私保护，成为目前隐私保护领域的一个研究热点。

本文首先分析 LBS 系统所面临的隐私泄露风险，然后对以上四类隐私保护技术的基本原理和实现方法进行综述，通过对已有研究成果的梳理，详细分析和比较这些技术的优缺点，最后探讨位置隐私保护技术在未来的研究方向。

2. LBS 的基本结构与威胁分析

2.1. LBS 的基本结构

LBS 的基本结构(如图 1)通常由 4 个部分构成：移动终端、定位系统、传输网络和位置信息服务器[10]。用户利用移动终端向 LBS 服务器发送位置服务请求(如“离我最近的加油站”)，服务请求则通过传输网络到达 LBS 服务器，LBS 服务器查询存储的位置信息，再将查询结果通过传输网络发送给用户的移动终

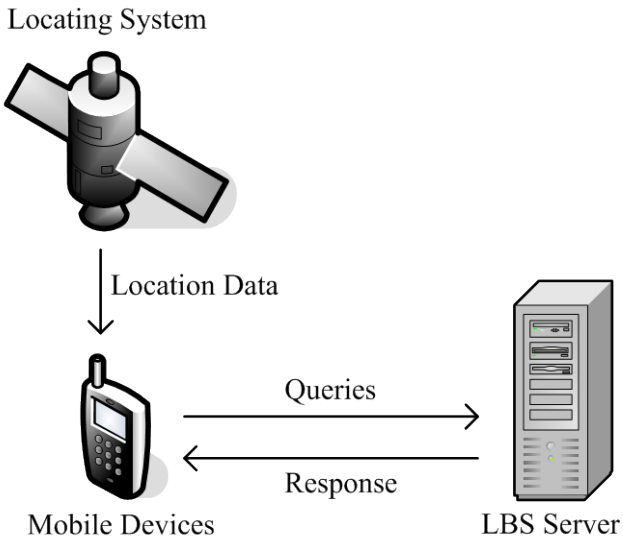

Figure 1. Basic framework of LBS system
图 1. LBS 系统的基本结构

端。用户及查询结果的位置信息由定位系统提供。可见，位置信息是 LBS 网络中传输的最核心数据，也是位置隐私保护的对象。

2.2. LBS 面临的隐私泄露威胁

在 LBS 体系结构中，传输网络和 LBS 服务器被认为是存在安全威胁的不受信任方。用户的请求信息有可能在传输网络中被截获和分析，也有可能在服务器端被非授权的泄露，从而对用户隐私造成威胁。位置隐私威胁并不仅仅指位置信息的泄露，更重要的在位置信息暴露后受到的与时间和空间相关的推理攻击[11]：攻击者可以根据用户的位置信息推断出用户的个人隐私信息。因此，根据隐私泄露程度不同，位置隐私面临威胁可分为以下三种[12]：

(1) 物理威胁：攻击者直接攻击传输网络或者服务器等物理设备获取用户最原始的位置信息；

(2) 推理威胁：攻击者在获得用户的位置信息后，利用观察、推理、挖掘等技术推断出关于用户的隐私信息[13]；

(3) 联合攻击：攻击者在获得用户位置信息后联合用户使用的其他移动应用等外部资源，对用户隐私进行更深度的挖掘。例如，攻击者可以联合用户的社交网络信息来挖掘用户朋友的隐私信息[14]。

显然，物理威胁只涉及到用户的物理位置信息，推理威胁会危及到用户的个人身份信息，联合攻击则影响到了用户的整个生活环境。位置隐私的泄露是导致以上威胁的根本原因。

3. 位置隐私保护技术

近年来，学术界对位置隐私保护的研究取得了丰富的成果，各种隐私保护理论与模型在位置隐私保护中得到应用。本节对已有的位置隐私保护技术进行梳理和比较。

3.1. 空间模糊化

空间模糊化技术[15]-[17]是指用户在进行位置信息请求时，提交给服务器的并不是用户的精确位置，而是一个空间范围，使得攻击者不能分辨用户的具体位置。*k*-anonymity 是空间模糊技术应用的最主要的隐私保护模型[18]。该模型要求数据集的任一记录都至少和其它 $k-1$ 个记录不可区分[18] [19]。Marco

Grureser [20]最先将 *k*-anonymity 的概念运用到位置隐私保护中，要求每个用户在某个时间和空间范围内与其它至少 $k-1$ 个用户不可区分，使攻击者不能从至少 k 个用户中识别攻击目标进而推断出其准确位置。*k*-anonymity 的缺点在于模型参数过于单一，仅通过 k 来确定隐私水平，在某些情况下会使隐私保护失效。例如在用户密集处，满足 *k*-anonymity 的狭小区域能够在一定程度上泄露用户的准确位置；而在用户稀疏处，则有可能用户数量达不到 k 的要求而匿名失败。

之后的研究对 *k*-anonymity 模型进行了改进以克服这些缺点。文献[21]提出的 Casper 模型在 *k*-anonymity 基础上引入了新的隐私参数 *Amin*，表示最小模糊区域。如图 2 所示，Casper 模型用金字塔的数据结构来进行位置信息存储，金字塔的每一层覆盖了同样的区域，并将这些区域划分为网格形式，每个网格记录着当前存在的用户数。当进行空间模糊时，利用网格的融合和分裂来满足匿名区域 k 和 *Amin* 的隐私水平。Casper 模型将空间模糊化的过程全部交给客户端和 LBS 服务器之间的可信任第三方(位置匿名器)来完成。用户将位置服务请求和自己的位置信息发送给位置匿名器，匿名器进行空间模糊化后发送给 LBS 服务器，服务器根据匿名后的请求查询候选结果，并返还给匿名器，匿名器进行结果筛选后将精确结果返回给用户。Casper 方案的缺点在于空间模糊化的过程比较复杂，因此对匿名器的性能要求较高。同时，匿名器作为一个必不可少的可信任第三方会带来一定的风险。另外，Casper 以真实活动的对象作为匿名参与者，在人员密度稀疏的地区就有可能导致模糊区域过大的问题，导致查询结果的可用性降低。

Casper 模型主要解决了查询快照中的位置隐私保护问题，但用户是不断运动的，连续位置之间的相关性往往可以被攻击者利用并进行相关位置攻击。相关位置攻击指攻击者利用用户运动位置之间的相关性推理用户精确位置的攻击方式。如图 3(a)所示，R_{A,t_i} 表示对象 A 在 t_i 时刻的模糊空间，$R_{A,t_{i+1}}$ 表示对象 A 在 t_{i+1} 时刻的模糊空间，假设攻击者已知 A 的最大运动速度，则可以推断 A 在 t_{i+1} 时刻可能达到的范围为 $MMB_{A,t,t_{i+1}}$，进而可以推理出 A 在 t_{i+1} 时刻的精确位置处在 $R_{A,t_{i+1}}$ 和 $MMB_{A,t,t_{i+1}}$ 相交部分。同样地，如图 3(b)所示，攻击者可以根据 A 在 t_{i+1} 时刻的模糊空间 $R_{A,t_{i+1}}$ 和 A 的最大运动速度推断出 t_i 时刻 A 可能的位置集合 MAB_{A,t_{i+1},t_i}，进而推理 A 在 t_i 时刻的精确位置处在 R_{A,t_i} 和 MAB_{A,t_{i+1},t_i} 的相交部分。因此攻击者可以利用相关位置攻击缩减攻击对象的模糊空间，降低隐私保护水平。

对此，文献[22]提出了 iCliqueCloak 模型解决相关位置攻击问题。如图 4 所示，对象 A 在 t_{i+1} 时刻生成模糊空间时要满足以下 2 个条件：(1)模糊空间 $R_{A,t_{i+1}}$ 必须包含在 t_{i+1} 时刻 A 可能到达的范围 $MMB_{A,t,t_{i+1}}$ 内；(2)以 $R_{A,t_{i+1}}$ 为基础推断出的 t_i 时刻 A 可能的位置集合 MAB_{A,t_{i+1},t_i} 必须包含 t_i 时刻 A 的模糊空间 R_{A,t_i}。iCliqueCloak 模型能够有效地防止攻击者利用相关位置攻击来缩减模糊空间，从而保持原有的隐私水平。

从实际应用效果来看，空间匿名技术能够为位置隐私保护提供了一个个性化的解决方案，但其缺点在于对区域的人口密度比较敏感。另外，匿名器的处理性能及其自身的安全性都会影响到空间匿名技术的应用效果。

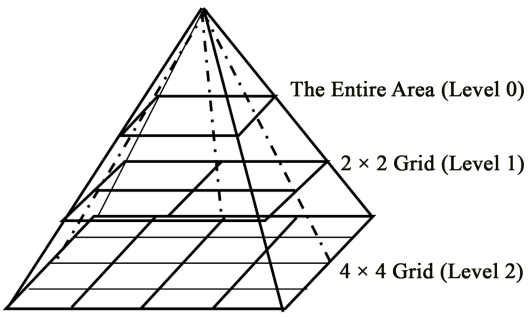

Figure 2. Data structure of Casper model
图 2. Casper 模型的数据结构

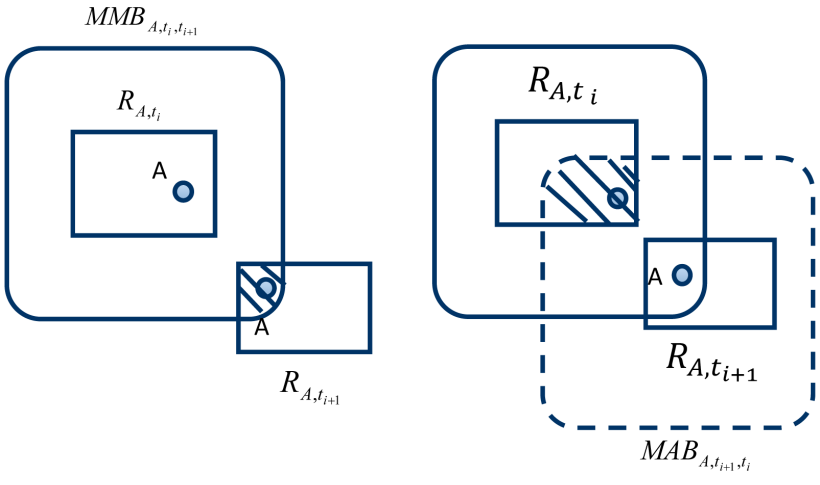

Figure 3. Relative location attack. (a) Forward deducing; (b) Backward deducing
图 3. 相关位置攻击。(a) 位置的正向推导；(b)位置的逆向推导

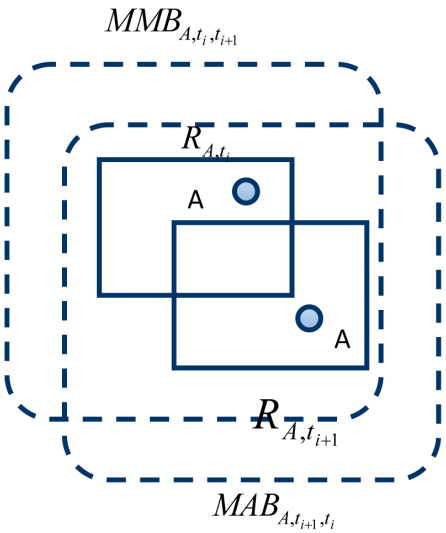

Figure 4. Requirements of iCliqueCloak
图 4. iCliqueCloak 模型的要求

3.2. 虚拟对象(Dummy)

Dummy [23]-[26]的保护方式简单来说就是用户在提交位置服务请求时，将自己的真实位置和几个虚假位置一起提交给 LBS 服务器，LBS 服务器针对所有提交的位置分别进行查询处理，将所有结果返还给用户，用户再根据自己的真实位置进行筛选。在 Dummy 的方法中，如何选择虚拟对象的位置是一个关键问题。对此，文献[24]提出了虚拟对象在分布上的一般性要求，包括分散性、稠密性和均匀性。

文献[26]提出了满足 k-anonymity 和最大模糊区域为 s 的两种虚拟位置的生成算法(Privacy-Area Aware Dummy GenerationAlgorithms, PAD)：基于圆和网格的 Dummy 生成方案。如图 5 所示，基于圆的 Dummy 生成方案是将用户真实位置 pos 绕着圆心 O 依次旋转 θ 角而成，每旋转一次生成一个 Dummy，并且满足：(1) $\theta = 2\pi/k$；(2) $0 \le d(\text{Dummy}, O) \le \sqrt{2s/(k \cdot \sin\theta)}$。基于网格的 Dummy 生成方案如图 6 所示，将用户真实位置沿着网格移动生成 Dummy，每个网格的边长 $g = \sqrt{s}/\left(\sqrt{k}-1\right)$。

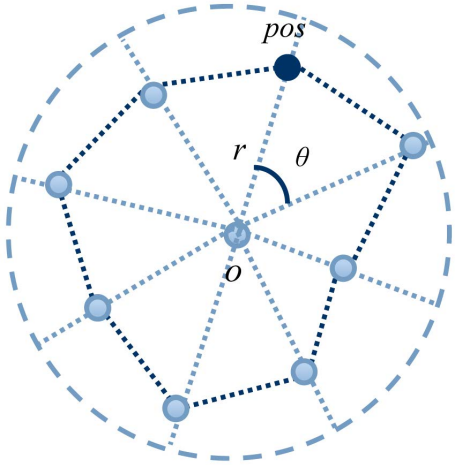

Figure 5. Circle-based dummy
图 5. 基于圆的虚拟位置

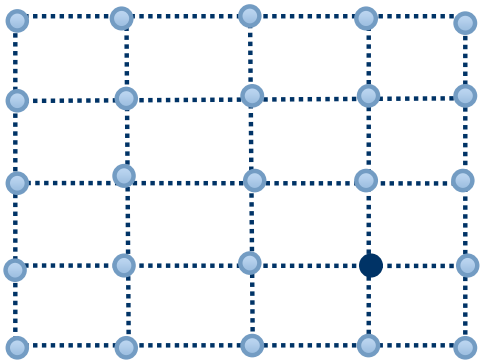

Figure 6. Grid-based dummy
图 6. 基于网格的虚拟位置

Dummy 的生成方式解决了 Casper 模型过于依赖区域人口密度的缺点，但其自身的缺点也不可忽视，PAD 算法中取得虚拟位置的方法过于规则化，而忽略了实际的地理特征。例如，按照这种规则化的方式选取的虚拟位置可能在现实环境中根本不可能有活动对象出现，那么这个虚拟位置就失去了混淆攻击者的功能。类似的，如果攻击者预先掌握了一些背景知识，例如地理环境、区域人口密度、运动最大速度等，即可实施背景知识攻击[27] [28]，将这些背景知识与获取的用户位置信息结合来推断用户的位置隐私和其他隐私信息。因此，为了阻止背景知识攻击，在选取虚拟位置时要使虚拟位置更接近真实对象的运动特点和规律。

文献[29]基于“越热门的地点越可能存在活动对象”的假设，应用熵理论建立了 DLS (Dummy-Location Selection) 模型来进行虚拟位置的选择。假设用 p_i 表示某个位置 i 被访问的概率，p_i 越大说明这个位置越热门，利用熵[30] [31]来衡量匿名程度:

$$H = -\sum_{i=1}^{k} p_i \cdot \log_2 p_i \tag{1}$$

H 值越大说明匿名效果越好。由式(1)可知，当所有参与匿名的 k 个位置的热门程度相同，即 p_i 为 $1/k$ 时，H 值达到最大。如图 7 所示，DLS 模型选取的虚拟位置分布在与真实位置热门程度接近的区域中。

使用 Dummy 的位置隐私保护机制的优点在于能够摆脱对现实环境的过度依赖，无论在人口密集区

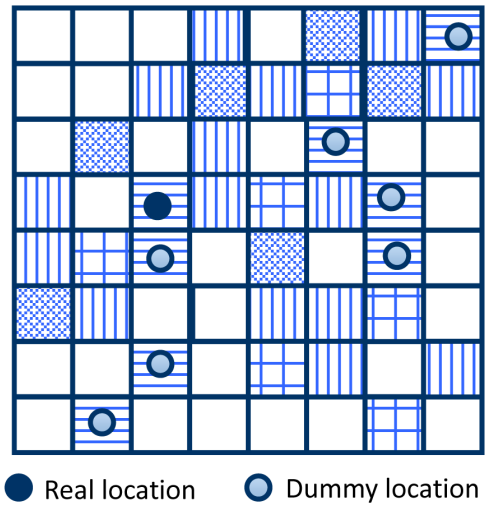

Figure 7. Dummy selection in DLS
图 7. DLS 模型中虚拟位置的选取

还是稀疏区都能较好地满足用户的隐私需求，提高了匿名化的成功率。但这些方法共同的不足在于，攻击者所掌握的背景知识是难以量化和准确建模的，因此在选取虚拟对象时往往忽略了对背景知识的考虑或者仅仅根据特定的背景知识假设提出针对性的解决方案，这样的保护机制无法应对基于新的背景知识的攻击。

3.3. 隐私信息检索

隐私信息检索(private information retrieval, PIR)是一种客户与服务器通信的安全协议，能够保证服务器无法识别客户在查询数据库时具体的查询对象，从而防止服务器端根据客户的查询对象来确定客户的兴趣点进而推断客户的隐私信息，因其能够提供了高水平的隐私保护，成为位置隐私保护的主要技术之一[32]-[35]。举例来说，PIR 要实现这样的功能：假设 Bob 拥有一个不可信任的数据库 DB，数据库包含 n 个记录，Alice 想要获取 DB[i]中的内容，PIR 协议能保证 Alice 不但能访问 DB[i]中的内容，而且不让 Bob 知道 i 的值[32]。PIR 主要分为 2 种实现形式即基于计算的 PIR [36]和基于硬件的 PIR [37] [38]。基于计算的 PIR 依赖广为人知的数学计算难题来保证检索的隐秘性，需要付出较高的计算代价，除此之外，基于计算的 PIR 在处理每个查询请求时都要线性浏览整个数据库，产生了高昂的处理成本。基于硬件的 PIR 技术目前已经在现实中开始采用(例如 IBM4758 安全协作器) [39]，具有很强的现实意义。基于硬件的 PIR 将一个受用户信赖的处理器内嵌入 LBS 服务器中，接受用户的访问请求，检索服务器中相应的数据库，并将结果加密，返还给用户。在整个检索过程中能防止服务器获知访问了哪些数据，并且用户的访问请求和检索结果被加密，能够防止信息在传输过程中被泄露。在位置隐私保护应用中，基于硬件的 PIR 系统结构如图 8 所示。

在服务器中储存了整个地区的地图和兴趣点(points of interest, POIs)信息，LBS 根据索引结构将 DB 划分为几个子数据库 DB_1, DB_2, …, PIR 处理器根据用户的请求对 DB_1, DB_2, …进行查询，并将结果返还给用户。在信息查询过程中，PIR 处理器就像一个黑盒子自动完成查询而不让服务器知道它访问了哪些子数据库。因此，这类方法的研究重点在于如何设计索引结构和访问顺序从而减少执行的检索复杂度和储存空间。文献[34]利用 Hilbert 空间曲线将 POI 的二维存储方式转化为 H 值的一维存储方式，减少了存储空间，并将 POI 根据 H 值的大小按照 B^+-tree 的结构来组织，以便简化检索次数。文献[33]则将存储区

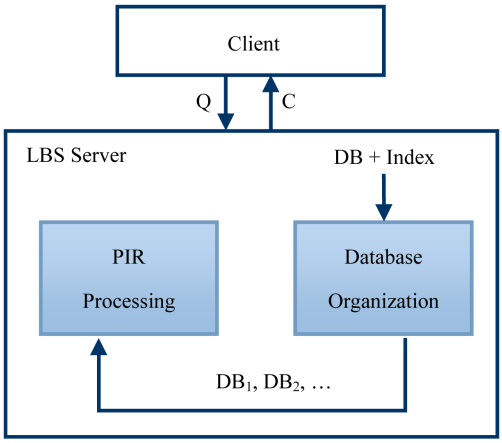

Figure 8. LBS with hardware-based PIR
图 8. 基于硬件 PIR 的 LBS 系统结构

域用网格表示，并用 Hilbert 值表示每个网格单元，同时建立了 3 个数据库 DB_1, DB_2, DB_3 来分别存储 POI 的不同信息。DB_1 按照 H 值的顺序存储每个网格单元中 POI 的数量信息，DB_2 按照 H 值存储每个 POI 的 ID、坐标和指向 DB_3 的指针，DB_3 存储了每个 POI 的其他详细信息。这样的存储结构能够在不遍历整个数据库的前提下高效地进行 *k*NN (*k* Nearest Neighbor)查询。首先根据用户的位置信息在 DB_1 中查找 *k*NN，然后在 DB_2 中确定 *k*NN 的 ID 和坐标，并根据指针在 DB_3 中获取 *k*NN 的详细信息。除此之外，还为每个查询建立了查询计划，保证每个查询都按照同样的顺序和次数进行检索，以避免外部的模式攻击。

PIR 协议能够对用户请求、信息检索及结果返回等整个通信过程都提供可靠的保密性，因此受到越来越多的关注。PIR 除运用在查询快照的位置隐私保护中，还被广泛应用在近邻查询和最短路径查询中[40]。PIR 有待继续研究的问题主要有两点：其一，PIR 的存储代价和计算代价较高，如何设计合理的检索计划和索引结构是应用 PIR 的主要挑战；其二，由于 LBS 服务器必须要存储整个区域的 POI 和地图信息，存储空间和检索效率的限制使得 PIR 目前只能用于区域范围较小的场合。

3.4. 差分隐私

差分隐私是由 Dwork 在 2006 年提出的一种新的隐私安全定义[41]。它能够保证数据集的查询结果对某个具体记录的变化不敏感，因此，一个记录存在于一个数据集里，就像它不存在于数据集里一样安全，攻击者无法通过观察和计算查询结果来推测用户的隐私信息[42]。差分隐私的定义[43]为：设随机算法 M，Range(M)为 M 所有可能的输出集合，对于任何两个邻近数据集 D1 和 D2，以及 Range(M)的子集 $S \in \text{Range}(M)$，若算法 M 满足：

$$\Pr[M(D_1) \in S] \le \exp(\varepsilon) \times \Pr[M(D_2) \in S]$$

则称算法 M 提供 ε-差分隐私保护，其中 ε 称为隐私保护预算。从原理上看， 隐私实质上是将数据集的精确查询结果转化为一个分布，使得对两个邻近数据集进行查询得到相同结果的概率几乎相同。Laplace 机制[44]和指数机制[45]是两种最基本的差分隐私实现机制。其中，Laplace 机制用于查询结果为数值型的情况，指数机制则用于保护非数值型查询结果。

由于差分隐私无需考虑攻击者掌握的任何背景知识，并能提供严格可证明的隐私保护，因此在隐私保护数据发布[46]-[49]和隐私保护数据挖掘[50]-[54]等方面得到广泛的研究和应用。显然，差分隐私更适用于保护多用户的聚合信息，在只涉及单个用户的位置隐私保护问题上并不合适。根据差分隐私的定义，用户位置的变化对查询结果的影响须微乎其微，这使得查询变得毫无意义。为解决这一问题，文献[55]

将差分隐私与 k-anonymity 结合起来，提出了一种混合模型，对于由 k 个位置构成的匿名集合，在提交位置时要求以相近的概率（小于 e^{ε}）输出 k 个位置中的任意一个。该模型的主要问题在于，匿名集合的选取对最终的隐私保护结果影响过大。

为此，文献[56]利用差分隐私的定义，提出了一种地域不可区分模型(Geo-Indistinguishability)。该模型基于位置隐私保护的现实，认为用户位置的微小变化应该对查询结果影响很小，但当用户位置变化较大时，查询结果可以有较大的变化，因此可以根据用户位置的变化程度来设定相应的隐私保护水平。Geo-Indistinguishability 的定义为：设 X 表示用户可能的位置集合，Z 表示可能发布的位置集合，$d(\cdot,\cdot)$表示欧氏距离，对于任意两个位置 $x_1, x_2 \in X$ 和 $z \in Z$ 并且 $d(x_1, x_2) \le r$，若算法 K 满足：

$$\Pr[K(x_1) = z] \le \exp(\varepsilon \cdot d(x_1, x_2)) \cdot \Pr[K(x_2) = z]$$

则称 K 在半径 r 内满足 ε-地域不可分，其中 ε 表示每单位距离的隐私保护水平。这一定义表明，对于两个非常接近的真实位置 x_1 和 x_2，它们产生相同新位置 z 的概率分布也越接近；反之，随着 x_1 和 x_2 距离增大，产生相同新位置 z 的概率分布则可以相差较大，两个概率分布之间的差异由隐私保护水平 $\varepsilon \cdot d(x_1, x_2)$ 来控制。如图 9 所示，用户 A 在以半径为 r 的圆形区域内能享受 εr 的隐私保护水平，r 越小则隐私保护水平越高，反之则隐私保护水平越低。Geo-Indistinguishability 可以通过向用户的真实位置添加二维 Laplace 噪声来实现。地域不可区分模型为差分隐私在位置隐私保护中的应用提出了一个切实可行的机制，成为一些后续研究的基础。

如何降低噪声量是差分隐私在应用中无法回避的问题。文献[57]认为，Geo-Indistinguishability 模型在保护单个位置(用户只进行一次查询)时是有效的，但一个用户往往会进行多次查询，连续的位置变化会形成轨迹，如果将 Geo-Indistinguishability 独立地应用到每个位置上，所产生的噪声量将是不可接受的。根据差分隐私中位数机制(median mechanism) [58]的基本思想，充分利用查询之间的关联关系，可以有效提高隐私保护预算的利用效率，因此，文中提出了一种针对位置保护的可预测差分隐私机制。该机制由预测函数、加噪机制和测试机制构成。预测函数根据先前提交的位置来预测当前须提交的新位置 $\tilde{z}$，然后由测试机制来测试 $\tilde{z}$ 与用户当前位置的距离是否在某个阈值之内，如果是，则直接提交 $\tilde{z}$，否则才调用加噪机制来产生新的位置。由于仅在调用加噪机制时才消耗隐私保护预算，所以可极大地提高预算利用率，降低噪声。另外，文献[59]针对降噪问题提出了一种面向位置数据发布的差分隐私保护算法

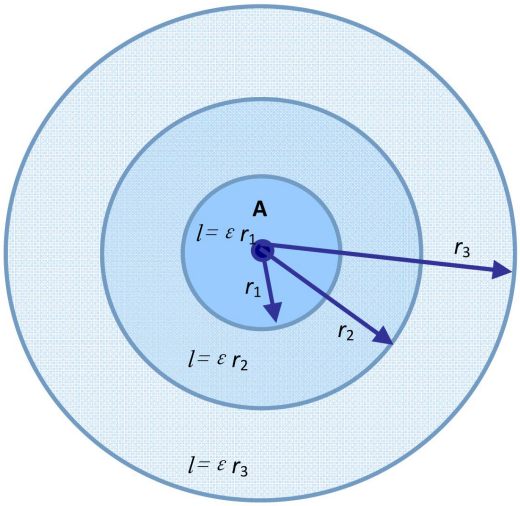

Figure 9. Privacy level varies with r
图 9. 隐私保护水平随 r 发生变化

(PriLocation)。位置数据发布的内容通常包括用户到过的位置集及其统计频次，如果直接应用差分隐私来保护发布的内容，将会因为位置频次的稀疏性导致噪声量过大。PriLocation 算法由位置聚类、权重干扰、位置选择等三个操作构成，首先根据距离将所有位置划分到 k 个簇中，每个位置则泛化为其所在的簇；然后将每个用户的位置统计频次转化为簇的频次统计，并用 Laplace 噪声进行干扰；最后利用指数机制从涉及的簇中选择位置作为用户到过的位置。由于簇的数量要远小于位置的数量，使得加入噪声的次数急剧减少，从而降低了噪声量。

差分隐私的主要优点在于它对攻击者所掌握的背景知识完全免疫，能够为用户提供强健的隐私保护。但从其在位置隐私保护中的应用效果来看，在有些方面还有待继续深入的研究，包括：(1)在处理高敏感度查询时，添加的噪声过大，会极大地降低数据的可用性；(2)给定的隐私预保护算会限制数据查询次数；(3)计算复杂度普遍较高。

3.5. 小结

由于无线通信技术和 LBS 服务模式的不断创新，位置隐私保护技术目前以及未来的一段时期仍将处于研究的高峰期。本节对现有的实现技术进行了分类梳理，将这些技术分为了四类，包括空间模糊化、虚拟对象、PIR 和差分隐私，并分别对其中的代表性成果进行了介绍和分析。每一类的技术都有其各自的优缺点和适用范围，它们在隐私保护水平、运行开销以及主要优缺点等方面的比较如表 1 所示。

总的来看，空间模糊化和虚拟对象技术相对成熟，能够较好地达到数据安全性和可用性的平衡，在目前来说，实用性相对较好；PIR 技术由于基于密码学基础，能够提供高水平的隐私保护，但计算代价高是其主要劣势，因此主要更适合于安全级别要求较高的场合；差分隐私能够提供可控的和可证明的隐私保护，但噪声大进而影响到数据可用性是有待继续研究的问题。

4. 未来的研究方向

位置隐私保护是一个相对年轻的研究领域，从目前的研究现状来看，在理论基础和实现技术等许多方面尚有待深入研究。同时，随着移动通信业务的不断推陈出新，位置隐私保护也必将面临更多的挑战，其未来的研究方向主要包括以下几个方面：

(1) 隐私保护参数的设置与优化

位置隐私保护技术在理论上都是基于一些隐私保护模型，例如 k-anonymity [19]、l-diversity [60]、t-closeness [27]、m-invariance [61]、p-confidentiality [62]、ε-DP [43]等，其隐私保护水平都是由相应的隐私保护参数来调节的。如何通过对这些参数的设置来达到隐私保护水平和服务水平的最佳平衡，即如何寻求隐私保护参数的最优解，是一个需要继续研究的问题，它可能涉及到对用户的调查、对行为和心理的评估，以及对现实环境的分析等。

Table 1. Comparison between location privacy preserving techniques
表 1. 位置隐私保护技术比较

种类	隐私保护水平	复杂度	性能	优点	缺点
空间模糊化	中	中	高	能通过调整隐私保护参数来达到较好的数据安全性和可用性的平衡	对用户密度过于敏感.
虚拟对象	中	中	高	能通过调整隐私保护参数来达到较好的数据安全性和可用性的平衡	不能抵抗背景知识攻击
PIR	高	高	中	适用于具有较强的隐私水平和一般实用性的特定应用程序	运行开销高，所需存储空间较大
差分隐私	高	高	中	提供可证隐私保护，能对抗背景知识攻击	大量的噪声导致可用性降低

(2) 个性化的位置隐私保护方案

在现实当中，对隐私保护的需求往往因用户或地域的不同而有很大的区别。但目前的位置隐私保护方案大多并没有考虑这些多样化的需求，隐私保护系统往往工作在某种统一的设置下。虽然有些研究已经意识到这个问题并提出了相应的解决方法[63] [64]，但这些方法大多工作在特定的环境下，还不具备一般通用性。设计细粒度的、支持不同层次的隐私水平的个性化隐私保护方案是未来的一个研究方向[65]。

(3) 社交网络中的位置隐私保护

社交网络的风靡对隐私保护提出了新的挑战[66]。在移动互联网中，位置数据与图片、文字、音频数据结合在一起，一般的结构化数据转变为非结构化数据，同时，采用实名认证的移动社交网络将个人身份信息与位置信息进行了绑定，社交网络中与用户之间的互动则导致隐私暴露的范围扩大。传统的隐私保护方法并不能适应这些新的变化，研究社交网络的位置隐私保护方法是未来的一个重要的研究方向。

5. 结束语

随着 LBS 的广泛应用，位置隐私保护问题受到了学术界、政府部门、消费者和产业界的多方关注。本文对 LBS 的一般体系结构和存在的位置隐私威胁进行阐述和分析，介绍了目前主要的位置隐私保护技术，并对各自的适用范围及优缺点进行了详细的分析和对比。最后，结合位置隐私保护的研究现状，指出了该领域在未来的研究方向。

基金项目

国家自然科学基金项目(61304067)；湖北省自然科学基金项目(2014CFB354)；中央高校基本科研业务费专项资金(31541511301)。

参考文献 (References)

[1] 陈永东. 2014 盘点: 全球移动应用使用增长 76% [EB/OL]. http://column.iresearch.cn/b/201501/693695.shtml, 2015-01-06.

[2] 新浪科技. Google Play 开发者和应用数量首次超 App Store [EB/OL]. http://tech.sina.com.cn/i/2015-01-15/doc-iawzunex8985555.shtml, 2015-01-15.

[3] 搜狐 IT. 百度地图日接受请求 35 亿次用户量突破 2 亿 [EB/OL]. http://it.sohu.com/20130822/n384819513.shtml, 2013-08-22.

[4] Beresford, A.R. and Stajano, F. (2003) Location Privacy in Pervasive Computing. *IEEE Pervasive Computing*, **2**, 46-55. http://dx.doi.org/10.1109/MPRV.2003.1186725

[5] Enck, W., Gilbert, P., Han, S., Tendulkar, V., Chun, B.-G., Cox, L.P., *et al.* (2014) TaintDroid: An Information-Flow Tracking System for Realtime Privacy Monitoring on Smartphones. *ACM Transactions on Computer Systems* (*TOCS*), **32**, 5.

[6] Mayer-Sch Nberger, V. and Cukier, K. (2013) Big Data: A Revolution That Will Transform How We Live, Work, and Think. Houghton Mifflin Harcourt, Boston.

[7] Hoh, B., Gruteser, M., Xiong, H. and Alrabady, A. (2006) Enhancing Security and Privacy in Traffic-Monitoring Systems. *IEEE Pervasive Computing*, **5**, 38-46.

[8] Matsuo, Y., Okazaki, N., Izumi, K., Nakamura, Y., Nishimura, T., Hasida, K., *et al.* (2007) Inferring Long-Term User Properties Based on Users' Location History. IJCAI, 2159-2165.

[9] Wicker, S.B. (2012) The Loss of Location Privacy in the Cellular Age. *Communications of the ACM*, **55**, 60-68. http://dx.doi.org/10.1145/2240236.2240255

[10] Shin, K.G., Ju, X., Chen, Z. and Hu, X. (2012) Privacy Protection for Users of Location-Based Services. *IEEE Wireless Communications*, **19**, 30-39. http://dx.doi.org/10.1109/MWC.2012.6155874

[11] Zheng, K., Shang, S., Yuan, N.J. and Yang, Y. (2013) Towards Efficient Search for Activity Trajectories. 2013 *IEEE* 29*th International Conference on Data Engineering* (*ICDE*), Brisbane, 8-12 April 2013, 230-241.

[12] 潘晓，肖珍，孟小峰. 位置隐私研究综述[J]. 计算机科学与探索, 2007(3): 268-281.

[13] 王璐, 孟小峰. 位置大数据隐私保护研究综述[J]. 软件学报, 2014(4): 693-712.

[14] Beresford, A.R., Rice, A., Skehin, N. and Sohan, R. (2011) MockDroid: Trading Privacy for Application Functionality on Smartphones. *Proceedings of the* 12*th Workshop on Mobile Computing Systems and Applications*. *ACM*, 49-54. http://dx.doi.org/10.1145/2184489.2184500

[15] Bamba, B., Liu, L., Pesti, P. and Wang, T. (2008) Supporting Anonymous Location Queries in Mobile Environments with Privacygrid. *Proceedings of the* 17*th International Conference on World Wide Web*. *ACM*, 237-246.

[16] Duckham, M. and Kulik, L. (2005) A Formal Model of Obfuscation and Negotiation for Location Privacy. In: Gellersen, H.-W., Want, R. and Schmidt, A., Eds., *Pervasive Computing*, Springer, Berlin, 152-170.

[17] Xue, M., Kalnis, P. and Pung, H.K. (2009) Location Diversity: Enhanced Privacy Protection in Location Based Services. In: Choudhury, T., Quigley, A., Strang, T. and Suginuma, K., Eds., *Location and Context Awareness*, Springer, Berlin, 70-87.

[18] Sweeney, L. (2002) k-Anonymity: A Model for Protecting Privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, **10**, 557-570. http://dx.doi.org/10.1142/S0218488502001648

[19] Sweeney, L. (2002) Achieving k-Anonymity Privacy Protection Using Generalization and Suppression. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, **10**, 571-588. http://dx.doi.org/10.1142/S021848850200165X

[20] Gruteser, M. and Grunwald, D. (2003) Anonymous Usage of Location-Based Services through Spatial and Temporal Cloaking. *Proceedings of the* 1*st International Conference on Mobile Systems, Applications and Services*, San Francisco, 5-8 May 2003, 31-42.

[21] Mokbel, M.F., Chow, C.-Y. and Aref, W.G. (2006) The New Casper: Query Processing for Location Services without Compromising Privacy. *Proceedings of the* 32*nd International Conference on Very Large Data Bases*, Seoul, 12-15 September 2006, 763-774.

[22] Pan, X., Xu, J. and Meng, X. (2012) Protecting Location Privacy against Location-Dependent Attacks in Mobile Services. *IEEE Transactions on Knowledge and Data Engineering*, **24**, 1506-1519. http://dx.doi.org/10.1109/TKDE.2011.105

[23] Niu, B., Zhang, Z., Li, X. and Li, H. (2014) Privacy-Area Aware Dummy Generation Algorithms for Location-Based Services. 2014 *IEEE International Conference on Communications* (*ICC*), Sydney, 10-14 June 2014, 957-962.

[24] Kido, H., Yanagisawa, Y. and Satoh, T. (2005) An Anonymous Communication Technique Using Dummies for Location-Based Services. *Proceedings of the International Conference on Pervasive Services*, 11-14 July 2005, 88-97.

[25] Kido, H., Yanagisawa, Y. and Satoh, T. (2005) Protection of Location Privacy Using Dummies for Location-Based Services. 21*st International Conference on Data Engineering Workshops*, 5-8 April 2005, 1248.

[26] Lu, H., Jensen, C.S. and Yiu, M.L. (2008) PAD: Privacy-Area Aware, Dummy-Based Location Privacy in Mobile Services. *Proceedings of the* 7*th ACM International Workshop on Data Engineering for Wireless and Mobile Access*, Vancouver, 13 June 2008, 16-23. http://dx.doi.org/10.1145/1626536.1626540

[27] Li, N., Li, T. and Venkatasubramanian, S. (2007) t-Closeness: Privacy beyond k-Anonymity and l-Diversity. *IEEE* 23*rd International Conference on Data Engineering*, Istanbul, 15-20 April 2007, 106-115.

[28] Ilarri, S., Mena, E. and Illarramendi, A. (2010) Location-Dependent Query Processing: Where We Are and Where We Are Heading. *ACM Computing Surveys*, **42**, Article No. 12. http://dx.doi.org/10.1145/1670679.1670682

[29] Niu, B., Li, Q., Zhu, X., Cao, G. and Li, H. (2014) Achieving k-Anonymity in Privacy-Aware Location-Based Services. *IEEE INFOCOM* 2014-*IEEE Conference on Computer Communications*, Toronto, 27 April-2 May 2014, 754-762.

[30] Niu, B., Zhu, X., Lei, X., Zhang, W. and Li, H. (2013) EPS: Encounter-Based Privacy-Preserving Scheme for Location-Based Services. 2013 *IEEE Global Communications Conference* (*GLOBECOM*), Atlanta, 9-13 December 2013, 2139-2144.

[31] Zhu, X., Chi, H., Niu, B., Zhang, W., Li, Z. and Li, H. (2013) MobiCache: When k-Anonymity Meets Cache. 2013 *IEEE Global Communications Conference* (*GLOBECOM*), Atlanta, 9-13 December 2013, 820-825.

[32] Khoshgozaran, A., Shahabi, C. and Shirani-Mehr, H. (2011) Location Privacy: Going beyond K-Anonymity, Cloaking and Anonymizers. *Knowledge and Information Systems*, **26**, 435-465. http://dx.doi.org/10.1007/s10115-010-0286-z

[33] Papadopoulos, S., Bakiras, S. and Papadias, D. (2010) Nearest Neighbor Search with Strong Location Privacy. *Proceedings of the VLDB Endowment*, **3**, 619-629. http://dx.doi.org/10.14778/1920841.1920920

[34] Ghinita, G., Kalnis, P., Khoshgozaran, A., Shahabi, C. and Tan, K.-L. (2008) Private Queries in Location Based Services: Anonymizers Are Not Necessary. *Proceedings of the* 2008 *ACM SIGMOD International Conference on Management of Data*, Vancouver, 9-12 June 2008, 121-132. http://dx.doi.org/10.1145/1376616.1376631

[35] Paulet, R., Kaosar, M.G., Yi, X. and Bertino, E. (2014) Privacy-Preserving and Content-Protecting Location Based

Queries. *IEEE Transactions on Knowledge and Data Engineering*, **26**, 1200-1210. http://dx.doi.org/10.1109/TKDE.2013.87

[36] Kushilevitz, E. and Ostrovsky, R. (1997) Replication Is Not Needed: Single Database, Computationally-Private Information Retrieval. *Proceedings of the* 38*th Annual Symposium on Foundations of Computer Science*, Miami Beach, 20-22 October 1997, 364-373.

[37] Asonov, D. and Freytag, J.-C. (2003) Almost Optimal Private Information Retrieval. In: R. Dingledine and P. Syverson, Eds., *Privacy Enhancing Technologies*, Springer, Berlin, 209-223.

[38] Smith, S.W., Safford, D. and Ord, D.S. (2000) Practical Private Information Retrieval with Secure Coprocessors.

[39] Iliev, A. and Smith, S. (2005) More Efficient Secure Function Evaluation Using Tiny Trusted Third Parties. Department of Computer Science, Dartmouth University, Dartmouth Computer Science Technical Report TR2005-551.

[40] Mouratidis, K. and Yiu, M.L. (2012) Shortest Path Computation with No Information Leakage. *Proceedings of the VLDB Endowment*, **5**, 692-703. http://dx.doi.org/10.14778/2212351.2212352

[41] Dwork, C. (2008) Differential Privacy: A Survey of Results. In: Agrawal, M., Du, D.Z., Duan, Z.H. and Li, A.S., Eds., *Theory and Applications of Models of Computation*, Springer, Berlin, 1-19.

[42] 熊平, 朱天清, 王晓峰. 差分隐私保护及其应用[J]. 计算机学报, 2014, 37(1) : 101-122.

[43] Dwork, C. (2011) A Firm Foundation for Private Data Analysis. *Communications of the ACM*, **54**, 86-95. http://dx.doi.org/10.1145/1866739.1866758

[44] Dwork, C., Mcsherry, F., Nissim, K. and Smith, A. (2006) Calibrating Noise to Sensitivity in Private Data Analysis. In: Halevi, S. and Rabin, T., Eds., *Theory of Cryptography*, Springer, Berlin, 265-284.

[45] Mcsherry, F. and Talwar, K. (2007) Mechanism Design via Differential Privacy. 48*th Annual IEEE Symposium on Foundations of Computer Science*, Providence, 21-23 October 2007, 94-103.

[46] Hay, M., Rastogi, V., Miklau, G. and Suciu, D. (2010) Boosting the Accuracy of Differentially Private Histograms through Consistency. *Proceedings of the VLDB Endowment*, **3**, 1021-1032. http://dx.doi.org/10.14778/1920841.1920970

[47] Chen, R., Mohammed, N., Fung, B.C., Desai, B.C. and Xiong, L. (2011) Publishing Set-Valued Data via Differential Privacy. *Proceedings of the VLDB Endowment*, **4**, 1087-1098.

[48] Cormode, G., Procopiuc, C., Srivastava, D. and Tran, T.T. (2012) Differentially Private Summaries for Sparse Data. *Proceedings of the* 15*th International Conference on Database Theory*, Berlin, 26-30 March 2012, 299-311.

[49] Li, C. and Miklau, G. (2012) An Adaptive Mechanism for Accurate Query Answering under Differential Privacy. *Proceedings of the VLDB Endowment*, **5**, 514-525. http://dx.doi.org/10.14778/2168651.2168653

[50] Li, N., Qardaji, W., Su, D. and Cao, J. (2012) PrivBasis: Frequent Itemset Mining with Differential Privacy. *Proceedings of the VLDB Endowment*, **5**, 1340-1351. http://dx.doi.org/10.14778/2350229.2350251

[51] Zeng, C., Naughton, J.F. and Cai, J.-Y. (2012) On Differentially Private Frequent Itemset Mining. *Proceedings of the VLDB Endowment*, **6**, 25-36. http://dx.doi.org/10.14778/2428536.2428539

[52] Friedman, A. and Schuster, A. (2010) Data Mining with Differential Privacy. *Proceedings of the* 16*th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Washington DC, 25-28 July 2010, 493-502.

[53] Mohammed, N., Chen, R., Fung, B. and Yu, P.S. (2011) Differentially Private Data Release for Data Mining. *Proceedings of the* 17*th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Diego, 21-24 August 2011, 493-501.

[54] Smith, A. (2011) Privacy-Preserving Statistical Estimation with Optimal Convergence Rates. *Proceedings of the* 43*rd Annual ACM Symposium on Theory of Computing*, San Jose, 6-8 June 2011, 813-822.

[55] Dewri, R. (2013) Local Differential Perturbations: Location Privacy under Approximate Knowledge Attackers. *IEEE Transactions on Mobile Computing*, **12**, 2360-2372. http://dx.doi.org/10.1109/TMC.2012.208

[56] Andr, S.M.E., Bordenabe, N.E., Chatzikokolakis, K. and Palamidessi, C. (2013) Geo-indistinguishability: Differential Privacy for Location-Based Systems. *Proceedings of the* 2013 *ACM SIGSAC Conference on Computer & Communications Security*, Berlin, 4-8 November 2013, 901-914.

[57] Chatzikokolakis, K., Palamidessi, C. and Stronati, M. (2014) A Predictive Differentially-Private Mechanism for Mobility Traces. In: De Cristofaro, E. and Murdoch, S.J., Eds., *Privacy Enhancing Technologies*, Springer International Publishing, 21-41.

[58] Roth, A. and Roughgarden, T. (2010) Interactive Privacy via the Median Mechanism. *Proceedings of the* 42*nd ACM Symposium on Theory of Computing*, Cambridge, 6-8 June 2010, 765-774.

[59] Xiong, P., Zhu, T., Pan, L., Niu, W. and Li, G. (2014) Privacy Preserving in Location Data Release: A Differential Privacy Approach. In: Pham, D.-N. and Park, S.-B., Eds., *PRICAI* 2014: *Trends in Artificial Intelligence*, Springer, Berlin, 183-195.

[60] Machanavajjhala, A., Kifer, D., Gehrke, J. and Venkitasubramaniam, M. (2007) *l*-Diversity: Privacy beyond *k*-Anonymity. *ACM Transactions on Knowledge Discovery from Data* (*TKDD*), **1**, Article No. 3.

[61] Xiao, X. and Tao, Y. (2007) M-Invariance: Towards Privacy Preserving Re-Publication of Dynamic Datasets. *Proceedings of the* 2007 *ACM SIGMOD International Conference on Management of Data*, Beijing, 11-14 June 2007, 689-700. http://dx.doi.org/10.1145/1247480.1247556

[62] Cicek, A.E., Nergiz, M.E. and Saygin, Y. (2014) Ensuring Location Diversity in Privacy-Preserving Spatio-Temporal Data Publishing. *The VLDB Journal*, **23**, 609-625. http://dx.doi.org/10.1007/s00778-013-0342-x

[63] Page, X. and Kobsa, A. (2011) Personality-Based Privacy Management for Location-Sharing in Diverse Subpopulations. *Proceedings of the* 2011 *iConference*, Seattle, 8-11 February 2011, 736-738.

[64] Li, M., Qin, Z. and Wang, C. (2014) Sensitive Semantics-Aware Personality Cloaking on Road-Network Environment. *International Journal of Security & Its Applications*, **8**, 133-146.

[65] Li, X.-Y. and Jung, T. (2013) Search Me If You Can: Privacy-Preserving Location Query Service. 2013 *Proceedings IEEE INFOCOM*, Turin, 14-19 April 2013, 2760-2768.

[66] Liang, X., Zhang, K., Shen, X. and Lin, X. (2014) Security and Privacy in Mobile Social Networks: Challenges and Solutions. *IEEE Wireless Communications*, **21**, 33-41. http://dx.doi.org/10.1109/MWC.2014.6757895