文章编号:1000-5641(2015)05-0028-18

LBS 的隐私保护:模型与进展

赵大鹏[1], 梁 磊[1], 田秀霞[1,2], 王晓玲[1]
(1. 华东师范大学 数据科学与工程研究院 上海市高可信计算重点实验室,上海 200062;
2. 上海电力学院 计算机科学与技术学院,上海 201300)

摘要: 近些年来,随着配备定位功能的移动终端数量迅速增加,基于位置服务(LBS)的应用呈现爆炸式的增长,例如查找最邻近的加油站、一公里范围内的所有餐厅等. 在用户享受着这些 LBS 服务为工作、生活带来方便的同时,许多隐私安全问题也逐渐引起了人们的关注. 全面了解基于位置服务中现有的隐私保护工作,有利于研究者把握该领域的研究现状、未来发展方向以及存在的挑战. 本文对 LBS 隐私保护领域中近些年的发展进行了研究总结,重点介绍了 LBS 隐私保护领域现有的攻击模型、隐私保护模型、度量模型以及数据集,并对现有攻击模型与隐私保护模型进行分类总结,根据其特点进行对比分析,最后探讨了 LBS 隐私保护目前存在的问题以及未来的发展方向.

关键词: 基于位置服务; 隐私保护; 攻击模型; 度量模型; 数据集
中图分类号: TP311.5 **文献标识码**: A **DOI**:10.3969/j.issn.1000-5641.2015.05.003

Privacy protection in location-based services: Model and development

ZHAO Da-peng[1], LIANG Lei[1], TIAN Xiu-xia[1,2], WANG Xiao-ling[1]
(1. *Shanghai Key Laboratory of Trustworthy Computing, Institute for Data Science and Engineering, East China Normal University, Shanghai* 200062, *China*;
2. *School of Computer Science and Technology, Shanghai University of Electric Power, Shanghai* 201300, *China*)

Abstract: In recent years, with the rapid increase in the number of GPS-enabled mobile devices, location-based services (LBS) applications grow explosively, such as finding the nearest gas station or restaurants within one kilometer and so on. Users benefit from convenience of LBS. However, many privacy issues draw people's attention gradually. Acomprehensive understanding of existing privacy protection work in the location-based services is important for researchers to grasp the present research status, the future development directionsand the challenges. We give a deep survey of the recent improvement in LBS, which mainly focus on existing attacking models, privacy protection model, measure model and datasets. What's more, we classifies the existing attacking model and privacy protection model and made comparisons based on different features. Finally unsolved problems and future development are also discussed.

Key words: location-based service; privacy protection; attacking model; measure model; dataset

收稿日期:2015-06
基金项目:国家自然科学基金(61170085,61472141);上海市重点学科建设项目(B412);上海市可信物联网软件协同创新中心(ZF1213)资助
第一作者:赵大鹏,男,硕士研究生,研究方向为隐私保护. E-mail: 51141500066@ecnu.cn.

0　引　　言

随着全球定位系统 GPS、蜂窝网、Wi-Fi 和无线射频识别等定位技术快速发展，基于位置的服务(LBS)[1-2]已经变得无处不在，例如商业上的位置查询(查找距离某用户一定范围内的餐厅)、电子营销(发电子优惠券给附近的顾客)、交通状况监控(根据一段时间内车辆的位置和速度来推测交通拥堵状况)、路线查找应用(查找两地之间的最短路径)等. 虽然 LBS 给用户提供了许多有价值的服务，但是许多隐私安全问题[3]也逐渐引起了人们的关注. 从位置信息中，不仅可以知道你在哪里，还可以进一步推断出其他敏感信息，如家庭住址、健康信息、宗教信仰等. 如何让人们享受服务的同时保护用户的隐私，是 LBS 领域面临的新挑战. 2014 年 7 月，苹果手机用户隐私泄露事件[4]引起了广大用户的关注，在用户完全不知情的情况下，苹果手机记录用户使用软件的时间地点，从而获得用户的轨迹信息. 2014 年 8 月，安全专家 John McAfee[5]在拉斯维加斯参加 Defcon 黑客大会时指出 Google 一直在对用户的邮件及搜索记录等个人内容进行扫描，并且 Google 等多家公司已经承认这一事实. 如果用户的隐私得不到妥善保护，将会使用户的权益和安全面临严重威胁，用户在使用 LBS 时，也会产生一些顾虑. 2010 年 7 月的两份调查结果[6]显示，超过 55%的 LBS 应用使用者担心自己的位置隐私被恶意攻击者窃取，美国 50%的社交网站用户同样担忧他们的隐私泄露问题. 因此，能否很好地解决隐私保护问题可以看作是公众可否放心地使用 LBS 服务的前提.

随着学术界与工业界对 LBS 的隐私保护重视程度不断增加，LBS 隐私保护技术在不断地快速发展. LBS 中的隐私主要包括用户的位置隐私[7]与查询隐私[8]. 位置隐私是指用户不愿意让他人知道的过去或当前的位置信息，根据位置信息，攻击者不仅可以知道用户在哪，还可以结合其他背景知识推断出与该用户相关的许多敏感信息，例如攻击者得知某个用户在某一时刻到达过位置 P，又知道 P 位于教堂之中，则可以推测该用户具有宗教信仰. 保护用户的位置隐私，使得攻击者无法获取用户的位置信息，也无法推断出与用户相关的其他敏感信息. 常用的位置隐私保护模型有：①位置模糊[9-12]，将用户的准确位置进行模糊化(例如，用匿名区域[9]代替用户的准确位置作为服务请求点)或者对用户准确位置进行干扰(例如，将用户周边具有代表性的一些位置[12]作为用户的服务请求位置)；②位置加密[13](例如，使用隐私信息检索技术来保护用户的位置隐私).

查询隐私是指用户不愿意让他人知道自己向 LBS 服务供应商提出过哪些 LBS 请求，包括请求内容与请求记录. 与位置隐私一样，查询隐私也是一种重要的个人隐私，攻击者通过窃取用户的查询记录，可以获知用户的兴趣爱好、健康状况等敏感信息. 例如，用户在某一敏感位置(如自己家中)提出了查询最近的艾滋病诊所的请求，攻击者窃取到这一请求内容以后，结合房屋归属信息，就可以推测出房屋主人可能患有艾滋病. 保护查询隐私的思路主要是将用户的身份与查询请求者分离，使攻击者无法轻易识别是哪个用户提出了查询请求. 隐私保护中使用最多的 k 匿名[9,14]思想就是基于这种思路，即要求匿名集中包含查询用户与至少 k-1 个其他用户，使得攻击者无法准确识别某个查询请求是由匿名集中哪个用户提出的，许多有效的 LBS 隐私保护方法都是基于这种 k 匿名思想.

本文对 LBS 隐私保护领域中近些年的发展进行综述，总结 LBS 隐私保护中的攻击模型、隐私保护模型等，主要贡献点包括以下三点：

(1) 系统地总结了现有工作中的攻击模型,从攻击者所知信息量角度对攻击模型进行分类比较.

(2) 总结 LBS 中现有的隐私保护模型,从四个方面对现有隐私保护模型进行概括,并综合各类模型特点进行对比分析.

(3) 总结现有工作中所使用的度量模型和试验中所使用的数据集,全面了解现有工作.

本文第 1 节介绍攻击模型;第 2 节详述现有的隐私保护模型;第 3 节介绍 LBS 隐私保护技术的度量模型;第 4 节介绍现有工作中常用的数据集;第 5 节对工作进行总结与展望.

1 攻击模型

攻击模型研究的是恶意攻击者根据哪些信息,以怎样的方式窃取用户的隐私信息,理解攻击模型能够帮助我们有针对性地提出相应的隐私保护方法. 攻击者所能获得的信息可以分为两类:时空信息和背景信息. 时空信息(t,P)是二维信息,其中,t 是时间信息,P 是位置信息,时空信息进一步分为单次时空信息和多次时空信息(轨迹信息). 背景信息(也称为背景知识),是指除了时空信息之外所有可以帮助攻击者窃取用户隐私的信息,如房屋归属、交通数据统计、地图等. 通常情况下,攻击者除了具有用户的时空信息之外,可能还有一些其他的背景知识,攻击者可以将这些背景知识与查询信息结合起来推断出用户的敏感信息. 攻击者知道查询用户信息的种类不同,推断出用户隐私信息的难易程度也不相同,因此本节根据攻击者所知道的信息种类,对 LBS 中的查询隐私与位置隐私攻击模型进行分类.

1.1 查询隐私攻击模型

查询隐私攻击模型是旨在窃取用户的查询隐私的攻击模型,具体方式为关联用户与他提出的查询请求或在用户未授权的情况下窃取用户提交的请求内容. 根据攻击者获得用户提交查询的次数,可以将查询隐私中的攻击模型分为基于单次查询的查询隐私攻击模型与基于多次查询的查询隐私攻击模型.

(1) 基于单次查询的查询隐私攻击模型

基于单次查询的查询隐私攻击模型是指攻击者仅依赖用户某一次的查询请求来窃取用户隐私信息,主要包括位置分布攻击、查询采样攻击、同质攻击等.

位置分布攻击(Location Distribution Attack)[15]是指攻击者利用用户在匿名区域中的位置分布特点,推断出服务请求者的身份. 例如,离群点攻击[16], 攻击者根据用户位置分布不均匀的特点,推断出密度稀疏区域用户(离群点)的查询隐私信息. 如图 1 所示,离群点用户 A 提交了一个匿名度 k 为 4 的查询请求,匿名区域为图中的大矩形区域,包含 A,B,C,D 4 个用户. 若是 B,C,D 用户提出了该查询请求,匿名区域应当为图中包含 B,C,D,E 的小矩形区域,攻击者根据这一推测很容易判断出查询用户是 A.

查询采样攻击(Query Sampling Attack)[15]是指攻击者通过结合同一时刻的不同匿名集,排除两个匿名集中相同的查询请求,找到用户与真实查询之间的对应关系. 例如,假设同一时刻系统中有若干用户 A,B,C,D,E 分别提出查询请求 Q_A,Q_B,Q_C,Q_D,Q_E,匿名服务器为他们建立了如图 1 所示的两个匿名区域. 其中 B,C,D 用户同时位于两个匿名区域中,攻击者通过关联并排除两个匿名区域中相同的查询请求 Q_B,Q_C,Q_D,从而将用户 A 与 Q_A 关联,将用户 E 与 Q_E 关联.

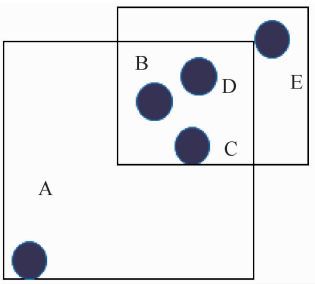

图 1　位置分布攻击

Fig. 1　Location distribution attack

查询同质攻击(Query Homogeneous Attack)[17]是指匿名集中包含的查询请求的语义类型很少甚至都相同，攻击者虽然不能关联用户与其提出的请求，但是可以推测出用户提出的请求内容. 例如，匿名集中所有用户都发布了与疾病相关的查询，则攻击者就可以根据请求内容推测出目标用户的健康信息.

敏感同质攻击(Sensitive Homogeneous Attack)[18]是指即使匿名集同时满足 k 匿名与 l 多样性，用户的隐私仍有可能泄露. 例如，匿名区域中的请求用户个数大于 k，并且类型数大于 l，但是其中的每一个请求都包含酒吧、医院等敏感语义，攻击者就可以推测出用户提出的请求包含敏感内容.

位置链接攻击(Location Linking Attack)[19]是指请求者用匿名提交查询请求，攻击者结合背景信息与某个特殊的位置来确定查询者的身份. 例如，用户在自己家中提交了一条关于艾滋病诊所的查询，如果攻击者知道房屋的拥有者是谁，就可以推测该房屋的拥有者可能患有艾滋病.

(2) 基于多次查询的查询隐私攻击模型

多次查询攻击是指在一段时间内，用户连续提交多个查询请求或者提交多次位置更新信息的情况下，攻击者根据这些信息对用户的查询隐私进行攻击. 基于多次查询的查询隐私攻击模型主要包括身份匹配攻击和查询追踪攻击等.

身份匹配攻击(Identity Matching Attack)[20]是指当用户连续使用不同的假名提交查询请求时，攻击者根据某些相互关联的属性将这些假名联系在一起，从而发现用户的移动轨迹. 例如，一个用户不断地用不同的假名提交对艾滋病医院的查询，攻击者可以根据这个查询内容关联这些假名.

查询追踪攻击(Query Tracking Attack)[21]是指当用户连续提交内容相同的查询请求，攻击者在用户多次提交的匿名集中取交集就可以推测出查询用户身份，并得到用户的移动轨迹.

计时攻击(Timing Attack)[14]是针对混合区域的攻击，混合区域[20]是为了防止攻击者连续跟踪用户轨迹提出的一种隐私保护方法. 该方法通过限制在混合区域中的用户发布 LBS 请求，使攻击者无法轻易将用户进入混合区域时的假名与离开时的假名关联起来. 但是，在实际情况中，先进入混合区域的用户有较大的概率会先离开区域，攻击者可以通过计算用户进入匿名区域的时间来将新旧假名关联起来，从而使得混合区的保护效果降低.

概率分布攻击(Probability Distribution Attack)[22]是指攻击者根据一些实际场景中用

户的分布情况和交通流量统计,窃取用户的查询隐私和位置隐私.例如,隐私保护系统在十字路口建立混合区域,但是根据交通数据统计,80%从 a 路口进入区域内的用户会从 d 路口出去,则对于一个从 a 进入隐藏区域的用户,攻击者就可以较为准确地将该用户进入混合区域与离开混合区域时使用的假名关联起来.

1.2 位置隐私攻击模型

位置隐私攻击模型是指旨在窃取用户的位置隐私的攻击模型,具体方式是在用户未授权的情况下获得用户在某一时刻的位置或者在一段时间内的连续轨迹,进而获取用户的其他敏感信息.位置隐私攻击模型主要包括位置同质攻击、位置依赖攻击等.

位置同质攻击(Location Homogeneity Attack)[23]是针对 k 匿名技术的攻击模型,如果匿名区域中的 k 个用户距离较近,都位于一个较小的区域(例如一个 5 m^2 的范围)内,则攻击者就可以知道所有用户的位置信息.

位置依赖攻击(Location Dependent Attack)[24]是指攻击者根据用户上一次提交的匿名区域所在位置以及他的最大移动速度,计算出最大移动边界,即用户最远可以到达的范围.若后一时刻用户的匿名区域仅有一部分位于最大移动边界中,则攻击者可以推断出查询用户的位置.如图 2 所示,t_1 时刻与 t_2 时刻用户提交的匿名区域分别为 R_1 和 R_2,攻击者根据用户的最大移动速度,计算出 t_2 时刻用户最远可达的边界(虚线框),则可以确定用户位于虚线区域的与 R_2 匿名区域的交叉区域内.

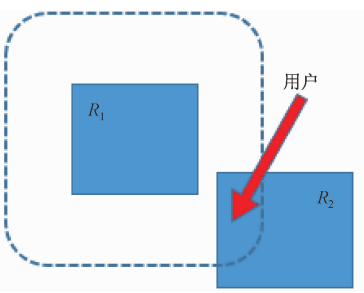

图 2 位置依赖攻击

Fig. 2 Location dependent attack

预测攻击(Prediction Attack)[25]是指假设攻击者拥有某区域内大量用户的历史移动轨迹,根据某个用户在一段时间内的签到轨迹,可以将该签到轨迹与历史移动轨迹进行匹配,进而预测出该用户的目的地,从而推测出用户的家庭住址等敏感信息.文献[25]中提出了一种基于子轨迹合成的目的地预测算法,大大加强了预测攻击的预测成功率.

LBS 社交发现中基于多个探针的攻击[19],Ding 等提出了一种新型攻击模型,利用微信等社交平台中搜索附近用户的功能,用户可以知道周围用户与自己的距离.文中设计了一种使用多个虚拟探针的方法,结合三角定位思想,可以将多个探针同时探测到的用户位置限定在一个较小的重叠区域内,从而窃取到用户的位置隐私.

1.3 攻击模型小结

对于不同攻击模型,攻击者所使用的信息不同,针对的场景也有所不同,如表 1 所示,攻击模型各有其特点.除了以上几种攻击模型以外,攻击者还可以将多种攻击模型结合起来进行攻击,这样更容易获得用户的隐私信息.例如对于一个连续用不同假名发布相同查询的用

户，攻击者可以先根据身份匹配攻击找到请求用户的轨迹，若轨迹中存在某个特殊位置点可以唯一确定用户的真实身份（如家庭住址），攻击者又可以使用位置链接攻击将用户的身份与该轨迹关联起来.

表 1　攻击模型

Tab. 1　Attacking model

类别	攻击模型	是否利用多次查询信息	是否利用背景信息
查询隐私攻击模型	位置分布攻击[15]	×	×
	查询采样攻击[15]	×	×
	查询同质攻击[17]	×	×
	敏感同质攻击[18]	×	×
	位置链接攻击[19]	○	×
	身份匹配攻击[20]	×	○
	查询追踪攻击[21]	×	○
位置隐私攻击模型	位置同质攻击[23]	×	×
	计时攻击[14]	×	×
	概率分布攻击[22]	○	×
	位置依赖攻击[24]	×	○
	推理攻击[9]	○	○
	基于多个探针的攻击[19]	○	○

2　LBS 隐私保护模型

近年来，针对不同的攻击模型，出现了许多保护用户的查询隐私或位置隐私的隐私保护模型，这些隐私保护模型主要分为四类：①基于匿名的隐私保护模型；②基于扰动的隐私保护模型；③基于密码学的隐私保护模型；④基于访问控制的隐私保护模型. 表 2 对比了四类技术的隐私保护强度、实现难度、LBS 服务质量、资源消耗（包括计算消耗、传输消耗、占用内存大小三个层面）的特点.

表 2　4 类隐私保护模型性能对比

Tab. 2　Performance comparisons of four privacy protection models

技术	隐私保护强度	实现难度	服务质量	资源消耗
匿名	中	中	中	中
干扰	中	中	中	中
密码学	强	大	中	大
访问控制	弱	小	高	小

2.1　基于匿名的隐私保护模型

匿名是用户为了防止攻击者将查询请求者与真实身份关联起来的一种基本的保护方法，从而使得用户真实身份与其查询分离，因此匿名技术主要保护的是用户的查询隐私. 当用户使用匿名技术请求 LBS 服务时，攻击者即使可以获得用户的位置，也无法直接识别该位置所对应的用户身份. 然而，仅仅使用匿名技术对于用户查询隐私的保护效果较弱，攻击者如果结合一些背景知识就能比较容易推断出用户的真实身份. 即使频繁更换假名，攻击者也可以使用身份匹配等攻击模型持续地跟踪用户，从而使匿名的作用失效，因此产生了更加有效的位置 k 匿名保护方法与混合区域保护方法.

2.1.1 位置 k 匿名

Gruteser 等[26]提出了位置 k 匿名的概念,最先把 k 匿名从关系数据库领域引入 LBS 隐私保护领域,之后大量的 LBS 隐私保护工作都基于位置 k 匿名展开. 位置 k 匿名要求当一个用户请求获得 LBS 服务时,用一个包含自己当前位置的匿名区域代替自己的准确位置上传给 LBS 服务器,并且匿名区域中必须包含该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不大于 $1/k$. 位置 k 匿名在保护了用户查询隐私的同时,用匿名区域代替用户的准确位置上传给 LBS 服务器,因此也保护了用户的位置隐私. 按系统结构模型来分,基于位置 k 匿名的隐私保护模型可以分为两类:集中式结构模型与分布式结构模型.

(1) 集中式结构模型

集中式结构模型[9,26-28]的系统框架如图 3 所示,在用户端与 LBS 服务提供商之间引入了第三方可信匿名服务器,由匿名服务器来完成匿名过程. 当用户提交 LBS 请求时,先将准确位置与查询请求发送给匿名服务器,服务器将用户的位置匿名处理后发送给 LBS 服务器,LBS 服务器进行基于匿名区域的查询,并将结果集返回给匿名服务器,再由匿名服务器根据用户准确的位置信息从候选结果集中筛选出正确的结果返回给终端用户. 目前大多数基于位置 k 匿名的隐私保护方法都采用集中式结构模型.

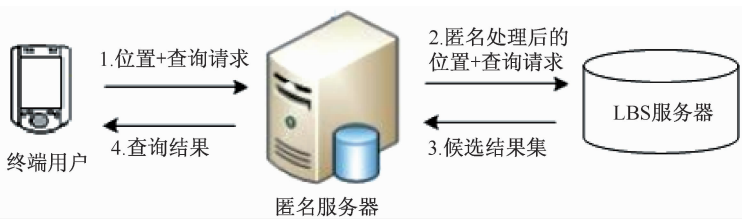

图 3 集中式结构模型

Fig. 3 Centralized architecture model

Gruteser 等在[26]中最先提出的 IntervalCloak(IC)就是基于集中式结构模型. IntervalCloak 设定一个全局匿名参数 k,通过递归地划分整个系统空间,直到查询用户所在的最小矩形中用户数目小于 k,将该区域的上一层矩形区域作为匿名区域传给 LBS 服务器. IntervalCloak 对于所有的用户均采用相同的匿名等级,无法满足不同用户在不同时间段的隐私要求. Mokbel 等[27]提出了可以满足用户个性化隐私保护要求的 Casper 模型,Casper 允许每个用户在各个时间段设定不同的匿名等级 k 与最小匿名区域大小 $A_{\min}$,并建立金字塔层次结构来寻找满足用户隐私要求的匿名区域. IntervalCloak 与 Casper 共同的缺陷是生成的匿名区域中包括的用户数往往大于系统或用户设定的 k 值,因此会导致不必要的计算开销. 由于系统需要不断地更新金字塔结构,因此维护开销较大. Gedik 等[9]提出的 CliqueCloak,将空间与时间相结合来构建匿名区域,并且用户可以通过设定最大匿名半径与匿名延迟容忍时间来保证服务质量. Kalnis 又提出了基于 Hilbert 曲线的 HilbertCloak(HC)模型,该模型首先利用 Hilbert 曲线将用户的二维位置坐标变换为一维的 Hilbert 序列,当一个用户请求服务时,计算出满足其 k 值的匿名区间,再将 Hilbert 序列中位于该匿名区间中的用户组成匿名区域上传给 LBS 服务器. HC 的时间复杂度仅为 $O(\log N + k)$,而且不易被攻击者识别查询者的身份. 为了防止敏感同质攻击,Xiao 等[52]提出了 p 敏感度概念,要求匿名区域中

的请求中，包括敏感内容的请求数目所占的百分比小于 p，建立了(k, p)敏感匿名模型来保证匿名区域满足 p 敏感度要求.

大多数基于位置 k 匿名的保护方法只适用于单次查询请求，而在第二节中提到的很多攻击模型都是针对用户在一段时间内提出多次查询的情况下进行攻击. 如果攻击者具有历史请求记录，利用查询追踪攻击可以识别出真实请求用户，进而窃取到用户的连续行为轨迹；利用位置依赖攻击等攻击模型可以缩小用户的位置范围，从而窃取到用户的位置隐私. 因此连续提交查询请求对于位置 k 匿名的有效性是一大挑战，Pan 等[24]针对这种位置依赖攻击提出了 ICliqueCloak 隐私保护模型. 为了保证连续更新请求中位置 k 匿名的有效性，ICliqueCloak 要求建立匿名区域时，前一个时刻建立的匿名区域的最大移动边界要覆盖当前匿名区域，并且当前匿名区域的最大移动边界也要覆盖前一个匿名区域. 为了减小每次上传的匿名区域面积，Xu 等[29]提出了一种使用历史足迹来构建 k 匿名区域的方法. 先将系统划分为网格，用户需要上传接下来的行驶轨迹给匿名服务器，匿名服务器从历史轨迹数据库中找出与该行驶轨迹最相近(公共网格最多)的 k-1 条历史轨迹，将这 k 条轨迹对应的点建立一个个连续的匿名区域. 这种做法构建的 k 匿名区域中，实际上包括的是真实用户与其他用户过去留下的足迹，如果攻击者可以实时观察到匿名区域中的真实状况时，就可以识别出真实请求者. 对此，Xu 等[30]提出的另一种做法是利用几次匿名区域中位置点的概率分布来计算新的匿名集合，并用信息熵来度量匿名集合的匿名程度是否达到了用户的等级. 这种做法的一个局限性是由于熵衡量的是不确定程度，如果匿名区域中的 k 个用户全部集中在同一个位置，就会导致隐私泄露.

集中式匿名服务器结构隐私保护效果好，能保证较高的服务质量，且移动终端用户所需的计算、传输开销较小，是目前最常用的系统结构. 但其缺点也较为明显，集中式结构模型过分依赖可信匿名服务器，一旦匿名服务器被攻破，将造成严重的隐私威胁，并且在人人垂涎于数据的大数据时代，难以找到真正可信权威的第三方服务器.

(2) 分布式结构模型

分布式结构模型[31]主要由移动终端和 LBS 服务器两部分组成，如图 4 所示. 终端用户之间遵循 P2P 协议并加入相应的组，当有用户提出查询请求时，组内用户相互合作建立匿名区域，再将匿名区域上传给 LBS 服务器；LBS 服务器中的查询处理器处理过后，将候选结果集返回给终端用户，终端用户再从中选出正确的结果.

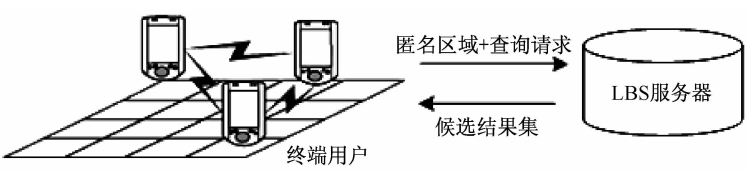

图 4　分布式结构模型

Fig. 4　P2P architecture model

分布式结构模型不依赖于第三方匿名服务器，多个终端相互协作建立隐私保护机制，但是终端用户的资源开销较大. 对于建立匿名区域的时机，目前主要存在的形式有两种：请求式和周期式. 请求式是在用户需要提出查询请求时，才向周围用户广播自己的查询请求，建立满足其隐私要求的匿名区域. 这种方法消耗的计算、传输开销较小，但由于是提出请求时

才建立匿名区域,所以响应时间较长.对于周期式,不论用户是否需要发布查询请求,用户都会定期地发送建立匿名区域的请求,这种方式增加了计算与传输的开销,但是当用户提出查询请求时,其响应时间相对较短.

Gabriel 等[31]提出了基于 Hilbert 曲线的方法 Prive,利用 Hilbert 曲线将二维坐标转化为一维序列,在一维序列中连续选取 k 个用户组成匿名区域.Prive 也引入了第三方服务器,第三方服务器不储存用户位置等敏感信息,仅仅用于身份验证(经过身份验证的用户视为安全用户)与告知当前系统中一些安全用户的 IP 地址,用户通过这些 IP 地址寻找相应的匿名组,组内轮流选取头结点负责处理成员的查询请求.Che 等[32]考虑到 map 中的位置语义信息,如商店、公路、湖等,文章中提出不同语义区域提交查询的概率是不同的,对不同位置语义赋予不同的权值,用户根据位置语义的权值计算 MaxDen 来确定自己的匿名区域.

分布式结构模型的优点是不用过分依赖第三方可信服务器,但是终端之间的合作机制会导致较高的计算与传输开销,并且难以保证所有的合作用户都是可信的,用户的隐私容易受到恶意用户的威胁.

2.1.2 混合区域

除了基于 k 匿名的隐私保护方法以外,有些学者提出了用混合区域实现匿名的方法.Levente 等[33]将路网划分成可观察区域和不可观察区域(混合区域),在可观察区域中车辆的身份是可见的,不可观察区域中车辆身份不可见,车辆只能在不可见区域中更换假名,从而使得攻击者无法将用户进出不可见区域前后所使用的假名关联起来.Palanisamy 等[14,34]提出在真实路网交叉口建立形状不规则的混合区域模型 Mix-zone,将用户真实发出请求的时间加入一个随机时间作为用户提交查询的时间,从而达到抵抗计时攻击的目的.Palanisamy 等[35]对 Mix-zone 进行改进,通过增加时间窗口的方法,使得 Mix-zone 的 k 值较大,增加其安全性,从而解决连续查询中用户查询隐私泄露的问题.

混合区域可以在一定程度上增强假名更换的有效性,但是其缺点也比较明显:①需要借助第三方匿名服务器;②位于混合区域中的用户无法提出查询请求,降低了系统的服务质量.

2.2 基于干扰的隐私保护模型

基于干扰技术主要思想是通过使用噪音对用户的准确位置进行扰动,或者用空间上距离较近的位置替代用户准确位置上传给服务器,使得攻击者不能获得用户的准确位置,主要技术包括:假数据和坐标转换.

(1) 假数据

假数据的类别主要有两种:假查询[8]和假位置[36].假查询是指用户在提交查询时,为保护用户的查询语义,用户通过某种机制生成一部分假查询,假查询的查询内容与真实查询不同.然后将用户的真实查询与假查询一起发送给服务提供商,使得服务提供商无法从查询集合中确定哪一个查询是用户的真实查询.文献[8]中提出了 DUMMY-Q 模型,利用轨迹预测模型与假查询来保护连续请求中的查询语义.假位置主要是指用户在提交查询时,为保护用户的位置隐私,用户根据相应的位置生成机制,生成一些假位置,将用户的准确位置与假位置一起发送给服务提供商,使得服务提供商无法确定位置集合中哪个位置是用户的真实位置.Lu 等[37]提出的隐私保护模型 PAD 就是利用假位置数据来建立匿名区域,并且指出,如果不考虑匿名区域内各个请求位置之间的距离,会导致匿名区域的隐私保护效果下降.例

如某个用户在人口密集的地方提出查询请求，周围有用户也提出查询请求的概率就比较大，如果仅考虑邻近性建立匿名集，最终生成的匿名区域的面积可能会很小，达不到很好的隐藏用户位置的效果. 为了获得一个较大的隐匿名区域，文中提出了两种寻找虚假请求位置点的方法，使得最终提交请求的各个位置点所形成的区域面积接近用户自定义的预设值. Niu 等[36]提出的添加假位置数据建立匿名区域的方法，不仅考虑了匿名区域的大小因素，还考虑到用户可能根据历史请求记录来排除虚假请求.

基于假数据的隐私保护方法优点是简单、易于实现，并且不依赖于任何第三方匿名服务器；其缺点是会增加移动终端的计算开销，因此对于终端的处理能力具有一定的要求，并且对于添加假数据实现 k 匿名的方法，由于组成匿名集的其他用户实际是不存在的，如果攻击者可以真实地观察到匿名区域中各个请求点的实际状况时，能够排除虚假请求的位置点，进而识别出请求者实际所在的位置.

(2) 坐标转换

坐标转换主要是指将用户的真实坐标转换为另一个坐标，作为用户的真实位置发送给服务提供商. 锚点是一种较常使用的坐标变换方法，是用户通过某种规则随机产生的一个位置，为自己真实位置附近的点. 攻击者只能获得锚点的位置，不能获得用户的真实位置. 文献[38]采用在用户真实位置附近随机选取点作为锚点的方法，对用户的真实坐标进行转换. 然而此方法没有考虑到位置语义，不能很好地保护用户的位置语义信息. Zhou[39]根据用户停留点的位置语义多样性对用户进行位置隐私保护，使用满足语义多样性的点作为锚点，更好地保护了用户的位置隐私信息. 使用锚点的位置转换方法避免了区域查询所造成的高计算和通信开销，但是隐私保护度不高.

(3) 差分隐私

差分隐私保护[40]是一种较新的隐私保护模型，被广泛认为是比较严格和强健的保护模型. 该保护模型的基本思想是对数据添加噪音来达到隐私保护效果. 该方法可以确保在某一数据集中插入或者删除一条记录的操作不会影响任何计算的输出结果，主要应用在数据发布的位置隐私保护中. Assam 等[41]提出一种向前缀树中添加噪音的方法保护轨迹数据，作者仅考虑了轨迹中的时间维度来保护轨迹数据. Chen 等[42]提出一种基于随机理论的差分隐私保护方法来保护轨迹数据，通过对数据进行干扰，公布简单轨迹的位置坐标(例如，一段时间内轨迹的坐标平均值). Cicek 等[40]通过将空间中的 POI 分组，然后将轨迹划分进不同 POI 组的方法，结合攻击者所知道的地理信息保护轨迹数据，然后通过 p 机密性衡量隐私保护的效果.

2.3　基于密码学的隐私保护模型

基于密码学的隐私保护模型是使用加密的方法保护用户的隐私，在 LBS 领域使用这些技术，可以在不泄露用户任何身份信息和查询信息的前提下给用户提供服务. 相对基于匿名隐私保护模型，该方法隐私保护效果更好，但是计算开销大.

Gabriel 等[13]通过隐私信息检索 PIR 的方法保护用户的位置隐私，PIR 加密方法依赖二次剩余假说思想. 用户需要根据当前位置推测出需要访问数据的位置，然后从服务端将信息传送给用户. Papadopoulos 等[43]提出 cPIR，对 PIR 的方法进一步优化，使得计算开销大幅度下降. Roman 等[44]通过引入半可信第三方和加密技术保护用户查询隐私，系统中的半可信第三方不知道用户的任何时空信息，仅用于验证查询结果. Lu 等[45]提出 PLAM 隐私

保护框架,通过使用同态加密技术保护用户隐私,但使用这种框架保护用户隐私所消耗的时间开销相对较大.

Khoshgozaran 等[46]提出了基于 Hilbert 曲线的加密方法,将用户的位置与用户兴趣点从二维坐标转移到一维加密空间,通过两条不同参数的 Hilbert 曲线转化而来的一维加密空间仍然保持了二维空间中的邻近性,因此在一维加密空间中同样可以进行 k 邻近查询与范围查询. 该方法引入的可信实体仅在离线阶段将用户兴趣点转化为 Hilbert 序列发送给 LBS 服务提供商,并将 Hilbert 转换规则提供给用户端,不涉及任何线上查询过程,因此不知道用户的位置信息. 当用户提出 LBS 请求时,将位置坐标转换成 Hilbert 序列值后,直接发送给 LBS 服务提供商进行查询,服务提供商再将满足其查询要求的对象返回给用户.

2.4 基于访问控制的隐私保护模型

传统的访问控制模型(如基于角色的访问控制)无法很好的解决 LBS 中的访问控制问题,在 LBS 中,访问控制机制面临新的挑战. 例如,一个提供电子图书的图书馆,只允许在图书馆内的用户访问电子图书,而禁止图书馆外的用户访问电子图书. 对于该场景,访问控制策略需要加入相应的时空维度来控制用户的访问权限,达到保护位置隐私和数据隐私的目的.

Carmen 等[47]提出了一种访问控制系统模型. 该系统为了更好地处理医疗保健中的一些突发事件,用距离作为访问控制的限制条件,从而达到保护病人的位置数据、病史数据等隐私的目的. Yingjun 等[48]提出基于路径的访问控制策略,首先根据用户的隐藏区域 A 和系统允许访问区域 B 的重叠面积 X 与用户隐藏区域面积 Y 计算出比值 $P(X/Y)$,然后通过比较 P 与系统设定阈值的大小来决定是否允许该用户访问数据. 若比值 P 大于阈值,则允许用户访问数据,反之则不允许用户访问数据.

2.5 LBS 隐私保护模型小结

本节中主要对现有工作中 4 类隐私保护模型进行介绍,每类隐私保护模型的应用场景都有所差异,隐私保护效果也不尽相同. 如表 3 所示,每种隐私保护模型都有其相应的特点.

3 LBS 隐私保护度量模型

第 2 章节总结了现有工作中的隐私保护模型,对于每一种隐私保护模型,都需要评估它在实际应用中的隐私保护效果,并结合其他性能指标,如 CPU 消耗、传输开销、系统吞吐量等,来综合考虑这种隐私保护方法是否适用于当前场景,能否满足用户的隐私要求. 本节将分别介绍查询隐私与位置隐私保护的度量模型.

3.1 查询隐私度量模型

查询隐私的度量模型是用于衡量 LBS 隐私保护模型对于用户查询隐私的保护效果,即衡量某种隐私保护模型在实际应用中,能否有效地避免攻击者将某个查询请求与实际请求者成功关联起来,如图 5,为目前常用的度量模型.

(1) 基于熵的查询隐私度量模型

熵(包括位置熵、成对熵等)的作用是衡量体系的不确定程度,熵的值越大,不确定度就越高. 基于熵的查询隐私度量模型是通过用户与查询请求之间的对应关系的不确定程度来度量隐私保护的效果. 攻击者关联某一查询请求与请求者之间的概率越小,熵越大,用户查询隐私暴露的可能性也就越小;反之,如果攻击者关联某一查询请求与请求者之间的概率越

大，熵越小，查询隐私暴露的可能性也就越大. 熵指标最早由 Serjantov 等[49]提出，并将此标准用于度量匿名集对于成员用户的隐私保护程度. 用户提交查询请求后，匿名系统为其建立匿名集，用匿名集作为通信单位. 攻击者的目标是使用某些攻击手段识别出某个请求是由匿名集中哪个用户发出的. 若攻击者判断出匿名集中每一个用户 i 发出该请求的概率是 P_i，则可以通过公式(1)来计算熵：

$$H(X) = -\sum_{i=0}^{k} p_i \times \log_2(P_i). \tag{1}$$

表 3　隐私保护模型汇总

Tab. 3　Summary of privacy protection model

类别	隐私保护模型	查询隐私	位置隐私	个性化需求	可信第三方	连续查询
集中式结构	IntervalClock[26]	○	○	×	○	×
	Casper[27]	○	○	○	○	×
	CliqueClock[9]	○	○	○	○	×
	PTreeCA[18]	○	○	○	○	×
	ICliqueClock[24]	○	○	○	○	○
	Xu[29]	○	○	○	○	○
	Xu[30]	○	○	○	○	○
独立式结构	PAD[37]	○	○	○	×	×
	Niu[36]	○	○	○	×	×
分布式结构	Prive[31]	○	○	○	×	×
	SALS[32]	×	○	○	×	×
混合区域	Levente[33]	○	○	×	○	○
	Mix-zone[6,14]	×	○	×	○	○
干扰	DUMMY-Q [8]	○	×	×	×	○
	Zhou[39]	×	○	×	○	×
	DLS[36]	×	○	○	○	×
	SpaceTwist[38]	×	○	○	×	×
密码学	PIR[13]	×	○	×	×	○
	cPIR[43]	×	○	×	×	○
	PLAM[45]	×	○	×	○	×
	Khoshgozaran[46]	×	○	×	×	×
访问控制	Carmen[47]	×	○	×	×	×
	Yingjun[48]	×	○	×	×	×

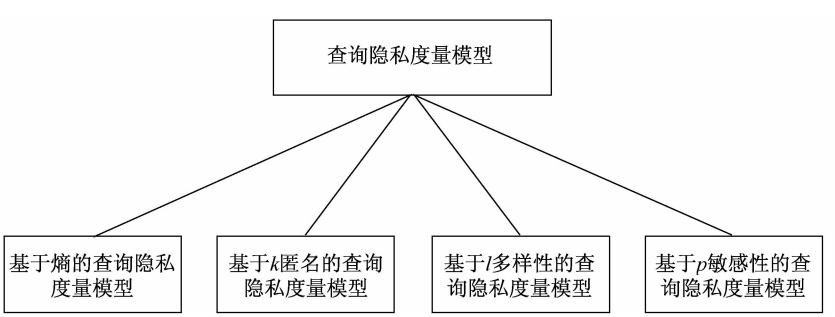

图 5　查询隐私度量模型

Fig. 5　Measure model of query privacy

LBS 隐私保护中的很多技术[24,50]通过熵来对用户身份与查询请求之间的不确定性进

行度量,从而达到保护用户隐私信息的目的.

(2) 基于 k 匿名的查询隐私度量模型

k 匿名是信息安全领域中一种普遍的度量标准,最初由 Sweeney 等[51]提出并应用在关系数据库领域中,其定义为:对于准标识符属性,要求单条记录无法与其他至少 k-1 条记录区分开来.准标识符是指生日、性别、邮政编码等可以与其它背景知识结合起来唯一关联到某人身份的属性,因此从理论上看,k 匿名可以使攻击者准确识别目标对象的概率不大于 $1/k$. 在 LBS 领域中,k 匿名的度量模型作为很多隐私保护模型[9-11]的隐私效果度量模型. Zhang 等[10]在 k 匿名的基础上提出了强 k 匿名概念,要求同一用户在连续多次查询请求时,组成匿名区域的 k 个用户要尽可能保持不变,这样可以防止攻击者通过计算用户多次提交的匿名区域的交集来准确识别出查询请求者.

(3) k 匿名查询隐私度量模型的衍生

k 匿名虽然在一定程度上可以保护用户的查询隐私,但是它仅仅利用空间上的邻近性来构建匿名区域,而没有考虑匿名集中各个查询请求之间的语义关系.如果匿名区域中的用户提交的查询请求类型都相同,可能会导致隐私保护效果的降低.例如,如果某个匿名集中所有的用户提出的查询请求都是寻找教堂,那么对于该匿名集中的任意一个用户,攻击者都可以知道其请求内容是寻找教堂,从而推断出用户具有宗教信仰.总的来说,k 匿名只能防止攻击者关联用户与他提出的请求,而不能防止攻击者关联用户与匿名集中的请求内容.针对这一问题,Liu 等[17]在 k 匿名的基础上进行扩展,提出了 l 多样性的概念,要求匿名集不仅需要满足 k 匿名,而且至少包含 l 种不同语义类型的请求内容(如饭店与医院就属于不同语义类型的查询请求).l 多样性对匿名区域中用户的查询语义作了约束,因此也被用于 LBS 中隐私保护模型的衡量指标.然而,本文作者提出,即使满足 l 多样性仍可能泄露用户的隐私,假如匿名区域中的请求语义类型数大于 l,但是其中的每一个请求都包括敏感内容,则攻击者可以认定用户提出了一个敏感请求.因此,Zhen 等[52]提出了 p 敏感度指标,要求匿名区域中,包括敏感内容的请求数目所占的比例小于 p.

3.2 位置隐私度量模型

位置隐私的度量模型是用于衡量 LBS 隐私保护模型对于用户位置隐私的保护效果,即衡量 LBS 隐私保护模型能否有效地防止攻击者获取用户的位置隐私.如图 6,目前常用的位置隐私度量模型包括基于位置熵、位置 k 匿名、位置 l 多样性的度量模型,也有一些技术采用距离来衡量位置隐私保护效果.

(1) 基于位置熵的位置隐私度量模型

熵不仅可以作为查询隐私的度量指标,也可以作为位置隐私的度量指标.基于位置熵[36]的位置隐私度量模型是通过用户与位置之间对应关系的不确定性来度量隐私保护的效果,位置熵越大,用户与位置之间的不确定性越高,隐私保护效果越好.Niu 等[36]基于位置熵的度量指标提出了一种 DLS 算法.先将区域划分为许多网格,基于历史查询记录,每个网格都有自己的请求概率,所有网格的请求概率值的和为 1. 假设用户提交请求的真实位置所在的网格的请求概率值为 P,为了使攻击者难以根据边信息识别出真实的用户请求位置,DLS 要求在与 P 值最相近的 $2k$ 个网格中选择 k-1 个网格发布假数据,使得这 k-1 个网格的请求概率值与用户的真实位置的 P 值所计算出来的熵最大.熵越大,攻击者识别出用户真实请求位置的难度越大.

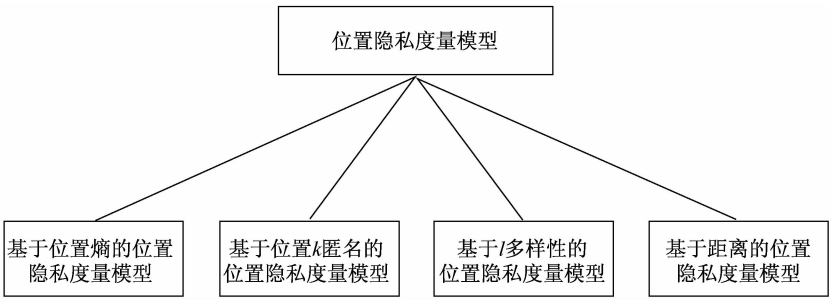

图 6　位置隐私度量模型

Fig. 6　Measure model of query location privacy

(2) 基于位置 k 匿名的位置隐私度量模型

在 LBS 位置隐私保护领域中，位置 k 匿名既能保护用户的查询隐私也能保护用户的位置隐私. 例如文献[9]提出的隐私保护方法中，用户提交查询时生成的匿名集合，需要包括用户的真实位置与至少 k-1 个假位置，攻击者能够准确识别用户真实请求位置概率不大于 $1/k$，从而保护用户的位置隐私.

(3) 基于位置 l 多样性的位置隐私度量模型

位置 k 匿名保护模型将同一个匿名集中的用户包括在一个匿名区域内，虽然在一定程度上可以对位置进行模糊化，但是构建匿名集过程中并未考虑各个用户请求位置之间的关系. 如果匿名区域中各个用户的位置非常相近，可能会导致 k 匿名对于位置隐私保护效果的降低. 例如，某个用户提出请求时，上传的匿名区域完全位于教堂范围内，那么攻击者就可以确定服务请求者一定位于教堂中，进而推测出用户具有宗教信仰. 针对这一问题，文献[40,53-54]在位置 k 匿名的基础上进行扩展，提出了位置 l 多样性的概念，不仅需要完成位置 k 匿名，并且匿名区域中需包含至少 l 种不同语义类型的位置地点. l 多样性对匿名区域中用户的位置类型作了约束，实际上也确保了某些用户位置之间存在一定间距，可以防止匿名区域面积太小.

(4) 基于距离的位置隐私度量模型

基于距离的度量模型[22,55]主要是针对位置模糊隐私保护模型的衡量指标. 攻击者通过分析用户过去以及当前位置(经过模糊处理)，可以估计用户当前所在的位置，一些文献使用攻击者估计出的位置与用户真实位置之间的距离作为衡量指标. 攻击者估计的用户位置与用户真实位置之间距离越大，隐私保护效果越好，反之，则隐私保护效果越差. Shokri 等[22,55]基于这种思想提出了一种应用广泛的位置隐私度量工具，采用攻击者的估计误差期望作为度量指标. 该标准与攻击者获得用户位置的准确程度相关联，期望越小，攻击者获得用户位置信息越准确，用户位置隐私暴露越大，反之亦然.

4　数据集

在大数据时代，拥有数据就等于拥有财富，由于多种原因，数据拥有者不愿意将其数据公布给研究者使用，只有小部分研究者拥有适合自己研究的数据集. 数据集一直受到广大研究者的关注，能够拥有适合自己研究的数据集可以有效提高研究的进度. 对于一些公共数据(例如，各种 POI 位置数据集和路网数据集)，一些公司会提供相应的 API 供研究者使用(例

如,百度 API)或者政府机构会公布一些公共数据供研究者使用,所以,现有工作中公共数据集常为真实数据.但是对于一些涉及用户隐私的数据(例如,用户的轨迹数据),不能随意将其公布给研究者使用,很多研究者较难获得一些个人的真实数据,所以在现有研究中常为模拟数据.实验中所需数据集的获取问题一直困扰着广大研究者们,成为实验成败的关键因素.实验中的数据集可以分为两类:真实数据集和模拟数据集.

4.1 真实数据集

真实数据集是指现实生活中真实存在的数据集.由于公共数据集(例如,路网数据、POI 位置数据等)较容易获得,现有工作中使用的公共数据集多数为真实数据集,地图数据和路网数据是研究者使用较多的真实数据(文献[46]中的饭店数据集 NAVTEQ(www.navteq.com)、文献[36]中的纽约市地图数据、文献[24]中的德国奥尔登堡路网数据和文献[37]中的学校数据集 SC(包括 162 896 个学校的位置数据)等).

对于有些研究者难以获得的数据,有时存在着某些真实数据可以近似替代,部分研究者采用替代的方法进行相应的实验.文献[8]中研究者没有用户的历史查询请求数据,直观上路网密度大是区域历史查询请求数量较多,所以研究者使用美国康乃迪克州的路网密度来替代网格区域中的用户查询请求数量,可以近似满足研究者的要求.

4.2 模拟数据集

模拟数据是指现实中并非真实存在的数据,而是研究者通过某种生成方法模拟生成的数据.由于很多数据集包含着大量的用户隐私和牵涉到数据拥有者的利益,数据拥有者不愿意将数据公布给研究者使用.对于大多数研究者,数据集难以获得,所以许多研究者在实验中使用模拟数据集进行实验.

很多研究者没有所需的数据,为了获得适合自己研究的数据集,他们常使用模拟器模拟生成相应的数据集.文献[24,31,47]中将真实路网作为输入,使用 Network-based Generator of Moving Objects[56]在真实路网上生成用户的移动轨迹,该移动轨迹生成器可以根据研究者的需要设定相应参数,从而满足研究者实验的个性化需求,也是当下使用最为广泛的移动轨迹数据生成器.文献[14,34-35]采用 GT Mobile simulator 在真实路网上生成车辆的移动轨迹.文献[57]使用 VANET[9]车辆移动轨迹模拟器模拟生成车辆的移动轨迹数据集.在另一些研究工作中,研究者根据数据的分布特征,采用随机生成数据的方法生成所需的数据集(例如,文献[44]中的 POI 位置数据集就是采用该方法随机生成的).

5 总结与展望

本文对 LBS 隐私保护领域现有工作进行总结,主要从四个角度进行论述:攻击模型、LBS 隐私保护模型、度量模型和实验中使用的数据集.

隐私保护的概念已经提出很多年,但 LBS 隐私保护研究才刚起步不久,发展还不够成熟,仍然存在着很多的问题需要进一步解决,例如:①对于用户连续请求的隐私保护问题目前还没有好的解决办法,现有的一些方法[21,30],从隐私保护效果和 LBS 服务质量两方面来看并不理想;②没有统一的度量模型量化机制,导致不能很好对现有的一些技术做出充分合理的评价;③现有的研究工作中,没有对攻击者所具有的知识多少作出量化.判断隐私保护模型在攻击者知道哪些信息的情况下是安全的,在攻击者知道哪些信息的情况下是不安全的;④在大数据时代,攻击者可以根据大数据之间的关联性,通过数据挖掘的手段挖掘用户

隐私信息. 所以在大数据时代给用户位置服务中的隐私保护带来了新的挑战和机遇；⑤如何给用户提供个性化的隐私保护机制仍然是一大难题，用户还不能够完全根据自己的需求定制个性化的隐私保护模型. 隐私保护是一门结合了多门学科的技术，隐私保护机制制定者需要了解包括政策、密码学、心理学和数学等多方面的知识，才能更好地研究，使得隐私保护技术快速发展.

[参　考　文　献]

[1] JUNGLAS I A, WATSON R T. Location-based services[J]. Communications of the ACM, 2008, 51(3): 65-69.

[2] ZICKUHR K. Location-based services[J]. Pew Research, 2013:1-25.

[3] BARKHUUS L, DEY A K. Location-based services for mobile telephony: A study of users' privacy concerns [C]//IFIP TC13 International Conference on Human-Computer Interaction. Zurich, Switzerland: DBLP, 2003: 709-712.

[4] 鲁中网-鲁中晨报. iphone 苹果手机定位服务涉嫌泄露用户隐私[EB/OL]. [2015-05-30]. http://news.lznews.cn/2014/0712/740174.html.

[5] cnBeta.cm. McAfee 称谷歌获取用户隐私给社会带来毁灭性影响[EB/OL]. [2015-05-30]. http://www.cnbeta.com/articles/317799.htm.

[6] Marist Poll. Half of Social Networkers Online Concerned about Privacy[EB/OL]. [2015-05-30]http://maristpoll.marist.edu/714-half-of-social-networkers-online-concerned-about-privacy/.

[7] 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.

[8] PINGLEY A, ZHANG N, FU X, et al. Protection of query privacy for continuous location based services[C]//INFOCOM, 2011 Proceedings IEEE. [s.l.]:IEEE, 2011: 1710-1718.

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

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

[11] CHOW C Y, MOKBEL M F, AREDF W G. Casper*: Query processing for location services without compromising privacy[J]. ACM Transactions on Database Systems (TODS), 2009, 34(4): 24.

[12] REBOLLO-MONEDERO D, PARRA-AMAU J, DIAZ C, et al. On the measurement of privacy as an attacker's estimation error[J]. International Journal of Information Security, 2013, 12(2): 129-149.

[13] GHINITA G, KALNIS P, KHOSHGOZARAN A, et al. Private queries in location based services: anonymizers are not necessary[C]//Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data. New York: ACM, 2008: 121-132.

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

[15] CHOW C Y, MOKBEL M F. Enabling private continuous queries for revealed user locations[M]//Advances in Spatial and Temporal Databases. Berlin: Springer, 2007: 258-275.

[16] XIAO P, ZHEN X. Survey of location privacy-preserving[J]. Journal of Frontiers of Computer Science and Technology, 2007, 1(3): 268-281.

[17] LIU F Y, HUA K A, CAI Y. Query *l*-diversity in location-based services[C]//Mobile Data Management: Systems, Services and Middleware. Tenth International Conference on IEEE. [s.l.]: IEEE Xplore, 2009: 436-442.

[18] 吴雷，潘晓，朴春慧，等. 基于位置服务中防止敏感同质性攻击的个性化隐私保护[J]. 计算机应用，2014，34(8): 2356-2360.

[19] DING Y, PEDDINTI S T, ROSS K W. Stalking Beijing from Timbuktu: A generic measurement approach for exploiting location-based social discovery[C]//Proceedings of the 4th ACM Workshop on Security and Privacy in Smartphones and Mobile Devices. New York: ACM, 2014: 75-80.

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

[21] PAN X, MENG X, XU J. Distortion-based anonymity for continuous queries in location-based mobile services [C]//Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. New York: ACM, 2009: 256-265.

[22] SHOKRI R, THEODORAKOPOULOS G, LE BOUDEC J Y, et al. Quantifying location privacy[C]//Security and Privacy (SP), IEEE Symposium on. [s. l.]:IEEE, 2011: 247-262.

[23] MACHANAVAJJHALA A, KIFER D, GEHRKE J, et al. *l*-Diversity: Privacy beyond *k*-anonymity[J]. ACM Transactions on Knowledge Discovery from Data (TKDD), 2007, 1(1): 3.

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

[25] XUE A Y, ZHANG R, ZHENG Y, et al. Destination prediction by sub-trajectory synthesis and privacy protection against such prediction[C]//Data Engineering (ICDE), 2013 IEEE 29th International Conference on. [s. l.]: IEEE, 2013: 254-265.

[26] GRUTESER M, GRUNWALD D. Anonymous usage of location-based services through spatial and temporal cloaking[C]//Proceedings of the 1st Iinternational Conference on Mobile Systems, Applications and Services. New York: ACM, 2003: 31-42.

[27] MOKBEL M F, CHOW C Y, AREF W G. The new Casper: Query processing for location services without compromising privacy[C]//Proc. of the 32nd Int'l Conf. on Very Large Data Bases. Seoul: VLDB Endowment, 2006:763-774.

[28] 田秀霞, 王晓玲, 高明, 等. 数据库服务——安全与隐私保护[J]. 软件学报, 2010, 21(5): 991-1006.

[29] XU T, CAI Y. Exploring historical location data for anonymity preservation in location-based services[C]//INFOCOM 2008. The 27th Conference on Computer Communications. IEEE. [s. l.]:IEEE, 2008.

[30] XU T, CAI Y. Location anonymity in continuous location-based services[C]//Proceedings of the 15th annual ACM international symposium on Advances in geographic information systems. New York: ACM, 2007: 39.

[31] GHINITA G, KALNIS P, SKIADOPOULOS S. PRIVE: anonymous location-based queries in distributed mobile systems[C]//Proceedings of the 16th international conference on World Wide Web. New York: ACM, 2007: 371-380.

[32] CHE Y, CHIEW K, HONG X, et al. SALS: semantics-aware location sharing based on cloaking zone in mobile social networks[C]//Proceedings of the First ACM SIGSPATIAL International Workshop on Mobile Geographic Information Systems. New York: ACM, 2012: 49-56.

[33] BUTTYAN L, HOLCZER T, VAJDA I. On the effectiveness of changing pseudonyms to provide location privacy in VANETs[M]//Security and Privacy in Ad-hoc and Sensor Networks. Berlin: Springer, 2007: 129-141.

[34] PALANISAMY B, LIU L. Mobimix: Protecting location privacy with mix-zones over road networks[C]//Proceedings of the 27th International Conference on Data Engineering. Hannover,Germany: IEEE, 2011: 494-505.

[35] PALANISAMY B, LIU L. Effective mix-zone anonymization techniques for mobile travelers[J]. GeoInformatica, 2014, 18(1): 135-164.

[36] NIU B, LI Q, ZHU X, et al. Achieving *k*-anonymity in privacy-aware location-based services[C]//IEEE INFOCOM 2014-IEEE Conference on Computer Communications. [s. l.]: IEEE, 2014: 754-762.

[37] LU H, JENSEN C S, YIU M L. Pad: Privacy-area aware, dummy-based location privacy in mobile services[C]//Proceedings of the Seventh ACM International Workshop on Data Engineering for Wireless and Mobile Access. New York: ACM, 2008: 16-23.

[38] YIU M L, JENSEN C S, MOLLER J, et al. Design and analysis of a ranking approach to private location-based services[J]. ACM Transactions on Database Systems (TODS), 2011, 36(2): 10.

[39] ZHOU C, MA C, YANG S, et al. A Location Privacy Preserving Method Based on Sensitive Diversity for LBS [M]//Network and Parallel Computing. Berlin: Springer, 2014: 409-422.

[40] 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.

[41] ASSAM R, HASSANI M, SEIDL T. Differential private trajectory protection of moving objects[C]//Proceedings of the Third ACM SIGSPATIAL International Workshop on GeoStreaming. New York: ACM, 2012: 68-77.

[42] CHEN R, FUNG B, DESAI B C, et al. Differentially private transit data publication: a case study on the montreal transportation system[C]//Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining. New York: ACM, 2012: 213-221.

[43] PAPADOPOULOS S, BAKIRAS S, PAPADIAS D. pCloud: A Distributed System for Practical PIR[J]. IEEE Trans Dependable Sec Comput, 2012, 9(1): 115-127.

[44] SCHLEGEL R, CHOW C Y, HUANG Q, et al. User-Defined Privacy Grid System for Continuous Location-Based Services[J]. Mobile Computing, IEEE Transactions on, 2015, 14(10): 2158-2172.

[45] LU R, LIN X, SHI Z, et al. PLAM: A privacy-preserving framework for local-area mobile social networks[C]//IEEE INFOCOM 2014-IEEE Conference on Computer Communications. [s. l.]: IEEE, 2014: 763-771.

[46] KHOSHGOZARAN A, SHIRANI-MEHR H, SHAHABI C. Blind evaluation of location based queries using space transformation to preserve location privacy[J]. GeoInformatica, 2013, 17(4): 599-634.

[47] VICENTE C R, KIRKPATRICK M, GHINITA G, et al. Towards location-based access control in healthcare emergency response[C]//Proceedings of the 2nd SIGSPATIAL ACM GIS 2009 International Workshop on Security and Privacy in GIS and LBS. New York: ACM, 2009: 22-26.

[48] ZHANG Y, CHEN K, LIAN Y. A path-based access control method for location obfuscation in mobile environment[C]//Electrical and Electronics Engineering (EEESYM), IEEE Symposium on. [s. l.]: IEEE, 2012: 570-573.

[49] SERJANTOV A, DANEZIS G. Towards an information theoretic metric for anonymity[C]//Proceedings of the Workshop on Privacy Enhancing Technologies. Berlin: Springer, 2003: 41-53.

[50] XU J, TANG X, HU H, et al. Privacy-conscious location-based queries in mobile environments[J]. Parallel and Distributed Systems, IEEE Transactions on, 2010, 21(3): 313-326.

[51] SWEENEY L. *k*-anonymity: A model for protecting privacy[J]. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2002, 10(05): 557-570.

[52] XIAO Z, XU J, MENG X. *p*-Sensitivity: A semantic privacy-protection model for location-based services[C]//Mobile Data Management Workshops, 2008. MDMW 2008. Ninth International Conference on IEEE. [s. l.]: IEEE, 2008: 47-54.

[53] YANG N, CAO Y, LIU Q, et al. A novel personalized TTP-free location privacy preserving method[J]. International Journal of Security and Its Applications, 2014, 8(2): 387-398.

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

[55] SHOKRI R, THEODORAKOPOULOS G, DANEZIS G, et al. Quantifying location privacy: The case of sporadic location exposure[C]//Privacy Enhancing Technologies. Berlin: Springer, 2011: 57-76.

[56] BRINKHOFF T. A framework for generating network-based moving objects[J]. Geoinformatica, 2002, 6(2): 153-180.

[57] YAO L, LIN C, LIU G, et al. Location anonymity based on fake queries in continuous location-based services [C]//Availability, Reliability and Security (ARES), 2012 Seventh International Conference on IEEE. [s. l.]: IEEE, 2012: 375-382.

（责任编辑　张　晶）