

## 基于通告距离特性的“悦跑圈”移动应用用户定位方法

张冰<sup>1,2</sup>, 王跻霖<sup>3</sup>, 时文旗<sup>2</sup>, 刘瑞婷<sup>3</sup>, 罗向阳<sup>2</sup>

(1. 郑州大学网络空间安全学院, 河南 郑州 450002; 2. 河南省网络空间态势感知重点实验室, 河南 郑州 450001;  
3. 中国航天系统科学与工程研究院, 北京 100037)

**摘要:** 开展移动应用用户定位研究对于分析敏感目标用户地理位置、线下打击网络犯罪具有重要现实意义。现有移动应用用户定位技术通常基于用户发布文本、社交关系或系统通告距离, 存在定位偏差大、需目标用户开启位置分享功能等不足。为此, 考虑到运动健身类移动应用往往默认开启用户位置分享服务, 且拥有庞大的用户规模, 提出一种针对主流运动健身应用“悦跑圈”的用户定位方法, 以期为用户定位提供一条新途径。首先, 设计一种基于距离递增的探针动态部署算法, 以识别“悦跑圈”通告距离对应的真实距离边界和分段临界区间; 其次, 在潜在区域等距部署探针以发现用户, 并设计一种基于象限判定探针部署的初步定位算法, 实现对用户的粗粒度定位; 最后, 为提升定位精度, 设计一种基于通告距离分段临界点查找的细粒度用户定位算法, 利用距离特性部署探针以确定用户所处的象限, 并结合二分法确定分段临界点, 基于三边测量推断敏感用户的具体位置。基于300个“悦跑圈”用户点位开展定位实验, 结果表明所提方法可实现“悦跑圈”用户的精确定位, 平均定位误差为1.86 m, 99%用户的定位误差小于10 m, 与基于空间划分、启发式数论等典型的用户定位方法相比, 平均定位误差降低85.08%。

**关键词:** 用户定位; 运动健身类应用; 悦跑圈; 距离递增; 探针部署; 三边测量

**中图分类号:** TP309

**文献标志码:** A

**DOI:** 10.11959/j.issn.2096-109x.2025055

## User geolocation method for Joyrun mobile application based on reported distance feature

ZHANG Bing<sup>1,2</sup>, WANG Jilin<sup>3</sup>, SHI Wenqi<sup>2</sup>, LIU Ruiting<sup>3</sup>, LUO Xiangyang<sup>2</sup>

1. School of Cyber Science and Engineering, Zhengzhou University, Zhengzhou 450002, China  
2. Key Laboratory of Cyberspace Situation Awareness of Henan Province, Zhengzhou 450001, China  
3. China Aerospace Academy of Systems Science and Engineering, Beijing 100037, China

**收稿日期:** 2025-03-23; **修回日期:** 2025-05-26

**通信作者:** 时文旗, shiwenqi1606@163.com

**基金项目:** 国家自然科学基金 (U23A20305, 62172435); 国家重点研发计划 (2022YFB3102900); 河南省重点研发专项基金 (221111321200)

**Foundation Items:** The National Natural Science Foundation of China (U23A20305, 62172435), The National Key Research and Development Program of China (2022YFB3102900), The Key Research and Development Project of Henan Province (221111321200)

**引用格式:** 张冰, 王跻霖, 时文旗, 等. 基于通告距离特性的“悦跑圈”移动应用用户定位方法[J]. 网络与信息安全学报, 2025, 11(5): 172-188.

**Citation Format:** ZHANG B, WANG J L, SHI W Q, et al. User geolocation method for Joyrun mobile application based on reported distance feature[J]. Chinese Journal of Network and Information Security, 2025, 11(5): 172-188.

**Abstract:** The implementation of mobile application user positioning was recognized as crucial for analyzing the geographic locations of sensitive target network users and combating cybercrime through offline means. Existing mobile application user positioning technologies were found to rely on users' published texts, social relationships, or system-reported distances, resulting in significant positioning deviations and requiring users to enable location-sharing functions. Given that mobile sports and fitness applications frequently had location-sharing enabled by default and possessed large user bases, a geolocation method targeting the mainstream fitness application Joyrun was developed to provide a novel approach to user positioning. A distance-incremental probe dynamic deployment algorithm was designed to identify the real distance boundary and the threshold interval corresponding to the Joyrun report distance. Next, probes were deployed equidistantly in potential areas to discover users, and a preliminary localization algorithm based on quadrant determination was formulated for coarse-grained user localization. To enhance accuracy, a fine-grained user localization algorithm based on report distance segmentation threshold finding was designed. Distance features were utilized to deploy probes for determining the user quadrant, and combined with the bisection method for identifying segmentation thresholds, the specific location of sensitive users was inferred through trilateration. Localization experiments were conducted based on 300 Joyrun users. The results demonstrate that the proposed method could achieve accurate localization of Joyrun users. The average localization error is measured as 1.86 m, with 99% of localization errors being less than 10 m. Compared to typical user positioning methods based on spatial segmentation and heuristic number theory, the average positioning error is reduced by 85.08%.

**Keywords:** user positioning, sports and fitness application, Joyrun, distance increment, probe deployment, trilateration

## 0 引言

移动互联网应用凭借其即时性与便捷性的显著优势,深刻重塑了人们的生活方式与社会互动模式。从社交娱乐到出行规划,移动应用无缝整合各类需求,构建起全方位的数字生活服务体系,为用户提供了极大便利。然而,在技术赋能的背后,网络安全问题日益凸显。近年来,依托移动应用实施的非法信息传播、网络诈骗、数据窃取等新型网络犯罪活动迅速增长,不仅严重威胁用户隐私安全,也对社会稳定构成潜在挑战<sup>[1-3]</sup>。在此背景下,移动应用用户定位技术应运而生,成为执法机构打击网络犯罪、维护网络空间安全的关键手段之一<sup>[4]</sup>。该技术能够将虚拟网络空间中的用户映射至现实地理空间,为网络犯罪的调查取证工作提供有力支持<sup>[5]</sup>。此外,在隐私保护领域,该技术还可用于评估移动应用平台位置隐私保护策略的有效性,从而提升用户隐私安全防护意识。

现有的移动应用用户定位方法主要分为基于用户发布文本的定位、基于社交关系的定位和基于通告距离的定位3类。其中基于用户发布文本的定位方法通过分析用户发布文本中提及的地

点和上下文信息,结合地理数据库,推断用户的地理位置<sup>[6-9]</sup>。基于社交关系的定位方法利用用户间的社交联系进行位置推断,其核心是假设频繁互动的用户在地理位置上往往接近<sup>[10-11]</sup>。然而,这两类定位方法普遍存在定位偏差大、定位精度不足的问题<sup>[12]</sup>。相比之下,基于通告距离的定位方法利用移动应用中基于位置的社交发现(location based social discovery, LBSD)服务<sup>[13]</sup>所公开的用户通告距离信息进行定位,通常具有更高的定位精度。根据定位策略的不同,现有基于通告距离的定位方法主要分为基于数论的定位方法和基于逐次逼近的定位方法两类。

基于数论的定位方法借助通告距离的带状分布特性,将通告距离与实际距离的映射关系抽象为数学模型,通过在目标区域策略性地部署探针,并对各探针获取的目标用户通告距离进行约束求解,实现用户位置的精确定位。王荣荣等<sup>[14]</sup>提出一种针对一维LBSD用户的带宽距离报告模式,并设计了高精度定位方法。Xue等<sup>[15]</sup>将一维数论定位思想扩展到二维空间,通过部署虚拟探针形成等距点晶格结构,

构建了二维数论定位模型。该方法在理想条件下可实现用户的高精度定位。Peng 等<sup>[16]</sup>针对通告距离的噪声干扰影响,提出一种基于启发式数论的定位 (heuristic number theory-based geolocation, HNBG) 算法。该算法通过在水平与垂直边界部署多组探针,选取各边界中通告距离最小的中间探针作为预测坐标。然而,此类算法虽在理论层面具备较高的定位精度,但由于缺乏对实际环境中 LBSD 通告距离噪声分布的系统性验证,实际定位效果与理论预期存在显著差距。

基于逐次逼近的定位方法通过迭代优化策略实现位置精确定位:首先界定目标用户的初始定位空间并生成中间定位解,继而通过动态调整探针部署,依据不同位置探针获取的通告距离变化量对中间解进行迭代修正,逐次逼近真实位置。Ding 等<sup>[17]</sup>提出的经典三边测量定位方法,利用 3 个已知位置获取与目标用户的相对距离范围,绘制 3 个圆环,其交叠区域的中心点即为定位结果。Li 等<sup>[18]</sup>提出的 SPBG (space partition based geolocation) 算法通过递归划分用户潜在空间,不断缩小定位范围,突破了最小通告距离的约束,实现了空间维度的逐次逼近。Wang 等<sup>[19]</sup>提出一种统计迭代的定位算法,该算法根据周边探针获取的目标通告距离来估计目标相对中心探针的方位,通过迭代计算逐步收敛至目标位置,提升了动态场景下的定位鲁棒性。Wang 等<sup>[20]</sup>提出了一种两步定位算法,该算法能够量化定位误差和查询复杂度:首先,基于迭代三边测量确定目标所在的小范围区域;其次,在该区域内应用最大似然估计确定目标的具体位置。

受经典三边测量思想<sup>[17]</sup>启发,Shi 等<sup>[21]</sup>基于用户序列的异常状态特征,提出了一种错序状态分析定位法 (missequence state analysis based geolocation, MSAG)。该方法通过分析邻近用户序列的异常状态,找到导致异常序列状态的最大距离差,进而推断目标用户的距离范围。时文旗<sup>[22]</sup>提出了一种基于加权最小二乘定位 (weighted least squares based geolocation, WLBG) 方法。该方法通过对真实距离边界进行加权,以构建目标函

数,并利用最小二乘法<sup>[23]</sup>求解最优位置。针对远距离用户定位场景,Shi 等<sup>[24]</sup>提出了一种基于探针动态调整的远程定位方法 (dynamic adjustment of probes based localization method, DAPL)。该方法通过动态部署探针快速逼近目标真实位置,并监测目标通告距离的异常变化以提高定位效率。

当前,基于移动应用的用户定位技术通常依赖于目标用户主动开启位置分享服务。以即时通信应用为例,大多数产品要求用户通过特定操作路径手动激活位置功能。例如,在微信中,目标用户必须主动点击“发现”页面的“附近”选项,并授权应用获取实时位置信息。在此前提下,定位者方有可能通过“附近的人”列表获取目标用户的通告距离,进而实施定位。因此,在多数情况下,由于用户未主动授权应用获取实时位置信息,定位者无法获得必要的距离信息,从而难以实现对目标用户的准确定位。

相比之下,运动健身类移动应用具有用户规模庞大、位置服务默认开启等特性,为用户定位提供了稳定的通告距离数据源。为实现运动数据记录、运动轨迹分享等主要功能,该类应用用户通常默认开启位置分享服务,定位者可通过“附近动态”等功能持续获取与周边活跃用户的相对距离数据,为定位提供基础支撑。

本文以运动健身应用“悦跑圈”为研究对象,该应用注册用户超 1.1 亿,活跃用户覆盖全球 2 500 个城市。通过对该应用通告距离与真实距离数据对的收集与分析,发现通告距离特性,进而获取通告距离对应的真实距离边界及分段临界区间。在此基础上,设计了基于象限判定的初步定位方法,并进一步提出基于通告距离分段临界点查找的细粒度定位方法,实现对“悦跑圈”用户的高精度定位。本方法适用于用户历史位置的确定,而非对用户实时位置的追踪。

本文的主要贡献如下。

1) 设计了一种基于距离递增探针部署策略的数据收集算法。通过在“悦跑圈”的“附近动态”服务中进行上万次数据采集,并结合分类分

析,该算法成功识别出“悦跑圈”通告距离对应的真实距离边界,并确定了通告距离的分段临界区间。

2) 设计了一种基于象限判定的探针部署的初步定位算法。该算法以三边测量定位为基础,通过合理部署探针,利用探针通告距离的变化确定用户所在的象限,有效避免了探针分布不合理的情况,显著缩小了定位交界区域,从而在一定程度上提高了三边测量定位的精度。

3) 设计了一种基于通告距离分段临界点查找的细粒度用户定位算法。该算法通过部署探针触发通告距离的分段变化,以确定用户所在的象限;在此基础上,利用二分法查找分段临界点,并结合三边测量定位技术,实现更加精准的用户定位。

## 1 “悦跑圈”位置保护机制分析

本节将简要介绍运动健身移动应用的“附近动态”服务和距离混淆策略,并描述本文所提定位方法的前提条件。

### 1.1 “悦跑圈”的“附近动态”服务简介

“悦跑圈”的“附近动态”服务是一项基于地理位置服务的社交化运动互动功能。用户可将个人运动记录分享给周边用户,同时也能浏览附近用户发布的动态。这一设计不仅强化了用户的参与感与互动体验,更通过社交属性提升了应用的用户黏性与活跃度。该服务依托GPS卫星定位、Wi-Fi热点定位、蜂窝网络定位等多元LBS(location-based services)技术,锁定用户的当前位置<sup>[25-26]</sup>。应用服务器会定期更新用户账户相关数据,并将其安全地存储在数据库中。

如图1所示,当用户发起周边动态查询时,系统会将包含其位置信息的请求发送至服务器;服务器接收请求后,记录查询者的当前位置信息,并从数据库中检索出与查询者地理位置相近的用户动态及对应的位置信息;检索结果按照时间由近及远的顺序即时反馈给查询者。基于“附近动态”的信息展示,用户可通过点赞、发表评论、发送私信等多种方式,轻松建立并深化彼此间的社交联系与互动体验。



图1 “悦跑圈”的“附近动态”场景图

Figure 1 Scene map of the nearby dynamics of Joyrun

### 1.2 “悦跑圈”距离混淆策略

“悦跑圈”在促进用户交流与分享运动成就的同时,也深刻意识到位置信息数据泄露所带来的潜在风险。鉴于此,“悦跑圈”提供了以下距离混淆策略。

1) 相对距离通告机制:仅通告用户间的相对距离,隐藏用户的实际经纬度坐标。该机制既能满足用户基于位置的社交需求,又能确保用户的准确位置信息不被直接泄露。

2) 距离段模糊处理:采用距离段表示用户间的相对距离,而非直接展示用户位置的精确值。具体而言,“悦跑圈”以100 m为单位量化用户间距离,查询结果均为100 m的整数倍。该粗粒度距离信息既能满足用户感知附近用户距离的要求,又有效规避了泄露精确位置信息的风险。

3) 距离范围严格限制:对用户动态展示实施严格的距离范围约束(“悦跑圈”仅展示与发布者通告距离在10 km以内的用户动态)。该措施有效限制了不法分子对用户行踪和日常活动的推测空间,进一步提升了用户信息的安全性。

4) 时间窗口动态限制:为防御长时间恶意追踪,对LBSD功能设置严格的时间窗口约束(“悦跑圈”仅展示近一周内的用户动态)。此机制既保障了信息时效性,又显著降低了信息泄露风险。

### 1.3 问题描述

本文聚焦于研究精确定位“悦跑圈”用户

的可行性。研究目标并非获取某区域内指定用户的实时位置，而是致力于实现对某一指定用户历史位置的精确定位。需明确的是，研究者被设定为普通“悦跑圈”用户，仅能依托普通用户可获取的公开信息对其他用户的历史位置进行推测。此外，研究者仅拥有与普通用户完全等同的操作权限，这意味着研究者不能通过任何非法手段（如渗透应用服务器）获取额外敏感数据。不过，研究者具备两项关键操作能力：一是可自由修改设备位置，通过在不同地理位置切换视角，查看并收集对应区域的附近历史动态信息；二是能够借助自动化手段爬取“悦跑圈”的“附近动态”界面的用户相关信息，具体涵盖用户名、用户间的大致距离范围（相对距离）、用户 IP 地址所指示的地理位置及动态发表时间等关键数据。

## 2 本文方法

由于运动健身类移动应用的“附近动态”服务会展示附近用户间的通告距离，因此可通过收集大量的<通告距离，真实距离>数据对，剖析通告距离与真实距离之间的对应关系。本文聚焦于通告距离的分段特性，旨在确定通告距离各分段临界点所对应的真实距离区间，从而实现对用户历史位置的精确定位。

### 2.1 概念与符号解释

在深入探讨定位方法之前，首先需要明确一些关键概念和符号。

1) 通告距离：“悦跑圈”的“附近动态”界面提供的用户间距离，是用户定位过程所需的关键数据。

2) 实际距离：真实地理空间中用户间的直线距离，该距离可通过球面距离公式等方法计算得出。

3) 参考探针：设置的“悦跑圈”用户，具有发布和查看“附近动态”的权限，并可自主修改位置。由于探针的坐标都是已知的，因此探针间的真实距离是可以通过计算得到的。

4) 通告距离分段特性：当用户间实际距离连续变化时，通告距离并非同步平滑波动，而是在一定范围内保持固定值。仅当用户间实际距离达到特定临界值，满足进入下一个通告距离区间的

条件时，通告距离才会跃升至下一个固定值区间。这种类似“台阶式”的分段变化特征，本文称之为“距离分段特性”。

5) 通告距离分段临界点：指通告距离发生变化时探针所处的位置。假设变化前的通告距离为 $(k - 100)$ ，变化后为 $k$ ，则该位置称为通告距离 $k$ 的分段临界点，通告距离 $k$ 的分段临界点所在的区间称为通告距离 $k$ 的分段临界区间。

本文涉及的符号定义如表 1 所示。

表 1 符号定义  
Table 1 Explanation of notation

| 符号     | 描述                           |
|--------|------------------------------|
| $d_t$  | 真实距离，使用球面距离计算公式计算出的距离        |
| $d_r$  | 通告距离，“悦跑圈”的“附近动态”界面显示的用户间的距离 |
| $U$    | 最小通告距离单位，“悦跑圈”为 100 m        |
| $P$    | 参考探针                         |
| $B$    | 通告距离对应的真实距离边界                |
| $I$    | 通告距离分段临界区间                   |
| $p$    | 通告距离分段临界点                    |
| $t$    | 数据收集阶段的探针移动步长                |
| $s$    | 探针部署时的移动距离                   |
| $Th_d$ | 潜在区域用户发现阶段的通告距离阈值            |
| $Th_p$ | 初步定位阶段的通告距离阈值                |
| $Th_s$ | 临界点获取阶段的二分法查找区间阈值            |
| $Th_f$ | 最终定位阶段的定位误差阈值                |

### 2.2 所提方法的基本原理和主要步骤

本文所提定位方法的原理示意图如图 2 所示。首先，系统对采集的“悦跑圈”附近动态的<通告距离，真实距离>数据对进行分类分析，识别出通告距离对应的真实距离边界，并统计分析通告距离发生变化的条件，确定通告距离临界区间；接着，在目标所处潜在区域等距部署探针搜寻目标用户，直至目标用户出现在探针的“附近动态”界面；然后，采用基于象限判定的探针部署策略，并结合三边测量方法得到初步定位结果；之后，设计基于通告距离临界区间的探针部署方案，以触发通告距离变化，通过二分法查找通告距离的分段临界点，从而确定真实距离，并获得目标用户的象限判定结果；最后，在三边测量方法的基础上，结合象限判定结果，得到目标

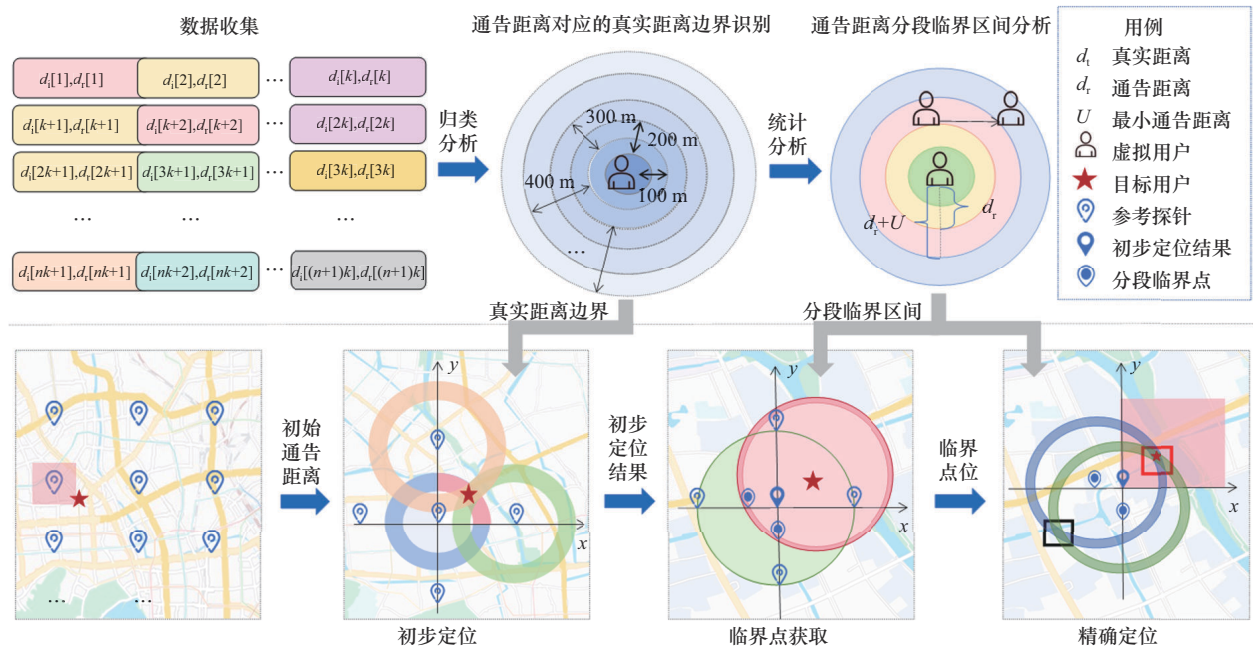


图2 基于通告距离特性的“悦跑圈”移动应用用户定位方法原理示意图

Figure 2 Schematic of the user geolocation method for Joyrun mobile application based on reported distance feature

用户的理论位置。该方法的主要步骤如下。

### 步骤1 数据收集分析

在这一步中，设计一种基于距离递增探针部署的数据收集算法，收集大量<真实距离，通告距离>数据对，分析获得通告距离对应的真实距离边界 $B$ 和通告距离分段临界区间 $I$ 。

首先，设定一个位置固定的参考探针作为目标用户。随后，部署其他参考探针充当查询者。当以固定步长逐步增大查询者与目标用户之间的真实距离时，查询者与目标用户之间的通告距离也会发生相应的变化。由于所有用户（包括查询者和目标用户）的位置都是已知的，因此可以使用球面距离计算公式计算其间的真实距离。基于此过程，可构建一个包含<通告距离，真实距离>数据对的数据集 $D$ ，其表达式如下。

$$D = \{d_t[i], d_r[i] | i = 1, 2, \dots, n\} \quad (1)$$

其中 $d_t[i]$ 表示第 $i$ 个点处查询者与目标用户的真实距离， $d_r[i]$ 表示第 $i$ 个点处查询者与目标用户的通告距离。

通告距离对应的真实距离边界识别：对于相同通告距离 $k$  ( $k = 100, 200, \dots$ )，找到对应的所有真实距离，得到集合 $S_k$ 。

$$S_k = \{d_t[i] | d_r[i] = k, i = 1, 2, \dots, n\} \quad (2)$$

所以通告距离 $k$ 对应的真实距离最小值为 $\min(S_k)$ 、最大值为 $\max(S_k)$ 。然而在区间 $(\min(S_k) - s, \min(S_k))$ 和 $(\max(S_k), \max(S_k) + s)$ 中也可能存在通告距离为 $k$ 的情况，只是由于移动步长的限制未在此区间内获得通告距离，假设移动步长为 $t_1$ ，则通告距离 $k$ 对应的真实距离边界 $B_k$ 为

$$B_k = [L_k, R_k] = \begin{cases} (\min(S_k) - t_1, \max(S_k) + t_1), & k \neq 100 \\ [0, \max(S_k) + t_1], & k = 100 \end{cases} \quad (3)$$

通告距离分段临界区间确定： $t_2$ 为数据收集时探针的移动步长，统计得到通告距离 $a$  ( $a = 100, 200, \dots$ )对应的最大真实距离 $\max(S_a)$ ，通告距离 $b$  ( $b = 100, 200, \dots$ )对应的最小真实距离 $\min(S_b)$ ，其中 $b - a = 100$  m，那么通告距离 $b$ 的分段临界区间 $I_b$ 为

$$I_b = \begin{cases} (\min(S_b) - t_2, \max(S_a) + t_2), & \max(S_a) \geq \min(S_b) \\ (\min(S_a), \min(S_b)), & \max(S_a) < \min(S_b) \end{cases} \quad (4)$$

### 步骤2 潜在区域用户发现

在目标用户潜在区域内, 等距部署参考探针搜寻目标用户, 直至目标用户出现在某一探针的“附近动态”中, 且此时的通告距离不超过阈值  $Th_d$ 。选择该探针作为初始探针  $P_0$ , 并将当前探针与目标用户的通告距离记录为  $d_l[0]$ 。考虑到现有其他定位方法将目标限定在  $1\ 000\text{ m} \times 1000\text{ m}$  的区域内, 为方便后续对比实验分析, 本文将  $Th_d$  设置为  $1\ 000\text{ m}$ 。

### 步骤3 目标用户初步定位

本文基于三边测量定位方法设计了基于象限判定的探针部署优化策略, 以提高初步定位精度。以初始探针  $P_0$  为坐标原点建立坐标系, 在坐标轴上距离原点  $s_1$  处共部署 4 个探针。取其中与目标用户通告距离最小, 且不在同一坐标轴上的两个探针, 将这两个探针与初始探针联立。结合每个探针与目标用户通告距离所对应的真实距离边界进行三边测量定位, 取交界区中心为定位结果。接着, 获取定位结果与目标用户的通告距离, 并将其与阈值  $Th_p$  进行比较。若未满足阈值, 则按照上述流程迭代定位, 直至满足阈值  $Th_p$ , 将最终符合要求的定位结果记作  $P'_0$ 。

本文将初步定位结果的通告距离阈值  $Th_p$  设置为  $100\text{ m}$ , 该阈值的取值依据将在后续的实验环节详细阐述。此外, 关于探针部署距离  $s_1$  的取值, 将在第 2.3 节中详细介绍。

### 步骤4 通告距离分段临界点获取

在该步骤中, 通过部署探针主动触发通告距离分段变化, 判断用户潜在象限, 并利用二分法查找位于探针与初步定位结果连线上的通告距离分段临界点。具体操作如下: 以初始定位结果  $P'_0$  为原点构建直角坐标系, 在坐标轴上距离原点  $s_2$  处部署 4 个探针, 并获取探针与目标用户的通告距离。若 4 个探针的通告距离均呈现增大趋势, 表明初始定位结果已极度接近目标用户位置, 满足最终定位阈值  $Th_l$  的要求, 此时直接将初始定位结果作为最终输出, 跳过后续步骤。若存在探针的通告距离未变, 则借助象限判定算法确定目标用户所在潜在象限, 并筛选出两个通告距离增大且不在同一坐标轴的探针, 对所选探针与原点的连线区间执行二分查找, 迭代搜索通告距离分

段临界点, 直至区间长度小于预设阈值  $Th_s$ , 取最终区间中点作为分段临界点  $p_1$ 、 $p_2$ 。

关于  $s_2$  的取值和分段临界点  $p_1$ 、 $p_2$  的二分查找相关内容, 将在第 2.4 节详细探讨。

### 步骤5 目标用户精确定位

利用步骤 4 获取的两个分段临界点  $p_1$ 、 $p_2$ , 结合目标用户所在象限信息进行精确定位。步骤 4 中已通过二分法确定了分段临界点的位置, 其最大误差为  $Th_s/2$ 。为确保定位结果包含目标用户实际位置, 将步骤 1 中获取的分段临界区间左右边界分别向两侧扩展  $Th_s/2$ 。

在以初始定位结果  $P'_0$  为原点的平面直角坐标系中, 以两个分段临界点为圆心, 分别以其对应的扩展后真实距离区间为内外半径, 构建两个圆环区域, 目标用户理论上位于这两个圆环的交界区内。尽管交界区可能不唯一, 但根据步骤 4 中获得的象限信息, 可以进一步缩小目标用户可能所在的区域, 取对应区域内的交界区中心作为目标用户的最终定位结果。该方法仅依赖两个临界点进行定位, 与经典的三点定位算法相比, 有效缩减了时间开销。由于利用了额外的方位信息(所处区域和象限), 因此对定位精度的影响不大。

在上述步骤中, 步骤 1、步骤 3 和步骤 4 是本文定位方法的关键所在, 其具体细节分别将在第 2.3 节、第 2.4 节和第 2.5 节详细阐述。

## 2.3 基于距离递增探针部署的数据收集

本节主要介绍数据收集阶段的探针部署算法。该算法通过逐渐增大已知参考探针与目标对象的真实距离, 并同步获取对应通告距离, 系统性地收集<真实距离, 通告距离>数据对, 旨在尽可能全面覆盖不同距离区间的映射关系, 为后续通告距离对应的真实距离边界识别和通告距离分段临界区间的确定提供充分的数据支撑。探针部署算法的流程如算法 1 所示。

**算法 1** 基于距离递增探针部署的数据收集算法

**输入** 起始位置  $P_0$ , 方向向量  $V$ , 距离间隔  $\Delta d$ , 最大通告距离  $d_{\max}$ , 通告距离获取函数 `getReportedDistance()`, 位置更新函数 `changeLocation()`

**输出** <真实距离, 通告距离>对数据集:  
 $D = \{(d_t[1], d_r[1]), (d_t[2], d_r[2]), \dots, (d_t[n], d_r[n])\}$   
 1) 初始化探针位置:  
 $A \leftarrow P_0$  //探针 A 固定在起始位置  
 $B \leftarrow P_0$  //探针 B 从起始位置开始沿方向  $V$  移动  
 2) 初始化真实距离:  $d_t \leftarrow 0$   
 3) 初始化数据集:  $D \leftarrow \emptyset$   
 4) Repeat 1:  
 5)  $d_r \leftarrow \text{getReportedDistance}(A, B)$   
 6)  $D \leftarrow D \cup \{(d_t, d_r)\}$   
 7)  $d_t \leftarrow d_t + \Delta d$   
 8)  $B \leftarrow A + d_t \cdot V$   
 9)  $\text{changeLocation}(B, B)$   
 10) 如果  $d_r \geq d_{\max}$ , 结束 Repeat1  
 11) Return  $D$

在上述算法流程中, 步骤5)为通告距离获取函数, 用于获取两个参考探针之间的通告距离; 步骤9)为位置更新函数, 用于将参考探针部署至指定位置。算法最终输出<真实距离, 通告距离>对数据集。

## 2.4 基于象限判定探针部署的初步定位

本节着重阐述目标用户的初步定位流程, 该流程由探针部署优化与三边测量迭代定位两部分有机结合, 实现精准且高效的定位效果。

1) 探针部署优化。采用基于象限判定的智能部署策略。算法依据目标用户的潜在位置信息, 动态规划探针的分布位置, 使其能够更高效地覆盖用户所在区域。这种方式避免了传统随机部署带来的资源浪费和定位误差, 显著提升了探针的利用率和定位的准确性。

2) 三边测量迭代定位。系统基于三边测量原理, 利用探针与目标用户之间的距离数据, 逐步调整定位坐标, 使其不断逼近真实位置。经过多轮计算与调整, 最终获得满足初始定位阈值要求的初步定位结果。

### 2.4.1 象限判定探针部署优化

在利用三边测量定位方法确定目标用户位置时, 探针的空间布局对定位精度具有显著影响。如图3(a)所示, 当3个探针位于共线或接近共线

时, 以各探针为圆心、目标点到探针的距离为半径形成的圆环区域会出现交界区不唯一的现象, 导致无法准确判断目标用户位于哪一交界区。如图3(b)所示, 当3个探针之间的距离过近时, 它们所确定的圆环相交区域较大且形状复杂(呈现狭长的椭圆或者不规则的长条形)。其原因在于: 探针间距过近使得各探针对目标点的位置约束无法从不同空间方向形成有效限制, 在空间上高度重叠。

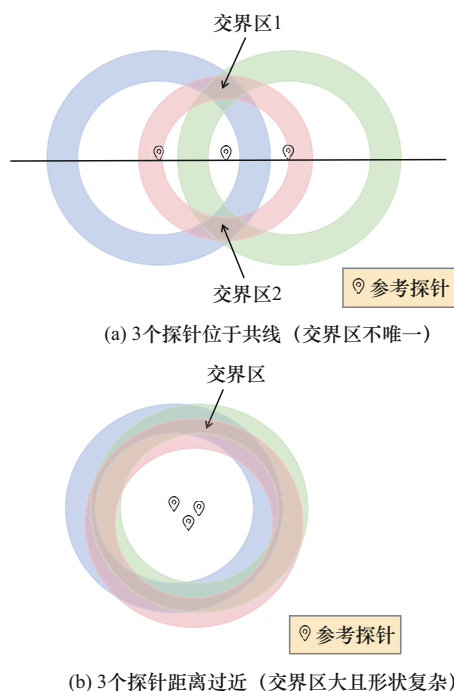


图3 探针选择异常实例  
Figure 3 Example of probe selection exception

在理想的三边测量定位场景中, 期望3个圆环的交界区具备唯一性且面积最小化, 此时以交界区中心作为定位结果时误差最小。在平面直角坐标系下, 若目标区域无特殊限制, 3个探针应分布在以目标点为原点的坐标系的不同象限, 如此可从多个角度对目标位置形成约束, 有效避免定位盲区。同时, 通过确定目标用户所在象限, 可规避探针与目标点距离过远导致的定位误差较大问题。

在用户发现阶段, 当目标用户被感知处于初始探针  $P_0$  的通告距离  $d_r[0]$  处时, 设其对应的真实距离边界为  $[d[0]_{\min}, d[0]_{\max}]$ , 即目标用户位

于以初始探针  $P_0$  为圆心、 $[d[0]_{\min}, d[0]_{\max}]$  为半径形成的圆环区域。此时，在以初始探针  $P_0$  为原点的坐标系中选取其余探针的位置时，探针与初始探针的间距大小直接影响目标所处象限判定的准确性。

为规避定位异常并提升定位精度，本文采用图 4 所示的探针部署策略。首先，以初始探针  $P_0$  为原点建立平面直角坐标系，以  $P_0$  为圆心、 $R$  ( $R = s_1$ ) 为半径作圆；然后，在圆与坐标轴的 4 个交点处分别部署探针  $P_1$ 、 $P_2$ 、 $P_3$  和  $P_4$ ；最后，获取各探针与目标用户的通告距离，选择通告距离较小且不在同一坐标轴上的两个探针，将其与初始探针  $P_0$  组合形成三边测量的 3 个参考点。

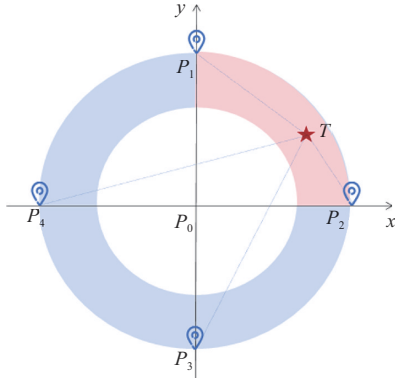


图 4 探针部署  
Figure 4 Probe deployment

设  $T$  的坐标为  $(x_T, y_T)$ 。根据两点间距离公式， $d = \left( (x_2 - x_1)^2 + (y_2 - y_1)^2 \right)^{\frac{1}{2}}$ ，可分别获得  $P_1$ 、 $P_2$ 、 $P_3$  和  $P_4$  与目标  $T$  的距离。

1)  $P_1$  点与目标  $T$  的距离  $|P_1T|$

$$|P_1T| = \left( x_T^2 + (y_T - R)^2 \right)^{\frac{1}{2}} = \left( x_T^2 + y_T^2 - 2Ry_T + R^2 \right)^{\frac{1}{2}} \quad (5)$$

2)  $P_2$  点与目标  $T$  的距离  $|P_2T|$

$$|P_2T| = \left( (x_T - R)^2 + y_T^2 \right)^{\frac{1}{2}} = \left( x_T^2 - 2Rx_T + R^2 + y_T^2 \right)^{\frac{1}{2}} \quad (6)$$

3)  $P_3$  点与目标  $T$  的距离  $|P_3T|$

$$|P_3T| = \left( x_T^2 + (y_T + R)^2 \right)^{\frac{1}{2}} = \left( x_T^2 + y_T^2 + 2Ry_T + R^2 \right)^{\frac{1}{2}} \quad (7)$$

4)  $P_4$  点与目标  $T$  的距离  $|P_4T|$

$$|P_4T| = \left( (x_T + R)^2 + y_T^2 \right)^{\frac{1}{2}} = \left( x_T^2 + 2Rx_T + R^2 + y_T^2 \right)^{\frac{1}{2}} \quad (8)$$

假设目标用户  $T$  位于以  $P_0$  为原点的坐标系的第一象限内，那么  $x_T > 0$ ， $y_T > 0$ 。根据  $R > 0$ ，可知  $Rx_T > 0$ 、 $Ry_T > 0$ 、 $-Rx_T < 0$ 、 $-Ry_T < 0$ ；同时可知  $Rx_T > 0 > -Rx_T$ 、 $Rx_T > 0 > -Ry_T$ 、 $Ry_T > 0 > -Rx_T$ 、 $Ry_T > 0 > -Ry_T$ 。据此可以得到以下关系。

$$|P_4T| = \left( x_T^2 + 2Rx_T + R^2 + y_T^2 \right)^{\frac{1}{2}} > \left( x_T^2 - 2Rx_T + R^2 + y_T^2 \right)^{\frac{1}{2}} = |P_2T| \quad (9)$$

$$|P_4T| = \left( x_T^2 + 2Rx_T + R^2 + y_T^2 \right)^{\frac{1}{2}} > \left( x_T^2 + y_T^2 - 2Ry_T + R^2 \right)^{\frac{1}{2}} = |P_1T| \quad (10)$$

$$|P_3T| = \left( x_T^2 + y_T^2 + 2Ry_T + R^2 \right)^{\frac{1}{2}} > \left( x_T^2 - 2Rx_T + R^2 + y_T^2 \right)^{\frac{1}{2}} = |P_2T| \quad (11)$$

$$|P_3T| = \left( x_T^2 + y_T^2 + 2Ry_T + R^2 \right)^{\frac{1}{2}} > \left( x_T^2 + y_T^2 - 2Ry_T + R^2 \right)^{\frac{1}{2}} = |P_1T| \quad (12)$$

化简得

$$\max(|P_1T|, |P_2T|) < \min(|P_3T|, |P_4T|) \quad (13)$$

因为  $x_T^2 \geq 0$ 、 $y_T^2 \geq 0$ ，所以  $x_T^2 + y_T^2 + 2Ry_T > 0$ 、 $x_T^2 + 2Rx_T + R^2 + y_T^2 > R^2$ 、 $x_T^2 + y_T^2 + 2Rx_T > 0$ 、 $x_T^2 + y_T^2 + 2Ry_T + R^2 > R^2$ ，进而可得  $|P_4T| = \left( x_T^2 + 2Rx_T + R^2 + y_T^2 \right)^{\frac{1}{2}} > R$ 、 $|P_3T| = \left( x_T^2 + y_T^2 + 2Ry_T + R^2 \right)^{\frac{1}{2}} > R$ 。若想通过通告距离的变化判断目标用户所处的象限，则必须使  $R \geq d[0]_{\max}$ 。本文令  $R = d[0]_{\max}$ ，所以  $|P_4T| > d[0]_{\max}$ 、 $|P_3T| > d[0]_{\max}$ ，即探针  $P_3$ 、 $P_4$  与目标用户的通告距离大于  $d_r[0]$ 。由此可知：当目标用户位于第一象限时，探针  $P_3$ 、 $P_4$  与目标用户的通告距离大于

$d_r[0]$ 。其逆否命题为：若探针 $P_3$ 或 $P_4$ 与目标用户的通告距离不大于 $d_r[0]$ ，则目标用户一定不在第一象限。同理，当目标用户位于其他象限时，通过上述过程可推断出类似结论：若探针 $P_1$ 或 $P_2$ 与目标用户的通告距离不大于 $d_r[0]$ ，则目标用户一定不在第三象限；若探针 $P_2$ 或 $P_3$ 与目标用户的通告距离不大于 $d_r[0]$ ，则目标用户一定不在第二象限；若探针 $P_1$ 或 $P_4$ 与目标用户的通告距离不大于 $d_r[0]$ ，则目标用户一定不在第四象限。

综上所述：若探针 $P_1$ 与目标用户的通告距离不大于 $d_r[0]$ ，则目标用户一定不在第三和第四象限；若探针 $P_2$ 与目标用户的通告距离不大于 $d_r[0]$ ，则目标用户一定不在第二和第三象限；若探针 $P_3$ 与目标用户的通告距离不大于 $d_r[0]$ ，则目标用户一定不在第一和第二象限；若探针 $P_4$ 与目标用户的通告距离不大于 $d_r[0]$ ，则目标用户一定不在第一和第四象限。因此，根据获取的探针与目标用户的通告距离变化来判断目标用户所处的象限，从而选定三边测量定位的探针，有助于实现合理定位。

#### 2.4.2 三边测量定位

根据第2.4.1节的探针部署优化方法，可以得到3个虚拟探针 $P_0$ 、 $P_1$ 、 $P_2$ ，同时获得这3个探针与目标用户的通告距离 $d_r[0]$ 、 $d_r[1]$ 、 $d_r[2]$ 。设这3个通告距离对应的真实距离边界分别为 $[d[0]_{\min}, d[0]_{\max}]$ 、 $[d[1]_{\min}, d[1]_{\max}]$ 、 $[d[2]_{\min}, d[2]_{\max}]$ ，采用三边测量定位算法可实现对目标用户的初步定位。以各探针位置为圆心，以各探针确定的目标用户真实距离的下限为内径、上限为外径，分别绘制圆环，取3个圆环交叉区域的中心作为本次三边测量的定位结果，如图5所示。

获取上述定位结果与目标用户的通告距离，若满足阈值 $Th_p$ 的条件，则将本次定位结果作为初始定位结果；若不满足阈值 $Th_p$ 的条件，则以该定位结果为原点，运用第2.4.1节的象限判定探针部署策略重新选择探针，并重复上述定位过程，直到定位结果与目标用户的通告距离满足阈值 $Th_p$ 的条件为止。

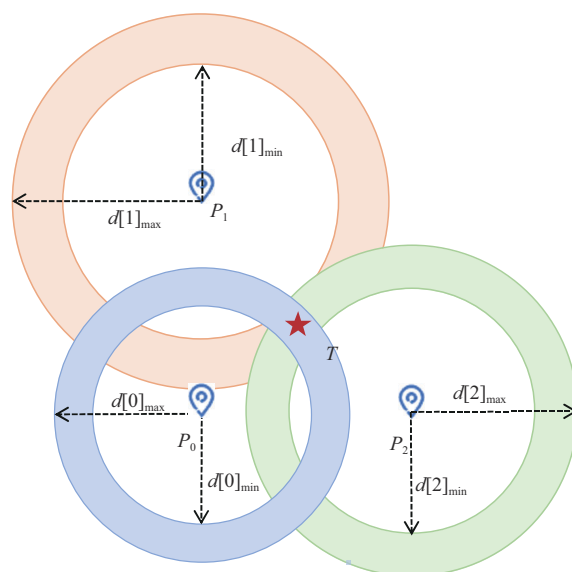


图5 三边测量定位

Figure 5 Triangulation positioning

### 2.5 基于距离分段特性确定通告距离分段临界点

本节聚焦于利用运动健身移动应用的通告距离分段特性，精准划定与目标用户的真实距离范围。首先，通过分析目标用户与初步定位结果的相对位置，确定探针移动距离，并结合探针通告距离变化判定用户所处象限，最终采用二分查找法确定通告距离分段临界点。

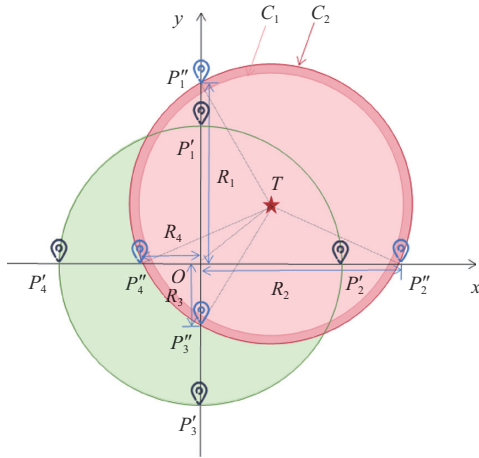
#### 2.5.1 基于距离分段特性的探针部署

当用户间真实距离连续变化时，运动健身应用的通告距离呈现分段式变化，其分段的转折点称为分段临界点。但在实际场景中，由于探针与目标用户的相对位置未知，无法通过静态部署直接获取分段临界点，需通过动态探针移动触发通告距离变化，进而定位分段临界点位置。

完成初步定位后，得到初步定位结果 $P'_0$ 及其100 m通告距离。为实现精确定位，本文以查找通告距离200 m的分段临界点为目标。采用“先大步移动距离 $s_2$ 触发通告距离改变，再二分法缩减区间”的策略。接下来对初始定位结果与目标用户的分布情况进行分析，以确定 $s_2$ 的取值。

首先，基于第2.3节的探针部署算法，以初始定位结果 $P'_0$ 为原点构建平面直角坐标系，以

距离  $s_2$  为半径画圆，在圆与坐标轴的 4 个交点处部署探针  $P'_1$ 、 $P'_2$ 、 $P'_3$  和  $P'_4$ ，获取各探针与目标用户的通告距离。以目标用户  $T$  为圆心，以通告距离 200 m 对应的分段临界区间边界为圆  $C_1$  和  $C_2$  的半径  $R_1$ 、 $R_0$ 。其中，位于圆  $C_1$  内的点与目标用户的通告距离均为 100 m，位于圆  $C_2$  外的点与目标用户的通告距离大于 100 m，圆  $C_1$  和  $C_2$  所形成的圆环即为通告距离 200 m 的分段临界区间。设圆  $C_2$  与坐标轴的 4 个交点分别为  $P''_1$ 、 $P''_2$ 、 $P''_3$  和  $P''_4$ ，记其与原点的距离为  $R_1$ 、 $R_2$ 、 $R_3$  和  $R_4$ ，如图 6 所示。因此，若要使  $P'_i$  处的通告距离发生改变，则须满足  $|OP'_i| \geq |OP''_i|$  ( $i = 1, 2, 3, 4$ )，即  $s_2 \geq R_i$  ( $i = 1, 2, 3, 4$ )。

图 6  $s_2$  取值(一般情况)Figure 6 Values of  $s_2$ (general cases)

(1) 一般情况 (初始定位结果不满足最终定位误差阈值  $\text{Th}_f$  的条件，表明其与目标用户存在一定距离)

假设目标用户  $T$  位于以初步定位结果为原点建立的平面直角坐标系的第一象限。如图 6 所示，分别将  $O$ 、 $P'_1$ 、 $P'_2$ 、 $P'_3$ 、 $P'_4$  与  $T$  连接，假设  $\angle TP'_1P'_3 = \alpha$  ( $0 < \alpha < 90^\circ$ )、 $\angle TP'_2P'_4 = \beta$  ( $0 < \beta < 90^\circ$ )、 $\text{dis}(O, T) = \text{error}$  ( $3 < \text{error} < R_0$ )。分析可得：

$$R_1 + R_3 = 2R_0 \cdot \cos \alpha \quad (14)$$

$$R_2 + R_4 = 2R_0 \cdot \cos \beta \quad (15)$$

因为  $0 < \alpha < 90^\circ$ 、 $0 < \beta < 90^\circ$ ，所以  $0 < \cos \alpha < 1$ 、 $0 < \cos \beta < 1$ 。据此可得：

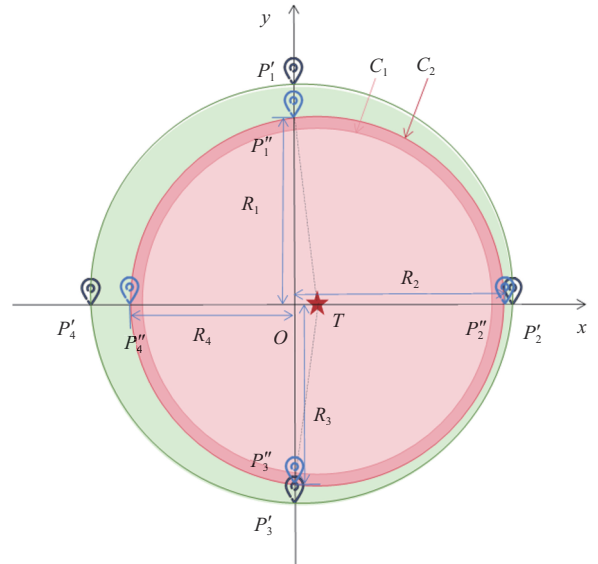
$$0 < R_1 + R_3 < 2R_0 \quad (16)$$

$$0 < R_2 + R_4 < 2R_0 \quad (17)$$

$R_1$  和  $R_3$ 、 $R_2$  和  $R_4$  每一对中至少有一条边的长度小于  $R_0$ 。因此，只要保证  $s_2 \geq R_0$ ，就能确保至少存在两个点，它们与坐标轴的交点的通告距离大于 100 m。

(2) 极端情况 (初始定位结果满足最终定位误差阈值  $\text{Th}_f$  的条件，表明其已经非常接近目标用户)

假设目标用户位于  $x$  轴正半轴，与原点相距  $\text{error}$  的位置 (此时直接将初始定位结果作为最终定位结果输出)。如图 7 所示，分别将  $P'_1$ 、 $P'_3$  与  $T$  连接， $|OT| = \text{error}$  ( $0 \leq \text{error} \leq \text{Th}_f$ )。分析可得：

图 7  $s_2$  取值(极端情况)Figure 7 Values of  $s_2$  (extreme cases)

$$R_1 = (R_0^2 - \text{error}^2)^{\frac{1}{2}} \quad (18)$$

$$R_2 = R_0 + \text{error} \quad (19)$$

$$R_3 = R_1 \quad (20)$$

$$R_4 = R_0 - \text{error} \quad (21)$$

此时， $R_2$  值最大，只要保证  $s_2 \geq R_2$ ，即可使 4 个点通告距离均改变。因为  $\max(R_2) = R_0 + \text{Th}_f$ ，所以  $s_2 \geq R_0 + \text{Th}_f$ 。

$s_2$  的取值并非越大越好， $s_2$  值过大会显著增加区间划分的时间开销。基于上述一般情况和极

端情况下对 $s_2$ 取值的分析, 本文令 $s_2 = R_o + Th_r$ , 获取4个探针( $P'_1$ 、 $P'_2$ 、 $P'_3$ 、 $P'_4$ )与目标用户的通告距离, 对这4个点位的通告距离情况进行分析。

1) 若4个探针的通告距离均改变, 表明初步定位结果的精度已经满足要求, 直接输出结果作为最终定位结果。

2) 若存在通告距离未改变的探针, 则在通告距离改变的探针中选择两个非共轴探针, 在其与原点的连线区间上寻找分段临界点。同时, 依据第2.4节的象限判定理论, 根据探针通告距离变化情况来判定用户所处的象限。

### 2.5.2 通告距离分段临界点获取

当初始定位结果未达到目标精度时, 需采用通告距离分段临界点技术实现用户精确定位。具体实施逻辑如下: ①以原点与探针的连线为基准, 对目标区间进行迭代二分操作。每次在二分点部署探针, 实时获取其与目标用户的通告距离。②根据通告距离的变化趋势, 判断目标用户所在的象限(基于二维坐标系逻辑), 以此缩小定位区间。若距离减小, 说明用户靠近探针方向; 若距离增大, 则说明用户远离探针方向。③持续执行上述二分一判定流程, 直至区间长度不超过预设阈值 $Th_s$ 。此时取区间中点作为分段临界点 $p$ , 完成精确定位。

以通告距离200 m的分段临界点获取为例, 假设临界点 $p$ 位于线段 $OA$ 上。在初始状态下,  $|OA| = dis$ ,  $O$ 点与目标用户的通告距离 $d_o = 100$  m,

$A$ 点与目标用户的通告距离 $d_A = 200$  m。首先二分 $OA$ , 并部署探针 $P_1$ , 若 $P_1$ 与目标用户的通告距离 $d_1 = 100$  m, 表明分段临界点 $p$ 在 $P_1A$ 上; 然后二分 $P_1A$ , 并部署探针 $P_2$ , 若 $P_2$ 与目标用户的通告距离 $d_2 = 200$  m, 说明分段临界点 $p$ 在 $P_1P_2$ 上; 依次循环迭代, 直到 $|P_{i-1}P| \leq Th_s$ , 此时取线段 $P_{i-1}P$ 的中点作为通告距离200 m的分段临界点 $p$ , 如图8所示。阈值 $Th_s$ 的设定须兼顾定位效率与精度, 值过小会增加探针移动及通告距离获取次数, 导致定位开销上升; 值过大会降低定位精度。经综合考虑, 本文将 $Th_s$ 设为5 m。

## 3 实验结果及分析

为验证本文提出的基于通告距离特性的用户定位方法的有效性, 本文在真实环境下对“悦跑圈”用户开展定位实验, 并与其他基于系统通告距离的典型定位方法进行对比分析。“悦跑圈”的“附近动态”服务的通告距离的最小粒度为100 m, 可展示周边10 km以内的用户动态。

### 3.1 实验设置

实验由4部分组成: 通告距离对应的真实距离边界识别实验、初步定位实验、分段临界区间确定实验和目标用户精确定位实验。具体实验设置见表2。实验采用安装“悦跑圈”应用的真机作为探针, 借助虚拟位置软件Fake Location实现探针位置修改。

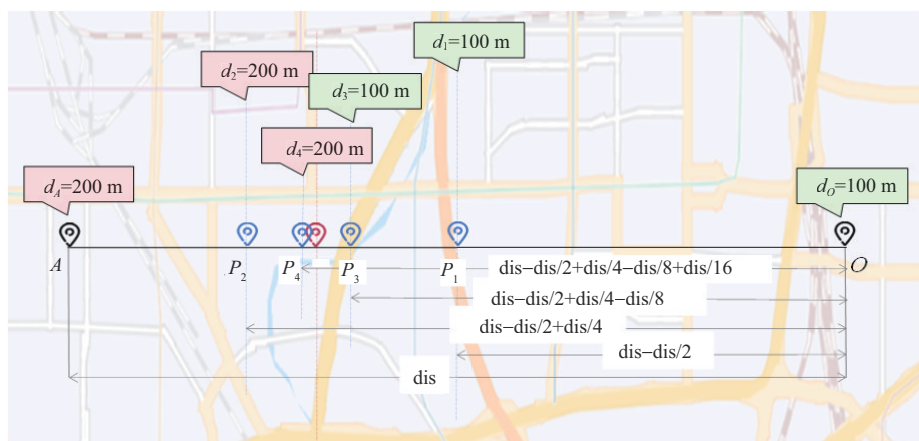


图8 二分法确定通告距离分段临界点

Figure 8 Dichotomy to determine report distance segmentation thresholds

表 2 实验设置

Table 2 Experience settings

| 序号 | 项目                       | 设置                    |
|----|--------------------------|-----------------------|
| 1  | 通告距离范围                   | 100~1 000 m           |
| 2  | “悦跑圈”账号数量                | 6个                    |
| 3  | 数据收集阶段探针移动步长 $t$         | 20 m, 5 m, 1 m        |
| 4  | 目标发现过程中相邻探针间隔            | 400 m                 |
| 5  | 潜在区域用户发现的通告距离阈值 $Th_d$   | 1 000 m               |
| 6  | 临界点获取阶段的二分法查找区间阈值 $Th_s$ | 5 m                   |
| 7  | 最终定位阶段的定位误差阈值 $Th_f$     | 3 m                   |
| 8  | 实验区域                     | 郑州市区、杭州市区             |
| 9  | 自动化工具                    | Appium1.22.3          |
| 10 | 测试用手机                    | 2部MI 9, 2部Redmi K60   |
| 11 | 手机系统版本                   | Android 10、Android 13 |
| 12 | “悦跑圈”版本                  | 5.41.0                |
| 13 | 虚拟位置软件及版本                | Fake Location 1.3.5   |

### 3.2 数据收集和初步定位相关实验

#### (1) 通告距离对应的真实距离边界识别实验

选取郑州市区、杭州市区两地作为实验区域,利用6个“悦跑圈”账号收集用户间通告距离和真实距离对,构建通告距离—真实距离区间的统计模型。在实验过程中,固定目标用户 $A$ 的位置并发表动态,使用多个探针账号调整位置获得与 $A$ 的通告距离。以探针 $B$ 为例,起始与 $A$ 同位置,随后以步长20 m移动,记录每一步的真实距离和通告距离,直到通告距离达1 000 m。重复上述过程30次,发现通告距离呈分段变化,且变化点对应的真实距离均在 $(100i + 20, 100i + 60]$  ( $i = 1, 2, \dots, 10$ )内。为提升数据采集效率,将探针的移动步长 $t_1$ 缩减到5 m,以真实距离为 $(100i + 20 + 5j)$  ( $i = 1, 2, 3, \dots, 10; j = 1, 2, \dots, 8$ )移动探针收集数据,共计进行50轮测试。在剔除异常数据后,对相同通告距离对应的真实距离数据进行归类排序,获取其最大值、最小值,确定通告距离对应的真实距离边界。最终获得100~1 000 m(最小粒度100 m)共计10种通告距离对应的真实距离边界,具体结果见表3。

表 3 通告距离对应的真实距离边界

Table 3 Real distance boundaries corresponding to report distances

| 通告距离/m | 对应最小真实距离/m | 对应最大真实距离/m | 对应真实距离边界/m   |
|--------|------------|------------|--------------|
| 100    | 0          | 150        | [0, 155)     |
| 200    | 150        | 245        | (145, 250)   |
| 300    | 250        | 345        | (245, 350)   |
| 400    | 350        | 450        | (345, 455)   |
| 500    | 450        | 550        | (445, 555)   |
| 600    | 555        | 650        | (550, 655)   |
| 700    | 650        | 750        | (645, 755)   |
| 800    | 755        | 845        | (750, 850)   |
| 900    | 850        | 950        | (845, 955)   |
| 1 000  | 950        | 1 045      | (945, 1 055) |

#### (2) 初步定位实验

将一用户账号作为目标用户随机设置在实验区域中,其余账号作为探针。首先采用潜在区域用户逼近方法逼近目标用户,探针在实验区域内每隔400 m查询一次“附近动态”,直到探针显示的通告距离 $d_r \leq Th_d = 1 000$  m,对目标用户进行初步定位。初步定位基于三边测量定位算法,一共进行500次初步定位实验,每50次定位为一组,计算定位误差,共计获得10组数据,具体见表4。在实际操作中,发现一次三边测量定位结果的通告距离基本能达到100 m,所以本文将初步定位过程的通告距离阈值 $Th_p$ 设置为100 m。

#### (3) 通告距离分段临界区间确定实验

根据初始定位结果可知,初步定位结果与目标用户误差最大值为102.48 m,其通告距离为100 m。所以该方法主要研究通告距离200 m的分段临界点。根据通告距离对应的真实距离边界识别实验数据,可知该分段临界点位于(145, 155) m区间。为了缩小区间长度,提高定位精度,将移动步长 $t_2$ 设为1 m。设置用户 $A$ 为目标用户,用户 $B$ 为探针,将用户 $A$ 、 $B$ 的真实距离以步长1 m从145 m逐步增大至155 m,进行50轮上述实验。统计发现,通告距离100 m对应的最大真实距离为150 m,通告距离200 m对应的最小真实距离为150 m,根据第2.2节步骤1中的公式,可得通告距离200 m的分段临界点所在

区间为(149, 151) m。所以探针移动距离  $s_2 = 151 + Th_f = 151 + 3 = 154$  m。

表4 初始定位误差统计

Table 4 Initial positioning error statistics

| 轮次 | 误差<br>均值/m | 误差<br>中位数/m | 误差<br>最小值/m | 误差<br>最大值/m |
|----|------------|-------------|-------------|-------------|
| 1  | 38.08      | 33.46       | 11.84       | 102.48      |
| 2  | 32.00      | 30.51       | 1.16        | 70.68       |
| 3  | 34.04      | 33.56       | 12.64       | 50.88       |
| 4  | 37.03      | 39.04       | 2.641       | 51.21       |
| 5  | 38.63      | 39.30       | 7.49        | 61.55       |
| 6  | 31.52      | 31.95       | 6.12        | 58.47       |
| 7  | 34.32      | 36.73       | 9.67        | 89.63       |
| 8  | 33.41      | 36.48       | 8.40        | 54.60       |
| 9  | 34.18      | 32.34       | 16.26       | 56.03       |
| 10 | 32.88      | 30.89       | 15.10       | 52.34       |
| 总计 | 34.61      | 35.12       | 1.16        | 102.48      |

### 3.3 定位结果对比分析

首先对现有的基于系统通告距离的移动应用用户定位方法的定位前提条件进行分析,判断其是否适用于“悦跑圈”的用户定位,结果见表5。然后从中选择合适的定位方法与本文所提方法从定位精度(定位误差)和定位效率两方面进行对比分析。其中,定位误差通过计算各定位结果(经纬度)与目标用户真实位置间的偏差得出;定位效率则以发现和定位目标用户期间,各探针针对目标用户通告距离的查询次数作为衡量指标进

行计算得出。

本文选择HNBG<sup>[16]</sup>、SPBG<sup>[18]</sup>和WLSG<sup>[22]</sup>这3种定位方法与所提方法进行对比分析。其中,从定位前提来看,HNBG要求通告距离具有带状特性,SPBG和WLSG均要求通告距离与真实距离具有一定的对应关系。对“悦跑圈”进行调研可知,“悦跑圈”的通告距离存在上述HNBG所要求的带状特性,以及SPBG和WLSG所要求的与真实距离的对应关系等特性;从应用场景来看,上述3种定位方法与本文所提定位方法均针对目标在特定区域内时对目标用户的真实位置的精确推断问题展开研究。因此选择上述3种定位方法开展对比实验。

#### (1) 定位精度对比分析

对300个“悦跑圈”用户点位进行定位计算,各种方法的定位误差如图9所示。从图中能够清晰地看出,即使在动态增强用户周边通告距离混淆的复杂条件下,本文所提方法依然能够对“悦跑圈”用户实现有效定位(最小误差低至0.22 m,平均误差为1.86 m)。与其他3种定位方法相比,本文方法的平均误差分别降低了85.08%、88.24%、86.10%,精度提升优势显著。

为科学评估各定位方法的效果,本文对不同方法的定位误差分布开展了统计分析,结果如图10所示。图中横坐标为定位误差的分布区间,纵坐标则是各误差区间内的定位结果在总定位结果中所占的比例。分析图10可知,与其他定位方法相比,本文提出的方法在定位精度分布

表5 基于系统通告距离的用户定位方法分析

Table 5 Analysis of user localization method based on system report distance

| 类别     | 定位方法                                | 定位前提               | 适用于“悦跑圈” |
|--------|-------------------------------------|--------------------|----------|
| 基于数论   | 一维数论定位 <sup>[14]</sup>              | 目标方位已知,通告距离无混淆     | 否        |
|        | 二维数论定位 <sup>[15]</sup>              | 通告距离无混淆            | 否        |
|        | 启发式二维数论定位(HNBG) <sup>[16]</sup>     | 通告距离的带状特性          | 是        |
| 基于逐次逼近 | 经典三边测量定位 <sup>[17]</sup>            | 通告距离无混淆            | 否        |
|        | 基于空间划分定位(SPBG) <sup>[18]</sup>      | 通告距离与真实距离具有一定的对应关系 | 是        |
|        | 基于统计迭代定位 <sup>[19]</sup>            | 查询次数无限制            | 否        |
|        | 两步定位 <sup>[20]</sup>                | 系统误差符合特定分布         | 否        |
|        | 基于用户错序状态定位(MSAG) <sup>[21]</sup>    | 用户顺序与真实距离的对应关系     | 否        |
|        | 基于加权最小二乘定位(WLSG) <sup>[22]</sup>    | 通告距离与真实距离具有一定的对应关系 | 是        |
|        | 基于探针动态调整的远程定位(DAPL) <sup>[24]</sup> | 可显示远距离下的通告距离       | 否        |

上优势更为显著：能够将约 99% 的目标定位在 10 m 误差范围内，同时实现 100% 的目标定位误差均控制在 20 m 以内（即 99% 的目标误差  $\leq 10$  m，1% 的目标误差在 10~20 m 区间）。

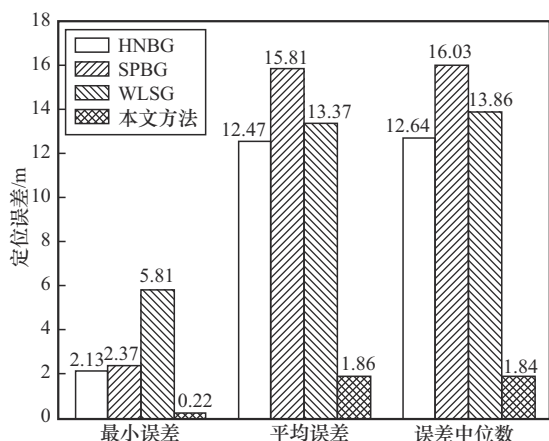


图 9 定位误差对比

Figure 9 Comparison of positioning error

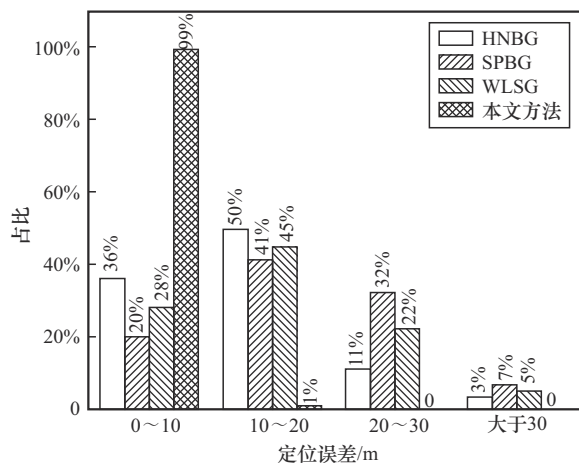


图 10 误差分布对比

Figure 10 Comparison of error distribution

## (2) 定位效率对比分析

本文所提方法和对比方法对目标用户通告距离的平均查询次数如图 11 所示。本文在计算查询次数时，由于 HNBG 方法讨论 1 000 m×1 000 m 区域内的目标用户定位，因此本文将目标用户潜在区域设定为 1 000 m×1 000 m。AQT-F (average query times of user finding) 表示在用户发现过程中的平均查询次数，AQT-G (average query times of user geolocating) 表示在用户定位过程中的平均查询次数，AQT-T (total average query times) 表示平均总查询次数。

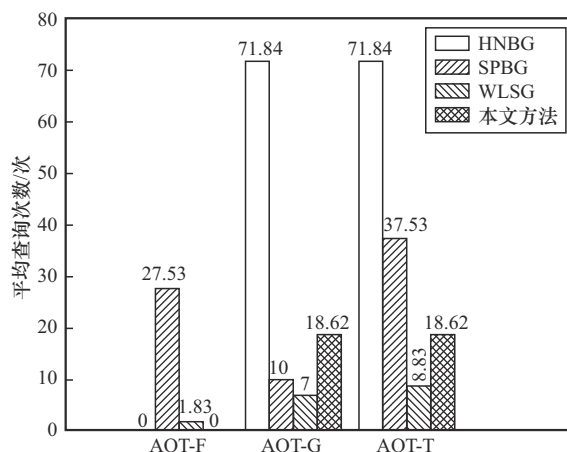


图 11 定位效率对比

Figure 11 Comparison of positioning efficiency

由图 11 可知，本文所提方法在查询效率上优势显著，AQT-T 仅为 18.62 次，显著低于 SPBG、HNBG 等对比方法。SPBG 和 WLSG 方法在用户发现阶段设置了通告距离约束条件。SPBG 要求目标用户必须满足最小通告距离才触发定位，导致查询次数增加；WLSG 同理须依赖特定通告距离条件。二者均因“条件触发机制”产生额外查询开销。当用户潜在区域为 1 000 m×1 000 m 时，本文所提方法和 HNBG 方法均省略用户发现过程。HNBG 方法的用户定位过程首先需要沿区域边界部署大量探针以确定目标用户的初始坐标，并且在精确定位阶段仍需高频查询目标通告距离，双重操作导致其查询次数为所有方法中最多。

综上，相较于现有典型的移动应用用户定位方法，本文提出的基于距离特性的“悦跑圈”用户定位方法具有更好的定位精度和定位效率。

## 4 结束语

针对现有移动应用用户定位方法存在的位置偏差大、高度依赖用户主动开启位置分享功能等问题，本文提出一种专为“悦跑圈”这类默认开启位置分享服务的运动健身应用设计的用户定位方法。该方法深度挖掘“悦跑圈”通告距离特性，通过科学划定目标用户与分段临界点的距离范围，精准判定目标象限，并有机结合三边测量定位算法，成功实现了对“悦跑圈”用户的高精

度定位。实验数据表明,与现有经典定位方法相比,本文提出的方法在定位精度和定位效率方面均取得了显著突破。从保障用户位置隐私安全的角度出发,应用开发者可采取一系列有效措施应对本文提出的定位方法,如限制通告距离获取频次、强化通告距离混淆机制,甚至直接取消通告距离显示功能等。本研究成果具有重要的现实意义:一方面,能够为执法机构打击违法犯罪活动提供有力的技术支持;另一方面,通过揭示公开数据中潜在的隐私风险,为应用平台优化数据展示策略提供科学参考。在此特别强调,本定位方法严禁用于任何侵犯个人隐私或危害社会安全的场景,必须严格遵守相关法律法规。

## 参考文献:

- [1] MINK J, YUILE A R, PAL U, et al. Users can deduce sensitive locations protected by privacy zones on fitness tracking Apps[C]//Proceedings of the CHI Conference on Human Factors in Computing Systems. New York: ACM, 2022: 1-21.
- [2] DHONDT K, POCHAT V L, VOULIMENEAS A. A run a day won't keep the hacker away: inference attacks on endpoint privacy zones in fitness tracking social networks[C]//Proceedings of the 29th ACM SIGSAC Conference on Computer and Communications Security. New York: ACM, 2022: 801-814.
- [3] RODRIGUES A, VILLELA M L, FEITOSA E. A systematic mapping study on social network privacy: threats and solutions[J]. ACM Computing Surveys, 2024, 56(7): 1-29.
- [4] 杨海峰,杜少勇,魏果,等.基于时空行为与静态属性分析的陌陌应用恶意用户检测方法[J].网络与信息安全学报,2024,10(4): 143-158.  
YANG H F, DU S Y, WEI G, et al. Malicious user detection in Momo application based on spatio-temporal behavior and static attribute analysis[J]. Chinese Journal of Network and Information Security, 2024, 10(4): 143-158.
- [5] LI H J, ZHU H J, DU S G, et al. Privacy leakage of location sharing in mobile social networks: attacks and defense[J]. IEEE Transactions on Dependable and Secure Computing, 2018, 15(4): 646-660.
- [6] ZHENG X, HAN J L, SUN A X. A survey of location prediction on twitter[J]. IEEE Transactions on Knowledge and Data Engineering, 2018, 30(9): 1652-1671.
- [7] POULSTON A, STEVENSON M, BONTCHEVA K. Hyperlocal home location identification of twitter profiles[C]//Proceedings of the 28th ACM Conference on Hypertext and Social Media. New York: ACM, 2017: 45-54.
- [8] OZDIKIS O. Locality-adapted kernel densities of term cooccurrences for location prediction of tweets[J]. Information Processing and Management, 2019, 56(4): 1280-1299.
- [9] CAILLAUT G, AUCLAIR S, GRACIANNE C, et al. Entity linking for real-time geolocation of natural disasters from social network posts[J]. PLOS ONE, 2024, 19(10): e0307254.
- [10] LUO X Y, QIAO Y Q, LI C L, et al. An overview of microblog user geolocation methods[J]. Information Processing & Management, 2020, 57(6): 1-18.
- [11] LIU Y M, LUO X Y, TAO Z Y, et al. UGCC: social media user geolocation via cyclic coupling[J]. IEEE Transactions on Big Data, 2023, 9(4): 1128-1141.
- [12] FEI G L, LIU Y, HU G M, et al. Online social network user home location inference based on heterogeneous networks[J]. IEEE Transactions on Dependable and Secure Computing, 2024, 21(6): 5509-5525.
- [13] JIANG H B, LI J, ZHAO P, et al. Location privacy-preserving mechanisms in location-based services: a comprehensive survey[J]. ACM Computing Surveys, 2022, 54(1): 1-36.
- [14] 王荣荣,薛旻辉,李祥学,等.基于位置社交网络的高效定位算法[J].华东师范大学学报(自然科学版),2016,2: 62-72.  
WANG R R, XUE M H, LI X X, et al. An effective localization attack in location based social network[J]. Journal of East China Normal University(Natural Science), 2016, 2: 62-72.
- [15] XUE M H, LIU Y, ROSS K W, et al. I know where you are: thwarting privacy protection in location-based social discovery services[C]//2015 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). Piscataway: IEEE Press, 2015: 179-184.
- [16] PENG J W, MENG Y, XUE M H, et al. Attacks and defenses in location-based social networks: a heuristic number theory approach [C]//2015 International Symposium on Security and Privacy in Social Networks and Big Data (SocialSec). Piscataway: IEEE Press, 2015: 64-71.
- [17] 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 & Mobile Devices. New York: ACM, 2014: 75-80.
- [18] LI M Y, ZHU H J, GAO Z Y, et al. All your location are belong to us: breaking mobile social networks for automated user location tracking[C]//Proceedings of the 15th ACM International Symposium on Mobile Ad-Hoc Networking and Computing. New York: ACM, 2014: 43-52.
- [19] WANG G, WANG B L, WANG T Y, et al. Whispers in the dark: analysis of an anonymous social network[C]//Proceedings of the 2014 Conference on Internet Measurement Conference. New York: ACM, 2014: 137-150.
- [20] WANG J L, CHENG H N, XUE M H, et al. Revisiting localization attacks in mobile app people-nearby services[C]//Proceedings of Security, Privacy, and Anonymity in Computation, Communication, and Storage-10th International Conference (SpaCCS 2017). Berlin: Springer, 2017: 17-30.
- [21] SHI W Q, LUO X Y, GUO J D, et al. Where are Wechat users: a

geolocation method based on user missequence state analysis[J]. IEEE Transactions on Computational Social Systems, 2021, 8(2): 319-331.

- [22] 时文旗. LBSD 类社交平台用户定位技术研究[D]. 郑州: 信息工程大学, 2022.

SHI W Q. Research on user geolocation technology of LBSD social platforms[D]. Zhengzhou: Information Engineering University, 2022.

- [23] LIU J, ZHANG Y, ZHAO F. Robust distributed node localization with error management[C]//Proceedings of the 7th ACM International Symposium on Mobile Ad Hoc Networking and Computing. New York: ACM, 2006: 250-261.

- [24] SHI W Q, LUO X Y, GUO J S, et al. An efficient geolocation method for malicious LBSD users based on dynamic adjustment of probes[J]. Wireless Communications and Mobile Computing, 2022, 2022(1): 1-15.

- [25] SADOON B, AL-BAYARI O. Location based services using geographical information systems[J]. Computer Communications, 2007, 30(16): 3154-3160.

- [26] ZANDBERGEN P A. Accuracy of iPhone locations: a comparison of assisted GPS, Wi-Fi and cellular positioning[J]. Transactions in GIS, 2009, 13(s1): 5-25.

#### [作者简介]



张冰 (2000-), 女, 河南南阳人, 郑州大学硕士生, 主要研究方向为移动应用用户位置安全、网络空间测绘等。



王跻霖 (1993-), 女, 黑龙江齐齐哈尔人, 硕士, 中国航天系统科学与工程研究院工程师, 主要研究方向为网络空间安全。



时文旗 (1993-), 男, 河南周口人, 博士, 河南省网络空间态势感知重点实验室讲师, 主要研究方向为移动应用用户位置安全、网络空间测绘。



刘瑞婷 (1982-), 女, 河南郑州人, 河南省网络空间态势感知重点实验室副教授, 主要研究方向为网络资源测绘、网络数据分析。



罗向阳 (1978-), 男, 湖北荆门人, 博士, 河南省网络空间态势感知重点实验室教授、博士生导师, 主要研究方向为网络空间资源测绘、多媒体安全。