软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software,2018,29(7):1852−1862 [doi: 10.13328/j.cnki.jos.005355]

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

基于用户分布感知的移动 P2P 快速位置匿名算法*

许明艳[1,2], 赵 华[1,2], 季新生[1,2], 申 涓[1]

[1](国家数字交换系统工程技术研究中心,河南 郑州 450002)

[2](移动互联网安全技术国家工程实验室,北京 100876)

通讯作者: 许明艳, E-mail: xumingyan886@126.com

摘 要: 针对移动点对点(P2P)结构下位置隐私保护匿名区形成存在着通信开销大、匿名效率低以及成功率低等问题,提出了一种移动 P2P 结构下用户分布感知方案,用户在邻域内共享邻域加权密度参数,获取邻域用户实时分布信息,根据用户分布特征为用户推荐隐私参数及候选用户查找半径,帮助用户快速形成匿名区.仿真结果表明,该算法通信开销小,在满足移动 P2P 网络移动设备节能需求的同时,匿名区生成时间平均在 500ms 以下,平均成功率达到 92%以上.

关键词: 位置隐私;移动 P2P 网络;*k*-匿名;用户分布感知;隐私参数推荐

中图法分类号: TP309

中文引用格式: 许明艳,赵华,季新生,申涓.基于用户分布感知的移动 P2P 快速位置匿名算法.软件学报,2018,29(7):1852−1862. http://www.jos.org.cn/1000-9825/5355.htm

英文引用格式: Xu MY, Zhao H, Ji XS, Shen J. Distribution-Perceptive-Based spatial cloaking algorithm for location privacy in mobile peer-to-peer enviroments. Ruan Jian Xue Bao/Journal of Software, 2018,29(7):1852−1862 (in Chinese). http://www.jos.org.cn/1000-9825/5355.htm

Distribution-Perceptive-Based Spatial Cloaking Algorithm for Location Privacy in Mobile Peer-to-Peer Enviroments

XU Ming-Yan[1,2], ZHAO Hua[1,2], JI Xin-Sheng[1,2], SHEN Juan[1]

[1](China National Digital Switching System Engineering and Technological R&D Center, Zhengzhou 450002, China)

[2](National Engineering Laboratory for Mobile Network Security, Beijing 100876, China)

Abstract: The mobile peer-to-peer environment is easier to implement in location privacy preserving research. The mobile users cooperate through P2P multi-hop routing to blur their accurate locations into a spatial cloaking region, but most existing spatial cloaking algorithm cannot work well because of the high communication overhead, time consumption and the lower success rate. This paper proposes an algorithm that can recommend user's privacy requirements by collecting users' weighed density information in their neighborhood, and therefore help mobile users to find enough collaborative users quickly. The approach shows great anonymization success rate by 92% through extensive simulation experiments for a range of P2P environment scenarios. It achieves lower communication cost and less than 500ms of searching time at the same time.

Key words: location privacy; mobile peer-to-peer network; *k*-anonymity; user-distribution-perceptive; recommendation of privacy parameters

* 基金项目: 国家自然科学基金(61521003); Research on the Fundamental Theories for Cyber-Space Mimic Defense

Foundation item: National Natural Science Foundation (61521003); Research on the Fundamental Theories for Cyber-Space Mimic Defense

本文由"面向隐私保护的新型技术与密码算法"专题特约编辑黄欣沂教授推荐.

收稿时间: 2017-05-18; 修改时间: 2017-07-13; 采用时间: 2017-08-22; jos 在线出版时间: 2017-10-17

CNKI 网络优先出版: 2017-10-17 13:37:58, http://kns.cnki.net/kcms/detail/11.2560.TP.20171017.1337.002.html

随着智能终端计算能力的不断提高和各种定位技术的发展,基于位置的服务(location based service,简称 LBS)已成为热点移动应用.LBS 与用户提出请求的位置有关,为了获得优质的位置服务,人们必须将自己精确的位置提交给应用服务器,同时提出查询请求.位置信息作为重要的个人隐私,暴露给网络及不信任的第三方,有可能导致严重的隐私泄露问题.如何在为用户提供 LBS 服务的同时保护位置及个人隐私安全是当前的一个研究热点.

LBS 位置隐私保护领域已有的研究成果是,通过对用户的位置信息以假位置、假名、模糊化等方式进行扰动后发送给服务器,切断用户标识与精确位置之间的关系.2003 年,Gruteser 等人首先将在关系数据库中的 k-匿名技术引入到 LBS 位置隐私保护技术中:当一个移动用户的位置扩展到无法与其他 k–1 个用户区分开来时,用户的标识也与这些用户无法区分,则称此位置满足位置 k-匿名[1].具体实施时,有两种主流的位置隐私保护的系统结构:集中式服务器结构与移动 P2P(peer to peer)结构.集中式服务器结构在用户与 LBS 服务器之间增加一个可信的第三方服务器(trusted third party,简称 TTP),负责收集用户的精确位置和 LBS 请求,根据用户的隐私需求,挑选邻近区域的 k 个用户组成匿名区[2−5].而移动 P2P 结构则通过用户间相互协作的方式组成协作用户组,形成 k-匿名空间[6−8].两种结构中,移动 P2P 结构因无需 TTP 支持,不存在单点失效的风险[9],是近年来的研究热点.移动 P2P 结构下 k-匿名与各种隐私保护加强方案的有效结合,如匿名区变换调整策略[10,11]、用户间信任方案[12,13]、用户位置扰动方案[14,15]和加密方案[16,17]等,显著提高了移动 P2P 结构下 k-匿名位置隐私保护方案的安全强度和抵御各种攻击的能力.

移动 P2P 结构下 k-匿名算法的核心问题是移动 P2P 用户如何以小的通信开销快速找到协作用户形成匿名区.Chow 等人提出的移动 P2P 环境下的 CloakP2P 算法[6],采用“逐跳洪泛”方式查找协作用户形成 k-匿名区.cloakP2P 算法简单,直到现在仍然是众多移动 P2P 网络下位置隐私研究的基础,其缺点是用户查找速度慢,成功率低,尤其是在用户量稀少的情况下,成功率仅为 60%.后续的研究大多通过获取移动 P2P 结构下用户分布信息来提高匿名成功率.如 Chow 的 Pro-Active 算法[6]与车延辙的 DA 算法[7],通过在匿名前由移动用户周期性地与周围用户交换位置信息,获取周围用户的分布信息,以提高匿名空间的生成效率;但是为了维护本地的候选位置信息,用户需要很大的通信开销.Chow 的 IS 算法[8]通过引入位置缓存机制,利用历史分布信息减小通信开销,但缓存的位置信息并不能保证候选用户位置的新鲜和有效,从而会导致匿名区域无效,降低了匿名区的成功率和安全性.Aniket 的 CAP 方案[18]提出根据路网密度与用户密度呈线性关系的原则对不同地区进行划分,推断应用场景及用户密度,而用户分布随时间及社交活动等行为呈现不同特征,路网密度不能实时反映用户分布信息.现有方案都未能实现以较低的通信开销获取移动 P2P 结构下用户的实时分布信息,从而提高位置 k-匿名形成的效率.

针对以上需求,本文提出了一种基于用户分布感知的移动 P2P 快速位置匿名算法(a distribution-perceptive-based spatial cloaking algorithm for location privacy in mobile P2P enviorments,简称 DPB).DPB 算法由用户邻域分布感知、匿名参数推荐以及匿名区快速查找这 3 部分组成.

(1) 首先设计了一种用户邻域分布感知方案,通过移动用户在邻居间更新和发布用户邻域加权密度参数,获取用户邻域范围内用户点的实时分布信息(见第 2 节);

(2) 然后,根据邻域用户的实时分布疏密情况,给出一种隐私参数推荐策略,为用户推荐适合应用上下文的隐私参数 k、匿名区大小及查找范围(见第 3 节);

(3) 最后,基于推荐参数改进传统匿名区查找方案,依据推荐的匿名区查找范围,快速找到协作用户形成匿名区(见第 4 节).

第 5 节的仿真结果表明,相比经典算法,DPB 算法在减小部分通信开销的基础上,提高了匿名区生成效率和成功率.本文第 2 节~第 4 节分 3 部分详细阐述 DPB 算法.第 5 节通过仿真对 DPB 算法进行评估与性能分析.在第 6 节对本文工作进行总结与展望.

1 系统结构

目前,基于移动 P2P 结构的 LBS 位置匿名算法普遍基于图 1 所示的系统模型.每个移动终端都具有定位功能,支持无线点对点协议及无线互联网两种通信方式.移动 P2P 结构下位置 *k*-匿名算法的基本工作模式为:当移动用户有位置匿名需求时(称为请求用户),通过无线点对点协议与其他移动用户通信,找到 *k*–1 个愿意参与匿名组的用户(称为协作用户或候选用户),形成共同的匿名区;然后,请求用户按某种策略挑选组中一个移动用户作为转发代理和 LBS 服务提供商进行通信.

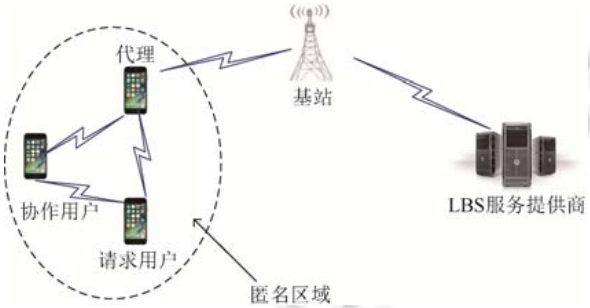


Fig.1 System model for location privacy preserving in mobile peer-to-peer enviorments

图 1 移动 P2P 位置隐私保护系统模型

本文的工作目标为:当请求用户有位置匿名需求时,以较小的通信代价实现协作用户的快速查找与匿名区的形成.在此,基于以上通用的移动 P2P 结构的系统模型,设定系统的工作前提为:(1) 系统内每个移动用户都可运行 DPB 算法参与位置隐私匿名;(2) 系统内用户之间相互信任,他们不会通过搜集协作用户的位置信息来对其他用户或系统进行攻击;(3) LBS 服务器可以完成基于匿名的查询服务.

如图 2 所示,DPB 算法由用户邻域分布感知、匿名区参数推荐和匿名区快速查找这 3 部分组成,其中,用户邻域分布感知是整个 DPB 算法的基础,它给出邻域用户密度参数的求解步骤,以此为参考,在匿名算法中为用户推荐隐私匿名参数 *k* 及匿名查找范围,并引出匿名区快速查找算法.

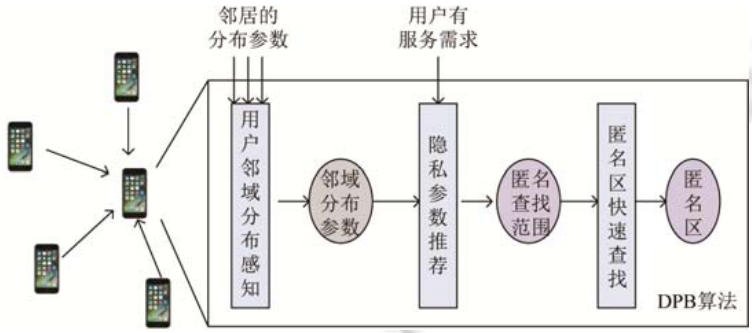


Fig.2 Algorithm structure of DPB

图 2 DPB 算法组成

2 用户邻域分布感知

2.1 预备知识

应用场景的用户空间分布特征一般以用户密度*d*描述,指的是“单位面积上的移动用户数”,表示用户分布的

疏密.根据极大似然估计基本原理,对于一个用户分布空间 D,以单位面积内的用户数 D_i 作为样本空间,若$\{D_1, D_2,...,D_n\}$是取自总体 D 的样本空间集,则整个用户空间的平均用户密度期望值 d^*为

$$d^* = \overline{D} = \frac{1}{n}\sum_{i=1}^{n} D_i \tag{1}$$

对于移动P2P结构下寻找匿名区的请求用户u,其邻域范围内的用户分布特征决定了匿名区的大小与形状,因此,相比全局的用户分布密度,用户 u 更关心以自己为中心的一定邻域范围内的用户分布密度.于是,本文在极大似然估计的基础上,提出一种用户根据距离远近进行加权的用户分布感知方案.为此,我们首先给出相关定义.

在移动 P2P 结构下,如果对于一个应用空间 D,任意两个用户通过单跳或多跳通信可达,则称此应用空间为连通空间 C,其中,用户间的跳数 h 表示用户间距离.用户单跳通信所覆盖的范围定义为单位面积,称为用户的直接感知视野.用户 h 邻域则指用户通过 h 跳通信可达的区域范围.

定义 1. C 为连通的网络空间,用户 $u \in C$,u 的 h 邻域定义为

$$D(u,h) = \{y \mid y \in C, \|u - y\| \leqslant h\} \tag{2}$$

在 h=1 时,$D(u,1)$所覆盖区域范围即为单位面积,其中的用户数称为用户 u 的局部样本 D_u.

2.2 邻域分布感知过程

定义 2. C 为连通的网络空间,对用户 $u \in C$,有邻域加权密度 d_u,表示以用户 u 为中心的 h 邻域内单位面积上的近似平均用户数($h \geqslant 1$),并定时地向 $D(u,1)$内的用户发布自己的 d_u,d_u 初始值为用户 u 的局部样本 D_u.假设用户 u 收集到直接感知视野 $D(u,1)$内用户集$\{...u_{i-1},u_i,u_{i+1}...\}$的有效密度参数 d_i 个数为 n_u,则用户 u 计算并更新 d_u 为

$$d_u = \frac{1}{n_u + 1}\left(D_u + \sum_{i=1}^{n_u} d_i\right) \tag{3}$$

为了能够实时地感知邻域 $D(u,1)$内的用户分布动态,每个用户始终将当前的局部样本值作为计算的基本项.同理,用户 $u_i \in \{...u_{i-1},u_i,u_{i+1}...\}$,其 d_i 根据其收集到的 $D(u_i,1)$内用户的邻域用户集合$\{...u_{j-1},u_j,u_{j+1}...\}$的 d_j 值以及 u_i 当前的 D_i,可以得到:

$$\begin{cases} d_i = \dfrac{1}{n_i + 1}\left(D_i + \displaystyle\sum_{j=1}^{n_i} d_j\right) \\ d_j = ... \end{cases} \tag{4}$$

依此类推,并把式(3)、式(4)中的 d_j,d_i 等逐级代入,并对每个用户的局部样本进行同类项合并,得到:

$$d_u = \sum_{m=1}^{n} \rho_m D_m,\ \ \rho_m = \sum_{M}\prod_{h=1}^{H}\frac{1}{n_h} \tag{5}$$

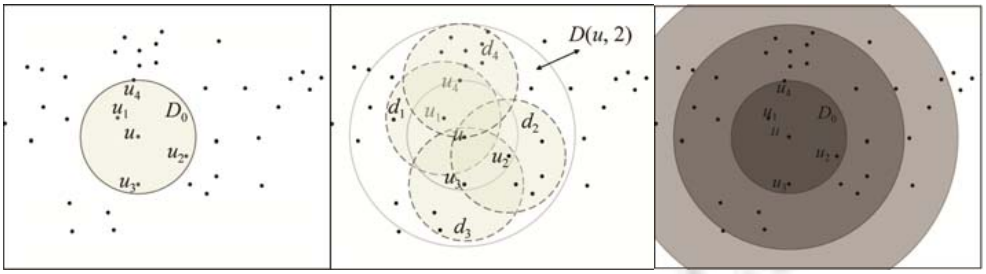
其中,n 为用户 u 在 $D(u,h)$内的总用户数,D_m 为每个用户 u_m 的局部样本,ρ_m 表示一个 D_m 通过 M 种途径到达用户 u,每种途径经过 H 次迭代后形成的系数,我们称其为每个用户 u_m 的分布影响因子.由式(5)可知,u_m 离用户 u 越远,其 D_m 到达用户 u 经过的迭代次数越多,分布影响因子 ρ_m 越小,ρ_m 与用户 u_m 到用户 u 的距离成反比,用户自己的 ρ_u 最大.

图 3 所示为用户 u 的邻域用户分布感知过程.图 3(a)表示了 u 的初始邻域加权密度 d_u 的感知范围 $D(u,1)$,其内有 5 个用户:u_1 到 u_5;图 3(b)表示进行了一次迭代更新后,借助 u_1 到 u_5 的直接感知视野,用户 u 的感知范围扩展到 $D(u,2)$;图 3(c)表示了用户 u 的 d_u 值随距离的增加其影响因子的变化情况,颜色越浅影响因子越小.

邻域用户分布感知过程的实质是每个用户通过与直接感知视野内的用户交换邻域加权用户密度参数,获得更远处用户的局部样本空间,逐步扩大感知视野的过程.对请求用户 u 而言,离得越远的用户,参与匿名的可能性越小,因此,在计算邻域用户密度参数时,越远的用户提供的样本空间所占比重越小,影响因子越小.

邻域用户分布感知过程中需要考虑的另一个重要因素是 d_u 更新与发布的时机和频率.周期性地更新和分享 d_u 能够保证用户实时感知周围用户分布的变动,但过多的消息发送会造成较大的通信开销.当网络空间内用

户分布相对稳定时,在用户的 d_u 到达稳态后仍不停分享 d_u 是对通信资源的浪费.因此,本方案中每个移动用户在本地维护直接感知视野内邻居用户的“邻域加权密度表”,存储当前邻居用户的最新邻域加权密度参数,同时设置检测周期,邻域分布参数的更新与发布采用动态触发机制,随着参数与用户的变化而触发的机制.



(a) u 的样本空间 (b) 1 跳迭代后 u 的感知视野 (c) 2 跳迭代后 u 的感知视野

Fig.3 Neiborhood-Distribution perception process

图 3 邻域分布感知过程

2.3 算法描述

邻域用户分布感知过程的伪代码见算法 1.用户 u 完成初始化后,在每个检测周期对 d_u 发布进行判断,当检测到自己的 d_u 发生变动或直接感知视野内有新的邻居用户加入时,向邻域用户发布“m_share”消息,其中携带自己的邻域加权密度参数 d_u(第 4 行~第 6 行);同时收集“m_share”消息,检测当前直接感知视野内的邻居用户数;读取每条消息的内容,更新存储的参数列表(第 9 行~第 12 行);当收到的参数值或邻居用户集合发生变化时,则通过公式(3)计算并更新本地的 d_u(第 13 行~第 15 行).

邻域用户分布感知的过程中,用户用来发布邻域加权密度参数的消息为“m_share=$\{d_i\}$”.

算法 1. 用户邻域分布感知.

输入:邻居的“m_share”消息集合;

输出:d_u 及发出的“m_share”消息.

1. 设置 d_u 初始值 D_u,扫描周期 t;
2. 初始化邻居参数列表为空;
3. WHILE (t 周期)
4. IF (d_u 变化或有新邻居)
5. 产生并发送“m_share”消息;
6. END IF
7. 收到的“m_share”集合为 M';
8. $D_u \leftarrow D(u,1)$;
9. WHILE (M'中每一个“m_share”消息)
10. 读取消息中 u_i 的 d_i;
11. 更新列表中相应邻居 u_i 的 d_i;
12. END WHILE
13. IF (d_i 有变化或参数列表有修改)
14. $d_u \leftarrow \left(\sum d_i + D_u\right) \Big/ |D_u|$
15. END IF
16. END WHILE

用户邻域分布密度感知算法的消息分享和收集仅限于直接感知视野内用户,消息简短,内容为用户当前邻域加权参数,用户之间交互消息量随用户量的增多而呈正比增加;邻域分布密度感知算法的时间复杂度为

$O(n)$,n 为用户数.

3 匿名参数推荐

k 匿名的评价标准决定了 k 值越大隐私保护安全性越高,但事实上,由于网络空间大小以及用户分布等的限制,无限地扩大搜索半径并不意味着能够找到更多的用户,尤其是在用户分布稀疏时,较大的 k 值往往导致匿名失败.Ahamed 在文献[19]中提到隐私 k 与应用场景的适配是匿名成功的关键因素,并且极大地提高了隐私保护的效果.基于上节的用户分布感知算法,用户可以根据邻域加权密度 d_u 判定邻域内用户分布的疏密程度,预估满足隐私需求 k 的匿名区大小及查找半径.

图 4 表示邻域加权密度 d_u 分别为 3、5、7 和 10 时,搜索范围与用户数之间的关系,曲线中每一点表示增加 1 跳的搜索范围后找到的用户数.由图 4 可知,每增加 1 跳可以找到更多的用户.设每增加 1 跳时找到的用户数为 N_h,而为扩大搜索范围付出的通信代价为 C_h,则算法每一跳的搜索增益为图中曲线在每一跳处的斜率 N_h/C_h.随着搜索范围的扩大,搜索增益不断减小,在搜索范围平均超过 4 跳后,搜索增益小于 1,每增加一个用户需要巨大的通信开销.因此,对于给定的 d_u,为保证最小的通信代价以及最大的匿名成功率,本文建议隐私参数 $k \leqslant 4 \times d_u$.

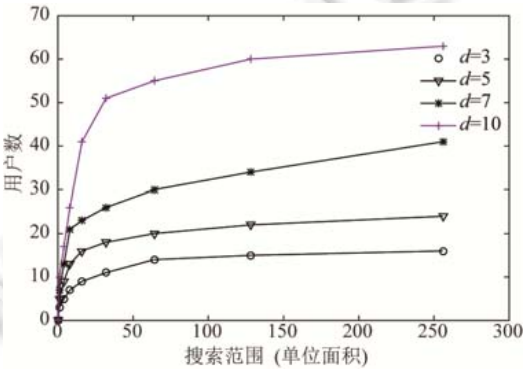


Fig.4 Relationship between searching region and cooperated peers

图 4 用户数与搜索范围的关系

在实际应用中,由于网络空间 C 的形状并不规则,用户可以分布在 C 的任何位置.理想情况下,用户呈随机均匀分布时,请求用户平均需要 $\sqrt{k/d_u}$ 跳数即可找到足够的候选用户;而在用户呈非均匀分布时,最坏的情况为用户分布呈直线特征,此时用户需要较大的跳数(k/d_u)才能找到足够的候选用户.结合两种情况,推荐的匿名空间大小及查找范围为

$$\begin{cases} h_{initial} = \alpha\sqrt{k/d_u} + \beta(k/d_u) \\ \alpha + \beta = 1, (\alpha, \beta > 0) \qquad (k \leqslant 4 \times d_u) \\ h_{end} = \min(8, k/d_u) \end{cases} \tag{6}$$

其中,α、β为应用场景的适配因子.当用户位于路网中时,用户分布多呈线型,此时选择$\beta>\alpha$;而当用户位于图书馆、商场等位置时,用户在空间中的分布更加随机、均匀,$\alpha>\beta$是更合理的选择.一般情况下,根据用户的移动速度能够在某种程度上基本判定用户的应用场景.如果用户移动速度较快,则位于路网中的可能性更大,而当用户移动速度为静止或步行时,随机均匀分布的可能性更大.对应用场景更精确的判断有助于确定α与β更加合理的取值,我们将作为下一阶段研究的目标之一.

4 移动 P2P 快速匿名算法

本算法基于 CloakP2P[6]算法“逐跳洪泛”的搜索方式对匿名区查找方法进行改进.由于对用户分布一无所

知,CloakP2P 算法从 1 跳范围内的用户开始查找,逐步扩大搜索范围.在本方案中,匿名区查找半径的推荐明确了用户的查找范围,匿名区查找直接从跳数 $h_{initial}$ 处开始,返回结果为用户邻域 $h_{initial}$ 跳内的协作用户;如果响应用户数 n 小于 k 且跳数小于 h_{end},则继续增加跳数;若返回用户数 n 已大于 k,则在用户中随机去掉 $n–k$ 个邻域加权密度参数较大的用户,使其余 k 个用户在匿名区中的分布尽量随机、均匀,按用户地理位置生成矩形匿名空间.

DPB 匿名区快速查找方案由用户请求和响应两部分组成,包括请求用户的“c_group_req”和响应用户的“c_group_rsp”两种消息类型.匿名区查找请求消息“c_group_req”,由请求用户发出,包括匿名区标识 cid、查找范围 h、隐私参数 k、匿名区范围 $A_{\min}$ 和 $A_{\max}$ 等:c_goup_req={cid,h,k,$A_{\min}$,$A_{\max}$}.匿名区查找响应消息“c_group_rsp”,由响应用户发出,包括用户接受的匿名区标识 cid 以及用户的位置坐标.其格式为 c_goup_req={cid,$loc(x,y)$}.

算法 2-1 和算法 2-2 分别描述了请求用户的匿名区查找过程和响应用户的回复及转发过程.其中,算法 2-1 中请求用户的匿名区查找过程分为匿名区查找和匿名区生成两步.

算法 2-1 步骤 1:协作用户查找确认.当请求用户 u_{req} 有匿名需求时,生成匿名区标识 cid,设置初始查找半径 $h_{initial}$ 以及匿名集 P;生成“c_group_req”消息,并将自己的隐私参数包含在消息中,向邻居用户发送(第 3 行~第 4 行);对每一个发送响应消息的用户,说明能够接受请求用户 u_{req} 的隐私要求并同意加入匿名集合(见算法 2-2),于是用户 u_{req} 将响应用户增加至匿名集合 P(第 6 行~第 8 行);如果匿名集用户数 n 小于隐私参数 k,而且查找半径小于 h_{end},则用户 u_{req} 将查找半径加 1,再次发送“c_group_req”消息(第 9 行,第 2 行),直到匿名集 P 大小满足隐私需求;如果对于某次请求收到的响应消息数为 0,说明查找范围已经到达网络边界,匿名失败(第 5 行).

算法 2-1 步骤 2:匿名区形成.如果用户数 $n>k$,则在用户中随机去掉($n–k$)个邻域分布密度较大的用户,使用户在匿名区中的分布尽量随机、均匀,其余 k 个用户生成矩形匿名空间.

算法 2-1. 匿名区快速查找(请求节点 u_{req}).

输入:无;

输出:成功(匿名区)或失败.

```
//步骤 1:匿名区查找
1.    生成 cid,h=h_initial,P←{u_req},n←|P|;
2.    WHILE (n<k)&&(h≤h_end)
3.        生成“c_group_req”并发送给邻居用户;
4.        收到的“c_group_rsp”消息为 M′;
5.        IF (收到消息个数为 0) THEN 转入步骤 2;
6.        WHILE (M′中每个“c_group_rsp”消息)
7.            响应用户进入集合 P,n←n+1;
8.        END WHILE
9.        h←h+1
10.   END WHILE
//步骤 2:匿名区形成
11.   IF (n≥k) THEN
12.       随机删除 n–k 个集合 P 中 d′值最大的用户;
13.       A←包含组内用户坐标的区域;
14.       RETURN A;
15.   ELSE RETURN −1;
16.   END IF
```

算法 2-2 描述了响应用户的行为:当响应用户收到“c_group_req”消息时,如果还未加入任何匿名组,而且请求用户的隐私需求在可接受范围内,则加入此匿名组,设置匿名组号并将自己的地理位置信息返回给请求用户(第 3 行~第 7 行);如果消息中跳数 $h>1$,表示查找过程还未结束,则响应用户将消息中跳数减 1,转发给邻域用户

(第 8 行~第 11 行).如果响应节点收到下一跳节点的“c_group_rsp”消息,则将其转发至上一跳节点(第 12 行~第 13 行).

算法 2-2. 匿名区快速查找方案(响应节点 u_{rsp}).

输入:收到的“c_group_req”或“c_group_rsp”消息;

输出:回复或转发消息.

```
1.    SWITCH (消息类型)
2.      CASE(“c_group_req”)
3.        读取消息中 cid,h,k 等参数
4.        IF (没有加入任何匿名组且可接受隐私要求)
5.          设置 cid←请求消息中的 cid;
6.          生成“c_group_rsp”消息并返回发送方;
7.        END IF
8.        IF (h>1) THEN
9.          h←h–1;
10.         转发“c_group_req”消息给下一跳用户;
11.       END IF
12.     CASE(“c_group_rsp”)
13.       转发“c_group_rsp”消息给上一跳用户;
14.   END SWITCH
```

匿名查找范围的推荐,加快了查找速度,提高了匿名成功率.在匿名区查找阶段,如果协作用户有不同隐私参数,当前大部分算法选择组内用户最大的 k 值作为整个匿名组的 k 值,导致匿名区过大,服务质量下降.本方案把用户之间隐私参数的协商和参与匿名的决定权交与响应用户,不会引起隐私参数的变动,保证了服务质量.

5 仿真与测试

本节对匿名区生成算法的性能进行仿真测试,根据 DPB 算法的适用环境,本文挑选同样适用于连通网络空间的两种算法,Chow 提出的 P2PCloak[6]以及浙江大学车延辙提出的双向主动探测 DA 算法[7],与 DPB 算法进行性能比较.作为经典算法,P2PCloak 算法随需触发,具有最小的通信成本,但是生成时间长、成功率低;DA 算法由于共享位置信息,生成时间短、成功率高,但是通信成本较高.

5.1 实验环境

算法采用 Java 实现,运行平台为 Intel i5 2.3GHz 处理器,8GB 内存的 Windows 10 系统.本文实验使用的数据集来源于 Thomas Brinkhoff[20]移动对象生成器,它建立于真实地图信息的基础上,以城市 Oldenburg 的交通路网(区域面积为 23.57km×29.92km)作为输入,生成模拟的移动用户数据.每个用户为一个线程,根据用户描述数据模拟请求与响应行为.仿真数据使用的参数值见表 1.

Table 1 Specification of experiments

表 1 实验参数

参数名称	默认值
移动用户数量	3000~9000
K 匿名要求	5~40
点对点传输距离	250m
消息处理时间	100ms
消息长度	64bytes

根据实验平台及开发环境的条件,测试移动用户数量在 3000~9000 之间,设置 3 000、5 000、7 000 和 9 000 这 4 种应用场景,用户随机发起匿名请求的概率为 10%,k 匿名需求在 5~40 间选取,并且假定移动设备的点对点

通信能力为250m,为便于计算通信量,假设移动用户处理一条消息的时间为100ms,每条消息长度为64字节.

5.2 用户分布感知算法性能分析

邻域用户分布感知方案是本文DPB算法的基础,它的性能决定了DPB算法的优劣.在用户分布相对稳定的情况下,邻域加权密度参数是否能够快速达到稳态说明了邻域参数计算的收敛性和有效性.

我们选择应用场景中用户相对稳定的时段,在邻域用户分布感知阶段跟踪邻域加权密度 d_u 的更新变化情况.通过仿真我们发现,用户参数 d_u 达到稳态的平均收敛速度与参数 d_u 的大小有关.图5所示为邻域用户密度参数 d_u 从初始值到稳态的过程中所需要的更新与发布的平均次数.参数 d_u 越小,到达稳态需要的次数越少,参数越大,到达稳态则越慢.但即使是较大的 d_u,在平均4次更新后即可达到稳态.

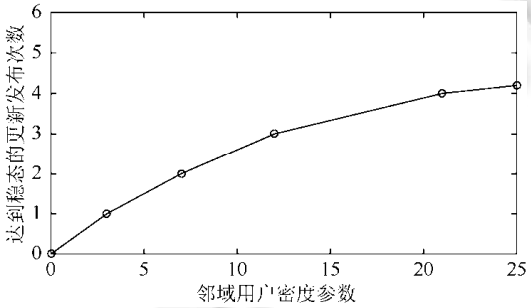


Fig.5 Convergent rate of neighborhood-distribution density

图5 邻域加权密度参数的收敛速度

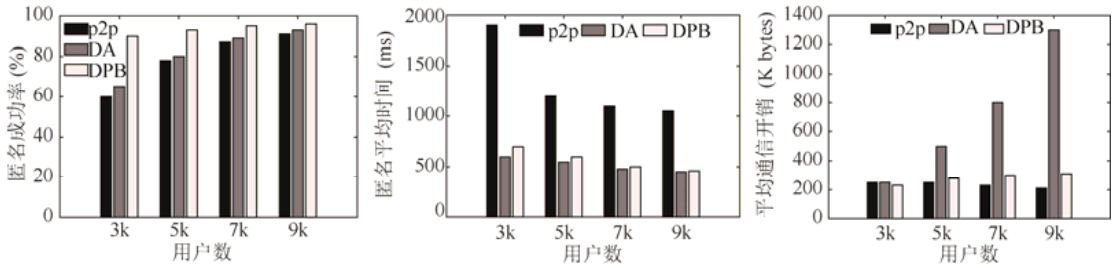
5.3 DPB算法性能分析

鉴于本应用场景为交通路网,因此在快速匿名算法中匿名查找范围的设置为α=0.4,β=0.6(见公式(6)).本节根据3个关键性能指标对算法进行评估.(1) 匿名空间生成成功率,指的是算法中发起请求的用户中成功申请到匿名空间的用户比例,成功的匿名空间是指真实的 k 个或多于 k 个用户共享一个匿名空间.(2) 平均生成时间,是指一个用户从发出匿名请求到匿名空间生成的平均时间.(3) 平均通信开销,是指用户从准备发起请求到生成匿名空间这段时间内,为了生成匿名空间直接或间接产生的通信量总和,包括用户主动发起给其他用户和从其他用户处接收到的各类用于联络以及请求响应的消息总开销.

图6(a)所示为不同应用场景下匿名成功率的比较情况.对于任何算法,随着用户数的增加,匿名成功率都呈上升趋势.DA算法根据用户交互得到的位置信息生成匿名空间,成功率比P2PCloak算法高;DPB与P2PCloak算法的匿名区生成机制相同,但隐私参数的推荐使DPB算法的成功率大幅度提高.尤其是在低用户分布的应用场景下,成功率提高25%,平均匿名成功率达到92%以上.实验中发现,DPB算法用户匿名失败的原因除网络分区的限制外,两个匿名区的同时发起使协作用户分流,导致两个请求用户的匿名都失败.

图6(b)表示各算法在不同应用场景下平均生成时间的性能比较.对于相同的 k 值,随着用户数的增加,平均生成时间逐渐降低.相比于P2PCloak算法,DPB算法推荐了匿名区查找半径,请求用户直接寻找 $h_{initial}$ 范围内的用户,查找速度提高,匿名平均生成时间比P2PCloak算法降低50%以上.而DA算法通过之前大量的位置信息交互,用户拥有邻域 n 跳内用户的位置,生成速度与本文的DPB算法相当,平均生成时间为500ms.

图6(c)表示各算法在不同应用场景下平均通信开销的比较.随着用户数的增加,P2PCloak算法查找协作用户需要的平均跳数减少,平均通信量呈下降趋势;对于DA算法,用户数的增加使之需要维护和广播的位置信息记录增加,因此平均通信量呈大幅上升趋势;DPB算法在邻域用户分布密度感知阶段随着触发的消息发布策略,大大降低了消息交互数量,平均通信量远远低于收集位置信息的DA算法,在用户数少的情况下甚至低于P2PCloak算法,随用户量的增加仅有小幅增长.



(a) 匿名成功率 (b) 平均匿名时间 (c) 平均通信开销

Fig.6 Algorithm performance evaluation against different metrics

图 6 匿名算法性能比较

6 总 结

在移动 P2P 结构下,本文提出一种基于用户邻域分布感知的快速位置匿名算法,通过邻域加权用户分布信息的获取为用户推荐匿名区大小和查找范围,从而实现匿名区的快速查找.与已有算法的模拟仿真测试比较表明,本算法在减少通信开销的基础上,提高了匿名空间生成效率,同时计算复杂度低,具有更好的可扩展性和实用性.未来的工作将集中于移动 P2P 网络下 *k*-匿名与密码体制的结合,加强 *k*-匿名隐私保护强度,开发基于 DPB 算法的位置隐私保护原型系统.

References:

[1] Gruteser M, Grunwald D. Anonymous usage of location-based services through spatial and temporal cloaking. In: Proc. of the 1st Int'l Conf. on Mobile Systems, Applications and Services. New York: ACM Press, 2003. 31–42.

[2] Gedik B, Liu L. Location privacy in mobile systems: A personalized anonymization model. In: Proc. of the 25th IEEE Int'l Conf. on Distributed Computing Systems. IEEE, 2005. 620–629.

[3] Bu GG, Liu L. A customizable *k*-anonymity model for protecting location privacy. In: Proc. of the ICDCS. 2004. 620–629.

[4] Mokbel MF, Chow CY, Aref WG. The new Casper: Query processing for location services without compromising privacy. In: Proc. of the 32nd Int'l Conf. on Very Large Data Bases. VLDB Endowment, 2006. 763–774.

[5] Duckham M, Kulik L. A formal model of obfuscation and negotiation for location privacy. In: Pervasive Computing. Berlin, Heidelberg: Springer-Verlag, 2005. 152–170.

[6] Chow CY, Mokbel MF, Liu X. A peer-to-peer spatial cloaking algorithm for anonymous location-based service. In: Proc. of the 14th Annual ACM Int'l Symp. on Advances in Geographic Information Systems. ACM, 2006. 171–178.

[7] Che Y, Yang Q, Hong X. A dual-active spatial cloaking algorithm for location privacy preserving in mobile peer-to-peer networks. In: Proc. of the 2012 IEEE Wireless Communications and Networking Conf. (WCNC). IEEE, 2012. 2098–2102.

[8] Chow CY, Mokbel MF, Liu X. Spatial cloaking for anonymous location-based services in mobile peer-to-peer environments. GeoInformatica, 2011,15(2):351–380.

[9] Mokbel MF, Chow CY. Challenges in preserving location privacy in peer-to-peer environments. In: Proc. of the 7th Int'l Conf. on Web-Age Information Management Workshops, WAIM 2006. IEEE, 2006. 1.

[10] Huang Y, Huo Z, Meng X. CoPrivacy: A collaborative location privacy-preserving method without cloaking. Chinese Journal of Computers, 2011,34(10):1976–1985 (in Chinese with English abstract).

[11] Zhang C, Huang Y. Cloaking locations for anonymous location based services: A hybrid approach. GeoInformatica, 2009,13(2): 159–182.

[12] Yang N, Cao Y, Liu Q, *et al*. A novel personalized TTP-free location privacy preserving method. Int'l Journal of Security and Its Applications, 2014,8(2):387–398.

[13] Solanas A, Martinez-Balleste A. A TTP-free protocol for location privacy in location-based services. Computer Communications, 2008,31(6):1181–1191.

[14] Hashem T, Kulik L. “Don’t trust anyone”: Privacy protection for location-based services. Pervasive and Mobile Computing, 2011, 7(1):44–59.

[15] Gao S, Ma J, Yao Q, *et al*. Towards cooperation location privacy-preserving group nearest neighbor queries in LBS. Journal on Communication, 2015,3:142–150 (in Chinese with English abstract).

[16] Ghaffari M, Ghadiri N, Manshaei MH, *et al*. P4QS: A peer to peer privacy preserving query service for location-based mobile applications. IEEE Trans. on Vehicular Technology, 2016,PP(99):1.

[17] Dargahi T, Ambrosin M, Conti M, *et al*. ABAKA: A novel attribute-based k-anonymous collaborative solution for LBSs. Computer Communications, 2016,85:1–13.

[18] Pingley A, Yu W, Zhang N, *et al*. CAP: A context-aware privacy protection system for location-based services. In: Proc. of the IEEE Int’l Conf. on Distributed Computing Systems. 2009. 49–57.

[19] Ahamed SI, Haque MM, Hasan CS. A novel location privacy framework without trusted third party based on location anonymity prediction. ACM SIGAPP Applied Computing Review, 2012,12(1):24–34.

[20] Brinkhoff T. A framework for generating network-based moving objects. GeoInformatica, 2002,6(2):153–180.

附中文参考文献:

[10] 黄毅,霍峥,孟小峰.CoPrivacy:一种用户协作无匿名区域的位置隐私保护方法.计算机学报,2011,34(10):1976–1985.

[15] 高胜,马建峰,姚青松,孙聪.LBS 中面向协同位置隐私保护的群组最近邻查询.通信学报,2015,3:142–150.

许明艳(1974－),女,河南新乡人,副研究员,主要研究领域为移动通信,移动互联网安全.

季新生(1968－),男,教授,博士生导师,主要研究领域为电信网安全,移动通信,移动互联网安全.

赵华(1979－),女,讲师,主要研究领域为移动通信,移动互联网安全.

申涓(1977－),女,讲师,主要研究领域为网络体系结构,移动互联网安全.