基于博弈论的用户相互协作的位置隐私保护方法

陈玉凤 刘学军 李 斌
(南京工业大学电子与信息学院 南京 211816)

摘 要 位置隐私保护正在受到越来越多人的关注与研究,目前基于用户相互合作的无中心服务器的位置隐私保护成为当前研究的重点。为了在不可信环境下更好地保护用户位置隐私,从技术上提出了一种基于博弈分析思想的用户协作的位置隐私保护方法 Privacy_l,此方法通过用户协作形成匿名组,以匿名组的密度中心作为锚点代替真实位置发起查询;通过安全求和来计算锚点,解决在现实不可信环境下不诚信合作的问题;同时根据用户的不同位置隐私需求,通过设置不同的隐私保护参数水平,达到不同的匿名保护效果,并且采用改进的增量查询方法提高近邻查询效率。仿真实验表明,此方法具有较好的性能,能够更好应用于现实环境。

关键词 博弈,用户协作,位置隐私,隐私保护
中图法分类号 TP311 **文献标识码** A

Collaborative Position Privacy Protection Method Based on Game Theory

CHEN Yu-feng LIU Xue-jun LI Bin
(College of Electrioncs and Information, Nanjing University of Science and Technology, Nanjing 211816, China)

Abstract Location privacy protection is arising people's more attention and research. Currently location privacy-preserving based on users of mutual cooperation without a central server is now the focus of the study. In order to better protect the privacy of the user's location in real untrusted environment, this paper provided a location privacy protection method Privacy_l which is based on the idea of user collaboration game. Anonymous group is formed by users collaborate, density center of anonymous group as the anchor instead of the true position to initiate the query. To calculate the anchor through secure sum, this method could eliminate the situations that do not cooperate in good faith in real implausible circumstances. Meanwhile, according to the users' different location-privacy requirements, different anonymous protective effect is achieved by setting different levels of privacy protection parameters. In addition improved incremental query method is used to improve the efficiency of nearest neighbor queries. Simulation results show that this method has better performance, and can be applied to reality better.

Keywords Game, Collaboration, Location privacy, Privacy-protection

1 引言

位置服务(LBS, Location Based Services)伴随 GPS 和无线上网技术的发展,需求呈大幅度增长趋势。但是用户在使用这些服务的同时,可能面临着隐私泄露的威胁。恶意的位置服务提供商或其他针对位置服务器的攻击者根据用户位置和查询内容鉴别出用户的身份,进而获得用户的隐私信息[1,2]。

为了解决位置服务中的隐私问题,文献[3]最早提出了位置 k-匿名模型。它的基本思想是在发布用户位置的时候,用一个覆盖其他 $k-1$ 个用户的匿名区域代替用户的真实位置,从而使位置服务提供商无法从 k 个用户中鉴别出某一个用户的身份。目前大多数基于 k-匿名模型的研究都采用基于中心服务器的结构[4-8]。然而采用这种方法存在着很多的不足,现在越来越多的研究采用无中心服务器结构的隐私保护方法[9-13]。文献[13]提出了 SpaceTwist 方法,即用户随机选取自己真实位置附近的点作为锚点,然后使用该锚点代替自己的真实位置向位置服务提供商发起增量近邻查询,再根据返回的结果和用户真实位置进行计算,得到精确的查询结果。然而 SpaceTwist 方法无法达到 k-匿名的效果,并且容易使攻击者通过分析用户查询,将用户限定在一个区域中。如果该区域只有一个用户发起查询,攻击者就有可能根据查询内容鉴别出用户,进而获得用户隐私。文献[1]提出了一种基于用户协作的隐私保护方法 Coprivacy,即用户之间通过相互协作,不需要中心服务器,并通过单跳和多跳协议形成匿名组,不需生成匿名区域就能达到 k-匿名的效果。组内用户使用该组形成区域的密度中心作为锚点,并使用该锚点代替自己的真实位置,采用增量近邻查询返回结果,进而根据用户真实位置得到精确的近邻查询结果。但是 Coprivacy 方法是基于所有协作用户都是可信的,无法解决协作用户不可信的情况。

到稿日期:2012-12-23 返修日期:2013-04-04 本文受国家自然科学基金(61073197),江苏省科技支撑计划(SBE201077457)资助。
陈玉凤(1988—),女,硕士生,主要研究方向为数据挖掘、隐私保护,E-mail:chen_yufeiye@163.com;刘学军(1971—),男,博士,副教授,主要研究方向为数据库、数据挖掘、传感器网络、隐私保护等;李 斌(1979—),男,硕士,讲师,主要研究方向为传感器网络、智能信息处理。

本文结合用户相互合作的方式与无中心服务器结构的优点，针对协作用户存在不可信的情况，提出了一种基于博弈论思想的用户相互合作的位置隐私保护方法。该方法主要是运用博弈论的思想进行研究，分析了用户之间选择什么样的合作方式来保护自己的某些隐私，以及他们是如何应对不可信环境下用户之间合作的。本文借鉴了文献[2]中的安全求和算法，采用多次求和后比较结果的方法来防止怀有恶意的站点提供不准确数据，以达到保护用户位置隐私的目的。

本文第2节主要介绍了相关研究工作；第3节介绍博弈论基础；第4节提出了一种位置隐私保护方法；第5节进行了实验仿真；最后总结了本文工作。

2 相关研究

基于位置隐私保护的研究已经引起了许多学者的关注，并取得了一定研究成果。基于位置隐私保护的研究技术目前主要有基于中心服务器的位置隐私保护方法和基于无中心服务器的位置隐私保护方法[1]，而后者成为当前位置隐私研究的重点。本文提出的方法是在文献[1]的基础上提出的基于无中心服务器的位置隐私保护方法。

基于无中心服务器的位置隐私保护主要分为用户相互协作与无用户协作两种。用户协作是指在不使用匿名方法的情况下达到匿名的效果，这种方法不牺牲位置信息的服务质量。无用户协作的方法一般通过使用假位置或加密的方法实现位置隐私保护，这种方法会降低位置信息的质量。本文的研究是基于用户协作的，当协作用户的协作收益大于获得别人隐私的收益时，用户之间更倾向于诚信合作；而当协作用户的协作收益远远小于获得别人隐私的收益时，就会出现欺骗性合作。那么用怎样的方法可以减少甚至杜绝这种现象的发生？运用文献[2]中安全求和的方法，可以解决位置隐私相互协作中节点不诚实合作的问题。

用博弈论思想研究位置隐私保护仍处于初步阶段。文献[14]是第一个开始研究基于博弈论思想的位置隐私保护的，这篇文章主要描述了在完全静态信息博弈下，用户之间是一种非合作博弈状态，当匿名改变的成本增加时，节点之间更倾向于合作；而当匿名改变的成本很小时，节点之间就会形成非合作行为。

文献[16]提出的基于博弈论的隐私保护分布式数据挖掘方法，主要研究了分布式数据挖掘中参与者的决策问题，不仅指出在准诚信攻击的假设下，参与者的非共谋策略不是一个纳什均衡策略，而且给出了该博弈的混合战略纳什均衡，并进行了案例分析；同时文中给出了分布式数据挖掘中参与者选择准诚信攻击策略的理论依据，因为准诚信攻击策略是一个帕累托最优的纳什均衡策略。

本文提出的基于博弈论的用户相互协作的位置隐私保护方法是在文献[1,2,14,16]的研究基础之上提出的。本文的主要贡献是：(1)提出了不可信环境下用户相互协作的位置隐私保护方法，扩大了文献[1]提出方法的应用范围；(2)借鉴文献[2]的安全多方求和方法，解决了位置隐私相互协作中节点不诚实合作的问题。本文提出了 Privacy_I(s,δ,k)算法，s 代表用户指定的匿名区域半径，k 代表匿名数量，δ 表示用户对于服务误差的限制。通过用户个人设定参数 s,δ,k 来达到其希望的隐私保护水平，因为有时用户可能不是完全可信或完全不可信的，或者是半可信的，这种情况下用户想要达到的隐私水平可以根据个人需求达到不同的效果。(3)本文最后对近邻查询处理方法做了一定的改进，提高了文献[1]中增量近邻查询方法的效率。

3 博弈论基础

博弈论(game theory)是研究冲突与合作的一系列数学模型的汇集，它所关注的是在一系列确定的环境和结果中，为个体决策者寻找到最佳的行为。在很多领域，博弈论均可作为重要的理论工具，解决其中的冲突与合作问题，如战争、政治学、经济学、社会学、心理学和生物学等，近年来也开始应用于无线通信与网络，包括功率控制、资源分配、负载均衡、流量拥塞控制、网络路由、媒体接入控制和 QoS 支持等。

完全信息静态博弈[16]是博弈论中一种最简单的博弈，这里“完全信息”指每个参与人对所有其他参与人的特征(战略空间、成本与收益等)有完全的了解，“静态”指的是所有参与人同时选择行动且只选择一次。纳什均衡是完全信息静态博弈解的一般概念，其正式定义如下[15]：

定义 1(纳什均衡) 有 n 个参与人的战略式表述博弈 $G=\{S_1,\cdots,S_n;u_1,\cdots,u_n\}$，战略组合 $s^*=(s_1^*,\cdots,s_i^*,\cdots,s_n^*)$ 是一个纳什均衡，如果对于每一个 i，s_i^* 是给定其他参与人选择 $s_{-i}^*=(s_1^*,\cdots,s_{i-1}^*,s_{i+1}^*,\cdots,s_n^*)$ 的情况下第 i 个参与人的最优战略，即 $u_i(s_i^*,s_{-i}^*)\geqslant u_i(s_i,s_{-i}^*)$，$\forall s_i\in S_i$，$\forall i$ 或者用另一种表达方式，s_i^* 是下述最大化问题的解 $s_i^*\in\arg\max_{s_i\in S_i}u_i(s_1^*,\cdots,s_{i-1}^*,s_i,s_{i+1}^*,\cdots,s_n^*)$，其中 $i=1,2,\cdots,n$。

在博弈中，如果一个战略规定参与人在每一个给定的信息情况下只选择一种特定的行动，则称该战略是纯战略。相反，如果一个战略规定参与人在给定信息情况下以某种概率分布随机地选择不同的行动，则称该战略是混合战略。下面给出混合战略纳什均衡的定义[15]。

定义 2(混合战略纳什均衡) 在 n 个参与人博弈的战略式表述 $G=\{S_1,\cdots,S_n;u_1,\cdots,u_n\}$ 中，假定参与人 i 有 k 个纯战略 $S_i=\{s_{i1},\cdots,s_{iK}\}$，则概率分布 $\sigma_i=(\sigma_{i1},\cdots,\sigma_{iK})$ 称为 i 的一个混合战略，这里 $\sigma_{ik}=\sigma(s_{ik})$ 是 i 选择 s_{ik} 的概率，对于所有的 $k=1,\cdots,k,0\leqslant\sigma_{ik}\leqslant1,\sum_1^K\sigma_{ik}=1$，混合战略组合 $\sigma^*=(\sigma_1^*,\cdots,\sigma_i^*\cdots,\sigma_n^*)$ 是一个纳什均衡，如果对于所有的 $i=1,2\cdots,n$，下式成立：$v_i(s_{ik},\sigma_{-i})\geqslant v_i(s_{ik}',\sigma_{-i})$，$\forall s'_{ik}\in S_i$。

4 位置隐私保护方法

本文研究的是在现实生活中用户对于位置隐私保护是怎样选择合作方式的，在合作状态下用户是怎样处理诚信与不诚信合作的。如果协作组内所有节点都是诚信的，对于某个想要对位置隐私进行保护的节点来说，则会选择诚信合作共同保护彼此的位置隐私；相反如果协作组内存在欺骗节点，对于某个想要对位置隐私进行保护的节点来说，是应该选择诚信合作还是欺骗性合作呢？基于利益最大化的考虑，用户之间的行为是一种博弈模型。

4.1 位置隐私保护中的博弈关系分析

本文提出的方法是基于用户合作的位置隐私保护方法。在用户合作的状态下，用户可以选择诚信合作或者欺骗合作。因为在协议的执行过程中，参与者通过欺骗合作可以获得除位置服务与位置隐私保护之外的利益，如诚实节点的隐私数据。借鉴文献[16]的结论，得出当欺骗合作收益小于诚信合

作收益时，参与者之间的最优纳什均衡策略是诚信合作。而当诚信合作收益小于欺骗合作收益时，博弈的纳什均衡中所有参与者的最优策略是欺骗性合作。由于额外收益的存在使得博弈双方互相不信任，从而都不诚信合作，这样的结果便不能获得正确的位置服务。因此，在算法设计上需要保证欺骗合作收益小于诚信合作收益，使用户最终选择诚信合作，保证位置服务和隐私保护的有效性。

在不完全信息下，用户以怎样的概率选择怎样的合作方式，这是一种混合战略均衡。海萨尼(Harsanyi)[15]证明，混合战略均衡等于不完全信息下的纯战略纳什均衡。因为完全信息只是一个理想状态，现实中每个人对其他人的目标函数总不可能是完全了解的。在不完全信息下，每个参与人在选择自己的战略时，面对的大部分对手的选择方式都是不确定的，尽管每个参与人可能选择的都是纯战略。在我们讨论的位置隐私保护博弈中，混合战略纳什均衡的本质特征不在于参与人随机的选择行动，而在于参与人互相不知道对方选择什么纯战略的前提下决定自己的战略是什么。根据文献[16]，有如下结论：在两个参与者 P1 与 P2 的情况下，P1 将以概率 $\alpha^*=\frac{\theta_1}{u+\theta_1-\theta_2}$ 选择合作，P2 将以概率 $\beta^*=\frac{\theta_2}{u+\theta_2-\theta_1}$ 选择合作，其中 u 代表两者诚信合作时的收益，θ_2 代表 P1 诚信合作 P2 欺骗合作时 P2 的收益，$-\theta_2$ 代表 P1 诚信合作 P2 欺骗合作时 P1 的收益，θ_1 代表 P2 诚信合作 P1 欺骗合作时 P1 的收益，$-\theta_1$ 代表 P2 诚信合作 P1 欺骗合作时 P2 的收益，并且 $u>0,\theta_1>0,\theta_2>0$。

那么如何从技术上解决不合作的问题？借鉴文献[2]提出的安全求和方法，并对其进行了一定的改进，使得算法更符合实际隐私需求。

4.2 位置隐私保护算法实现

文献[1]提出了一种用户相互协作的匿名隐私保护方法 Coprivacy，然而该方法是基于所有用户可信的情况下达到的位置隐私保护。本文改进了 Coprivacy 方法，借鉴文献[2]的安全多方求和方法，提出了一种在不可信环境下用户之间的协作算法。在现实中，可能部分用户是不完全可信的，而且每一个用户也可能有一些私心，不想别人知道自己隐私太多，并且对于一定范围内的服务误差也是能够接受的。因此，本文提出的安全求和算法 Privacy_I(s,δ,k)改进了文献[2]中的安全求和算法。Privacy_I(s,δ,k)算法中：s 代表某个用户开始寻找组建合作团队的广播的一个范围，k 代表匿名数量，δ 表示用户对于服务误差的限制。

4.2.1 算法基本概念

定义 1(查询 Q) 用户发出的查询 Q 可以表示为以下形式：$Q=\{id,t,con,s,\delta,k\}$，

式中，id 表示用户发出查询时的唯一标识码，t 表示用户查询发出的时刻，con 表示用户输入的查询内容，s 表示用户指定的匿名区域半径，δ 表示用户对于服务误差的限制，k 表示用户指定的匿名参数。

参数 id,t 是由用户发出查询时系统自动生成的；参数 con,s,δ,k 是需要用户指定的内容。参数 k 和 s 越大，用户隐私保护水平越高，同时处理时间越长；δ 越大说明用户对于误差的限制越小，相反则越大。

定义 2(k-匿名组 K-S) 可以形式化表示为

$K\text{-}S=\{gid,k,s,anchor\}$

式中，gid 表示该匿名组的标识符，k 表示匿名组中含有的成员数量，s 表示匿名组区域的最小半径，$anchor$ 表示该匿名组的锚点，也就是每个成员发出查询时使用的位置。

定义 3(合作协议 Co-Agreements) 形成匿名组后，用户之间的通信方式可以形式化表示为

$\text{Co-Agreements}=\{id,l,k,\delta\}$

式中，id 表示用户发出查询时的 id 号，$l=(x,y)$ 表示用户协作的位置，x 表示位置的经度，y 表示位置的纬度，可以通过 GPS 导航获得，k 代表匿名组中成员个数，δ 表示用户对于服务误差的限制。

定义 4 欧氏距离 $dist(p,q)$ 表示点 p 与点 q 在二维平面上的距离

$$dist(p,q)=\sqrt{(p_x-q_x)^2+(p_y-q_y)^2}$$

4.2.2 算法描述

在本文描述的算法中，需要位置隐私保护的用户有 3 种状态：1)不在任何匿名组中。不在任何匿名组中的某个用户首先发出匿名组成立请求的广播，通过单跳或多跳的方式发现近邻用户，如果发现近邻用户的数目大于等于 k，再比较匿名区域是否满足用户对区域半径的最低要求 s，如满足，则形成匿名组，否则扩大匿名区域直到满足为止；2)已在匿名组中且未获得锚点。匿名组内的用户通过合作协议，也就是安全求和的方式进行匿名组锚点的计算。匿名组发起者通过与组内用户合作算出锚点，然后将锚点以广播的方式发送给组内的每一个用户；3)已在匿名组中且已获得锚点。组内的用户用这个锚点代替自己的真实位置向位置服务提供商发起查询，查询结束后匿名组解散。具体来说，上述过程分为 4 个步骤：匿名组的建立、锚点计算、广播锚点和位置近邻查询处理。

步骤 1 匿名组的建立。采用文献[1,9]提出的方法建立匿名组，过程如下：不在任何匿名组中的某个用户 r_q 首先发起组建匿名组的请求，然后广播消息 FORM_GROUP，消息内容为参数$\langle h,id,gid\rangle$，接着监听网络并等待邻居节点的响应。接收节点 p 接受到消息 FORM_GROUP$\langle h,id,gid\rangle$，首先检查消息是否与之前的相同，如果相同则抛弃，并向发送节点 r_q' 发送一个 ACK 确认。然后 p 检查自己的组编号是否为空或与接收到的组编号是否相同，如果满足则响应，否则不做处理。接着 p 检查 h，如果 $h>1$，则 p 将 h 减 1 并将发现消息 FORM_GROUP 进行广播，监听网络并等待邻居节点的响应。等收集完邻居节点的响应后，将整个消息再发送给原来的发送节点 r_q'；如果 $h=1$，则 p 将自己的信息直接加入到发现节点的集合中去，并将消息发送给发送方 r_q'。如果 r_q 以 h 跳收集到匿名组 T 后，发现匿名组不满足 $k-1$ 个节点，则 p 将广播跳数 h 加 1，重新广播发现的消息 FORM_GROUP。r_q 重复广播发现消息，直到发现 $k-1$ 个匿名节点。最后 r_q 还要考虑匿名区域大小是否满足其要求的最小区域 s。如果不满足，则 r_q 需要以其为圆心原点向外扩大区域半径，直到满足要求为止。

步骤 2 锚点计算。此处使用的算法通过分析可达到不可信环境下保护用户位置隐私的目的，根据用户的不同需求相应地协调合作成本，从而达到诚信合作，杜绝不诚信合作的情况，其合作行为在此形成博弈行为。

请求发起者 r_q 建立匿名组后，将匿名组中的成员按其 id 号建立一个通信环，并分别标上序号 $0,1,\cdots,k-1$。r_q 在组建匿名组的过程中，估算出匿名组中的不诚信成员数量 k'，并在建立通信环的过程中将 k' 转发给环中的每一个成员，同时 r_q 置为组内的第一个成员 p_0(算法 1 第 1－2 行)。环中每一成员先都将自己的真实位置分成 k' 份进行合作求和(算法 1

第3行)。首先 p_0 将自己的第一份数据加上一个随机数 r_m，发送给下一个节点 p_1，p_1 将收到 p_0 的数据加上自己的第一份数据又传送给下一个节点 p_2，直到组内所有成员都将自己的第一份数据加进来为止，组内的最后一个成员 p_{k-1} 将数据发送给 p_0，得到 $sum_1.[0]$ 的值(算法1第4—9行)，将 $sum_1.[0]$ 减去 r_m 赋值给 $sum_2.[0]$，接着将随机数改为 r_{m+1}，再进行一次累积求和，最后数值减去 r_{m+1} 得到新的 $sum_1.[0]$(算法1第10—14行)。比较 $sum_1.[0]$ 和 $sum_2.[0]$ 的差值是否满足用户定义的最小误差范围 $\frac{\delta}{k'}$，如果不满足则这份数据需要重新计算(算法1第15—16行)。同理可得到最终正确的 $sum_1.[1],\cdots,sum_1.[k'-1]$，最后可计算得到总和数据 sum(算法1第17行)。接着计算匿名组的密度中心，也就是锚点 $anchor=\frac{sum}{k}$(算法1第20行)。

算法1 计算锚点 Anchor_acquired(n,k)

1. r_q 将收到的节点按其 id 号组建成一个通信环
2. 估计可能的欺骗站点数目 k′，并广播给组内的每一成员 $p_0 \leftarrow r_q$
3. p_i 将 l(x,y)随机分成 k′份，$l_0,l_1,\cdots,l_{k-1}$
4. $n\leftarrow 0, sum_1.[k']\leftarrow 0, sum_2.[k']\leftarrow 0, sum\leftarrow 0$
5. for(i=0;i<k′;i++)
6. for(j=0;j<k;j++)
7. if(j=0) $sum_1.[i]=p_j.[i]+r_m$
8. else $sum_1.[i]+=p_j.[i]$
9. end if
10. if(n=0 and j=k−1)
11. $sum_2.[i]=sum_1.[i]-r_m$,m++,j=0,n++
12. end if
13. end for
14. $sum_1.[i]-=r_m$
15. if($|sum_2.[i]-sum_1.[i]|>\frac{\delta}{k}$)
16. n−−,m++,i−−
17. else sum+=$sum_1.[i]$,n←0
18. end if
19. end for
20. anchor=$\frac{sum}{k}$

如图1、图2所示，当节点0找到足够数量的匿名节点并且匿名区域满足最小要求后，0节点便以自己的 id 号为首建立一个匿名环，同时估计出不可信节点的数量，这里环中节点的总数量为12，不可信节点的数量为2，用户隐私精确度为1，匿名区域为2.5。节点0首先将自己的数据分成2份，先将第一份数据加上随机数 r_1，同时将这份数据传给节点1，节点1也将自己的数据分成两份，将自己的第一份数据加到从节点0刚刚接收到的数据上，再传给节点3，一直加到最后一个节点11，累积和减去 r_1 后为 $sum_1.1$。接着节点0再次将自己的第一份数据分别加上随机数 r_2，然后将这份数据传给节点1，再次进行上次的循环，直到接收到节点11发送来的数据，减去 r_2 得到 $sum_2.1$。然后判断 $|sum_1.1-sum_2.1|\leqslant\frac{1}{2}$ 是否成立，如果都满足说明达到用户对服务误差限制的最低要求，接着以同样方法算出第二份数据的累计和 $sum_1.2$，最后将数据 $sum_1.1$ 与 $sum_1.2$ 相加便得出整个累计和 $sum=sum_1.1+sum_1.2$。如果有哪次不满足，例如 $|sum_1.1-sum_2.1|>\frac{1}{2}$，则节点0需将自己的第一份数据重新进行累计和的计算。

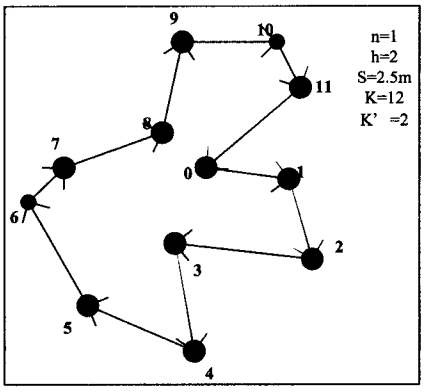


图1 匿名组建立

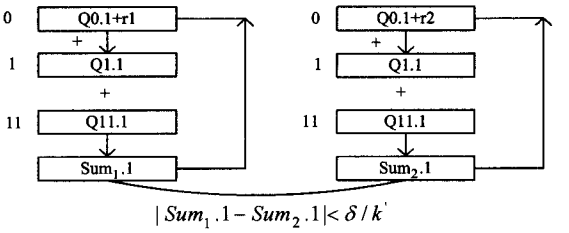


图2 锚点计算过程

步骤3 广播锚点和位置近邻查询处理。用户 r_q 计算得到锚点 $anchor$ 后，通过通信环将锚点 $anchor$ 发给组内的每一个用户。组内用户获得锚点后，首先对比自己的隐私需求参数 $p_i.k$ 和组内用户个数 k。如果 $p_i.k\geqslant k$，用户所在的匿名组满足用户 k-匿名需求，用户可以直接使用获得的锚点向位置服务提供商发起位置近邻查询，如果 $p_i.k<k$，用户所在的匿名组不满足用户 k-匿名需求，为了满足用户 k-匿名需求，还需要向位置服务提供商发起 $p_i.k-k$ 次假查询[1]。

位置近邻查询处理算法见算法2。首先位置服务提供商将离锚点最近的 n 个位置以离锚点距离从小到大的顺序排序后放入数组 $W[n]$ 中，并将 $W[n]$ 结果发送给用户请求者(算法2第2行)。用户对 $W[n]$ 中的数据分别比较其与锚点之间的距离，当 $dist(p_i,w_i)+dist(p_i,q)\leqslant dist(q,w_{i+1})$ 满足时，表示在 $W[n]$ 中的第 i 个位置是离用户最近的(算法2第3—7行)。如果在 $W[n]$ 中没有找到此点，则扩大搜索范围，用户向位置服务商发起增量查询，也就是在原来基础上再找 m 个离锚点最近的位置，位置服务商将这 m 个位置连同上次的 n 个位置一起再次发送给用户，用户从第 n 个位置继续搜索剩下的 m 个数据直到找到离自己最近的位置服务为止(算法2第8—12行)。

算法2 位置近邻查询处理

1. //节点 p 查询，锚点设为 q
2. W[n]←位置服务提供商返回给 p 离 q 最近有序的 n 个位置
3. for(i=0;i<n;i++)
4. if((dist(p,q)+dist(q,W[i]))>dist(q,W[i+1]))
5. continue
6. else return i
7. end if
8. if(i=n−1)
9. p 向位置服务提供商请求 n+m 个最近位置
10. n←n+m
11. end if
12. end for

如图 3 所示，A 点为发起站点，B 为锚点。服务提供商返回给 A 节点离 B 最近的 5 个节点，并按照离 B 点的距离由小到大的顺序发送给 A 点，则可以知道当 $dist(p_i,A)+dist(A,B)\leqslant dist(p_{i+1},B)$ 满足时，p_i 为离 A 点最近的点。

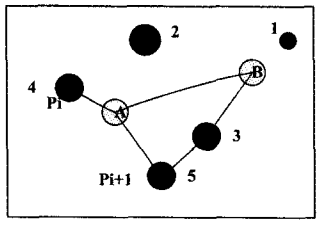


图 3　近邻查询点

证明：假设 p_i 不是离 A 点最近点，而 p_{i+1} 为离 A 点最近的距离，则有 $dist(p_i,A)>dist(p_{i+1},A)$，则 $dist(p_i,A)+dist(A,B)>dist(p_{i+1},A)+dist(A,B)$，而 $dist(p_{i+1},A)+dist(A,B)\geqslant dist(p_{i+1},B)$，所以有 $dist(p_i,A)+dist(A,B)>dist(p_{i+1},B)$，与已知条件 $dist(p_i,A)+dist(A,B)\leqslant dist(p_{i+1},B)$ 相矛盾。则可知 p_i 为离 A 点最近的点。

同样对于节点 p_{i+j} $(j>1)$ 肯定有 $dist(p_{i+j},B)\geqslant dist(p_{i+1},B)$，因为位置服务提供商提供的是按离 B 从小到大有序的位置节点，则必定满足 $dist(p_j,A)+dist(A,B)\leqslant dist(p_{i+j},B)$，用以上方法同样可以得出 p_i 为离 A 点最近的点。

5　实验仿真

本文算法使用 C++编程语言环境实现，编程环境为 Microsoft Visual C++6.0，实验硬件环境为 2.9GHz 处理器，4GB 内存。操作系统平台是 Windows 7。

5.1　实验数据集和参数设置

实验数据用由移动数据管理研究界认可的 Thomas Brinkhoff 路网数据生成器[17]生成，它以城市 Oldenburg 的交通路网作为输入，生成模拟移动用户数据。实验中使用数据的参数值如表 1 所列。

表 1　实验中使用到的数据

参数名称	平均值
移动用户数量	4000
匿名参数需求 k	10
隐私保护区域半径 s	500m
用户服务误差限制 δ	1
位置服务提供商对每次查询提供的对象个数 n	10
用户扩大查询对象个数 m	5

实验使用 NS2 进行一个简单的网络节点模拟，带宽为 1Mbps，移动用户使用该信道进行 P2P 通信。匿名处理方式算法与近邻查询处理只是在 Microsoft Visual C++6.0 下编译的程序，并且还假设移动用户与位置服务提供商之间使用 3G 网络通信，带宽为 2Mbps。

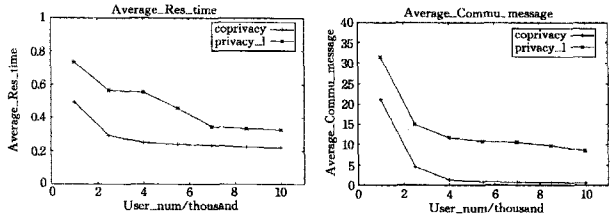
5.2　算法衡量标准

本实验在模拟数据集上对本文提出的方法的平均响应时间、匿名成功率、平均通信消息量以及近邻查询处理大小等进行了实验，并与 Coprivacy 方法中的各结果进行了对比。

匿名成功率的值越高，表明算法越好，这里有基于匿名数量 k 的匿名成功率的比较；平均响应时间越小说明算法效率越高，这里有基于用户数量的平均响应时间比较与基于用户匿名需求 k 的平均响应时间比较；平均通信量越低说明算法的通信效率越高，这里有基于用户数量的平均通信量比较与基于用户匿名数需求 k 的平均通信量比较；近邻查询越小说明算法效率越高，这里只有基于用户匿名数需求 k 的近邻查询大小比较。

5.3　实验结果分析

本文将 Privacy_1 方法中的数据设置为同 Coprivacy[1] 中的数据一致。

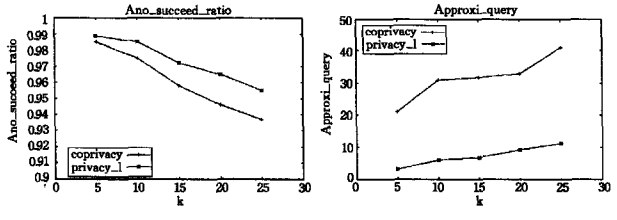


(a) 平均响应时间　　(b) 平均通信量

图 4

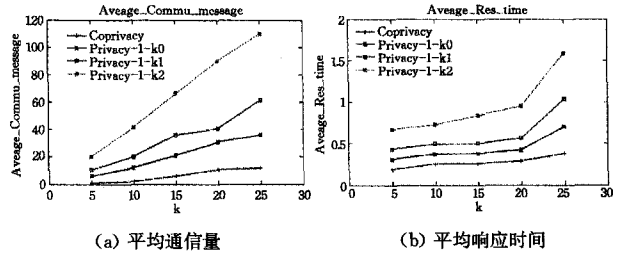
基于用户数量的比较，其数值变化范围从 1000 到 10000，同时这里假设用户都是可信的，也就是欺骗用户数量为 0。其他参数为表 2 中的默认参数。也使用 Oldenburg 数据集评估两种方法。由图 4(a)与(b)可知，平均响应时间与平均通信量随着用户数量的增加都有所降低，却比 Coprivacy 方法的时间要大点。这里虽然假设用户都是可信的，用户合作时不需要将数据分片，但相对 Coprivacy 仍有 for 循环用来合作计算匿名环的锚点，而 Coprivacy 中计算锚点时对于位置信息在建立匿名组时已经都知道，所以相对而言效率会低一点。本文提出的 Privacy_1 方法虽然需要花更多的时间来协商形成可靠的环境，但能够解决不可信的环境下的合作。

基于匿名需求 k 的比较，其数值变化范围在 5 到 25 之间。由图 5(a)的分析可知随着匿名数量增加，用户的匿名成功率有所下降，因为随匿名需求 k 的增加在用户数量不变的情况下匿名要求也就增加了，且本文提出的方法的匿名成功率要高于 Coprivacy，原因就是本文添加了对匿名最小区域的限制。随着 k 匿名数的增加，近邻查询大小呈增大趋势(见图 5(b))，且本文提出的方法的近邻查询大小要小于 Coprivacy，原因是本文不需要建立大顶堆的排序方式，而是采取一种更直接的比较方法，不需要将查询的所有结果进行比较，可以就前面的节点尽快找到最近节点，所以近邻查询效率有所提高。



(a) 匿名成功率　　(b) 近邻查询

图 5



(a) 平均通信量　　(b) 平均响应时间

图 6

对本文提出的 Privacy_l 方法，其中用 Privacy_l_k0 表示欺骗用户数量为 0，用 Privacy_l_k1 表示欺骗用户数量为 1，用 Privacy_l_k2 表示欺骗用户数量为 2，一起与 Coprivacy 方法进行了比较。由图 6(a)与(b)可知，随着匿名数 k 的增加，平均通信量与平均响应时间都有所增加，且本文提出的方法中随着欺骗用户数 k' 的增加，平均通信量与平均响应时间也都要增加，并且都要大于 Coprivacy，因为本文中的锚点计算方法在完全诚信环境下，也就是欺骗用户数为 0 时，其通信量与响应时间相对 Coprivacy 都要大点，且 Coprivacy 锚点位置信息在匿名组建立时就知道了，而本文却需要进一步合作通信才能知道。同时也可以看到随着本文假设的欺骗用户数量的增加，其相应的平均通信量与平均响应时间都要增加，因为欺骗节点数量越多，每个用户越需要将各自的数据更多地分片来合作求锚点，从而保证诚信合作时不至于泄露自己的隐私给恶意用户。

从以上分析可知，不诚信成员数量 k' 越多，平均通信量与平均响应时间越会增加，因为这里随 k' 的增加需要循环求和的次数也有所增加。同样重复求和的次数越多，平均通信量和平均响应时间也会越大，这里比较的结果与基于不诚信数量 k' 的结果相似，就不再作图表示。

结束语 传统的位置隐私保护方法大多采用可信第三方的结构，这往往需要大量的计算，且容易使第三方成为系统瓶颈和集中攻击的目标。Coprivacy 方法使用用户相互协作的匿名隐私保护方法来保护组内用户的隐私，然而其只是基于所有用户可信的情况下达到的位置隐私。本文主要的贡献就是提出了一种在不可信环境下用户之间协作的方法，先从理论上说明用户之间的博弈行为，以怎样的方式选择什么样的合作可以达到各自的最大利益；再从技术上提出解决用户不可信环境下的协作方式，杜绝了不可信的合作方式，提出了一套完整的算法结构，以更好地达到用户位置隐私需求的保护。虽然锚点计算算法效率稍有降低，但是可用于更接近现实的匿名情况。并且本文中的近邻查询处理相对文献[1]有一定的改进，算法效率稍有提高。然而由于本文中的方法是假设各个用户之间合作计算锚点时是静止的，因此未来研究的工作可以在合作计算锚点时用户是移动的环境下展开。

参 考 文 献

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

[2] 张国荣，印鉴. 基于博弈论的安全多方求和方法[J]. 计算机应用研究，2009，26(4)：1497-1499

[3] Gruteser M, Grunwal D. Anonymous usage of location based services through spatial and temporal cloaking [C]// Proceedings of the International Conference Mobile Systems, Applications, and Services (MobiSys'03). New York, USA, 2003: 163-168

[4] Gedik B, Liu L. A customizable k-anonymity model for protecting location privacy[C]// Proceedings of the IEEE International Conference on Distributed Computing Systems (ICDCS'05). Columbus, Ohio, USA, 2005: 620-629

[5] Mokbel M F, Chow C Y, Aref W G. The new casper: Query processing for location services without compromising privacy[C]// Proceedings of the International Conference on Very Large Data Bases (VLDB'06). New York, USA, 2006: 763-774

[6] Xiao Z, Meng X, Xu J. Quality-aware privacy protection for location-based services[C]// Proceedings of the International Conference on Database Systems for Advanced Applications (DASFAA'07). Bangkok, Thailand, 2007: 434-446

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

[8] Bamba B, Liu L, Pesti P, et al. Supporting anonymous location queries in mobile environments with privacy grid[C]// Proceedings of the International World Wide Web Conference (WWW'08). Beijing, China, 2008: 237-246

[9] Chow C, Mokbel M F, Liu X. A peer-to-peer spatial cloaking algorithm for anonymous location-based services [C]// Proceedings of the Annual ACM International Symposium on Advances in Geographic Information System (GIS'06). Virginia, USA, 2006: 171-178

[10] Ghinita G, Kalnis P, Skiadopoulos S. PRIVE: Anonymous location based queries in distributed mobile systems[C]// Proceedings of the International Conference on World Wide Web (WWW'07). Banff, Alberta, Canada, 2007: 1-10

[11] Solanas A, Martínez Balleste A. Privacy protection in location-based services through a public-key privacy homomorphism[C]// Proceedings of the European PKI Workshop, Theory and Pratice. Lecture Notes in Computer Science. Palma de Mallorca, Spain, 2007: 362-368

[12] Solans A, Martínez-Balleste A. A TTP-free protocol for location privacy in location-based services [J]. Computer Communication, 2008, 31(6): 1181-1191

[13] Yiu M L, Jensen C S, Huang X, et al. Space Twist: Managing the trade-offs among location privacy, query performance, and query accuracy in mobile services[C]// Proceedings of the IEEE International Conference on Data Engineering (ICDE'08). Cancun, Mexico, 2008: 366-375

[14] Freudigery J, Manshaeiy M H. On Non-Cooperative Location Privacy: A Game-Theoretic Analysis[C]// Proceedings of the International Conference on Computer Control System (CCS'09). Chicago, Illinois, USA, 2009: 1-14

[15] 张维迎. 博弈论与信息经济学[M]. 上海：上海人民出版社，2004

[16] 葛新景，朱建明. 基于博弈论的隐私保护分布式数据挖掘[J]. 计算机科学，2011，38(11)：161-166

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