基于敏感位置多样性的 LBS 位置隐私保护方法研究

周长利，马春光，杨松涛
（哈尔滨工程大学 计算机科学与技术学院，黑龙江 哈尔滨 150001）

摘 要：针对 LBS 查询服务中构造的匿名框或选取的锚点仍位于敏感区域而导致的位置隐私泄漏问题，提出了基于敏感位置多样性的锚点选取算法。该算法根据用户访问数量和访问高峰时段，对不同敏感位置进行定义和筛选，选择具有相似特征的其他敏感位置构成多样性区域，并以该区域形心作为查询锚点，提高用户在敏感位置出现的多样性。以该锚点为查询标志，提出一种均衡增量近邻兴趣点查询算法 HINN，在无需用户提供真实位置坐标的条件下实现 K 近邻兴趣点查询，同时改进了 SpaceTwist 方法中存在的查询兴趣点围绕锚点分布的缺陷，提高了查询准确度。实验表明，本方法实现了用户在敏感区域停留时的位置隐私保护目标，同时具有良好的兴趣点查询质量和较低的通信开销。

关键词：位置隐私；基于位置的服务；增量近邻查询；l 多样性
中图分类号：TP311 **文献标识码**：A

Research of LBS location privacy preserving based on sensitive diversity

ZHOU Chang-li, MA Chun-guang, YANG Song-tao
(School of Computer Science and Technology, Harbin Engineering University, Harbin 150001, China)

Abstract: Before getting location-based query service, constructing a cloaking region or picking an anchor which is still in a sensitive area is vulnerable to lead location privacy exposure. An algorithm of selecting anchor is proposed based on sensitive location diversity. By defining sensitive locations and filtering different ones according to users' visiting number and peak time, locations with similar features are chosen to form a diversity zone, and its centroid is taken as the anchor' location which raises user' location diversity. Based on the anchor, a query algorithm HINN for points of interest (PoI) is proposed referring to SpaceTwist, and query results can be deduced without providing any user's actual location. The defect in SpaceTwist that PoI are found around the anchor is modified, which improves querying accuracy. The experiments show that users' location privacy is protected well when the user is staying at a sensitive place, and the method has good working performances.

Key words: location privacy; location-based service; incremental nearest neighbor query; l-diversity

1 引言

基于位置的服务(LBS, location based service)给人们生活带来便利的同时，也带来了位置隐私泄漏的风险。基于位置的查询是一项广泛应用的位置服务，用户将当前所在位置信息发送给位置服务商(LSP, location based service provider)，以获取查询服务，如查询周围的餐厅、最近的医院等。由于位置数据的时空敏感特性，其泄漏可能会给用户带来身份、住址、习惯爱好及健康状况等隐私信息的暴露。

收稿日期：2014-07-31；**修回日期**：2014-10-23
基金项目：国家自然科学基金资助项目(61170241, 61472097);高等学校博士学科点专项科研基金资助项目 (20132304110017); 黑龙江省杰出青年基金资助项目 (JC201117); 黑龙江省教育厅科学技术研究基金资助项目(12541788,12541788).
Foundation Items: The National Natural Science Foundation of China (61170241, 61472097); Specialized Research Fund for the Doctoral Program of Higher Education (20132304110017); Excellent Youth Foundation of Heilongjiang Province (JC 201117); Science and Technology Research Project of Heilongjiang Education Department (12513049,12541788)

因此，需要对查询过程中的位置数据加以保护，确保用户位置隐私不被泄漏。

移动用户在行进过程中的点可以分为2类：经过点和停留点。移动用户在行进过程中，随时可能基于当前位置提交LBS查询请求。用户经过的位置并不能表示用户与该位置存在特别关联关系，只表示运行经过的轨迹；而用户停留位置在某种程度上能够说明用户与该位置存在某种语义联系，尤其是停留在敏感区域，更为攻击者提供了深入挖掘的可能，如一些攻击者可以在敏感位置(医院、酒吧及商业区)设置信息窃听装置，来分析在这些地区出现用户的身份信息，以达到某种商业目的。

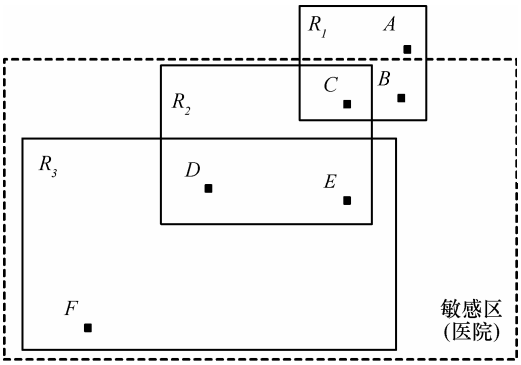


图1 匿名框构造

位置模糊与泛化是查询过程中广泛使用的方法。一类思想是将用户当前的位置进行扩展，将具体位置点扩展成为一个匿名框，实现位置 k 匿名(k-anonymity)[1, 2]。如图1所示，用户将当前位置数据提交给中心匿名服务器，匿名服务器寻找到该位置附近的 $k-1$ 个用户，构造匿名框 R_1，将匿名框 R_1 提交给LSP获取查询服务，并将查询后的结果求精后返回给用户。用户可以通过匿名度 k、匿名框大小及匿名容忍时间等参数自行定制隐私需求。但是这类方法存在两个问题，一是匿名服务器在面对大量用户时，需要为每个用户构造匿名框，服务器负担重也存在遭受异常点攻击[3]的可能。二是当用户停留在某一敏感区域时(如医院、教堂及酒吧等)，如图1中的虚线框区域。这种长时间内停留或反复短距离移动(如在 B、D 及 F 点)，会使查询所构造的匿名框 $R_1 \sim R_3$ 始终限制在该敏感区域内，依然会暴露该用户的位置隐私。因此需要保证匿名框中用户位置的多样性(l-diversity)[4,5,6]，尤其是用户处于敏感区域时，需要增加用户在其他敏感位置出现的可能，提高匿名质量。

另一类查询中的模糊泛化思想就是使用假位置(Dummy)[7,8]，如用户向LSP发送自身位置的同时发送多个假位置来隐藏真实位置，但是LSP会根据所有位置坐标进行相应的查询操作，并返回给用户，用户从中选择真实位置的查询结果，这给LSP和用户均带来了不必要的额外负担。假位置的方法还包括采用标志对象(significant object)[9]的方法，用户向LBS服务器发送标志对象位置替代真实位置，但这种方法存在查询结果不够精确的缺陷。Yiu 等[10]借鉴了标志对象的方法，提出了以锚点代替真实位置的客户端运行查询算法SpaceTwist，该方法以服务器增量近邻查询(INN, incremental nearest neighbor)返回的结果候选集为基础，计算出自己查询的 K 个近邻兴趣点(PoI, place of interest)。SpaceTwist在保护用户位置隐私的同时解决了查询效率低和查询结果精确度低的问题，但是处于敏感位置用户在选取锚点时，如果锚点位置依然位于该敏感区域内，仍会造成隐私泄漏。

用户在敏感区停留，会对用户隐私产生直接的语义关联，在连续多个敏感区的停留(轨迹)，会揭示该用户更为丰富的隐私信息[11,12]。因此，本文主要关注用户在敏感区停留或小范围活动时的位置隐私保护问题。本文的研究内容及创新性如下：

1) 针对用户利用敏感位置查询兴趣点带来的位置语义隐私泄漏问题，提出基于敏感位置多样性的锚点位置选取方法。该方法基于用户群体访问时空特征来定义并区分邻近敏感位置，使用来查询的锚点具有语义多样性，增加用户实际所处位置的不确定性。

2) 针对用户以真实位置查询可能带来的隐私泄露问题，提出了用户端运行的均衡增量近邻查询算法HINN，以锚点代替真实位置查询，并根据LSP返回候选结果计算所需的 K 个近邻兴趣点。该方法利用需求空间再扩大的方法，解决了SpaceTwist方法中存在的兴趣点查询结果不准确的问题。

2 相关工作

为了能够更好地从数据产生源头保护用户的位置隐私，用户在查询过程中需要对自己的真实位置进行模糊处理，以达到保护隐私的目的。Gruteser 等[1]将数据库中 k 匿名思想引入到位置隐私保护中来，提出了时空匿名方法，通过对用户位置进行时空模糊，降低位置的可识别度，从而增加攻击者将位置与用户真实身份关联的难度。Gruteser 等[1]首

次提出了中心匿名服务器架构(central anonymity server)，该架构在 LSP 与用户之间设置可信位置匿名服务器，利用该服务器为用户构造匿名框并对 LSP 返回的结果进行求精处理。Chow 和 Mokbel 等[13,14]利用无中心服务器的 P2P 匿名方法，避免了中心结构的服务瓶颈问题，但是用户端负担较重。

上述匿名方法虽然能够有效实现 k 匿名，但是依然存在 k 个用户在同一敏感区(如 k 个用户都在医院)而导致位置隐私泄漏的问题。为此，Bamba[4]等借鉴数据发布中的多样性思想(l-diversity)，提出了基于用户位置多样性的位置隐私保护方法。徐建[5]及杨晓春[15]等分别提出了路网环境下的位置匿名方法，此类该方法以实现匿名区域位置多样性为目标，文献[15]提出了路网环境下隐匿环与森林构造方法，该方法构造包含多条路段的隐匿环来提高用户位置隐私的保护质量。孟小峰[12]等在轨迹数据发布中提出了敏感位置多样性的保护思想，该方法对用户的轨迹隐私进行模糊处理。

在提交查询请求过程中除了扩大用户位置范围之外，还可以通过假位置方法实现隐私保护。Kido[7]等提出同时将多个假位置发送给 LSP，并在查询结果中选择自己实际位置的查询结果以达到混淆的目的，但这种方法存在 LSP 查询处理负担重的问题。Niu[8]等针对 LBS 服务端不可信问题，利用同样的假位置思想来提高真实位置的匿名性。Hong[9]等利用标志对象替代真实位置进行查询，但是这种方法存在查询结果不够精确的问题，Yiu[10]等借鉴标志对象的方法，提出了能够精确查询的 SpaceTwist 方法，该方法采用增量近邻查询方式，利用锚点(anchor)来实现对兴趣点的精确查询。本文正是借鉴该思想，提出了停留在敏感区域用户查询过程中的位置隐私保护方法。孟小峰[16]和 Gong[17]等针对 SpaceTwist 没有实现 k 匿名等问题，分别提出了相应的解决方法，这些方法均借鉴增量近邻查询思想，实现了对兴趣点的高效查询。

通过分析发现 SpaceTwist 方法仍然可以改进加强，来获取更精确的查询结果。引用 SpaceTwist 中的查询处理实例，如图 2 所示，来说明可以改进的问题。SpaceTwist 中的查询可分为图 2 的 3 个过程，处于位置 q 的用户以锚点 q' 为中心向 LSP 发起 INN 查询，逐步获取兴趣点。当找到第二个兴趣点 P_2 时，供应空间(supply space)继续扩大为半径是 τ 的深灰色区域，需求空间(demand space)则继续缩小为半径是 γ 浅灰色区域，以 q' 为中心的供应空间随兴趣点被找到逐步增大，而需求空间逐渐缩小，最后直到供应空间完全覆盖需求空间，查询结束，处于位置 q 的用户找到距离自己最近的 K 个兴趣点。

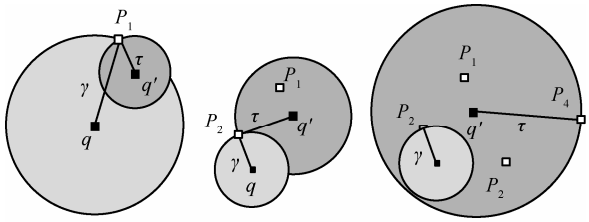


图 2　查询处理实例

上述查询过程存在可以改进的内容，当图 2(c)中兴趣点 P_3 出现后，如果下一个出现的兴趣点 P_x 到锚点 q' 的距离比 P_4 小，且 P_x 在用户真实位置 q 附近，如图 3(a)所示，此时由于供应空间覆盖需求空间，查询立即结束，兴趣点 P_4 不会再被算法查出。这是由于以锚点 q' 为中心查找导致 q 附近且在锚点 q' 相反方向出现某个兴趣点 P_x 时，需求空间半径会立即缩小并被供应空间覆盖，q 周围的其他兴趣点由于 P_x 出现导致查询结束而不再会被查出。同样，SpaceTwist 方法中所有查找到的兴趣点均分布在 q' 周围，如图 3(b)所示，而用户实际位置 q 在虚线左下方没有兴趣点被查询到。称这种由于兴趣点在真实位置附近出现而导致的 INN 查询快速结束，及找到的兴趣点围绕锚点分布而导致的真实位置所需的兴趣点分布不均匀的问题，称为查询不均衡问题。本文将在 SpaceTwist 算法基础上设计新的查询算法，使查找到的兴趣点围绕用户 q 分布，解决查询不平衡问题，提高查询精准度。

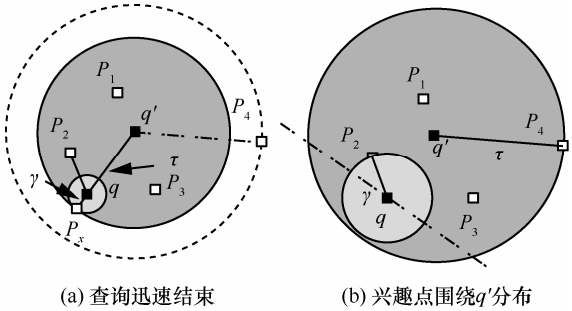


图 3　查询不平衡

3　系统架构

如前所述，LBS 中的隐私保护方法主要基于 2

种架构设计：有中心服务器架构和无中心服务器架构[3,16,18]。本文采用有中心服务器的系统架构，该架构在文献[1]中首次提出，并在之后的研究中被其他学者广泛采用[4,5,13,15,17]，该架构在用户和LBS服务提供商(LSP)之间添加可信第三方中心服务器(CS, central server)，如图4所示，一定数量的CS由可信认证机构负责部署和在线维护，可信认证机构可以是政府组织或是具有社会公信力的商业机构。每个CS根据本文提出的敏感位置多样性算法为附近用户计算锚点、组织用户实现协作 k 匿名并提供用户查询请求转发等功能。相关定义如下。

定义 1 网络内实体可以表示为三元组 $<\mathrm{U, CS, LSP}>$，其中U表示移动用户集合，用户通过无线智能终端实现通信获取智能服务，资源和能力受限；CS 表示中心服务器集合，部署在用户访问频繁的位置，本方法中主要用来为用户计算查询用的锚点，资源和能力较强；LSP表示基于位置的服务器，能够根据用户提供的位置增量查询该位置附近的兴趣点集合，分析处理能力极强。CS 可信的，LSP是半可信的。

用户为了降低位置被关联的可能，利用假名向CS 向位置服务提供商 LSP 发起相关请求。当 CS 收到锚点计算请求后，首先对用户身份进行认证，身份合法则会根据用户位置 loc 为其计算锚点 loc_{anchor}，并返回给用户。

定义 2 用户向LSP发起INN查询可以表示为 $<uid_{i1}^{'}, loc_{anchor}, C, \beta>$，其中 $uid_{i1}^{'}$ 是由用户真实身份 uid_i 生成的一个假名，不同假名之间无直接关联，C 表示查询内容，β 是每次返回查询兴趣点个数。

LSP获取该查询要求后，以锚点位置为圆心，以逐次递增方式，向用户发送查询兴趣点。

定义 3 用户所在平面空间位置点用集合 $P=\{P_1, P_2, \cdots, P_n\}$ 表示，其中 $P_i \in P$ 可以用来表示兴趣点、敏感位置等任意位置。由于兴趣点均具有一定的敏感性，只是程度不同，本文将兴趣点看作是敏感位置的子集。

4 敏感区域位置隐私保护方法

本节主要包括 CS 为用户生成多样性锚点和用户如何利用该锚点发起查询2个过程。第一个过程包括提出基于用户访问时空特征的敏感位置定义方法和基于敏感位置多样性的锚点选取算法；第二个过程提出利用该锚点的兴趣点查询算法 HINN。

4.1 锚点选取过程

首先 CS 基于用户访问数量及高峰时段2个特征，计算管辖区内不同位置的敏感权值，然后利用该权值为用户选取满足敏感位置多样性锚点，提高锚点出现在多个敏感位置的可能。

4.1.1 基于用户访问特征的敏感位置定义方法

社会性是符合人类普遍行为规律特征的基本属性。位置的访问热度在一定程度上反映了人类行为的趋向性，是某一位置社会属性的具体体现。频繁被访问的位置具有一定的时空语义特征，能在一定程度上反映出人类的行为习惯，在这些位置停留的用户会在某种程度上暴露其与该位置的语义关系。如在每个工作日早上被大量用户访问的位置可能是某单位，用户在此处停留可能暴露其工作单位，某位置在周日被频繁访问，可能是商场酒吧或是教堂等，用户在此处停留可能暴露其生活习惯及宗教信仰。本节以用户访问数量及高峰时段作为衡量位置敏感性的主要因素，同时利用高峰时段出现的时间差异性，获取敏感位置多样性。

考虑到敏感位置多基于路网分布，我们将分布在路网两侧的敏感位置定义如下。路网有向图为 $G=(\boldsymbol{V}, \boldsymbol{E})$，每个顶点 $v_i \in \boldsymbol{V}$ 被移动用户访问频率的权值用 $R(v_i)=\lambda_i$ 表示，$e_{ik} \in \boldsymbol{E}$ 是顶点 v_i 与 v_k 之间

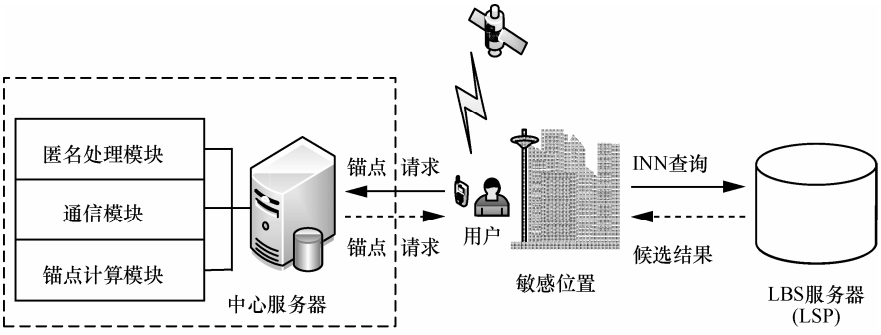


图4 系统架构

的有向边。如果 $\forall v_i, v_k \in \boldsymbol{V}$ 之间不存在其他顶点 $v_x \in \boldsymbol{V}/\{v_i, v_k\}$ 且存在唯一通路，则从 v_i 前往 v_k 的移动用户数服从到达率为 $\lambda_{ik} > 0$ 泊松分布，且 $e_{ik} = \lambda_{ik}$，否则 $e_{ik} = 0$。顶点 v_i 的访问频率权值为

$$R(v_i) = \lambda_i = \lambda_i' + \sum_{v_j \in \boldsymbol{V}, k \neq i} e_{ki} = \lambda_i' + \sum_{v_j \in \boldsymbol{V}, k \neq i} \lambda_{ki} \quad (1)$$

其中，λ_i' 表示从非顶点位置出发的移动用户至顶点 v_i 的累积到达率。假设到达顶点的移动用户在该顶点的每条出边离开的概率是相等的，则 v_i 某条出边的访问频率权值为 $R(v_i)/\deg_{out}(v_i)$，其中 $\deg_{out}(v_i)$ 是 v_i 的出度。考虑兴趣点以街道地址表达的实际情况，在形式化分析道路上行进的用户数量时，可以认为所有敏感位置分布在路网图边上，如图 5 边上的黑色方点。因此，处于相邻顶点 v_i 与 v_k 之间边上的某个位置 P_i 可能被来自该边 2 个顶点的移动用户访问，则某时间段内在路段 v_iv_k 上行进的用户数权值可以看作来自 2 个顶点出边的访问频率权值之和。

$$M = [\mathcal{R}(v_i)/\deg_{out}(v_i)] + [\mathcal{R}(v_k)/\deg_{out}(v_k)] \quad (2)$$

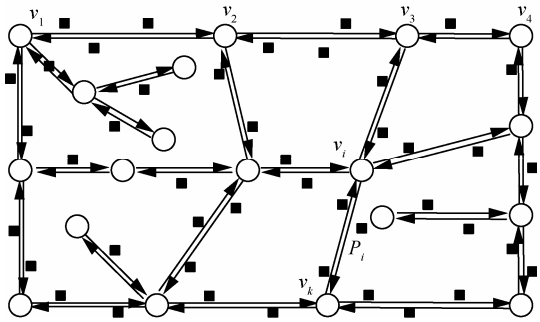


图 5　路网图

用户并不是在路段 v_iv_k 上每个位置都停留，假设位置 P_i 一天某时间段内用户经过数为 n，其中停留访问该位置概率为 p，则该时间段访问此位置的停留用户数服从参数 $\mu_i = np$ 的泊松分布，则在该位置停留用户数 X 大于某个阈值 X_T 的概率为

$$\begin{aligned} P(X > X_T) &= 1 - P(X \leqslant X_T) \\ &= 1 - \frac{e^{-\mu}}{0!} - \frac{\mu e^{-\mu}}{1!} - \ldots - \frac{\mu^{X_T} e^{-\mu}}{X_T!} \end{aligned} \quad (3)$$

则位置 P_i 的敏感权值可以定义为下式。

$$R(P_i) = M \cdot P(X > X_T) \quad (4)$$

通过式(4)可以为 v_i 与 v_k 之间的其他位置赋予敏感权值，该权值以访问用户数为主要因素。如 $R(P_i) > R_T$ 则为敏感位置，其中 R_T 为系统自定义的敏感阈值。通过访问用户数量的多少，可以给每个位置 P_i 赋予权值。为了使敏感权值 $R(P_i)$ 能更好的反映该点用户访问情况，可以在一天内多个时间段分别计算 $R(P_i)$ 并取全天的平均值 $\overline{R(P_i)}$ 作为该位置的敏感权值，同时能够获取不同时间段某个敏感位置 P_i 的用户访问分布特点。为下文基于用户访问频率的敏感位置多样性比较提供依据。

4.1.2　基于敏感位置多样性的锚点选取方法

本文采用选取用户附近多个其他敏感位置共同构成多边形，并取其形心作为锚点的方法。此时 CS 计算锚点的请求，可能来自位多个敏感位置的用户，增加了用户处于敏感位置的多样性。

选取用户附近的其他它敏感位置，最简单的方法可以利用式(4)位置的敏感权值来区分，但敏感权值只能反映全天用户访问的平均状况，不同位置可能全天用户访问量较为接近，但用户访问高峰时间段(简称高峰时段)不同，式(4)无法细致区分该情况。因此，敏感位置多样性获取一方面要根据式(4)找到距离用户较近的其他类型的敏感位置，另一方面要考虑不同敏感位置高峰访问时间段的相似性，获取那些具有相似高峰时段的敏感位置点(即用户此时间段也可能出现在该敏感位置的点)，避免攻击者利用背景知识分析，如酒吧和医院高峰时段不同，如果用户白天在医院发起查询时选择了酒吧作为满足多样性的混淆位置，显然此时用户出现在酒吧的概率很低，可以排除。同时还要考虑高峰时段类似的两个点可能是同类位置，如上述例子中选取的高峰时段相似的点可能依然是个医院，其位置邻近且高峰时段几乎一致，只不过规模与之不同，可能是更小一些的医院，此时用户位置的多样性就无法保障。

通过分析发现，CS 在为用户计算锚点时，可以将与其邻近的敏感位置点分为 3 类：

A. 高峰时段相似性差距较大的位置，这类位置由于用户频繁访问时间段差距较大，可能导致用户处于不同敏感位置的概率不均匀，所以不能选择此类点作为多样性混淆位置；

B. 高峰时段相似性高度一致，并表现出极高线性相关度的点，这类点可能是与用户所处敏感位置同类的点(如两所位置邻近的酒吧)，为了满足敏感位置多样性，这类点也不能选取；

C. 高峰时段相接近，且满足位置多样性的点，这类点能够保证用户敏感位置的多样性，同时避免高峰时段不同而引发的推断攻击。

当然，区分位置多样性的方法及衡量指标还有很多，本文主要研究受用户访问数量及高峰时段变化影响敏感位置区分方法。

CS 可以获取其覆盖区域敏感位置 P_i 在不同时间段的访问用户数列表 $V_{Pi}=(N_i^{t1},N_i^{t2},\cdots,N_i^{tn})$，其中 N_i^{tn} 表示在时间段 *tn* 内访问用户数。对于 A 类敏感位置，可以比对 2 个用户在不同时段的用户访问列表 V_{Pi} 及 V_{Pk}，在时间维度上找到用户访问频率的差异，间接找到高峰时段的差异，并以此作为度量敏感位置高峰时段相似性的依据。这种时间维度上访问数量变化趋势的相似性的比对，采用余弦相似度作为比对手段，这是由于余弦相似度能够较好的反映两组数据的变化趋势相似度。

$$sim(P_i,P_k)=\frac{\overrightarrow{V_{Pi}}\cdot\overrightarrow{V_{Pk}}}{\left\|V_{Pi}\right\|\cdot\left\|V_{Pk}\right\|}=\frac{\sum_{j=1}^{n}N_i^{tj}N_k^{tj}}{\sqrt{\sum_{j=1}^{n}(N_i^{tj})^2}\sqrt{\sum_{j=1}^{n}(N_k^{tj})^2}} \quad (5)$$

其中，$sim_1(P_i,P_k)\in[0,1]$，值越大表示相似度越高，两个敏感位置高峰时段越接近。根据式(4)可以排除与用户所处敏感位置的高峰时段差距较大的相邻敏感位置。但是式(4)只能反映用户访问量变化趋势的相似性，如在 2 个高峰时段 P_i 与 P_k 的访问用户数量分别为(2 000, 1 000)和(1 000, 500)，这 2 组数据具有相同的变化趋势，都减少了 50%，只是用户访问数量差距较大，且用户变化趋势表现出了一定的线性相关性，此时这两个位置为同类敏感位置的可能性较大，即 B 类位置。为了满足敏感位置多样性，需要排除 B 类节点，我们利用皮尔森相关系数(Pearson Correlation Coefficient)来分析变化趋势相似的位置集合中访问用户数量存在近似线性相关的位置有哪些，并排除这些位置：

$$r(P_i,P_k)=\frac{1}{n-1}\sum_{j=1}^{n}\left(\frac{N_i^{tj}-\overline{N_i}}{s_{N_i}}\right)\left(\frac{N_k^{tj}-\overline{N_k}}{s_{N_k}}\right) \quad (6)$$

其中，$\overline{N_i}$ 和 S_{N_i} 分别为均值和标准差，皮尔森相关系数 *r* 可以描述两组数据的线性相关强弱程度，本文将同类敏感位置看作在相似时间内访问用户数表现出一定的线性相关性的点。当 $|r|$ 趋近于 1 时，线性相关度越高，因此需要将相关系数高的敏感位置点去掉，避免同类敏感位置降低多样性。由于经过式(5)的筛选，2 个变量之间负相关性(r<0)的概率为 0，因此不存在访问用户数量增减负线性相关带来的高峰时段差异较大的可能。

通过式(5)、式(6)，能够筛选掉用户访问高峰时段差距大的敏感位置，同时能够去掉可能的同类敏感位置，余下的敏感位置既能满足多样性需求，也能防止高峰时段差异带来的推断攻击。而这些余下的候选敏感位置，依然需要按照与用户所处敏感位置之间的差异度排序，便于服务器为用户选取最好的多样性敏感位置。本文依然利用访问用户数量作为度量主要因素，以 2 个敏感位置访问用户数间的欧氏距离作为度量排序公式。

$$Dist(P_i,P_k)=\sqrt{\sum_{j=1}^{n}(N_i^{tj}-N_k^{tj})^2} \quad (7)$$

欧氏距离越大，意味着这个候选敏感位置与用户所处的敏感位置访问用户数量差距大，类似时段的访问人数差距越大，且不具备线性相关性，那么这个候选敏感位置表现出的多样性就越强。

当 CS 收到锚点计算请求后 $<uid_{i1}^{'},loc,l,C>$，执行如下锚点生成算法。

算法 1 CS 端敏感位置多样性判断算法

1) **Procedure**：中心服务器 CS_i 收到用户锚点计算请求 $<uid_{i1}^{'},loc,l,C>$

2) **if** $loc\in P_i$ // 表明位置 *loc* 属于某个敏感区域

3) 生成空数组 *Array*

4) 以 *loc* 为中心随机获取第一个邻居网格区域 Z_1，并记录指向该网格区域向量为 $\boldsymbol{v}_1$

5) **for** 每个 $P_k\in Z_1$ **do**

6) 计算 $sim(P_i,P_k)$

7) while $sim(P_i,P_k)>\xi_s$ **do** // ξ_s 表示相似度的某个阈值

8) 计算 $r(P_i,P_k)$

9) **if** $r(P_i,P_k)<\xi_r$ **then** // ξ_r 表示线性相关性的某个阈值

10) 计算 $Dist(P_i,P_k)$

11) $Array\leftarrow P_k$ // P_k 插入数组时按照值 $Dist(P_i,P_k)$ 从大到小排列

12) **return** *Array*

13) **if** $\|Array\|<l$ // Z_1 内敏感位置未满足用户多样性 *l*

14) 顺时针获取下一个邻居网格区域 Z_2，并记录指向该网格区域向量为 $\boldsymbol{v}_2$

15) **while** $\theta(\boldsymbol{v}_1,\boldsymbol{v}_2)<180^\circ$ **do**

16) 重复 5~13

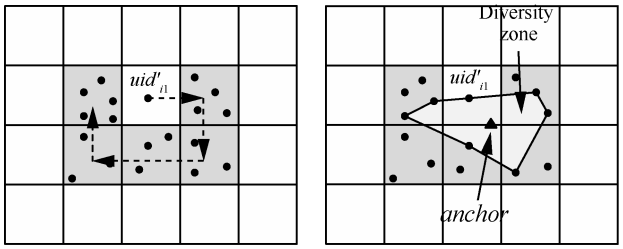
17) 连接前 l 个 $P_k \in Array$ 构成 $Zone_{\mathrm{div1}}$ // 取前 l 个高权值 $Dist(P_i,P_k)$ 构成多样性区域

18) 获取 $Zone_{\mathrm{div1}}$ 形心 loc_{anchor}

19) 将形心 loc_{anchor} 作为锚点，发送给用户 uid'_{i1}

20) **End Procedure.**

在算法 1 中，CS 首先根据其位置 loc 利用式(4)判断用户是否处于敏感位置，是则基于网格选取邻近网格内的多样性位置点，如图 6(a)所示。服务器首先随机选择一个初始方向网格区域，对其中的每个敏感位置执行式(5)、式(6)，分别筛选掉高峰时段相似性差距较大的敏感位置及线性相关度超过一定阈值的同类敏感位置，其次根据式(7)对剩余敏感位置排序，并判断该网格内筛选后的敏感位置是否大于用户定义多样性需求 l，如果满足则按照排序从高到低选取敏感位置，否则暂时不选取。



(a)多样性敏感位置选择　　(b)多样区域及锚点位置选择

图 6　敏感位置多样性转化

然后，按照顺时针方向计算下一个网格，直到满足多样性需求结束，若网格方向与初始方向夹角大于 180° 仍未找到 l 个敏感位置，则宣布查询失败。小于 180° 主要是为了避免用户位置被敏感位置包围，直接被攻击者确定中心网格即为用户所在区域的可能。最后，将敏感位置联接成封闭的多样性区域，将区域的形心作为锚点位置，发送给用户用于查询，如图 6(b)所示。

4.2　均衡增量近邻查询算法

针对相关工作中提出的 SpaceTwist 查询方法中存在的的兴趣点总是以锚点为中心分布、用户在锚点相反方向上查找到的兴趣点极少甚至没有的兴趣点的查全率下降等问题。本节提出用户端执行的均衡增量近邻查询算法(HINN, homogeneous incremental nearest neighbor)。LSP 根据锚点位置逐步将兴趣点候选集发送给用户，直至用户找到围绕自身分布的 K 个兴趣点。

通过分析发现，在图 2(c)及图 3(a)的查询结束时，兴趣点均围绕锚点 q' 分布，此时供应空间完全覆盖了需求空间使得查询结束，但用户在锚点反方向上的兴趣点并没有被查询到。解决该问题一方面需要以为 q' 中心的 INN 查询继续进行，获取用户位置另一侧的兴趣点；另一方面需要确保用户 q 获取距离最近的 K 个兴趣点后查询能够及时结束。

算法 2　用户端均衡增量近邻查询算法 HINN

1) **Procedure**：用户自定义邻居兴趣点数量 K、自身位置 q 及锚点位置 q'，其中 $q' \leftarrow loc_{anchor}$，查询内容 C 及单次返回兴趣点个数 β

2) 按照兴趣点 P_i 与用户位置 q 距离建立小顶堆 W_K

3) 初始化小顶堆 W_K，插入 K 组 $<\mathrm{NULL},\infty>$

4) $\gamma \leftarrow W_K$ 堆底元素 // 需求空间半径赋初值

5) $\tau \leftarrow 0$ // 供应空间半径赋初值

6) 以锚点 q' 为中心向 LSP 发起增量近邻查询 INN

7) **while** $\gamma + dist(q,q') > \tau$ **do**

8) 将 $<uid'_{i2},q',C,\beta>$ 发送给 LSP

9) $S \leftarrow$ 从 LSP 获取兴趣点数据分组

10) $\tau \leftarrow$ 获取 S 中最大的 $dist(q',P_i)$ // 扩大供应空间

11) **for** 每个 $P_i \in S$ **do**

12) **if** $dist(q,P_i) < \gamma$ **then**

13) 用 $P_i, dist(q,P_i)$ 更新 W_K

14) $\gamma \leftarrow dist(q,\mathcal{P}_i)$ // 缩小需求空间

15) $\gamma \leftarrow$ 获取 W_K 中第 K 个兴趣点与 q 的距离 $dist(q,\mathcal{P}_k)$ // 扩大并固定需求空间大小

16) **while** $\gamma + dist(q,q') > \tau$ **do**

17) $S \leftarrow$ 继续从 LSP 获取以 q' 为中心的 INN 查询兴趣点数据分组

18) $\tau \leftarrow$ 获取 S 中最大的 $dist(q',\mathcal{P}_h)$ // 继续扩大供应空间

19) **if** $dist(q,P_h) < \gamma$ **then**

20) 用 $P_h, dist(q,\mathcal{P}_h)$ 更新 W_K

21) 结束 INN 查询

22) **return** W_K 中的前 K 个兴趣点

23) **End Procedure.**

在算法 2 中，用户基于算法 1 计算出的锚点 q'，先建立小顶堆存储距离自身位置 q 最近的兴趣点。然后，用户以 q' 为中心向 LSP 发起增量查询请求，LSP 逐步增量返回兴趣点候选集，用户通过判断返

回兴趣点与自己的距离获取精确结果，算法 6~15 行。如图 2(c)所示，当 P_4 被找到时供应空间覆盖了需求空间，用户找到 K=3 的 3 个最近兴趣点，分别为 $\{P_2,P_1,P_3\}$，按照 SpaceTwist 算法描述查询结束，但 3 个兴趣点是围绕 q' 分布的。为了确保 LSP 端查询能够继续进行，并使用户能够获取围绕 q 的 K 个近邻兴趣点，我们首先扩大需求空间半径 γ，如图 7(a)所示，使需求空间至少覆盖 K 个兴趣点，确保用户在查询结束时至少能够获取 K 个兴趣点(算法第 15 行)。然后，在需求空间不再变化的条件下，继续向 LSP 发起增量查询，使供应空间进一步扩大，直至供应空间再次覆盖需求空间，此时查询结束，算法第 16-21 行。如图 7(c)所示用户共获取了 8 个兴趣点，当 K=3 时，用户获取的 3 个最近兴趣点分别为 $\{P_2,P_5,P_3\}$，与 SpaceTwist 获取兴趣点集合不同，但围绕用户位置 q 分布，更准确。

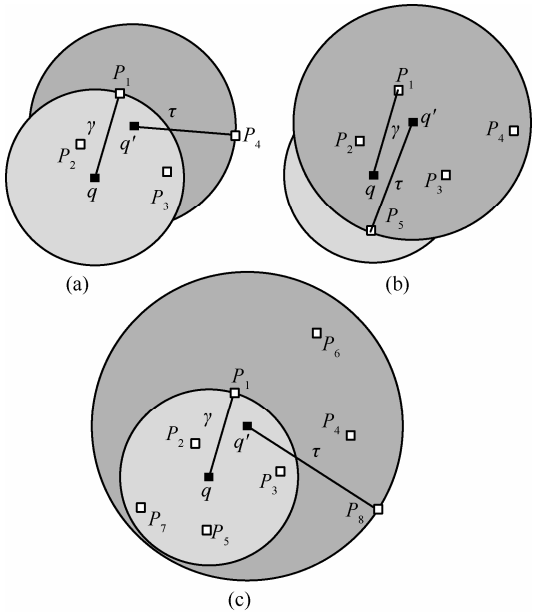


图 7 均衡增量近邻查询过程(HINN)

综上，通过敏感位置相似性的比较，获取满足敏感多样性的锚点，以该锚点位置为中心发起增量查询，通过算法 2 可以获取围绕用户分布的 K 个最近的兴趣点，确保了查询的准确性。

4.3 性能分析

本节主要对锚点选取及利用该锚点进行查询 2 个过程的安全性、查询效率及准确性进行分析。

1) 安全性

处于敏感区的用户无需构造匿名框，减少了中心服务器为用户构造匿名框负担的同时也避免了匿名框仍包含在敏感区内而暴露用户位置隐私的可能。在锚点选取过程中，如果锚点位置选取不当仍存在上述问题。基于敏感位置多样性的锚点选取方法增加了锚点出现在多个敏感位置的可能。经过式(5)余弦相似度的筛选去掉访问频率差异度大的敏感位置，式(6)进一步筛除同类的敏感位置，确保敏感位置的多样性，式(7)可以将最好的 $l-1$ 个敏感位置选取出来混淆用户所在的敏感位置。通过式(5~式 7)，用户出现在 l 个敏感位置的概率均为 $p(x_i)=1/l$，用户单次利用此锚点查询的信息熵为

$$H(q)=\sum_{i=1}^{l}p(x_i)\log\frac{1}{p(x_i)}=\sum_{i=1}^{l}\frac{1}{l}\log l=\log l \quad (8)$$

此时单次信息熵最大，意味着攻击者的不确定性达到理论最高值。用户以这个锚点作为查询标志，攻击者很难将锚点位置与用户所在的敏感区域联系起来。

当 LSP 将兴趣点逐步发送给用户时，用户只需要根据自己的位置计算出所需要的兴趣点即可。任何其他实体都不知道用户的确切位置及所在的敏感区域，用户被包含在半径为 τ 的供应空间内，由于算法 2 第 16~21 行的设计使用户被唯一锁定的概率为 $1/X$， X 为供应空间内用户数。

2) 查询效率及准确性

对于查询效率，假设区域内兴趣点均匀分布，用户以近邻增量查询的方式最多搜索至半径为 γ 的区域内即可获得 K 个兴趣点。则单位面积上的兴趣点可以表示为 $K/\pi\gamma^2$，则 LSP 查询半径为 τ 的圆形区域需要查询 $N_1=\pi\tau^2(K/\pi\gamma^2)=K(\tau/\gamma)^2$ 个兴趣点，如图 7(c) 所示，当算法 2 的查询结束时 $\tau=\gamma+dist(q,q')$，则 $N_1=K(1+dist(q,q'))^2$，由上述假设可知 γ 为固定值，因此锚点 q' 位置选取影响查询效率。图 7(c)显示本方法中 $dist(q,q')<\gamma$，因此有 $N_1<4K$。相比利用 n 个假位置查询兴趣点的方法，LSP 需要的查询 $N_2=nK$ 个兴趣点，当用户选取假位置数高于 4 时，效率明显低于本方法。而本方法 LSP 需要查询的兴趣点个数上限值为 $4K$，而在实际情况中 LSP 在未查询到该数量兴趣点时，查询已经完成。

对于准确性，在返回的兴趣点候选集中，用户利用自己的真实位置进行比较，找到最近的 K 个兴趣点，算法 2 第 15 行扩大并固定需求空间时已经保证至少找到了 K 个兴趣点(以锚点 q' 为中心分布)，当供应空间再次覆盖需求空间时，用户获得了围绕自己分布的 K 个兴趣点，这样用户总能够以接

近100%的准确率获取距离自己最近的K个兴趣点。

通过上述分析，本方法在锚点选取和查询过程中，能够有效利用敏感位置多样性混淆锚点位置以达到保护处于某个敏感区域用户位置隐私的目标，并且在查询过程中用户无需向不可信的 LBS 服务端提供自己的真实位置信息即可获取兴趣点精确查询结果，因此本方法实现了隐私保护和查询质量之间的平衡。

5　实验

本节主要围绕用户获取 LBS 查询服务过程中锚点计算成功率、数据通信量及响应时间 3 个指标在真实数据集和模拟数据集上的实验情况说明本方法的实际性能。

5.1　环境配置

实验在 Windows 7 系统上用 JAVA 语言实现，运行环境为 3.2 GHz Intel Core i5 处理器，内存大小为 2 GB。实验设置了数据集 GDS 和数据集 TDS 2 组对比。GDS 数据集来自美国地名委员会提供的地理数据集。

表 1　　**实验默认参数配置**

参数名	取值范围	默认值
区域内用户总数 U	$2.5 \leqslant U \leqslant 20$	15 万
全局兴趣点数 N	$2.5 \leqslant N \leqslant 15$	5 万
敏感位置用户数阈值 X_T	$100 \leqslant X_T \leqslant 1\,000$	200
敏感位置相似度阈值 ξ_s	$0 \leqslant \xi_s \leqslant 1$	0.4
LSP 单次返回 PoI 数 β	$1 \leqslant \beta \leqslant 11$	6
用户查询 PoI 个数 K	$1 \leqslant K \leqslant 15$	5
用户与锚点距离 $dist(q,q')$	$200 \leqslant dist(q,q') \leqslant 6\,000$	1 000

TDS 数据集为广泛使用的 Thomas Brinkhoff 路网移动节点数据生成器，它以德国奥尔登堡市交通路网数据为基础，城市区域面积约为 24 km×27 km 的矩形区域，用户可以自定义数据集的相关属性。用户与中心服务器 CS 的通信带宽为 3 Mbps。在服务端，每个数据集由 1 kbyte 的 R-tree 结构索引。表 1 为其他参数配置情况。

5.2　锚点计算成功率

实验在 2 类数据集 GDS 和 TDS 上进行，分别比较在敏感位置停留用户数阈值变化及敏感位置相似度阈值的变化对锚点计算成功率的影响。

由图 8 可以看出，当敏感位置停留用户阈值 X_T 逐渐变大时，锚点计算成功率逐渐降低，稳定在80%附近，这是由于 X_T 变大时意味着停留人数相对较少的位置将不再被 CS 认为是敏感位置，CS 在为用户计算锚点时刻利用的敏感位置点的数量减少，因此造成了锚点计算成功率下降。图 9 显示当敏感位置余弦相似度阈值 ξ_s 逐步增大时，某个敏感位置的高峰时段与用户所在敏感位置的相似度必须达到一定程度才能被选作混淆敏感位置，因此会造成部分敏感位置被筛选掉，造成了锚点成功率下降。

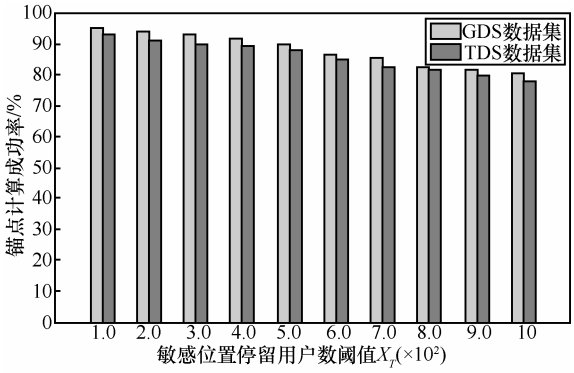


图 8　锚点计算成功率(阈值 X_T 变化)

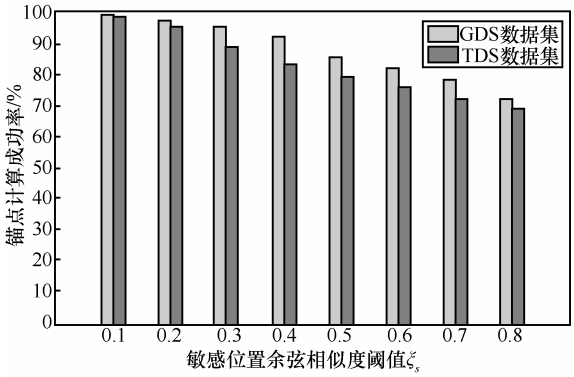


图 9　锚点计算成功率(阈值 ξ_s 变化)

5.3　HINN 与经典方法对比

分别在数据集 GDS 和 TDS 上进行实验，比较 HINN 与 SpaceTwist 在用户查询兴趣点个数 K 变化、锚点与用户距离 $dist(q,q')$ 变化过程中平均数据通信量（简称数据通信量）的表现。

如图 10~图 11，通过 2 类数据集上对比试验，发现当用户查询兴趣点个数 K 增加时，2 种查询方法数据通信量均表现出了增长的趋势，而 HINN 查询方法数据分组整体数据通信量略高，这是由于算法 2 第 15 行需求空间二次扩大造成的，LSP 服务需要继续 INN 查询并发送给用户，这使查询兴趣点围绕用户分布，查询效果更加准确，而通信量增长并不显著。因此，虽然 HINN 扩大了查找范围，可

以达到与 SpaceTwist 类似较少的数据通信量。

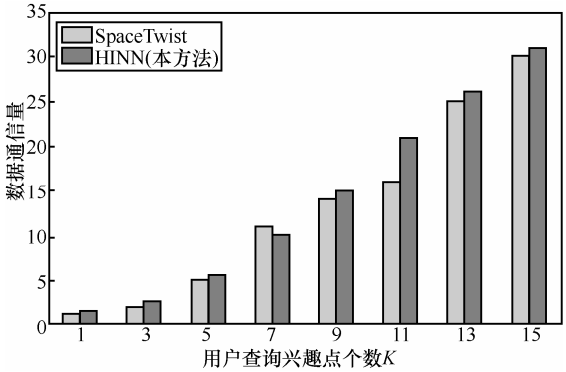


图 10 GDS 数据集通信量随 K 变化趋势

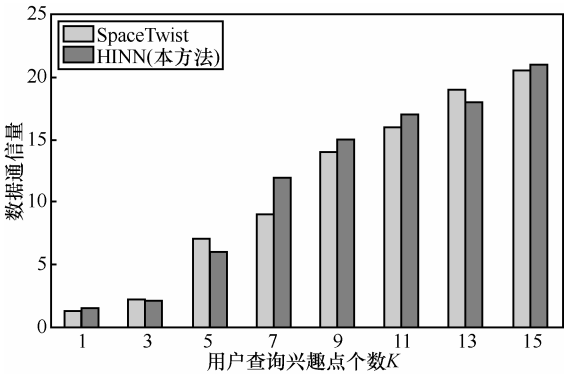


图 11 TDS 数据集通信量随 K 变化趋势

如图 12~图 13，分别在数据集 GDS 和 TDS 上对 4 种利用锚点查询方法的用户通信量进行了比对。当锚点位置远离用户时，意味着 LSP 需要检索更大范围的区域才能获取足够的兴趣点，因此两图中 4 种查询方法的数据通信量均表现出了明显的增长，KAWCR 与 Coprivacy 通信量相对较低，本方法由于需要搜索更大区域，因此通信量具有相对小幅增长，但如 4.3 节中分析 HINN 能够以接近 100%的准确率获取查询结果，这种通信量小幅提高确保了较高的查询服务质量。

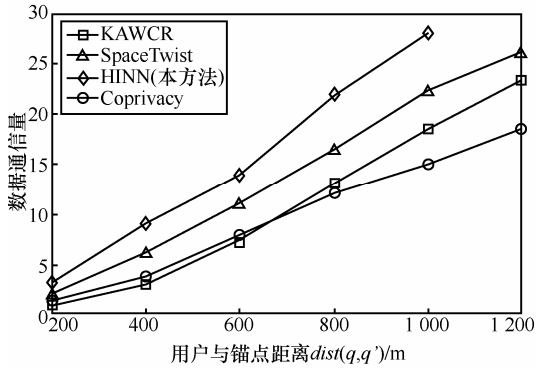


图 12 GDS 数据集通信量随 $dist(q,q')$ 变化趋势

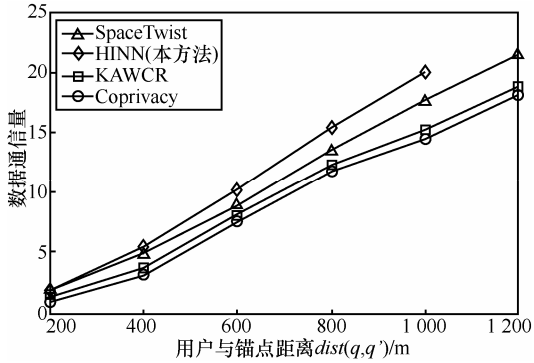


图 13 TDS 数据集通信量随 $dist(q,q')$ 变化趋势

另一方面，还考察了利用锚点查询方法中的查询效率和查询结果精度的变化情况。本文主要考虑在锚点位置逐渐远离用户的过程中服务响应时间和兴趣点 KNN 查询成功率的变化情况。如图 14 所示，K 近邻兴趣点查找成功率在 $dist(q,q')$ 较小时，3 种方法成功率都较低，这是由于锚点选取与用户较近时会导致查询快速结束，用户查询不到 K 个目标兴趣点，因此查询服务质量下降。当 $dist(q,q')$ 超过 800 m 后，查询成功率稳定在 80%以上，本文提出的方法在 90%以上，因此在锚点选取过程中，其与用户的距离应不低于 800 m，否则影响查询结果精准度。然而，锚点越远离用户虽然能够带来较高的查询成功率，但是用户服务体验会受到影响，主要体现在随着增加，服务响应时间延长，LBS 查询效率下降，这是由与用户需要搜索更大空间带来的时延，如图 15 所示，假设用户对服务响应时间的容忍度为不超过 2.5 s，则由该图可以看出，用户选择与锚点距离上限为 4 000 m。综上，用户选择与锚点的合理区间为[800,4 000] m。

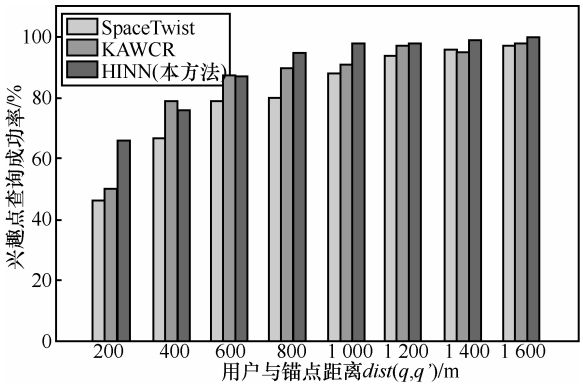


图 14 兴趣点查找成功率随 $dist(q,q')$ 变化趋势

同时，还将本方法与其他位置隐私保护经典方法 PrivacyGrid[4]和 P2P[14]进行了比较，实验在 TDS 数据集上进行。

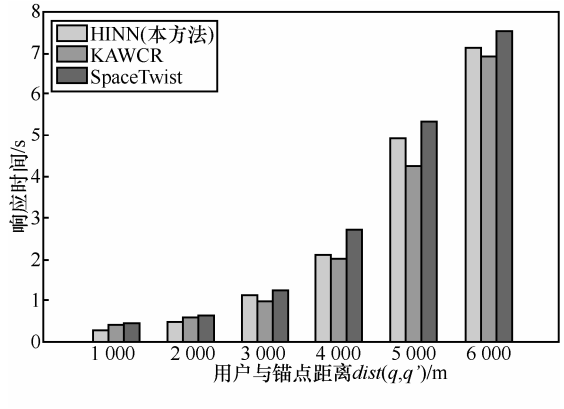


图 15　响应时间随 $dist(q,q')$ 变化趋势

如图 16 可知，当全局用户数量逐渐增加时，3 种方法的数据通信量均有所增加，P2P 方法增长明显，这是由于 P2P 方法需要用户自行组织构造匿名组，每个用户需要与其他多个用户通信。本方法通信量变化并不明显，这是因为本方法利用锚点实现兴趣点查询，用户数量增加只会带来 CS 端计算锚点造成的少量数据通信量增加，而用户分布式查询兴趣点的平均数据通信量不会有明显改变。

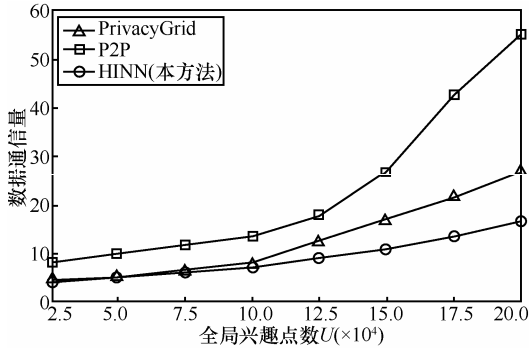


图 16　通信量全局移动用户数量 U 变化趋势

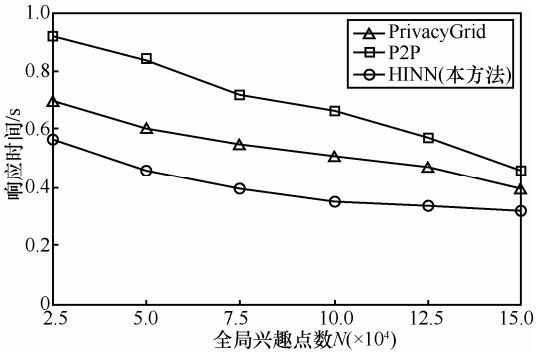


图 17　响应时间随全局兴趣点数量 N 变化趋势

在响应时间方面，如图 17 所示，当全局兴趣点数量N增加过程中，平均响应时间不断随之降低，降幅为 1/10 s 级，这是由于兴趣点数量增多会使用户在 KNN 查找过程中更快找到目标兴趣点，因此响应时间会随之缩短。

综上，本方法与同类位置隐私保护方法、其他经典位置隐私保护方法相比均具有良好工作性能。

6　结束语

本文首先针对 LBS 查询过程中选取的锚点仍位于敏感区域易造成的用户位置隐私泄漏问题，提出基于敏感位置多样性的锚点选取方法，该方法通过提高锚点出现在多个敏感位置的可能性来增加用户所处位置的不确定性。在查询过程中，用户无需提供真实位置，以该锚点为标志向 LSP 发起查询服务，获取兴趣点集。该方法同时改进了 SpaceTwist，使查询兴趣点围绕用户分布，提高了查询的准确性。2 类数据集上的实验表明，本方法在锚点选取和兴趣点查询 2 个过程中均表现出了良好的性能。由于本文主要考虑用户在敏感位置停留或短距离运动带来的隐私保护问题，因此本文并未针对速度运动较快物体的连续查询隐私保护。

同时，本文也存在一些不足，如敏感位置定义因素首要考虑的是用户访问时间和数量等，将在接下来的研究中考虑更多影响因素。

参考文献：

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

[2] NIU B, LI Q, ZHU X, *et al*. Achieving *k*-anonymity in privacy-aware location-based services[A].Proc IEEE INFOCOM[C]. 2014.754-762.

[3] GHINIta G. Privacy for location-based services[J]. Synthesis Lectures on Information Security, Privacy, & Trust, 2013, 4(1): 1-85.

[4] BAMBA B, LIU L, PESTI P, *et al*. Supporting anonymous location queries in mobile environments with privacygrid[A]. Proceedings of the 17th International Conference on World Wide Web[C]. ACM, 2008. 237-246.

[5] 徐建, 徐明, 林欣等. 路网限制环境中基于匿名蜂窝的位置隐私保护[J]. 浙江大学学报 (工学版), 2011, 3: 006.
XU J, XU M, LIN X, *et al*. Location privacy protection through anonymous cells in road network[J]. Journal of Zhejiang University (Engineering Science), 2011, 3: 006.

[7] KIDO H, YANAGISAWA Y, SATOH T. An anonymous communication technique using dummies for location-based services[A]. Pervasive Services, ICPS'05, Proceedings. International Conference[C]. IEEE, 2005. 88-97.

[8] NIU B, ZHANG Z, LI X, *et al*. Privacy-area aware dummy generation algorithms for Location-Based Services[A].Communications (ICC), 2014 IEEE International Conference[C]. IEEE, 2014. 957-962.

[9] HONG J I, LANDAY J A. An architecture for privacy-sensitive ubiquitous computing[A].Proceedings of the 2nd International Conference on Mobile Systems, Applications, and Services[C]. ACM, 2004. 177-189.

[10] YIU M L, JENSEN C S, HUANG X, *et al*. Spacetwist: Managing the trade-offs among location privacy, query performance, and query accuracy in mobile services[A].Data Engineering, ICDE 2008, IEEE 24th International Conference[C]. IEEE, 2008. 366-375.

[11] 杨松涛,马春光,周长利.面向 LBS 的隐私保护模型及方案[J].通信学报,2014,35(8):116 -124.
YANG S T, MA C G, ZHOU C L. LBS-oriented location privacy protection model and scheme[J]. Journal of Communications, 2014,35(8):116 -124.

[12] HUO Z, MENG X, HU H, *et al*. You can walk alone: trajectory privacy-preserving through significant stays protection[A]. Database Systems for Advanced Applications[C]. Springer Berlin Heidelberg, 2012. 351-366.

[13] MOKBEL M F. Towards privacy-aware location-based database servers[A].Data Engineering Workshops, Proceedings. 22nd International Conference[C]. IEEE, 2006. 93-93.

[14] CHOW C Y, MOKBEL M F, LIU X. Spatial cloaking for anonymous location-based services in mobile peer-to-peer environments[J]. GeoInformatica, 2011, 15(2): 351-380.

[19] 毛典辉, 蔡强, 李海生等.路网条件下基于用户协作的 LBS 隐私保护[J]. 高技术通讯, 2013,23(11):1148-1153.
MAO D H, CAI Q, LI H S, *et al*. A collaborative LBS privacy protective method in road-network[J]. Chinese High Technology Letters, 2013,23(11):1148-1153.

[16] 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.
HANG Y, HUO Z, MENG X F. CoPrivacy: A collaborative location privacy-preserving method without cloaking region[J]. Chinese Journal of Computers, 2011, 34(10):1976-1985.

[17] GONG Z, SUN G Z, XIE X. Protecting privacy in location-based services using *k*-anonymity without cloaked region[A]. Mobile Data Management (MDM), 2010 Eleventh International Conference[C]. IEEE, 2010. 366-371.

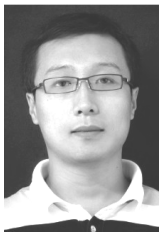
[6] DONDI R, MAURI G, ZOPPIS I. The l-diversity problem: Tractability and approximability[J]. Theoretical computer science, 2013, 511: 159-171.

[15] 薛姣, 刘向宇, 杨晓春, 王斌. 一种面向公路网络的位置隐私保护方法[J]. 计算机学报, 2011, 34(5): 865-878.
XUE J, LIU X Y, YANG X C, WANG B. A location privacy preserving approach on road network[C]. Chinese Journal of Computers, 2011, 34(5): 865-878.

[18] CHOW C Y, MOKBEL M F. Trajectory privacy in location-based services and data publication[J]. ACM SIGKDD Explorations Newsletter, 2011, 13(1): 19-29.

[20] ZHU Z, CAO G. Applaus: A privacy-preserving location proof updating system for location-based services[A]. INFOCOM, 2011 Proceedings IEEE[C]. IEEE, 2011.1889-1897.

作者简介：



周长利（1985-），男，黑龙江哈尔滨人，哈尔滨工程大学博士生，主要研究方向为位置隐私保护、网络与信息安全。

马春光（1974-），男，黑龙江双鸭山人，哈尔滨工程大学教授、博士生导师，主要研究方向为密码学、网络与信息安全。

杨松涛（1972-），男，黑龙江佳木斯人，哈尔滨工程大学博士生，哈尔滨工程大学副教授，主要研究方向为位置隐私保护。