doi:10.11959/j.issn.1000-436x.2016142

连续服务请求下基于假位置的用户隐私增强方法

刘海，李兴华，王二蒙，马建峰
（西安电子科技大学网络与信息安全学院，陕西 西安 710071）

摘 要：基于假位置的隐私保护方案在为用户提供准确位置服务查询结果的同时，还无需第三方和共享密钥。然而，当用户连续请求位置服务时，由于现有保护单次查询的假位置方案未考虑相邻位置集合间的时空关系，使攻击者能推断出假位置，降低用户的位置隐私保护等级。针对上述问题，采用现有假位置方案生成候选假位置，并通过连续合理性检查和单次隐私增强对其进行筛选，提出一个适用于连续请求的假位置隐私保护增强方法。安全性分析表明，所提方法能保证连续请求中形成的移动路径在时空上不可区分，有效保护连续请求中的用户位置隐私。大量实验表明，所提方法在不增加用户计算开销的同时，与采用的候选假位置生成方案相比，还能提高用户单次查询的隐私保护等级。

关键词：基于位置的服务；隐私增强；连续请求；假位置；连续合理性检查
中图分类号：TP309 **文献标识码**：A

Privacy enhancing method for dummy-based privacy protection with continuous location-based service queries

LIU Hai, LI Xing-hua, WANG Er-meng, MA Jian-feng
(School of Cyber Engineering, Xidian Universality, Xi'an 710071, China)

Abstract: Without need for the third party and sharing key, the dummy-based privacy protection scheme enabled users to obtain the precise query result in location-based services. However, in continuous queries, since the existing dummy-based privacy protection schemes ignored the spatio-temporal relevance of the submitted neighbor location sets, the adversary could infer dummies, making that the protection degree of users' location privacy was reduced. To solve this problem, a dummy-based privacy protection enhancing method toward continuous queries was proposed. In the proposal, the candidate dummies were first generated by the existing dummy-based schemes, and could be filtered through the check of continuous reasonability and single privacy enhancement. Security analysis shows that, in the proposed method, the formed movement paths are indistinguishable in time and space, so that protecting the user's location privacy effectively in continuous queries. Moreover, extensive experiments indicate that its computation cost is limited, and compared with the scheme adopted to generate candidate dummies, the user's privacy protection is also enhanced in snapshot query.

Key words: location-based service, privacy enhancing, continuous queries, dummy, check of continuous reasonability

1 引言

基于位置的服务[1~4]（LBS, location-based service）是指服务提供商为用户提供其指定位置的地理信息，或与其指定位置信息相关的其他业务。随着无线通信技术的迅猛发展和移动智能终端设备的普及，LBS 已成为人们日常生活不可或缺的重要组成。

然而，位置服务提供商在为用户提供便捷服务的同时，还可能搜集并滥用用户提交的数据，从而

收稿日期：2015-10-25；**修回日期**：2016-05-20
通信作者：李兴华，xhli1@mail.xidian.edu.cn
基金项目：国家自然科学基金资助项目（No.61372075, No.U1405255, No.61202389, No.61472310）
Foundation Item: The National Natural Science Foundation of China (No.61372075, No.U1405255, No.61202389, No.61472310)

非法获取他们的个人隐私信息。因此，位置服务中的用户隐私保护得到了国内外学者的广泛关注[5~14]。基于假位置的方法[5, 6]作为一种最常用的位置隐私保护方法，与其他位置隐私保护方法相比，它具有以下优点：1) 不依赖第三方；2) 可让用户获得准确的查询结果；3) 无需用户存储密钥。这就使基于假位置的方法已被广泛用于保护单次查询中的用户位置隐私[5, 6, 15~18]。

在现实生活中，除了使用单次 LBS 外，用户还会连续地发送 LBS 请求。例如，在驾车前往度假的途中，司机可能会不断地查询最近加油站的位置；到陌生城市旅行时，游客也会不断地查询周围的景点和美食，从而更好地享受假期。在连续 LBS 请求中，用户提交的相邻位置集合在地理时空上存在紧密的关联性。如果将现有保护单次查询的假位置方案直接用于连续请求场景时，攻击者就可利用相邻位置集合间的地理时空关系，通过识别虚假移动路径的方法正确推测出假位置，降低用户的位置隐私保护等级，甚至推测出用户的真实位置。

针对上述问题，本文利用现有保护单次查询的假位置方案生成候选假位置，从时间可达性检查、方向相似判断和出入度这 3 个方面评估候选假位置的连续合理性，对其进行筛选，提出了一种适用于连续 LBS 请求的基于假位置的用户隐私保护增强方法。此外，本方法还基于个人查询熵和位置分散度，从单次查询隐私保护的角度再次对剩余的候选假位置进行筛选，使最终生成的位置集合在有效保护连续 LBS 请求中用户位置隐私的同时，还能为用户的单次请求提供更好的隐私保护。本文的主要贡献如下。

1) 通过实验发现，现有保护单次查询的假位置方案直接用于连续 LBS 请求场景时，攻击者可通过相邻位置集合间的地理时空关系，正确推测出假位置，从而降低用户的位置隐私保护等级，甚至推测出用户的真实位置。

2) 从时间可达性、方向相似性和出入度这 3 个方面考虑连续请求中相邻位置集合间的地理时空关系，提出一个基于假位置的连续 LBS 请求位置隐私保护增强方法。安全性分析表明，本方法能保证在连续请求中形成的移动路径在地理时空上具有不可区分性，有效保护用户在连续请求中的位置隐私。

3) 大量实验对比表明，与采用的候选假位置生成方案相比，本方法在不降低单次请求中个人查询熵的同时，能增大位置分散度，为用户提供更高的单次查询隐私保护等级。此外，该方案还具有较低的计算开销。

2　相关工作

本节主要介绍现有连续 LBS 请求中的用户位置隐私保护方案和现有保护单次查询的假位置方案。

2.1　连续请求下的位置隐私保护

现有连续请求下的用户隐私保护大多采用 K 匿名的方法。Xu 和 Cai[19]提出利用匿名集合大小来度量 LBS 连续请求中的用户的位置隐私保护需求。在他们的方案中，用户通过在最小边界圆（minimum bounding circle）中查找不在上一匿名区内其他用户的方法来构造匿名集合。文献[20]则通过让用户制定公共区域来表示自己的隐私保护需求，提出了基于感觉（feeling）的连续 LBS 请求用户隐私保护模型。不幸的是，上述 2 个方案均不能避免 LSP 利用查询追踪攻击[21]推测出用户的真实位置。为了抵抗查询追踪攻击，潘晓等[22]通过预测用户未来发送 LBS 请求的位置，让用户在连续请求的最初时刻为所有请求生成统一匿名区，来保护连续请求中的用户位置隐私。Wang 等[23]指出，在连续 LBS 请求中，用户可能具有不同隐私保护需求。他们让匿名服务器查找能满足用户所有隐私需求的历史足迹来构造匿名区。Li 等[24]指出，在利用其他用户历史足迹构造匿名区时，会出现匿名区域过大、服务质量降低的问题。他们通过抑制用户的少量 LBS 请求，在满足用户连续请求中个性化位置隐私保护需求的同时，通过删除最远足迹缩小匿名区面积，提高服务质量。然而，一旦用户在预定/预测位置之外进行 LBS 请求，上述方案仍不能抵抗查询追踪攻击。

Schlegel 等[24]基于密文匹配的思想，提出了首个基于密码学方法的连续 LBS 请求位置隐私保护方案。然而，在他们的方案中，一旦半可信第三方与 LSP 进行合谋，就能缩小用户真实位置所属的区域，降低用户的隐私保护。

与基于假位置的方法相比，现有连续请求位置隐私保护的方法均引入第三方，这会导致用户与第三方之间存在通信瓶颈。并且，在基于 K 匿名的方法中，还要求第三方是完全可信的。然而在现实环境中，完全可信的第三方难以找到，这都降低了现有连续请求隐私保护方案的实用性。

2.2 单次请求下基于假位置的用户隐私保护

基于假位置的用户隐私保护方案最早是由 Kido 等[5, 6]于 2005 年提出的。其基本思想是指用户根据自己的隐私保护需求生成假位置，并将这些假位置与真实位置一同发送给服务提供商，使服务提供商无法正确推测出用户的真实位置。Lu 等[15]指出如果生成的假位置过于集中，会降低用户的位置隐私保护等级。在他们的方案中，用户根据自己的隐私需求，将匿名区域等份划分，使每个位置（包括用户的真实位置）均位于不同的圆半径或矩形顶点上。Niu 等[16]指出，上述方案均未考虑攻击者所拥有的背景知识。一旦攻击者掌握如地图信息等某些背景知识，这些方案生成的如位于河流中和山峰顶的假位置就能被攻击者轻易地识别。因此，他们提出在假位置生成过程中，首先应让每个假位置的查询频率与其真实位置的查询频率尽可能地相等，避免生成不合理的假位置；其次，在确保查询频率不发生改变的同时，让位置分散度变大，从而更好地保护用户的位置隐私。随后，他们还对文献[15]未考虑攻击者背景知识的缺陷进行改进[17]，通过平移原方案生成的各个假位置，使最终生成的位置集合的查询信息熵达到最大。此外，Niu 等[18]又提出利用接入热点的缓存来存储用户历史查询结果的方法，避免用户在同一位置频繁地生成假位置和向 LBS 服务器发送请求，从而降低其通信开销和计算开销。

从上述介绍中可知，现有基于假位置的隐私保护方法的研究仅集中在单次 LBS 请求场景。因此，将这些方案直接用于连续请求场景时，攻击者就能利用相邻位置集合间的地理时空关系正确推测出假位置，甚至直接推测出用户的真实位置。

3 基于假位置的隐私增强方法

针对现有保护单次查询的假位置方案并不适用于连续请求场景的问题，本文采用现有假位置方案生成候选假位置，并利用连续合理性检查和单次请求隐私增加对候选假位置进行筛选，提出了一个适用于连续 LBS 请求的基于假位置的用户隐私增强方法。该方法主要由 2 个部分组成：连续合理性检查算法和单次请求隐私增强算法。其基本框架如图 1 所示。

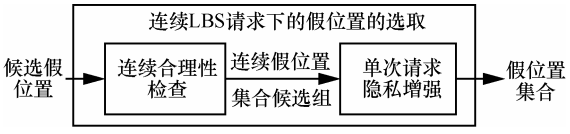


图 1 连续 LBS 请求下基于假位置的用户隐私增强流程

1) 候选假位置

本文的目的是针对现有基于假位置的用户隐私保护方案提出一个适用于连续请求情形的用户隐私保护增强算法。在本方法中，候选假位置可由现有保护单次查询的假位置方案直接生成。对于任意的第 Q_i 次连续 LBS 请求，首先生成的 n 个候选假位置，用集合 $C_i=\{c_1^i,c_2^i,\cdots,c_n^i\}$ 表示。其中，$c_l^i\in C_i$ 表示用户在第 Q_i 次请求中生成第 $l(1\leqslant l\leqslant n)$ 个候选假位置；$n>K_i$，K_i 表示用户在第 Q_i 次连续 LBS 请求中对当前真实位置 c_{real}^i 的隐私保护需求。

2) 连续合理性检查

将用户第 Q_i 次连续请求生成的 n 个候选假位置 $c_1^i,c_2^i,\cdots,c_n^i$ 与用户在第 Q_{i-1} 次请求中最终提交的 K_{i-1} 个位置 $S_{i-1}=\{c_1^i,c_2^i,\cdots,c_{K_{i-1}-1}^i,c_{\text{real}}^i\}$ 进行连续合理性检查，包括时间可达性检查、方向相似性判断和出入度评估，最终得到满足连续合理性检查的连续假位置集合候选组。

3) 单次请求位置隐私增强

对于每一个满足连续合理性检查的连续假位置集合候选组，对其为用户提供的单次隐私保护等级进行比较，选择提供单次隐私保护等级最高的连续假位置集合候选组作为最终生成的假位置集合 $\tilde{C}_i$，与真实位置一起提交给服务提供商。

当用户第一次进行 LBS 请求时，仍由原始方案初始生成 n 个候选假位置，并对 $\binom{n}{K_1-1}$ 个假位置集合进行单次请求的位置隐私增强检查后，得到最终的假位置集合。

3.1 连续合理性检查算法

该算法主要是对利用现有保护单次查询的假位置方案生成的候选假位置进行筛选，使连续请求中形成的移动路径在地理时空上具有不可区分性，具体流程如图 2 所示。候选假位置在经过连续合理性检查后，可得到适用于连续 LBS 请求的连续假位置集合候选组。如果经过连续合理性检查后未产生满足用户位置隐私需求的候选组，即筛选后剩下的假位置个数小于用户的隐私保护需求，则扩大生成的候选假位置个数，并重新生成候选假位置。

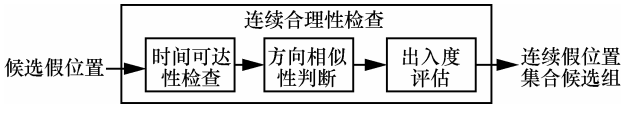


图 2 连续请求的可达性检查流程

3.1.1 时间可达性检查

时间可达性检查的目的是为了让用户在相邻 LBS 请求中形成的虚假移动路径能在请求时间间隔内可达。这将使攻击者无法利用城市交通地图等公共信息，通过识别虚假移动路径的方法，正确推测出某些假位置，从而降低连续请求中用户的位置隐私保护等级。

假设用户相邻2次LBS请求分别为第 Q_{i-1} 次和第 Q_i 次请求，可利用有向图 $G_T =< V_T, E_T >$ 来表示在第 Q_i 次请求中满足时间可达性检查的候选假位置。

1) $V_T = S_{i-1} \cup C_i' \cup \{c_{\text{real}}^i\}$。其中，$C_i' \subseteq C_i$ 表示集合 C_i 中 n 个候选假位置在经过时间可达性检查后，筛选剩余的候选假位置集合，它满足

$$C_i' = \{c^i \mid \forall c^i \in C_i, \exists c^{i-1} \in S_{i-1}: \Delta T \leqslant \sigma_T time < c_{\text{real}}^{i-1}, c_{\text{real}}^i >\}$$

其中，$\Delta T = \left| time < c^{i-1}, c^i > - time < c_{\text{real}}^{i-1}, c_{\text{real}}^i > \right|$ 表示虚假移动路径 $< c^{i-1}, c^i >$ 与真实移动路径 $< c_{\text{real}}^{i-1}, c_{\text{real}}^i >$ 的可达时间差；$time < c_{\text{real}}^{i-1}, c_{\text{real}}^i >= t_i - t_{i-1}$ 表示相邻 2 次请求的时间间隔，t_i 和 t_{i-1} 分别表示发送第 Q_{i-1} 次和第 Q_i 次请求的时间；σ_T 是用户设定的时间可达性检查阈值。

2) $E_T = \{< c^{i-1}, c^i > \cup < \tilde{c}^{i-1}, c_{\text{real}}^i > \mid \tilde{c}^{i-1} \in S_i \setminus \{c_{\text{real}}^{i-1}\} \wedge \Delta T' \leqslant \sigma_T time\ (< c_{\text{real}}^{i-1}, c_{\text{real}}^i >)\}$，表示经过时间可达性检查后，筛选剩下的候选假位置与第 Q_i 次请求最终提交的位置间形成的可达的虚假移动路径。其中，$< \tilde{c}^{i-1}, c_{\text{real}}^i >$ 表示用户在第 Q_{i-1} 次请求中提交的假位置 $\tilde{c}^{i-1}$ 与当前真实位置 c_{real}^i 形成的可达虚假移动路径；$\Delta T' = \left| time(< \tilde{c}^{i-1}, c_{\text{real}}^i >) - time(< c_{\text{real}}^{i-1}, c_{\text{real}}^i >) \right|$。

3.1.2 方向相似性判断

方向相似性判断的目的是为了避免在相邻请求中形成的可达虚假路径的移动方向与用户真实路径的移动方向相差过大，防止攻击者通过移动方向识别出可达虚假移动路径，推测出某些假位置，从而降低连续请求中用户的位置隐私保护等级。

在方向相似性判断中，利用可达的虚假移动路径与真实移动路径间的方向夹角作为判断标准。同样地，利用有向图 $G_D =< V_D, E_D >$ 来表示第 Q_i 次请求下满足方向相似性判断的候选假位置。

1) $V_D \subseteq V_T$，$E_D \subseteq E_T$；

2) $V_D = S_i \cup C_i'' \cup \{c_{\text{real}}^i\}$。

其中，$C_i'' \subseteq C_i'$ 表示经过方向相似性判断后，筛选剩下的候选假位置集合，其满足

$$C_i'' = \{\hat{c}^i \mid \forall \hat{c}^i \in C_i', \exists\ c^i \in S_i: direction\ (< c^{i-1}, \hat{c}^i >, < c_{\text{real}}^{i-1}, c_{\text{real}}^i >) \leqslant \sigma_D\}$$

其中，$direction\ (< c^{i-1}, \hat{c}^i >, < c_{\text{real}}^{i-1}, c_{\text{real}}^i >)$ 表示可达的虚假移动路径 $< c^{i-1}, \hat{c}^i >$ 和真实移动路径 $< c_{\text{real}}^{i-1}, c_{\text{real}}^i >$ 间的方向夹角；σ_D 是由用户指定的方向相似性判断阈值。

3) $E_D = \{< c^{i-1}, \hat{c}^i > \cup < \hat{c}^{i-1}, c_{\text{real}}^i >\}$，表示经过方向相似性判断后，筛选剩下的候选假位置以及当前真实位置 c_{real}^i 与第 Q_{i-1} 次请求提交的位置集合间形成的方向相似的可达虚假移动路径。其中，$< \hat{c}^{i-1}, c_{\text{real}}^i > \in E_T$ 表示用户在第 Q_{i-1} 次服务请求中提交的假位置 $\hat{c}^{i-1}$ 与当前真实位置 c_{real}^i 形成的与真实路径 $< c_{\text{real}}^{i-1}, c_{\text{real}}^i >$ 移动方向相似的可达虚假移动路径。

3.1.3 出入度评估

本文借用出入度的概念来度量相邻请求中形成移动路径的数量，即用各个位置的出度来表示以它为起点的移动路径的数量，用入度来表示以它为终点的移动路径的数量。出入度评估是为了避免在相邻请求形成的方向相似的可达移动路径中，用户真实位置的出入度值与假位置的出入度值相差过大，使攻击者能以较大概率直接推测出用户的真实位置。为了实现上述目的，分别将用户真实位置的出度值和入度值作为期望值，利用方差来度量生成的候选假位置集合中假位置出度值与入度值的波动情况。

本文继续用有向图 $G_N =< V_N, E_N >$ 来表示第 Q_i 次请求中满足出入度评估后，筛选剩下的候选假位置。

1) $V_N \subseteq V_D$，$E_N \subseteq E_D$；

2) $V_N = S_i \cup C_i''' \cup \{c_{\text{real}}^i\}$。

其中，$C_i''' = \breve{C}_i^1 \cup \breve{C}_i^2 \cup \cdots \cup \breve{C}_i^m$ 表示有 m 个连续假位置集合候选组。$\breve{C}_i^j = \{c_{l_1}^j, c_{l_2}^j, \cdots, c_{l_{K_{i+1}-1}}^j\}$ 表示经过出入度评估后，从筛选剩下的候选假位置中任意选择 $K_i - 1$ 个假位置，组成的第 j 个连续假位置集合候选组，它满足

$$\begin{cases} \forall\ c_l^j \in \breve{C}_i^j,\ c_l^j \in C_i'' \\ Var\_in(c_{\text{real}}^i, S_{i-1}, \breve{C}_i^j \cup \{c_{\text{real}}^i\}) \leqslant \sigma_{\text{in}-D} \\ Var\_out(c_{\text{real}}^{i-1}, S_{i-1}, \breve{C}_i^j \cup \{c_{\text{real}}^i\}) \leqslant \sigma_{\text{out}-D} \end{cases}$$

其中，$Var\_in(c_{\text{real}}^{i}, S_{i-1}, \breve{C}_i^j \cup \{c_{\text{real}}^{i}\})$ 表示以 c_{real}^{i} 的入度为期望值，计算得到的集合 $\breve{C}_i^j \cup \{c_{\text{real}}^{i}\}$ 的入度方差；$Var\_out(c_{\text{real}}^{i-1}, S_{i-1}, \breve{C}_i^j \cup \{c_{\text{real}}^{i}\})$ 表示以 c_{real}^{i-1} 的出度为期望值，计算得到的集合 S_{i-1} 的出度方差；$\sigma_{\text{in}-D}$ 和 $\sigma_{\text{out}-D}$ 是由用户指定的入度和出度评估阈值。

3) $E_N = \{< c^{i-1}, c_l^i > \cup < c^{i-1}, c_{\text{real}}^{i} > | < c^{i-1}, c_{\text{real}}^{i} > \in E_D\}$ 表示经过出入度评估后，筛选剩下的候选假位置与第 Q_i 次请求最终提交的位置间形成的连续合理的虚假移动路径。

综上所述，连续合理性检查算法如下所示。

算法 1 连续合理性检查算法

输入 Q_{i-1} 次请求最终生成的位置集合 $S_{i-1} = \{c_1^{i-1}, \cdots, c_{K_{i-1}-1}^{i-1}, c_{\text{real}}^{i-1}\}$；

Q_i 次请求的候选假位置候选集合 $C_i = \{c_1^i, c_2^i, \cdots, c_n^i\}$；

输出 Q_i 请求的连续假位置集合候选组 $C_i''' = \breve{C}_i^1 \cup \breve{C}_i^2 \cup \cdots \cup \breve{C}_i^m$；

1) for each $c^i \in C_i$ do

2) if $\exists\ c^{i-1} \in S_{i-1}$，$|time < c^{i-1}, c^i > - time < c_{\text{real}}^{i-1}, c_{\text{real}}^{i} > | \leqslant \sigma_T$ and

$direction\ (< c^{i-1}, \tilde{c}^i >, < c_{\text{real}}^{i-1}, c_{\text{real}}^{i} >) \leqslant \sigma_D$ then

3) $C_i'' \leftarrow c^i$;

4) end if

5) end for

6) if $|C_i''| < K_i - 1$ then

7) exit;

8) end if

9) 从集合 C_i'' 中随机选取 $K_i - 1$ 个假位置组成 $m' = \begin{pmatrix} K_i - 1 \\ |C_i''| \end{pmatrix}$ 个位置集合 $\breve{C}_i^1, \breve{C}_i^2, \cdots, \breve{C}_i^{m'}$；

10) for each $j \in [1, m']$ do

11) if $Var\_in(c_{\text{real}}^{i}, S_{i-1}, \breve{C}_i^j \cup \{c_{\text{real}}^{i}\}) \leqslant \sigma_{\text{in}-D}$ and

$Var\_out(c_{\text{real}}^{i-1}, S_{i-1}, \breve{C}_i^j \cup \{c_{\text{real}}^{i}\}) \leqslant \sigma_{\text{out}-D}$ then

12) $C_i'' \leftarrow \breve{C}_i^j$;

13) end if

14) end for

15) return C_i'''

3.2 单次请求的位置隐私增强

当生成的候选假位置经过连续合理性检查算法筛选后，可得到满足连续合理性的连续假位置集合候选组。当把任意一个连续假位置集合候选组与当前真实位置 c_{real}^{i} 一同提交给服务提供商时，由于它们与第 Q_{i-1} 次服务请求中最终提交的位置集合所形成的移动路径在地理时空上具有不可区分性，因此攻击者将无法推测出假位置，从而有效保护连续请求中用户的位置隐私，并且，为了避免降低用户单次请求中的位置隐私保护等级，本文还利用个人查询熵和位置分散度[4,16]，对上述生成的连续假位置集合候选组再次进行筛选，使最后剩下的连续假位置集合候选组还能从单次请求隐私保护的角度，为用户提供最高的位置隐私保护等级。其流程如图 3 所示。

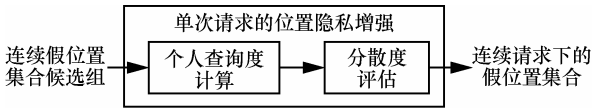


图 3 单次请求的位置隐私增强流程

当攻击者长期收集某用户发送的服务请求时，他能计算出这个用户在每个位置上发送服务请求的次数。若用户在生成的假位置上的个人查询次数与其在真实位置上的个人查询次数相差较大，则攻击者就能推测出用户此次请求的假位置，从而降低用户真实位置的隐私保护等级。为了避免上述问题，利用个人查询熵对经过连续合理性检查后形成的连续假位置集合候选组进行度量，确保在最终筛选出的连续假位置集合候选组中，用户在每个假位置的个人查询次数与其真实位置的个人查询次数尽可能相同。个人查询熵越大，则表示该连续假位置集合候选组中每个假位置的个人查询次数与真实位置的个人查询次数就越相近。因此，用户会将筛选出的个人查询熵最大的连续假位置集合候选组作为最终的假位置集合，与用户真实位置一同发送给服务提供商。

此外，若存在多个连续假位置集合候选组的个人查询熵相等且均是最大值时，将利用位置分散度对它们再次进行筛选，得到最终的假位置集合。假位置集合的位置分散度越大，则表示其所形成的区域面积就越大，能避免由于生成的假位置过于集中，使攻击者能推测出用户真实位置所属区域的隐私泄露问题。具体的单次请求位置隐私增强算法如下所示。

算法 2 单次请求位置隐私增强算法

输入 Q_i 次请求的假位置集合候选组 $C_i''' = \breve{C}_i^1 \cup$

$\breve{C}_i^2 \cup \cdots \cup \breve{C}_i^m$；

输出　Q_i 请求的最终假位置集合 $\tilde{C}_{i+1}=\{c_1^i, c_2^i, \cdots, c_{K_i-1}^i\}$；

1) for each $\breve{C}_i^l \in C_i'''$ do

2)　$A_l = \breve{C}_i^l \cup \{c_{\text{real}}^i\}$；

3)　$F[l] \leftarrow H(A_l)$；　//计算集合 A_l 的个人查询熵

4) end for

5) $A' \leftarrow \arg\max\_entropy(F[\cdot])$；

6) if $|A|=1$ and $\breve{C}_i^{l'} \in A$ then

7)　$\tilde{C}_i = \breve{C}_i^{l'}$；

8)　return $\tilde{C}_i$；

9) end if

10) for each $\breve{C}_i^{l'_q} \in A'$ do

11)　$A''[q] \leftarrow D(\breve{C}_i^{l'_q})$；　//计算集合 $\breve{C}_i^{l'_q}$ 的位置分散度

12) end for

13) $A''' \leftarrow \arg\max\_dispersion(A''[\cdot])$；

14) $\tilde{C}_i \leftarrow A''' \setminus \{c_{\text{real}}^i\}$；

15) return $\tilde{C}_i$

3.3　安全性分析

在本文所提的基于假位置的连续 LBS 请求用户隐私保护增强方法中，候选假位置是由现有保护单次查询的假位置方案直接生成的。因此，本文不再从单次查询的角度对本方法能否抵抗相应的推测攻击进行阐述，而仅通过安全性分析说明本方法能保证在连续请求中形成的可达移动路径在地理时空上具有不可区分性，使攻击者无法通过识别可达的虚假移动路径的方法推测出假位置，有效保护连续请求中用户的位置隐私。

与现有基于假位置的隐私保护方案一样，假设攻击者的目的是得到用户的准确位置，从而获得与用户位置信息紧密相关的个人隐私信息。本文假定攻击者具有如下的背景知识：1) 攻击者能攻陷 LBS 服务器，即他能知道用户提交的 LBS 请求的时间和位置信息。其中，位置信息包括用户的真实位置和生成的假位置。此外，攻击者还会长期收集并统计该用户在各个位置发送请求的频率；2) 攻击者具备地图背景知识，即他知道任意 2 个位置的距离和地图显示的可达时间。

在本文所提的连续合理性检查算法中，通过时间可达性检查、方向相似性判断和出入度评估对候选假位置进行筛选，最终生成假位置集合候选组。从时间可达性检查描述中可以发现，在利用时间可达性检查，对第 Q_i 次请求生成的候选假位置进行筛选后，可保证筛选剩下的候选假位置与第 Q_{i-1} 次请求提交的位置间至少存在 1 条在 $\sigma_T time < c_{\text{real}}^{i-1}, c_{\text{real}}^i >$ 时间内可达的虚假移动路径。随着 σ_T 的变小，所形成的可达虚假移动路径的可达时间与用户真实移动路径的可达时间越相近。并且，方向相似性判断又对经时间可达性检查筛选后剩下的候选假位置再次进行筛选，使筛选后剩下的候选假位置与第 Q_{i-1} 次请求最终提交的位置间形成的可达虚假移动路径与用户真实移动路径间的方向夹角不大于 σ_D。随着 σ_D 的变小，所形成的可达虚假路径在移动方向上与用户真实路径的移动方向越相似。而在最后的出入度评估中，所形成的连续假位置集合候选组能避免用户真实位置的出入度值与假位置的出入度值相差过大，使攻击者能以较大概率推测出用户的真实位置。随着用户指定的评估阈值 $\sigma_{\text{in}-D}$ 和 $\sigma_{\text{out}-D}$ 的不断变小，在位置集合 S_{i-1} 与 $\breve{C}_i^j \cup \{c_{\text{real}}^i\}$ 形成的方向相似的可达移动路径中，S_{i-1} 中各假位置的出度与 c_{real}^{i-1} 的出度越接近，而 $\breve{C}_i^j$ 中各假位置的入度则与 c_{real}^i 的入度也越接近。因此，当用户与攻击者具有相同的连续合理性检查知识时，本方法能保证在连续请求中形成的移动路径在地理时空上具有不可区分性，使攻击者无法通过识别虚假移动路径的方法推测出假位置，从而有效保护连续请求中用户的位置隐私。

4　实验仿真

为了说明现有假位置隐私保护方案不适用于连续请求场景，本文进行了大量的实验仿真。在此，本文选取 2014 年 INFOCOM 会议上的方案[16]作为实例，说明将该方案直接用于连续请求场景时，并不能有效保护用户的位置隐私。并通过与其进行比较，表明本方法在不增加用户计算开销的同时，不仅能有效保护连续请求中用户的位置隐私，还能为用户的单次请求提供更高的隐私保护等级。

4.1　实验数据及平台

本文首先选取 enhanced-DLS 算法[16]为本方案生成候选假位置。为了保证候选假位置的筛选成功率，首先生成 $4K$ 个候选假位置，其中，K 表示用

户的位置隐私保护需求。当经过时间可达性检查、方向相似性判断和出入度评估后，若筛选剩下的候选假位置个数少于 $K-1$，就再次使用 enhanced-DLS 算法重新生成 $8K$ 个候选假位置。

本文采用基于网络的移动对象生成器（network-based generator of moving objects）[26]生成实验数据。该生成器是以德国城市 Oldenberg（面积大约为 16 km×16 km）的城市交通路线图为基础，通过设置移动速度，可生成一系列用户的移动路径。选取该地图中心部分作为实验区域，面积为 8 km×8 km，将其划分为 6 400 个网格，其中，每个网格的面积为 100 m×100 m。并且，本文使用上述生成器的默认设置，随机生成 38 000 条移动路径（包含约 150 000 个不同的位置信息）作为用户历史 LBS 请求数据，用于统计用户在各个网格中所发送的历史服务请求次数。虚假移动路径的可达时间是通过查询 Google 地图获得。另外，在本实验中，用户的位置隐私需求 K 的变化范围为 3~20，并针对不同的 K 值，分别做了 10 次连续 LBS 请求假位置生成实验。本实验的算法均采用 C++编程语言实现，实验环境为 3.00 GHz Core 2 Duo CPU，4 GB DDR3-1600 RAM，操作系统为 Windows 7-64 bit。

4.2 连续请求对现有基于假位置方案带来的影响

为了证明现有保护单次查询的假位置保护方案并不适用于连续请求场景，本文选用 enhanced-DLS 算法[16]为用户不同的隐私保护需求，分别连续生成 10 次服务请求位置集合。具体实验过程如下：1) 针对不同的隐私保护需求，使用缺省设置，利用移动对象生成器为分别生成 1 条包含 10 个位置的移动路径，从而模拟用户连续 10 次发送服务请求时的真实位置；2) 根据每条移动路径中的 10 个真实位置的经纬度，确定其所属的网格，并统计相应网格的用户历史查询次数；3) 利用 enhanced-DLS 算法为每个真实位置生成相应的假位置集合；4) 利用 Goolge 地图获取相邻请求是可达时间，并将该时间作为用户相邻 2 次请求的时间间隔；5) 构造用户相邻请求各位置间的移动路径，设置时间可达性检查阈值 $\sigma_T=\frac{1}{2}$ 对它们进行筛选，获得用户相邻请求中形成的可达移动路径，若在第 Q_{i+1} 次请求生成的位置集合中的每个位置至少能与第 Q_i 次请求生成的位置集合中的任意一个位置构成可达的移动路径，则继续进行方向相似性判断；6) 设置方向相似性判断阈值 $\sigma_D=75°$ 对上述形成的可达移动路径进行判断，若无法筛选出均由假位置构成的可达虚假移动路径，就继续进行出入度评估；7) 统计相邻请求中，形成的方向相似的可达移动路径，并基于此计算出各位置的出入度，对假位置进行识别。在本文中，其他部分的实验过程基本与本部分的实验相似，唯一不同的就是各阈值的设定。因此，在之后的实验中不在对具体的实验过程进行描述。

针对不同的隐私保护需求 K 值，在这 10 次连续 LBS 请求中，最多只有 4 次连续请求生成的位置集合能通过连续合理性检查，如图 4 所示。而在其他请求中，由于生成的位置集合因在经过连续合理性检查后，剩下的假位置数量小于用户的隐私保护需求，从而无法保护连续请求中的用户位置隐私。

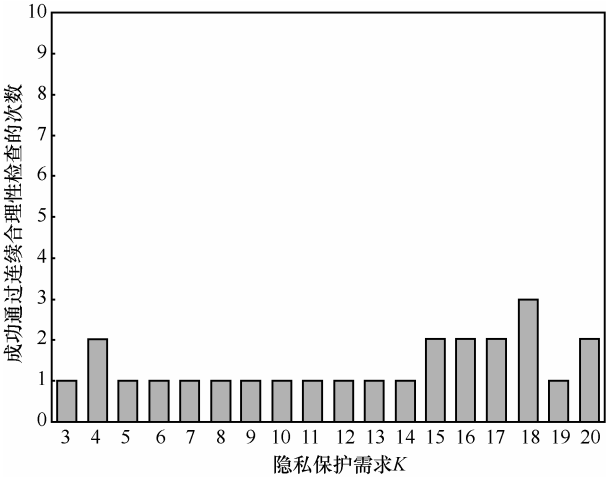


图 4 enhanced-DLS 假位置算法应用在连续请求情形下的通过次数

因此，enhanced-DLS 算法并不能有效保护连续请求中的用户位置隐私。而造成这一问题的根本原因就是用户在连续请求场景下生成假位置时，并未考虑生成的位置集合与上次请求生成的位置集合间的地理时空关系，从而使 LSP 能通过识别虚假移动路径的方法推测出生成的假位置。

4.3 本方法提供的隐私保护等级评估

1) 连续请求场景中的隐私保护等级

在连续 LBS 请求中，利用连续请求中所形成的连续合理的移动路径数量对用户位置隐私保护等级进行度量。当设定用户连续发送 10 次 LBS 请求、时间可达性检查阈值 $\sigma_T=\frac{1}{2}$、方向相似性判断阈值 $\sigma_D=75°$、出入度评估阈值 $\sigma_{\text{in}-D}$ 和 $\sigma_{\text{out}-D}$ 分别选择

各次实验中的计算得到的最小方差值时，本方案在相邻 2 次请求中，所生成的可达移动路径数量如图 5 所示。

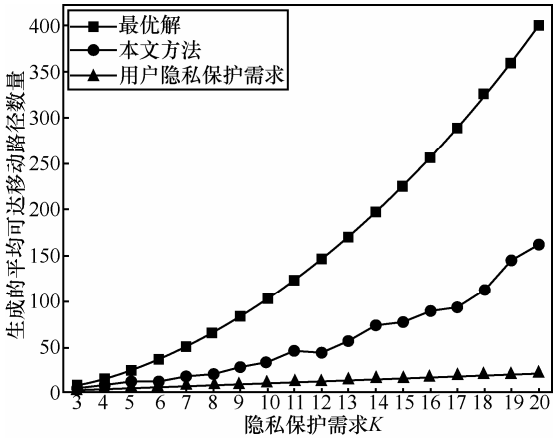


图 5　连续合理的移动路径数量与理想情况的对比

本文用“最优解”来表示的理想状态下应形成的移动路径数量。通过图 5 可以发现，随着用户位置隐私保护需求的提高，理想状态下应形成的移动路径数量与本文方法生成出的位置集合间形成的连续合理的移动路径的数量均随之增多。以用户在第 Q_i 次和第 Q_{i+1} 次请求时的隐私保护需求 $K_i = K_{i+1} = 6$ 为例，在理想状态下应形成 36 条连续合理的移动路径。然而，由于用户的下次服务请求具有很强的随机性，如果将出入度评估的阈值设置为 $\sigma_{\text{in}-D} = \sigma_{\text{out}-D} = 0$，将难以生成满足该条件的连续假位置集合候选组。虽然，本文在实验中将出入度评估阈值 $\sigma_{\text{in}-D}$ 和 $\sigma_{\text{out}-D}$ 分别选择为各次实验中的计算得到的最小方差值，使本方案在相邻 2 次请求中形成的连续合理的移动路径数量少于理想状态，但是仍能满足用户的隐私保护需求。

2) 单次请求场景中的隐私保护等级

基于个人查询熵和位置分散度，从单次位置隐私保护的角度，对本文提出的隐私增强方法与 enhanced-DLS 算法进行对比，以说明本方法在有效保护连续请求中用户位置隐私的同时，也能为用户提供更高的单次隐私保护等级。在本实验中，仍设定用户连续发送 10 次 LBS 请求、时间可达性检查阈值 $\sigma_T = \frac{1}{2}$、方向相似性判断阈值 $\sigma_D = 75°$、出入度评估阈值 $\sigma_{\text{in}-D}$ 和 $\sigma_{\text{out}-D}$ 分别选择各次实验中的计算得到的最小方差值。

本方法与 enhanced-DLS 算法生产的位置集合的个人查询熵如表 1 所示。从该表中可以发现，随着用户隐私保护需求 K 的增大，enhanced-DLS 算法和本文提出的适用于连续请求下的基于假位置的用户隐私增强算法所产生的信息熵均在增大。当用户隐私保护需求从 $K = 3$ 变化到 $K = 7$ 时，所生成位置集合的个人查询熵值完全相等；而当 $K = 20$ 时，本方法所生成的位置集合的个人查询熵仅比 enhanced-DLS 算法降低了 0.003。这表明本方法在利用连续合理性检查对 enhanced-DLS 算法生成的候选假位置筛选后，从个人查询熵的角度来说，并没降低单次请求时用户的隐私保护等级。

表 1　　**平均信息熵值的对比**

K	enhanced-DLS	本文
3	1.098 61	1.098 61
5	1.609 44	1.609 44
7	1.945 91	1.945 91
11	2.397 89	2.397 02
15	2.708 05	2.706 39
17	2.833 21	2.831 00
20	2.995 73	2.992 73

下面将对本文提出的隐私增强方法与 enhanced-DLS 算法在单次请求情形下生成的位置集合的匿名区面积和位置分散度进行比较，分别如图 6 和图 7 所示。

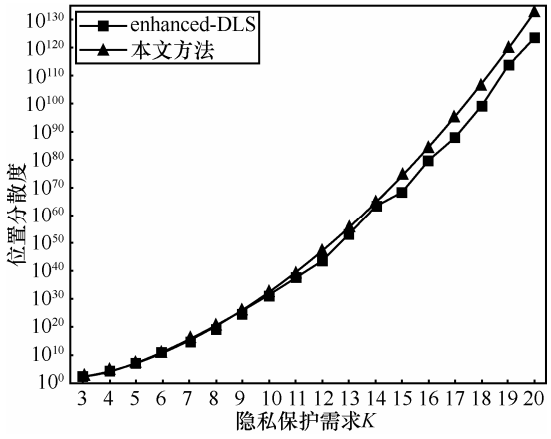


图 6　位置分散度比较

图 6 和图 7 表明，随着用户隐私保护需求 K 的提高，本方法与 enhanced-DLS 算法所生成的位置集合的位置分散度和形成的匿名区面积均呈增大趋势。位置分散度反映了各位置之间距离的乘积。分散度越大，各位置间的距离就越大，

所形成的匿名区域面积也就越大。虽然在图 7 中，匿名区域面积的增大存在一定的波动，但是造成这个波动的原因是由于本方法与 enhanced-DLS 算法均考虑假位置与真实位置间的查询频率。通过对比可以发现，本方法所生成的位置集合的分散度与形成的匿名区域面积始终大于 enhanced-DLS 算法，这表明与 enhanced-DLS 算法相比，本方法能为用户提供更好的单次查询隐私保护等级。

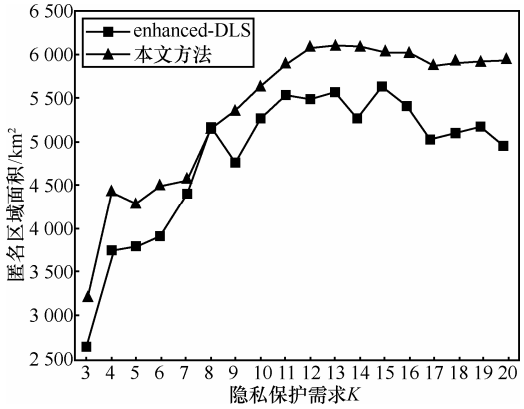


图 7 匿名区域面积比较

4.4 计算开销

首先分析时间可达性检查阈值 σ_T 与方向相似性判断阈值 σ_D 的设定对本方案在计算开销上的影响。由于在发送当前服务请求时，用户无法预测下次进行服务请求时的真实位置，这就使在连续请求所形成的方向相似的可达移动路径中，用户真实位置的出入度具有很强的随机性。因此，在此并不考虑出入度评估阈值 $\sigma_{\text{in}-D}$ 和 $\sigma_{\text{out}-D}$ 的设定对本方案在计算开销上的影响。

在这部分实验中，设定用户的隐私保护需求 $K=6$，连续发送 10 次 LBS 请求，出入度评估阈值 $\sigma_{\text{in}-D}$ 和 $\sigma_{\text{out}-D}$ 分别选择各次实验中的计算得到的最小方差值。即在利用出入度对候选假位置集合进行筛选时，选择出入度方差最小的候选假位置集合进行单次隐私增强筛选。在对时间合理性判断阈值设定对本文所提方案的性能影响实验中，设定方向相似性的阈值 $\sigma_D=75°$；而在对方向相似性的判断阈值变化对本文所提方案的性能影响实验中，设定时间可达性检查阈值 $\sigma_T=\frac{1}{2}$。具体的实验结果如图 8 所示。

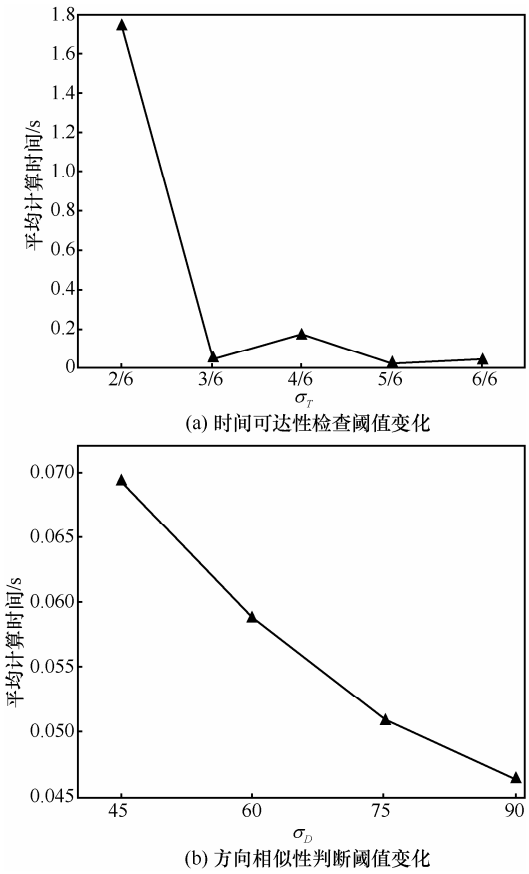


图 8 本方法性能测试

通过图 8 可以发现，随着时间可达性检查阈值 σ_T 与方向相似性判断阈值 σ_D 的不断变大，本方法所需要的计算时间也不断变小。其主要原因是：随着阈值的变大，满足时间可达性检查和方向相似性判断的候选假位置逐渐增多，减少了由于筛选后剩下的候选假位置个数不满足用户隐私保护需求情况的出现，从而避免重新生成更多的候选假位置。以时间可达性检查阈值 σ_T 的设定为例，当 $\sigma_T=\frac{2}{6}$ 时，连续 10 次请求生成最后位置集合的平均时间为 1.75 s，而当 $\sigma_T=\frac{3}{6}$ 时，连续 10 次请求生成最后位置集合的平均时间为 0.05 s。

下面给出 K 值变化对本文所提方案的运行时间的影响，从而说明所提方案具有较好的实用性，如表 2 所示。在这部分实验中，设定用户连续发送 10 次 LBS 请求、时间可达性检查阈值 $\sigma_T=\frac{1}{2}$、方向相似性判断阈值 $\sigma_D=75°$、出入度评估阈值 $\sigma_{\text{in}-D}$ 和 $\sigma_{\text{out}-D}$ 分别选择各次实验中的计算得到的最小方差值。在此要强调的是，仅罗列出用户的计算开销，

表 2　　本方法所需的计算时间

K	连续合理性检查/ms			单次请求位置隐私增强/ms	
	时间可达性检查	方向相似性判断	出入度评估	个人查询熵计算	位置分散度评估
3	0.414	13.5	27.2	2.6	0.8
4	0.168	11.9	4.0	8.0	1.4
5	0.277	15.3	5.6	26.7	1.2
6	0.314	387.1	13.1	125.2	2.3
7	0.385	27.2	46.4	394.6	4.8

而不考虑访问地图接口的时间（受网速和接口的限制等客观原因的约束）。

5　结束语

本文首先通过实验证明现有保护单次查询的假位置方案并不适用于连续请求场景。造成上述问题的根本原因是用户提交的相邻位置集合间具有较为紧密的地理时空关系，使攻击者能通过识别虚假移动路径的方式正确推测出某些假位置，甚至直接推测出用户的真实位置。针对该问题，分别从时间可达性、方向相似性和出入度 3 个方面考虑相邻请求中位置集合间的时空关系，对利用现有保护单次查询的假位置方案生成的候选假位置进行筛选，使用户在相邻请求中形成的连续合理的移动路径数量远大于其位置隐私保护需求，有效保护连续查询中的用户位置隐私。并且，本文还基于个人查询熵和分散度，从单次查询隐私保护的角度再次对剩余的候选假位置进行筛选，使本方案与采用的候选假位置生成方案相比，在不降低单次请求中个人查询熵的同时，能增大位置分散度，为用户提供更高的单次隐私保护等级。同时，还通过大量实验证明了方案的有效性和实用性。

参考文献：

[1] KRUMM J. A survey of computational location privacy[J]. Personal and Ubiquitous Computing, 2009, 13(6): 391-399.

[2] TIWARI S, KAUSHIK S, JAGWANI P, et al. A survey on LBS: system architecture, trends and broad research areas[C]//LNCS 7108: The 7th International Conference on Databases in Network Information Systems. Heidelberg: Springer. c2011: 223-241.

[3] 王璐, 孟小峰. 位置大数据隐私保护研究综述[J]. 软件学报, 2014, 25(4): 693-712.
WANG L, MENG X F. Location privacy preservation in big data era: a survey[J]. Journal of Software, 2014, 25(4): 693-712.

[4] 王宇航, 张宏莉, 余翔湛. 移动互联网中的位置隐私保护研究[J]. 通信学报, 2015, 36(9): 2015167.
WANG Y H, ZHANG H L, YU X Z. Research on location privacy in mobile Internet[J]. Journal on Communications, 2015, 36(9): 2015167.

[5] KIDO H, YANAGISAWA Y, SATOH T. An anonymous communication technique using dummies for location-based services[C]//The International Conference on Pervasive Services 2005. Washington, IEEE, c2005: 88-97.

[6] KIDO H, YANAGISWA Y, SATOH T. Protection of location privacy using dummies for location-based services[C]//The 21st International Conference on Data Engineering Workshops. Washington, IEEE, c2005: 1248.

[7] GRUTESER M, GRUNWALD. Anonymous usage of location-based services through spatial and temporal cloaking[C]//The 1st International Conference on Mobile, Systems, Applications and Services. New York, ACM, c2003: 31-42.

[8] TALUKDER N, AHAMED S I. Preventing multi-query attack in location-based services[C]//The 3rd ACM Conference on Wireless Network Security. New York, ACM, c2010: 25-36.

[9] BERSFORD A R, STAJANO F. Mix zones: user privacy in location-aware services[C]//The 2nd IEEE Annual Conference on Pervasive Computing and Communications Workshops. Washington, IEEE, c2004:127-131.

[10] GAO S, MA J F, SHI W S, et al. TrPF: a trajectory privacy-preserving framework for participatory sensing[J]. IEEE Transactions on Information Forensics and Security, 2013, 8(6):874-887.

[11] ARDAGNA C, CREMONINI M, DAMIANI E, et al. Location privacy protection through obfuscation-based techniques[C]//The 21st Annual IFIP WG 11.3 Working Conference on Data and Applications Security, Berlin: Springer. c2007: 47-60.

[12] DAMIANI M L, BERTINO E, SILVESTRI. The probe framework for the personalized cloaking of private locations[J]. Transactions on Data Privacy, 2010, 3(2):123-148.

[13] GHINITA G, KALNIS P, KHOSHGOZARAN A, et al. Private queries in location based services: anonymizers are not necessary[C]//The 2008 ACM SIGMOD International Conference on Management of Data. New York, ACM, c2008: 121-132..

[14] MASCETTI S, FRENI D, BETTINI C, et al. Privacy in geo-social networks: proximity notification with untrusted service providers and curious buddies[J]. VLDB Journal, 2011, 20(4): 541-566.

[15] LU H, JENSEN C S, YIU M L. Pad: privacy-area aware, dummy based location privacy in mobile services[C]//The Seventh ACM International Workshop on Data Engineering for Wireless and Mobile Access. New York, ACM, c2008: 16-23.

[16] NIU B, LI Q, ZHU X et al. Achieving *k*-anonymity in privacy-aware location-based services[C]//The 33rd Annual IEEE International Conference on Computer Communications. Washington, IEEE, c2014: 754-762.

[17] NIU B, ZHANG Z Y, LI X Q, et al. Privacy-area aware dummy generation algorithms for location-based services[C]//The 2014 IEEE International Conference on Communication. Washington, IEEE, c2014: 957-962.

[18] NIU B, LI Q, ZHU X, et al. Enhancing privacy through caching in location-based services[C]//The 34th Annual IEEE International Conference on Computer Communications. Washington, IEEE, c2015: 1017-1025.

[19] XU T, CAI Y. Exploring historical location data for anonymity preservation in location-based services[C]//The 27th Conference on Computer Communications. Washington, IEEE, c2008: 13-18.

[20] XU T, CAI Y. Feeling-based location privacy protection for location-based services[C]//The 16th ACM Conference on Computer and Communications Security. New York, ACM, c2009: 348-357.

[21] CHOW C Y, MOKBEL M. Enabling private continuous queries for revealed user locations[C]//The 10th International Symposium on Spatial and Temporal Databases. Berlin, Springer, c2007: 258-275.

[22] 潘晓, 郝兴, 孟小峰. 基于位置服务中的连续查询隐私保护研究[J]. 计算机研究与发展, 2010, 47(1): 121-129.
PAN X, HAO X, MENG X F. Privacy preserving towards continuous query in location-based services[J]. Journal of Computer Research and Development, 2010, 47(1): 121-129.

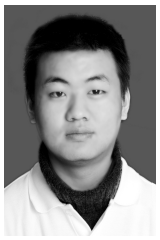
[23] WANG Y, XU D B, HE X, et al. L2P2: location-aware location privacy protection for location-based services[C]//The 31st Annual IEEE International Conference on Computer Communication. Washington, IEEE, c2012: 1996-2004.

[24] LI X H, WANG E M, YANG W D, et al. DALP: a demand-aware location privacy protection scheme in continuous location-based services[J]. Concurrency and Computation: Practice and Experience, 2016, 28(4): 1219-1236.

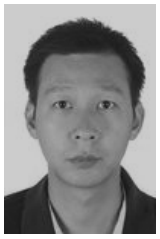
[25] SCHLEGEL R, CHOW C Y, HUANG Q, et al. User-defined privacy grid system for continuous location-based services[J]. IEEE Transactions on Mobile Computing, 2015, 14(10): 2158-2172.

[26] BRINKOFFff T. A framework for generating network-based moving objects[J]. Geo Informatica, 2002, 6(2): 153-180.

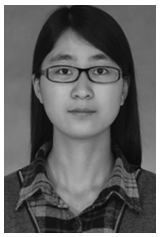
作者简介：



刘海（1984-），男，贵州贵阳人，西安电子科技大学博士生，主要研究方向为隐私保护和理性密码协议。



李兴华（1978-），男，河南南阳人，博士，西安电子科技大学教授、博士生导师，主要研究方向为隐私保护、网络与信息安全。



王二蒙（1990-），女，陕西渭南人，西安电子科技大学硕士生，主要研究方向为位置隐私保护。

马建峰（1963-），男，陕西西安人，博士，西安电子科技大学教授、博士生导师，主要研究方向为信道编码、网络与信息安全、密码学。