doi:10.11959/j.issn.1000-436x.2017131

LBS 隐私保护中基于查询范围的匿名区构造方案

裴卓雄，李兴华，刘海，雷凯跃，马建峰，李晖
（西安电子科技大学网络与信息安全学院，陕西 西安 710071）

摘 要：由于 *k*-匿名方法不仅能降低用户的计算开销，还能提供准确的查询结果，已被广泛用于位置隐私保护。然而，现有方案在匿名区构造过程中均未考虑位置服务提供商（LSP，location-based service provider）的查询区域面积，导致 LBS 查询服务质量降低。为了解决上述问题，将用户的查询范围引入到匿名区的构造中，匿名服务器首先生成满足用户隐私保护需求的初始子匿名区，再以 LSP 的查询区域面积为判定标准进行子匿名区合并。安全性和实验分析表明，所提方案在保护用户隐私的同时，能有效降低 LSP 的查询区域面积，从而提高 LBS 查询的服务质量。

关键词：基于位置的服务；*k*-匿名；服务质量；查询范围；匿名区
中图分类号：TP309 **文献标识码**：A

Anonymizing region construction scheme based on query range in location-based service privacy protection

PEI Zhuo-xiong, LI Xing-hua, LIU Hai, LEI Kai-yue, MA Jian-feng, LI Hui
(School of Cyber Engineering, Xidian University, Xi'an 710071, China)

Abstract: Since *k*-anonymity method can reduce the users' computation cost and provides the precise query results, it has been widely used to protect the user's privacy in location-based service. However, the existing schemes did not consider the size of the querying region for location based service provider (LSP) during the construction of the anonymizing region, which led that the quality of service was low. To solve this problem, the user's querying range was introduced to present a novel anonymizing region construction scheme. In the proposal, the anonymity server first generated the original anonymizing sub-regions according to the user's privacy requirements, and then merged these sub-regions to construct the anonymity region submitted to LSP based on the size of corresponding querying regions. The security and experiment analysis show that the presented scheme not only protects the user's privacy effectively, but also decreases LSP's querying regions, thereby improving the quality of service.

Key words: location-based service, *k*-anonymity, quality of service, query range, anonymizing region

1 引言

随着移动设备的普及和定位技术的发展，基于位置的服务（LBS，location-based service）得到广泛应用。它是指用户通过移动设备获取与其指定位置相关的信息查询及娱乐服务，如 Facebook Places、Google Latitudes 等。然而，位置服务提供商在为用户提供便利的 LBS 的同时，还可能会搜集并滥用用户的服务信息，从而非法获取用户的隐私信息，如家庭住址、工作单位和健康状况等。因此，LBS 中的位置隐私保护受到了研究者的广泛关注[1,2]。

k-匿名[3]作为最常用的 LBS 位置隐私保护方法，其基本思想是当用户进行 LBS 查询时，先将自己的真实位置和查询内容发送给可信的匿名服务器，匿名服务器去除用户的标识信息，并为其生成包含其他 $k-1$ 个用户的匿名区域，并连同查询内容一起发送给 LSP。此时，LSP 以不超过 $\frac{1}{k}$ 的正确率

收稿日期：2016-12-19；**修回日期**：2017-04-06
基金项目：国家自然科学基金资助项目（No.U170820014, No.61372075, No.U1135002）
Foundation Item: The National Natural Science Foundation of China (No.U170820014, No.61372075, No.U1135002)

将本次服务查询与用户进行关联，从而保护用户的位置隐私。与其他 LBS 隐私保护方法（如假位置[4,5]、模糊化[6, 7]、差分隐私[8, 9]和基于密码学的方法[10]）相比，*k*-匿名具有以下优点：1) 用户能获取准确的查询结果；2) 用户的计算开销和通信开销较小；3) 能混淆用户与 LBS 查询间的关联性。这就使该方法被广泛应用于 LBS 隐私保护中[3,11~17]。

在现实生活中，查询附近的兴趣点，如查询周围 500 m 内的餐馆、医院等是用户最常使用的一种 LBS 查询。然而，当采用 *k*-匿名方法保护上述查询中用户的隐私时，如果匿名服务器生成的匿名区面积过大，将会增加 LSP 的查询开销，导致服务质量降低。为了解决该问题，现有的方法[16,17]均是通过去除区域内不包含用户的部分，从而得到 n 个不相交的子匿名区域，使匿名区面积减少，从而提高服务质量，如图 1 所示。然而，在基于 *k*-匿名的 LBS 查询中，服务质量不仅与匿名区的大小有关，也与用户的查询范围有关。如果使用现有方法构造匿名区，不仅不能有效地提高服务质量，甚至还会出现进一步降低服务质量的情况。如图 2 所示，当采用现有匿名区划分方法对初始匿名区进行划分，LSP 会对部分区域内的兴趣点进行重复查询，从而降低了服务质量，其中，r 为查询半径。本文也通过实验证明了这一观点。

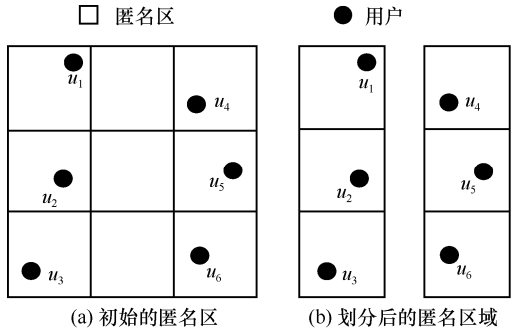

图 1 现有的匿名区域的划分方法

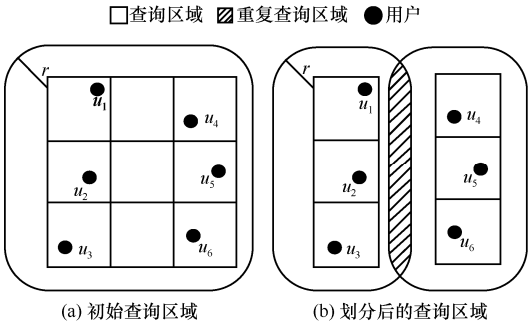

图 2 附近兴趣点的查询区域

为解决上述问题，本文提出了基于查询范围的匿名区构造方案。在本文方案中，匿名服务器首先根据用户的隐私保护需求生成 k 个初始子匿名区域，并根据其对应的查询区域进行匿名区域合并，使最终提交给 LSP 的匿名区在不降低用户隐私保护等级的同时，减少 LSP 的查询开销，提高服务质量。这是第一个基于用户查询范围构造匿名区的 *k*-匿名隐私保护方案。本文的主要贡献如下。

1) 从理论上分析得出，现有匿名区域划分方法不能降低 LSP 的查询开销，提高服务质量，并通过实验进行证明。

2) 以查询区域面积为子匿名区合并判断标准，提出一种基于用户查询范围的匿名区构造方案。安全性分析表明，本文方案构造的匿名区能有效地保护用户的位置隐私。

3) 大量实验表明，本文方案在不给匿名服务器带来较大计算开销的同时，有效地降低 LSP 的查询开销，从而提高 LBS 查询服务质量。

2 相关工作

Grutese 等[3]首次将 *k*-匿名的思想应用于位置隐私保护领域。匿名服务器采用四叉树结构划分区域，节点中存储对应区域中的用户。当用户请求服务时，从用户位置对应的叶子节点开始向上检索四叉树，如果当前叶子节点不满足 *k*-匿名，则检索其父节点，直至查找到不少于 $k-1$ 个其他用户，从而得到匿名区域。然而，如果叶子节点中用户的数量不满足隐私需求，需要检索其父节点，这会造成匿名区域呈 4 倍增长的态势，从而降低用户的服务质量。为解决上述问题，Mokbel 等[11,12]对文献[3]中匿名区构造方法进行改进。在他们的方案中，若当前叶子节点不满足 *k*-匿名，首先检索其兄弟节点，若此时仍不满足用户的隐私需求，再检索父节点。为进一步解决基于四叉树结构构造的匿名区域过大，导致服务质量降低的问题，Li[13]提出通过抑制用户的部分请求和删除最远足迹的方法来缩小匿名区面积，提高服务质量。

随后，学者们又分别提出 CliqueCloak [14]和 Hibert Cloak[15]方案，通过寻找最近满足隐私需求的 $k-1$ 个用户来构造匿名区。在文献[14]方案中，用户可自定义隐私保护需求。然而，该方案利用无向图构建匿名区域，方案开销过大，会出现超出匿名期限仍未成功构造出匿名区的情形。在文献[15]方

案中，匿名服务器根据 Hilbert 曲线将二维空间的所有用户映射到一维数组上，并根据 k 值将这些用户划分为若干个集合。当某一用户进行服务请求时，就利用该用户所属的用户集合来构造匿名区。

在上述所有方案中，若用于构造匿名区的 k 个用户分布较远，仍会出现匿名区过大，服务质量降低的问题。Tan 等[16]首次将区域划分的思想应用于匿名区的构造，其通过 Hilbert 空间填充曲线将匿名区域中的用户划分到不同的群组。当用户进行服务器请求时，匿名服务器将利用其所属群组中其他用户的位置构造匿名区。随后，Li 等[17]同样利用区域划分的方法对减小匿名区面积，提高服务质量进行研究。在他们的方案中，匿名服务器首先构造一个包含 k 个用户的匿名区，随后根据各用户位置间的关系，去除不包含用户的匿名区域，形成多个互不相交的子匿名区。

然而，现有的匿名区构造方案均忽略了用户的查询范围对 LBS 查询服务质量的影响。当采用现有方法构造匿名区时，LSP 会对部分区域内的兴趣点进行重复查询，从而降低 LBS 查询服务质量。

3　预备知识

3.1　系统架构

本文采用集中式系统结构[18]，由用户、匿名服务器和 LSP 这 3 个部分组成。当用户请求服务时，匿名服务器将用户的真实位置模糊化为一个匿名区域发送给 LSP，该区域内不仅包含真实用户，还包含其他至少 $k-1$ 个用户。此时，LSP 以不超过 $\frac{1}{k}$ 的概率将查询与用户相关联，从而实现 k-匿名。其系统结构如图 3 所示。

假设用户与匿名服务器存在一条安全的通信信道。当用户进行查询附近的兴趣点时，首先利用安全信道将查询请求 $q=<ID,L(x,y),r,POI,p>$ 通过安全信道发送给可信的匿名服务器。其中，ID 表示用户的身份；$L(x,y)$ 表示用户的位置坐标；r 表示用户的查询半径；POI 表示用户查询的兴趣点；$p=(k,A_{\min})$ 表示用户当前查询的隐私保护需求，k 表示匿名服务器生成的匿名区中至少包含 $k-1$ 个其他用户，$A_{\min}$ 表示匿名服务器生成的匿名区最小面积。

可信的匿名服务器收到用户请求后通过认证确定其身份，并根据用户的隐私保护需求 $p=\{k,A_{\min}\}$，寻找其他 $k-1$ 个用户从而生成面积不小于 $A_{\min}$ 的匿名区，并将匿名化处理得到的查询请求 $Q=<CR,r,POI>$ 发送给半可信的 LSP。其中，CR 表示匿名服务器为当前进行服务查询的用户生成的匿名区。

LSP 在收到匿名服务器发送的匿名查询请求后，在数据库中进行检索，将所有的查询候选结果返回给匿名服务器。匿名服务器在收到 LSP 发送来的查询候选结果后，根据用户的位置 $L(x,y)$ 对查询结果进行精拣，最后将准确的查询结果返回给用户。

在该系统模型中，本文直接将 LSP 视为攻击者，其攻击目的如下：1) 从匿名服务器发送来的匿名区域中识别出用户的真实位置；2) 推测出真实请求的用户。

此外，在上述模型中，LBS 查询服务质量主要受以下 4 个因素的影响：1) 匿名服务器生成匿名区所需的时间；2) 匿名服务器将匿名查询请求发送给 LSP 所需的时间；3) LSP 根据匿名服务器发送的匿名查询请求检索数据库所需的时间；4) LSP 将检索结果发送给匿名服务器所需的时间以及匿名服务器精拣查询结果所需的时间。由于 LSP 将检索结果发送给匿名服务器所需的时间以及匿名服务器精拣查询结果所需的时间受兴趣点分布的影响，而匿名服务器将匿名查询请求发送给 LSP 所需的时间受传输带宽的影响，因此，本文仅通过匿名服务器生成匿名区所需的时间和 LSP 检索数据库所需的时间来评估基于 k-匿名的 LBS 查询的服务质量。

3.2　LSP 的查询区域

LSP 收到来自匿名服务器发送的匿名查询请求 $Q=<CR,r,POI>$ 后，首先将根据匿名区 CR 和查询

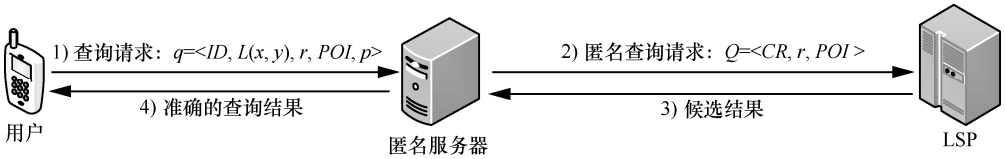

图 3　系统架构

半径 r 计算得到查询区域 QAR，再在 QAR 区域中检索用户查询的兴趣点。匿名区对应的查询区域如图 4 所示。

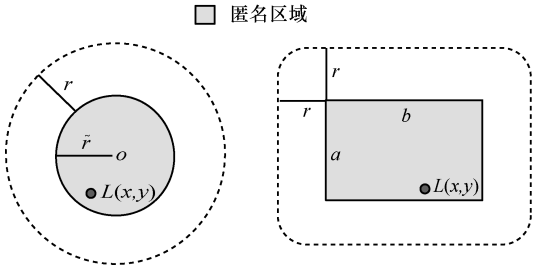

图 4 LSP 的查询区域

当匿名区为圆形时，查询区域 QAR 的面积为 $\pi(\tilde{r}+r)^2$。其中，$\tilde{r}$ 表示圆形匿名区的半径，$\pi\tilde{r}^2$ 是匿名区 CR 的面积。

当匿名区为矩形时，查询区域 QAR 的面积为 $ab+2(a+b)r+\pi r^2$。其中，a、b 是矩形匿名区的边长，ab 是匿名区 CR 的面积。

从上述分析中可以得知，当收到来自匿名服务器发送的匿名查询请求后，LSP 检索数据库所需的时间不仅与匿名服务器生成的匿名区 CR 的大小有关，还与用户的查询半径 r 有关，即由查询区域决定。然而，现有匿名区的划分方案中仅考虑了匿名区的大小，这就使这些方案并不能有效提高 LBS 查询服务质量，甚至还会出现进一步降低服务质量的情形。

4 基于查询范围的匿名区构造方案

在本文方案中，匿名服务器首先根据用户的隐私保护需求生成 k 个初始的子匿名区域，并通过计算其对应的查询区域，进行子匿名区域合并，最终得到子匿名区域集合。该方案在不降低用户隐私保护等级的同时，能减少 LSP 的查询开销，提高服务质量。

1) 初始子匿名区的生成

匿名服务器在收到用户发送的服务请求后，首先根据用户的隐私保护需求 $p=\{k,A_{\min}\}$，查找其余 $k-1$ 个用户，并获取其位置信息 $L_1(x_1,y_1),\cdots,L_{k-1}(x_{k-1},y_{k-1})$。随后，生成 k 个互不相交的初始子匿名区 $AR_0,AR_1,\cdots,AR_{k-1}$ 使

$$\begin{cases} centre(AR_i)\neq L_i(x_i,y_i) \\ Area(AR_i)\neq A_{\min} \end{cases}$$

其中，AR_i 表示包含第 i 个位置 $L_i(x_i,y_i)$ 的初始子匿名区，$0\leqslant i\leqslant k-1$；$L_0(x_0,y_0)$ 表示发送服务查询请求的用户的位置；$centre(AR_i)$ 表示初始子匿名区 AR_i 的中心点位置；$Area(AR_i)$ 表示初始子匿名区 AR_i 的面积。

2) 子匿名区的合并

当匿名服务器生成 k 个初始子匿名区 AR_0，$AR_1,\cdots,AR_{k-1}$ 后，根据用户的查询半径 r 分别计算其对应的查询区域 $QAR_0,QAR_1,\cdots,QAR_{k-1}$ 的面积 $Area(QAR_0),Area(QAR_1),\cdots,Area(QAR_{k-1})$，并根据查询区域面积判断是否需要将子匿名区进行合并。为了有效降低 LSP 检索兴趣点的开销，当子匿名区合并完成后，应确保

$$\min\sum_{i=0}^{l}Area(QAR_i')$$

即匿名区 CS 中各子匿名区 AR_i' 所对应的查询区域 QAR_i' 的面积之和最小。其中，AR_i' 表示子匿名区合并完成后形成的第 i 个子匿名区；$0\leqslant i\leqslant l$；$0\leqslant l\leqslant k-1$。

匿名服务器在进行子匿名区合并的过程如下。

① 筛选需要进行区域合并的子匿名区

为了确保实现匿名区域查询面积最小，匿名服务器选择合并后对应的查询区域面积最小的子匿名区 AR_i 和 AR_j 进行匿名区合并。即选择合并的子匿名区 AR_i 和 AR_j 满足

$$\forall i,j\in[0,k-1],$$

$$QAR_{i,j}=\arg\min\{Area(QAR_{i,j})\}_{i\neq j}$$

且

$$centre(QAR_{i,j})\neq L_i(x_i,y_i),$$

$$centre(QAR_{i,j})\neq L_j(\mathrm{x}_j,\mathrm{y}_j)$$

② 选择子匿名区进行合并

对于 2 个子匿名区 AR_i和AR_j，$AR_{ij}=Gen(AR_i,AR_j)$ 表示由子匿名区 AR_i,AR_j 合并后形成的新的子匿名区，其对应的查询区域为 QAR_{ij}。如果 $Area(QAR_i)+Area(QAR_j)\leqslant Area(QAR_{i,j})$，则子匿名区 AR_i 和 AR_j 不合并。如果 $Area(QAR_i)+Area(QAR_j)>Area(QAR_{i,j})$，则将子匿名区 AR_i 和 AR_j 合并成新的子匿名区 $AR_{i,j}$。

③ 重复上述过程，直至无需再进行子匿名区合并。此时，匿名服务器就得到匿名集合 $CS=\{AR_0',AR_1',\cdots,AR_l'\}$。

在本文方案中，由于匿名服务器每次进行子匿

名区合并时，均是选择合并后对应的查询区域面积最小的子匿名区 AR_i 和 AR_j 进行合并。因此，在每次子匿名区合并后，均会得到查询面积最小的子匿名区 $AR_{i,j}$。那么，当匿名服务器使用本文方案生成最终的匿名区 $CS=\{AR'_0,AR'_1,\cdots,AR'_l\}$ 时，就能确保其对应的查询面积最小，即

$$\min\sum_{i=0}^{l}Area(QAR'_i)$$

本文方案以 k 个用户的位置 $L_0(x_0y_0)$，$L_1(x_1,y_1),\cdots,L_{k-1}(x_{k-1},y_{k-1})$ 为输入，在为其生成子匿名区 $AR_0,AR_1,\cdots,AR_{k-1}$ 后，进行子匿名区合并，最终得到提交给 LSP 的匿名区为子匿名区域集合 CS。具体如算法 1 所示。

算法 1　基于查询区域划分的子匿名生成算法

输入　k 个位置 $L_0(x_0,y_0),L_1(x_1,y_1),\cdots,L_{k-1}(x_{k-1},y_{k-1})$；查询半径 r；隐私需求 $A_{\min}$；

输出　匿名区集合 CS；

1) for　$i=0$　to　$k-1$　do

2)　$AR_i\leftarrow Gen(L_i(x_i,y_i))$

3)　$centre(AR_i)\neq L_i(x_i,y_i)$，$Area(AR_i)=A_{\min}$；

4)　$CS\leftarrow(AR_i)$；

5)　$QAR_i\leftarrow Gen(AR_i,r)$，计算 $Area(QAR_i)$；

6) end for

7) for　$i,j=0$　to　ij　and　$i\neq j$　do

8)　$AR_{i,j}\leftarrow Gen(AR_i,AR_j)$；

9)　$CR\leftarrow AR_{i,j}$；

10)　$QAR_{i,j}\leftarrow Gen(AR_{i,j},r)$，计算 $Area(QAR_{i,j})$；

11)　if　$Area(QAR_i)+Area(QAR_j)\leqslant Area(QAR_{i,j})$ then

12)　$CS\leftarrow CS\setminus\{AR_{i,j}\}$；

13) end if

14)　if $Area(QAR_i)+Area(QAR_j)>Area(QAR_{i,j})$，$QAR_{i,j}=\arg\min\{Area(QAR_{i,j})\}_{i\neq j}$　$centre(QAR_{i,j})\neq L_i(x_i,y_i)$　and　$centre(QAR_{i,j})\neq L_j(x_j,y_j)$ then

15)　$CS\leftarrow CS\setminus\{AR_i,AR_j\}$；

16) return CS

5　方案分析

5.1　安全性分析

在本文方案中，首先根据 k 个的真实位置 $L_0(x_0,y_0),L_1(x_1,y_1),\cdots,L_{k-1}(x_{k-1},y_{k-1})$ 构造面积为 $A_{\min}$ 的互不相交的初始子匿名区 $AR_0,AR_1,\cdots,AR_{k-1}$，并且保证每个真实位置 $L_i(x_i,y_i)$ 并不位于初始子匿名区的中心位置，即 $centre(AR_i)\neq L_i(x_i,y_i)$。若匿名服务器直接将 $AR_0,AR_1,\cdots,AR_{k-1}$ 发送给 LSP，此时 LSP 将无法正确地从匿名区 AR_i 中识别出用户的真实位置 $L_i(x_i,y_i)$。若匿名服务器采用本文方案对子匿名区进行合并，将最终形成的匿名区 $CS=\{AR'_0,AR'_1,\cdots,AR'_l\}$ 发送给 LSP 时，由于每个子匿名区 AR_i 的面积 $Area(AR_i)=A_{\min}$，因此，合并后形成的每个子匿名区 AR'_i 的面积 $Area(AR'_i)\geqslant A_{\min}$，且对于任意的 $L_i(x_i,y_i)\in AR'_i$，$centre(AR'_i)\neq L_i(x_i,y_i)$。此外，在匿名服务器发送的匿名查询请求中，已去除用户的身份标识。因此，当 LSP 收到匿名查询请求 $Q=<CR,r,POI>$ 后，也无法获知进行服务请求的用户身份。因此，本文方案能有效保护用户的隐私信息。

5.2　计算复杂度分析

当匿名服务器收到用户发送的服务查询请求，并采用本文方案为其生成匿名区时，首先根据用户的隐私保护需求生成 k 个初始子匿名区。因此，生成初始子匿名区的计算复杂度为 $O(k)$。当匿名服务器判断任意 2 个子匿名区 AR_i 和 AR_j 是否需要进行合并的过程中，首先要利用子匿名区 AR_i 和 AR_j 生成新的匿名区 $AR_{i,j}$，此时共需要生成 $C_k^2=\frac{k(k-1)}{2}$ 个新的匿名区 $AR_{i,j}$，其计算复杂度为 $O(k^2)$；随后，匿名服务器计算每个新生成匿名区 $AR_{i,j}$ 对应的查询区域面积，共需要计算 $\frac{k(k-1)}{2}$ 次，其计算复杂度为 $O(k^2)$；最后，通过对比子匿名区 AR_i 和 AR_j 与新生成的匿名区 $AR_{i,j}$ 的查询区域面积，判断初始子匿名区 AR_i 和 AR_j 是否需要进行合并。此时，需要 $\frac{k(k-1)}{2}$ 次计算，其计算复杂度为 $O(k^2)$。因此，对初始子匿名区进行合并的计算复杂度为 $O(k^2)+O(k^2)+O(k^2)=O(k^2)$。当初始子匿名区 AR_i 和 AR_j 合并成新的子匿名区 $AR_{i,j}$ 后，还需要判断该子匿名区是否需要与其他子匿名区再次进行合并。然而，在匿名区合并过程中，最好的情况是 k 个初始子匿名区均不进行合并。此时执行本文方案所需的计算复杂度为

$$O_{\text{best}} = O(k) + O(k^2) = O(k^2)$$

最坏的情况是：k 个初始子匿名区均要进行合并，且最终合并成一个匿名区。此时，需重复进行上述匿名区合并判断 $k-1$ 次，执行本文方案所需的计算复杂度为

$$O_{\text{worst}} = O(k) + O((k-1)k^2) = O(k^3)$$

5.3 实验分析

本文采用基于网络的移动对象生成器（NGMO，network-based generator of moving objects）[19]生成实验数据。该生成器是现有 LBS 位置隐私保护研究中常使用的一种，它以德国城市 Oldenberg 地图为基础，通过设置移动对象数量等参数模拟生成用户的位置信息。本文设定隐私需求 k、生成的初始子匿名区域为矩形，且其面积 $A_{\min} = 160\,000\ \text{m}^2$。此外，为了评估 LSP 的查询开销，本文模拟构造了餐馆、酒店、医院和停车场等 500 000 个兴趣点，并采用 R 树结构[20]存取这些兴趣点。R 树作为目前最好的存储高维数据的平衡树，能有效提高在高维空间中的搜索效率。实验环境为 3.20 GHz Core(TM) i5 CPU，4 GB 内存。算法由 C++ 编程实现，程序运行在 Windows 7 环境下。

5.3.1 现有匿名区构造方案存在的问题

为了证明现有匿名区构造方案并不能有效提高 LBS 查询服务质量，本文分别选用 Casper 方案[11]和 Fragment 方案[17]进行附近兴趣点查询。Casper 作为最常用的匿名区域构造方法，生成一个至少包含所有 k 个用户的匿名区。Fragment 是对 Casper 方案生成的匿名区域进行处理，根据匿名区域内用户的位置，通过去除不包含用户位置的部分来减少匿名区面积。

本文通过对比上述 2 个方案中匿名服务器生成匿名区所需时间、匿名区的面积和 LSP 的查询区域面积，证明当存在重合的查询区域时，现有的匿名区的划分方法会进一步降低服务质量。在这部分实验中，本文设定用户的查询半径 $r = 500\ \text{m}$，实验结果如图 5 和图 6 所示。

由图 5 可知，与 Casper 方案相比，虽然 Fragment 方案采用区域划分方法缩小了匿名区面积，如 $k = 25$ 时，Casper 方案生成的匿名区面积为 $5.73\times10^7\ \text{m}^2$，生成匿名区所需时间为 177.275 ms，而 Fragment 方案生成的匿名区面积为 $3.40\times10^7\ \text{m}^2$，生成匿名区所需时间为 185.331 ms。但是，在 Casper 和 Fragment 方案中，LSP 查询兴趣点所需的时间却分别为 10.140 s（其对应生成的查询区域面积为 $7.314\times10^7\ \text{m}^2$）和 10.721 s（其对应生成的查询区域面积为 $7.379\times10^7\ \text{m}^2$）。因此，当匿名服务器采用 Casper 和 Fragment 方案生成匿名区时，用户获取查询结果的时间（不考虑传输时延）分别为 $10.140+0.177=10.317$ s 和 $10.721+0.185=10.906$ s。也就是说，当匿名服务器采用现有区域划分方法构造匿名区时，用户的获取查询结果的时间反而增加了 0.589 s。造成这一问题的根本原因就是 LSP 查询兴趣点的时间不仅受匿名区大小的影响，更受其查询区域的影响。这就说明了现有匿名区构造方案并不能有效提供 LBS 查询服务质量。

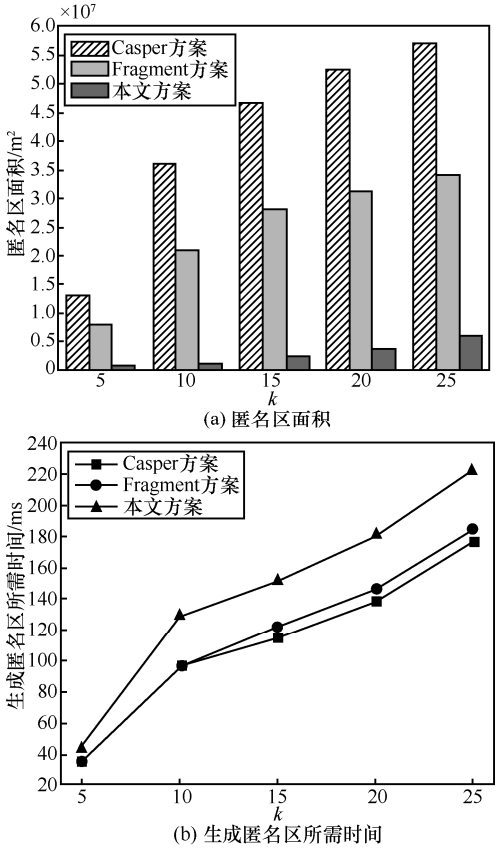

图 5 匿名服务器计算开销

5.3.2 本文方案的有效性

下面，将本文方案与 Casper 方案进行对比，表明本文方案能显著地降低 LSP 的查询开销，从而提高 LBS 查询服务质量。实验结果如图 5 和图 6 所示。

从图 5 和图 6 中可知，与现有方案相比，本文方案生成的匿名区面积和查询区域面积显著降低，LSP 查询兴趣点所需时间明显减少。以 $k = 25$ 为例，

本文方案生成的匿名区面积为 $5.93\times10^6\,\mathrm{m}^2$，比 Casper 方案生成的匿名区面积减少了 $3.40\times10^7-5.93\times10^6=2.80\times10^7\,\mathrm{m}^2$，而生成匿名区的时间仅从 185.331 ms 上升至 222.697 ms，增加 37.366 ms。而本文方案生成的查询区域面积为 $2.439\times10^7\,\mathrm{m}^2$，比 Casper 方案的查询面积减少 $4.875\times10^7\,\mathrm{m}^2$；LSP 的查询处理时间从10.140 s 降低为 2.286 s。因此，与 Casper 方案相比，当匿名服务器采用本方案生成匿名区时，能有效地提供 LBS 查询服务质量。

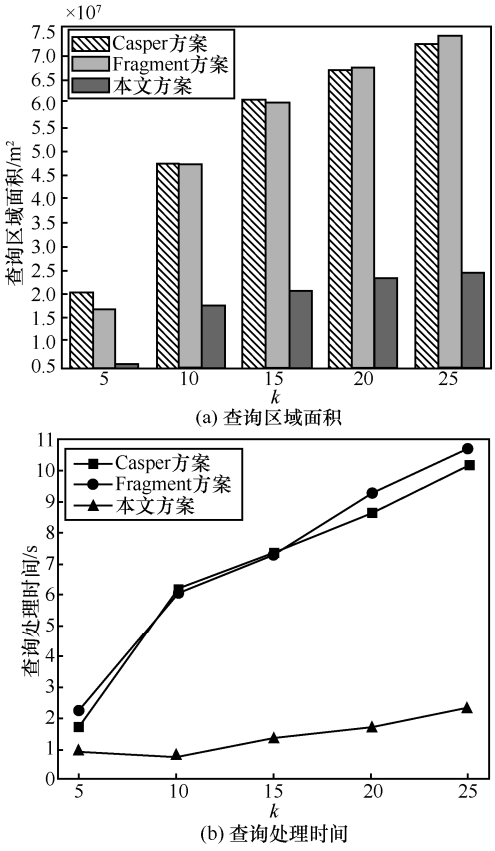

图 6　LSP 的查询开销

由图 5(b)与图 6(b)可知，与 Casper 方案相比，随着 k 值增大，本文方案额外处理的时间越大(即本方案与 Casper 方案生成匿名区所需时间的差值)，但查询处理时间缩短的也越明显。同时，匿名服务器生成匿名区域所需的时间为毫秒级，对用户时延影响较小，而查询处理的时间为秒级，对用户时延影响较大，可见，LSP 的处理时间极大程度决定了用户的服务质量。

下面，简要分析用户指定的查询半径 r 对本文方案的影响。实验结果如图 7 所示。随着用户查询半径 r 的增加，需要合并的子匿名区数量也随之增加，从而增加了匿名服务器的计算开销。此外，当用户查询半径 r 增加时，查询区域面积也不断增大，这也就使 LSP 的查询处理时间也不断增大。综上所述，随着 r 的增加，LBS 的查询服务质量呈降低态势。

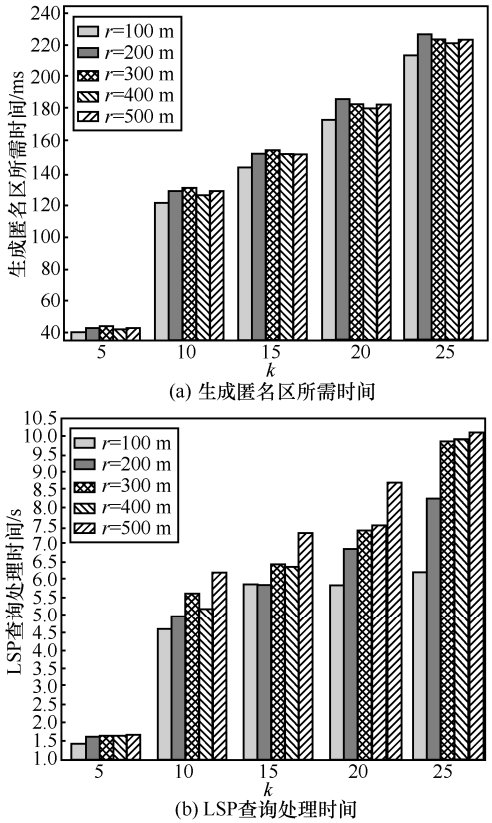

图 7　用户查询半径对本文方案的影响

综上所述，本文方案在不给匿名服务器带来较大计算开销的同时，显著地减少了 LSP 查询区域面积，降低其查询时间，从而有效地提高了 LBS 查询服务质量。

6　结束语

本文通过理论分析及实验证明，在 k 匿名位置隐私保护研究中，现有基于区域划分思想的匿名区构造方案并不能有效地提高 LBS 查询服务质量。造成这一问题的根本原因是 LBS 的服务质量不仅与匿名服务器构造的匿名区面积有关，更与 LSP 的查询范围有关。为了解决上述问题，本文将用户的查询范围引入到匿名区的构造过程中，匿名服务器首先根据用户隐私保护需求生成 k 个初始子匿名区，再以查询区域面积为判定标准进行子匿名区合并。

方案分析表明，本文方案在有效保护用户隐私的同时，能有效降低 LSP 的查询开销，提高服务质量。

参考文献：

[1] GAMBS S, KILLIJIAN M O, CORTEZ M N P. Show me how you move and I will tell you who you are[J]. Transactions on Data Privacy, 2011, 2(4):103-126.

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

[3] GRUTESER M, GRUNWALD D. Anonymous usage of location-based services through spatial and temporal cloaking[C]//The First International Conference on Mobile Systems, Applications, and Services. New York: ACM, 2003: 163-168.

[4] 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, 2014:754-762.

[5] 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, 2014: 957-962.

[6] DUCKHAM M, KULIK L. A formal model of obfuscation and negotiation for location privacy[C]//The 3rd International Conference on Pervasive Computing. Berlin: Springer, 2005: 152-170.

[7] MASCETTI S, FRENI D, BETTINI C, et al. Privacy in geo-social networks: proximity notification with untrusted service providers and curious buddies[J]. Journal of VLDB, 2010, 20(4): 541-566.

[8] ANDRES M E, BORDENABE N E, CHATZIKOKOLAKIS K, et al. Geo-indistinguishability: differential privacy for location-based systems[C]//The 2013 ACM SIGSAC Conference on Computer and Communications Security. New York: ACM, 2013:901-914.

[9] XIAO Y, LI X. Protecting locations with differential privacy under temporal correlations[C]//The 22nd ACM SIGSAC Conference on Computer and Communications Security. New York: ACM, 2015: 1298-1309.

[10] 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.

[11] MOKBEL M F, CHOW C Y, AREF W G. The new casper: privacy-aware location-based database server[C]//The 23rd International Conference on Data Engineering. Washington: IEEE, 2007: 1499-1500.

[12] MOKBEL M F, CHOW C Y, AREF W G. Casper: query processing for location services without compromising privacy[J]. ACM Transactions on Database Systems, 2009, 34(4):24-48.

[13] 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.

[14] 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.

[15] KALNIS P, GHINITA G, MOURATIDIS K, et al. Preventing location-based identity inference in anonymous spatial queries[J]. IEEE Transactions on Knowledge and Data Engineering, 2007, 19(12): 1719-1733.

[16] TAN K W, LIN Y. Spatial cloaking revisited: distinguishing information leakage from anonymity[C]//The 11th International Symposium on Advances in Spatial and Temporal Databases. Washington: IEEE, 2009:117-134.

[17] LI T C, ZHU W T. Protecting user anonymity in location-based services with fragmented cloaking region[C]//2012 IEEE International Conference on Computer Science and Automation Engineering. Washington: IEEE, 2012: 227-231.

[18] CHENG R, ZHANG Y, BERTINO E, et al. Preserving user location-privacy in mobile data management infrastructures[C]//The 6th International Workshop on Privacy Enhancing Technologies. Washington: IEEE, 2006: 393-412.

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

[20] HADJIELEFTHERIOU M, MANOLOPOULOS Y, THEODORIDIS Y, et al. R-trees-a dynamic index structure for spatial searching[C]//Encyclopedia of GIS. Berlin: Springer, 2008: 993-1002.

作者简介：

裴卓雄（1993-），男，山西运城人，西安电子科技大学硕士生，主要研究方向为位置隐私保护。

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。

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

雷凯跃（1990-），女，天津人，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，西安电子科技大学教授、博士生导师，主要研究方向为密码学、无线和移动安全等。

李晖（1968-），男，河南灵宝人，博士，西安电子科技大学教授、博士生导师，主要研究方向为密码学、无线网络安全、云计算安全、信息论与编码理论。