DOI: 10.3785/j.issn.1008-973X.2011.03.005

路网限制环境中基于匿名蜂窝的位置隐私保护

徐 建[1],徐 明[1],林 欣[2],郑 宁[1]

(1.杭州电子科技大学 计算机学院,浙江 杭州 310018;2.华东师范大学 信息科学技术学院,上海 200241)

摘 要:为了解决路网限制上下文环境的位置隐私保护问题,提出一个基于匿名蜂窝的位置隐私保护算法.根据路网环境特殊的点、线网络结构,事先将道路网络处理成由道路交叉点为圆心的圆形基本匿名区域——匿名蜂窝组成的网络.使用 Hilbert 曲线对匿名蜂窝进行空间编码,在算法中使用 B^+ 树对其进行索引.讨论了基于匿名蜂窝对用户查询进行匿名处理的 2 种成组方法:相邻结点优先成组和相邻路径结点优先成组,并对它们进行分析比较.在具体描述该算法的同时,进行了理论分析.通过实验验证了算法的可行性.实验结果表明:基于匿名蜂窝的相邻路径结点优先成组方法能够较好的增强算法应对推理攻击的鲁棒性.

关键词:道路网络;位置隐私;匿名蜂窝;基于位置的服务

中图分类号:TP 391 **文献标志码**:A **文章编号**:1008-973X(2011)03-0429-06

Location privacy protection through anonymous cells in road network

XU Jian[1], XU Ming[1], LIN Xin[2], ZHENG Ning[1]

(1. *College of Computer, Hangzhou Dianzi University, Hangzhou* 310018, *China*;
2. *School of Information Science and Technology, East China Normal University, Shanghai* 200241, *China*)

Abstract: To solve the problem of location privacy protection under restricted context in road network, this work developed and demonstrated a privacy protection algorithm. According to the characteristics of points and lines structure in the road network, the network is divided into anonymous cells with each road intersection as the center in advance. Cells are coded with Hilbert curve and indexed by a B+ tree. Two cellular-based methods for query anonymous processing nearest neighbor first and nearest road first were introduced. Then the two methods were analyzed and compared. Data from a real city map simulation show the effectiveness of the algorithm. The results prove that the nearest road first cloaking method can achieve more location privacy without loss of efficiency than the other, therefore improve the robustness against inference attacks.

Key words: road network; location privacy; anonymous cell; location-based services

随着 GPS 和其他定位技术的发展,在日益广泛的无线网络覆盖区域,越来越多基于位置的服务(location-based services, LBS)变为现实.这些基于位置的服务,都需要用户提供准确的、实时的位置信息,但如果向不可信的第三方披露这些信息,用户将面临巨大的风险[1-2].

位置隐私是一种特殊的个人信息隐私.在位置匿名处理系统中,使用最多的是 K-anonymity 模型,Sweeney [3]率先在数据库信息发布的隐私保护中讨论了 K-anonymity 模型,Gruteser 等 [4]将其引入位置隐私保护研究领域,提出了位置 K-anonymity.Kido 等[5-7]在基于空间的匿名模型中都采用了 K-anonymity 匿名的概念.在极端的情况下,k 个移动用户如果都位于同一物理位置很容易造成位置隐

收稿日期:2009-12-21. 浙江大学学报(工学版)网址:www.journals.zju.edu.cn/eng

基金项目:国家自然科学基金资助项目(61003195,60903169);浙江省自然科学基金资助项目(Y106427).

作者简介:徐建(1975-),男,浙江缙云人,副教授,从事分布式计算机系统的研究.E-mail:jian.xu@hdu.edu.cn

私泄漏. Bamba 等[8]将数据库信息发布中的 l-diversity 思想引入了位置隐私保护，同时将 l-diversity 思想扩展到了查询内容的类型上，认为匿名组查询时还应该包括不同内容类型的查询，以增加恶意攻击的难度.

Gedik 等[9]介绍了一种可定制匿名参数的模型并开发了相应的算法. Xu 等[10]提出了一种基于圆形匿名空间、能够防止跟踪分析的匿名算法，他认为圆形的匿名空间相对基于矩形能够生成较小查询结果集，从而提高算法的工作效率. Kyiakos 等[11]讨论了路网限制环境下基于道路路段的匿名查询处理.

目前已有的隐私匿名保护技术[4-6]都是考虑一个平面的欧几里德空间用户位置的移动；然而事实上，大部分的移动物体，例如车载的移动系统，它们的移动是受到道路网络限制的. 在这种情况下直接使用现有的隐私保护技术会产生新的问题，因为从路网环境可以提取很多的背景信息使得对用户当前位置隐私信息进行推理攻击成为可能.

通过分析可以知道，一种直观的解决思路就是每个匿名区域选择时应该包含尽可能多的道路交叉结点，以增加匿名区域内道路分布的均衡性，从而增强攻击者推理攻击的难度. 本文基于匿名区域信息熵的匿名衡量标准，提出适用于路网环境的 KD-Anonymity 算法，在满足 K-Anonymity 的条件下，达到更好的位置匿名效果，同时对算法进行了分析和优化，通过实验验证了算法的有效性.

1 系统模型

很多系统的设计[2-5,11]都采用了以下的策略：当一个用户发起一次位置相关的查询请求时，它仅向一个可信的匿名服务器(anonymizer)发送包含该用户位置信息的查询请求 $q(\mathrm{ID},\mathrm{Loc}[x,y],\mathrm{Query})$，其中：ID 表示用户的标识，Loc 表示用户的位置，Query 为查询的内容. 匿名服务器将查询者所在位置附近的多个查询，通过某个模糊化的匿名算法，合并成一个匿名空间区域(anonymizing spatial region，ASR)$([x_1,y_1],[x_2,y_2])$，向 LBS 的数据库发起查询请求. 匿名服务器从 LBS 处检索到与这个匿名空间区域相关的结果集(result candidate set，RCS)，这个结果集包含了用户原始的查询结果. 匿名服务器从结果集中选择用户需要的结果返还给用户. 整个过程如图 1 所示.

定义 1 用户查询请求 一个用户在 t 时刻的查询请求 $q(\mathrm{ID},\mathrm{Loc}[x,y],\mathrm{query})$.

图 1 系统模型

Fig. 1 System model

匿名服务器收到用户的查询后，使用一个链表 T 组织所有的用户查询，显然，表 T 具有属性 A^{ID}、A^{Loc}、A^{Query}，以及用户的查询时间 t^A. 随着时间的流逝，表 T 不断地被更新. 例如，新查询的插入和查询的删除. 匿名服务器根据一定的原则和用户的需求生成发往 LBS 的查询 Q，在结果集 RCS 中取得查询 q 的结果，并返回给用户后，删除查询 q.

定义 2 匿名服务器查询请求 匿名服务器生成一个发往 LBS 的匿名查询请求 $Q(\mathrm{ID}',L'([x_1,y_1],[x_2,y_2]),\mathrm{Query}')$. 其中 ID' 是匿名服务器对该查询的标识. 假设 Q 包含了 m 个用户的查询，那么对于 m 个查询中的第 i 个查询，有 $[x_i,y_i]\in([x_1,y_1],[x_2,y_2])$ 和 $\mathrm{query}_i\in\mathrm{Query}'$.

显然，可以认为匿名服务器根据链表 T 构造了一张新表 T^*，具有属性 ID'、L'、Query' 和发送查询 Q 的时间 t'.

为简化分析，使用无向图 $G=(V,E)$ 为路网环境建模，结点集 V 和边集合 E 表示路网中相应结点和结点之间的道路. 使用 $d(v)$ 表示结点 v 的度，说明结点相连的路径数目，当 $d(v)=1$ 时，认为是图 G 中某条路径的端结点；当 $d(v)=2$ 时，是某 2 条路径的邻接结点；当 $d(v)=3$ 时，是某 3 条路径的交叉结点.

考虑用户沿着图 G 中某条路线移动的情况，那么在某一个时刻用户 i 的位置 $\mathrm{Loc}[x_i,y_i]$ 总是位于某条路径 e 或者某个结点 v. 用户移动的路线 P 可以表示为 $P=(v_0,e_1,\cdots,e_n,v_n)$，其中：$v_0$ 表示用户的出发地，$e_1,\cdots,e_n$ 是用户移动经过的路径，v_n 表示用户的目的地. 在用户沿着 P 移动的过程中，匿名服务器生成的 T^* 所包含的每个 Q 都必须满足一定的限制要求，例如，Q 要满足一定的时间精度，包含的查询时间 $t_Q<\min\{t_1,\cdots,t_m\}+t_{\mathrm{Thr}}$，即匿名服务器必须在收到一个用户的查询请求 q 后在 t_{Thr} 向 LBS 转发查询请求. 又如空间精度，Q 包含的匿名框也不能超过用户能够接受的最大限制以影响最终的查询结果.

在时间精度和空间精度的基础上，本文主要考虑路网受限环境中用户 2 种类型的隐私保护策略：位置的 K-Anonymity 和 KD-Anonymity.

定义 3　K-Anonymity[3]　在匿名服务器向 LBS 发送的查询中包含的用户位置具有 K-Anonymity 的特征，也就是说查询 Q 匿名空间区域 $L'([x_1,y_1],[x_2,y_2])$ 中至少包含了 $(K-1)$ 个其他用户的位置信息.

当用户局限在路网环境中移动时，由于匿名空间区域的选取没有考虑区域内的地理信息，例如路网信息，使基于路网信息的推理攻击成为可能. 例如在图 2 中，匿名区域为 A 时，因为区域内包含了多条道路，每条道路都是用户可能处于的位置，但是当匿名区域为 B 时，区域内道路数目减少，用户可能的位置也就相应地减少；当匿名区域为 C 时，区域内只有一条道路，用户的位置将被确定在这条道路上行驶，此时增加匿名区域内的用户数对于位置匿名效果将不再起作用. 例如：当用户以 30～50 km/h 城市道路速度移动时，在用户追踪应用背景中，用户在某条道路这种类型的位置信息已经构成了用户位置信息的泄漏. 显然，仅有 K 匿名是不够的. 受 Bamba [8] 提出的 l-diversity 思想启发，本文引入了第二个位置隐私的衡量标准.

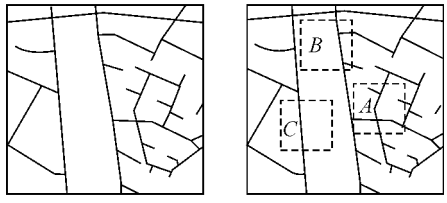

图 2　Oldenburg 地图中 cloak 区域示意

Fig. 2　Cloak regions in Oldenburg map

定义 4　KD-Anonymity　在匿名服务器向 LBS 发送的查询中包含的用户位置是 KD-Anonymity，那么也就是说查询 Q 匿名空间区域 $L'([x_1,y_1],[x_2,y_2])$ 中至少包含了 $(K-1)$ 个其他用户的位置信息的前提下，包含的路径数目 D 应尽可能大，较大的 D 值可以增大匿名区域的信息熵，增强恶意用户攻击的难度. KD-Anonymity 定义了路网环境中查询 Q 中用户位置分布的变化.

定义 5　位置匿名过程　在用户 u 沿着路线 P 移动的整个过程中，匿名服务器从用户处接收到的查询 q，按照用户对查询的服务质量要求 QoS，对其加以匿名处理，生成符合 QoS 的 Q，从 LBS 处查询获得相应的结果，并将结果返回给用户.

2　算法分析

基于蜂窝图(cellular graph)的位置匿名模型 KD-Anonymity 为了能在效率和攻击耐受性之间取得最佳平衡，首先将路网环境构造成以结点为中心的蜂窝图；其次，在单一的匿名蜂窝内没有达到用户所需的服务质量时，例如匿名的 K 系数没有达到用户的要求，算法可以动态调整蜂窝的大小或者合并相邻的蜂窝以适应用户的需求.

基于匿名蜂窝的 KD-Anonymity 算法由 3 个阶段组成：

1)路网环境的匿名空间区域预处理，也就是匿名蜂窝构造或者路网环境蜂窝化. 匿名空间区域预处理：给定一个路网 $G=(V,E)$，其中有结点 v 及邻接的路径 $e_1,e_2,\cdots,e_m$，构造一个相应的匿名蜂窝网络 $G_c=(V_c,E_c)$，以 l 表示路径 e 的长度，即把以 v 为圆心，以 $r=\max\{l_{e1}/2,l_{e2}/2,\cdots,l_{em}/2\}$ 为半径的圆形作为算法的基本匿名区域，在 G_c 中使用 v_c 表示，如图 3 所示. 显然，匿名蜂窝所包含路径的数目等于结点 v 的度 $d(v)$.

本文使用 Hilbert 曲线[12] 对匿名蜂窝进行空间编码. 一个 4×4 分区的 Hilbert 曲线如图 4 所示. 匿名蜂窝 Hilbert 编码如图 3 所示，蜂窝在算法中使用 B^+ 树进行索引.

2)相邻查询，按照所属的蜂窝根据用户要求进行分组.

(1)查询成组：成组方法为相邻结点优先成组和相邻路径结点优先成组，在后续的章节中对它们进行了分析比较.

(2)相邻查询优先成组(nearest neighbor first，NNF)：给定一个用户 u 的查询 q，u 对查询的服务质量要求为 (u_K,u_s,u_t)，其中：u_K 为匿名组大小，u_s 为匿名空间大小，u_t 为匿名过程所需时间. 在每个成组周期，从 u 的位置出发，在其路径前方匿名蜂窝内，沿着其前进方向的路径进行搜索，将相邻的查询进行匿名区域的撮合构造. 构造过程在满足 u_K 条件后结束.

(3)相邻路径查询优先成组(nearest road first，

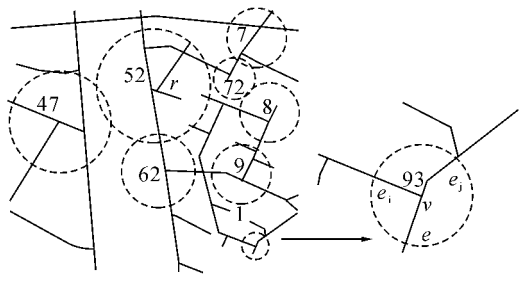

图 3　匿名蜂窝网络及蜂窝编码

Fig. 3　Anonymous cells network and coding

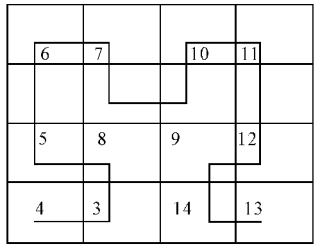

图 4 4×4 分区 Hilbert 曲线

Fig. 4 4×4 portion Hilbert curve

NRF):对于给定的查询 q,在其路径前方匿名蜂窝内,沿着顺时针方向在其相邻的路径上进行随机选取组成员.如图 3 所示,q 当前所处的位置在边 e_i,其所属的匿名区域为 v,算法对 q 进行匿名的时候,沿着顺时针方向,先在 e_j 随机选取组内的第 2 个查询,然后在 e_k 选取第 3 个查询,接着再在 e_i、e_j、e_k 直到满足条件,使匿名组内的查询分布尽可能均匀,增加对匿名查询推理攻击的难度.

3)蜂窝的调整和合并.匿名空间区域自适应调整.当基本匿名空间内的移动结点分布范围较小时,可以根据情况调整 r 的大小;当基本匿名空间内的查询数目没有达到用户的匿名要求时,可以对相邻的匿名蜂窝进行合并.

一个完整的查询匿名过程包含以下步骤:①用户查询预处理.接受查询后,通过检索匿名蜂窝 B^+ 索引,插入相应匿名蜂窝结点的队列,将匿名蜂窝结点标记为"脏结点"插入待处理列表.②匿名处理.在 t_{Thr} 时间内对用户查询进行 NNF 或者 NRF 匿名处理,如果该蜂窝结点达到用户 K-anonymity 要求,匿名处理完成;如果没有达到要求,沿着 B^+ 索引的叶子链表(如图 5 所示)找到相邻匿名蜂窝进行合并,直到达到匿名要求.③查询.将匿名成组后的查询插入匿名查询队列进行处理.④结果分拣.当 LBS 返回查询结果时,从结果集中分拣出查询具体结果返回给用户.

经分析可知,算法中路网匿名蜂窝的初始化、查询的插入和匿名成组过程的时间复杂度为 $\log(|V|)$,因此,算法能够很好地扩展至较大范围

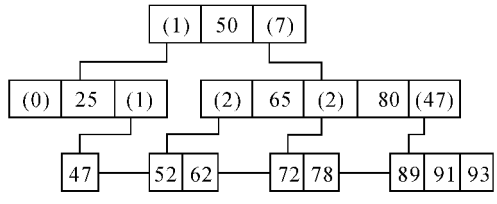

图 5 匿名蜂窝 B^+ 索引

Fig. 5 B^+ tree for anonymous cells

的匿名区域和较多数目的移动用户.在后续的实验部分本文使用了一个中等规模的区域 San Joaquin 对此进行了验证.

算法基本流程如下:

```
Initialize dirty cell list: List L
Initialize anonymous query list: Queue Q
Accept user query
While true do
    If L≠θ then
        v←List L
        Q←Cloak(v)
        Add Q to Queue Q
    While Queue Q≠θ do
        Q←Queue Q
        Retrieve
        Return results
Cloak(v)
    NNF or NRF
    Adjust
```

3 分析和模拟

3.1 推理攻击耐受力分析

为简化分析,本文仅考虑非连续查询条件下的位置隐私保护问题,连续查询条件下的位置隐私保护问题将在后续的论文中进行讨论.

3.1.1 攻击者假设 考虑恶意攻击者的目标是要将攻击者的身份(用户 ID)和其当前所处的位置联系起来的情况.并且有以下假设:

假设 1 用户和匿名服务器之间的通讯是安全的,但是恶意攻击者可以通过窃听匿名服务器发出的所有查询快照(Snapshot)并且明文解析其内容,或者从 LBS 处直接获得快照内容.

假设 2 恶意攻击者拥有与匿名服务器同样的路网地图知识,同时了解匿名服务器某个时刻所采用的具体匿名算法基本步骤.

3.1.2 攻击方法具体步骤 攻击者将所获得的快照按照匿名器发出的时间和快照中某个查询的具体匿名空间组成一组列表,仅考虑某个匿名空间在一个周期内所有查询的列表 SSList.用 s 表示列表的一个快照,则对于指定的快照 s:$<$ ID′, $L'([x_1,y_1],[x_2,y_2])$, Query′$>$,s. ID′、s. L'、s. Query′分别代表快照中相应的字段.把事件"具有 ID′的用户 u_i 在快照 s 中、在某个位置发送了 q_j"记为

$$u_i \xrightarrow[\text{Loc}]{s} q_j. \tag{1}$$

匿名攻击算法的目标就是通过计算

$$P(u \xrightarrow[\text{Loc}]{s} q_j) \tag{2}$$

找出某个用户 u_k 在某个具体位置(道路)的概率，也就是令 u_k 取 P 值最大时的 Loc. 其中 $i,j=1,2,3,\cdots,k$. 分析可知有限制条件如下：

$$\sum_{i=1}^{k} P(u \xrightarrow[\text{Loc}]{s} q_j) = 1, \tag{3}$$

$$\sum_{j=1}^{k} P(u_i \xrightarrow[\text{Loc}]{s} q_j) = 1. \tag{4}$$

在路网限制的情况下，假设匿名区域内用户的位置都位于某条道路，用 r 表示用户位于某条道路的情况，有下式：

$$\sum_{l=1}^{m} P(u_i \xrightarrow[r_l]{s} q_j) = 1. \tag{5}$$

式中：m 表示匿名区域内道路的数目.

仅考虑某个匿名空间在一个周期内的所有查询，也就是一个用户已经在某个匿名区域内. 攻击算法有如下步骤：

1)计算用户 u_i 发送查询 q_x 给匿名服务器的概率 $P(u_i \longrightarrow q_x)$. 为了简化复杂度，假设一个用户发送 q_1-q_k 查询的概率相等，即有

$$P(u_i \longrightarrow q_x)=1/k; x=1,2,3,\cdots,k. \tag{6}$$

2)求匿名服务器匿名后攻击者估计的条件概率 $P(u_i \xrightarrow{s} q_j)$. 设“$u_1-u_k$ 分别发送 q_1-q_k 之一”为事件 A，“u_i 发送 q_j 为事件 B”，$P(u_i \xrightarrow{s} q_j)$ 即指在事件 A 发生的条件下，发生事件 B 的条件概率 $P(B|A)$. 使用 $A(k)$ 表示 $1-k$ 所有排列的集合，α 为 $A(k)$ 中任意一种排列，α_n 为排列 α 中第 n 个元素. 结合约束条件(3)和(4)，则

$$P(u_i \xrightarrow{s} q_j) = P(B \mid A) = \frac{P(AB)}{P(A)}, \tag{7}$$

即

$$P(u_i \xrightarrow{s} q_j) = \frac{\sum\limits_{\forall a\in A(k)\text{且}a_i=j} \prod\limits_{w=1,2,\cdots,k} P(u_w \longrightarrow q_{a_w})}{\sum\limits_{\forall a\in A(k)} \prod\limits_{w=1,2,\cdots,k} P(u_w \longrightarrow q_{a_w})}. \tag{8}$$

对于恶意结点来说，s 中用户和查询的对应关系有 $P_k^k=k!$ 种，式(8)分母即为 $P(A)$，加上约束条件 $a_i=j$，式(7)中分子即表示事件 A 和 B 同时发生的概率 $P(AB)$.

3)攻击者使用背景知识计算 $P(u_i \xrightarrow[\text{Loc}]{s} q_j)$.

显然

$$P(u_I \xrightarrow[r_l]{s} q_j) = (r_l/\sum_{l=1}^{m} r_l) \cdot P(u_i \xrightarrow{s} q_j). \tag{9}$$

在 $m=1$ 的情况下，攻击者只要通过其他背景知识判定用户 u_i 确实在匿名区域内，马上就可以推断用户位于的道路，这是单点攻击在路网受限环境下的特例. 而如果 m 数字较大，例如 2 或者 3，用户所处的道路位置攻击者即使在第 1)、2)步得逞，已经知道某个用户在某个区域发送了具体的某个查询，也不能马上断定用户的位置.

相邻查询优先成组：如果某一道路具有较高的车流量，恶意用户通过事先赋予该路径较高的概率分布，就可以推断用户经过该路径的概率.

相邻路径查询成组：因为匿名过程中，匿名成组时各条道路的查询成组概率相似，恶意用户就不能主观的推断用户在某条道路上的概率，增加了匿名组的扰动性，使恶意推理攻击的难度增大.

3.2　实验模拟

匿名算法的实验主要通过攻击耐受力和匿名成组成功率来说明算法的性能. 攻击耐受力用匿名空间中查询所处的道路位置分布信息熵值表示，主要衡量算法对恶意攻击的鲁棒性；匿名成组成功率主要说明算法的计算效率，一个相对高效的算法在移动环境中对大量查询处理方面显然更有优势.

实验使用 2 幅实际的道路地图：一幅是德国的 Oldenburg，包含 6 105 个结点和 7 035 条边；另外一幅是美国的 San Joaquin，包含 18 496 个结点和 24 123条边. 通过这 2 幅地图分别表示较小规模和较大规模地理范围的路网状态. 在 Brinkhoff [13] 研发的模拟器基础上，使用 Java 语言开发实现本文的算法，并模拟不同交通状况条件下的查询. 模拟硬件平台为 TyanPSC 桌面高性能计算系统.

实验时，2 幅地图移动对象数量都设为 10 100，移动对象的移动速度分类为模拟器的默认速度分类，移动对象在实验进行期间在每个时间单位内以标准正态分布的概率发送查询，其所要求的匿名空间最小值取匿名蜂窝的平均值，在 cloak 时取基本 cloak 区间的大小，可容忍的最长匿名时间 t_{Thr} 为 5 个时间单位.

3.2.1　匿名区域道路分布　论文统计了在基于 cloak box 的 K-anonymity 算法中 Oldenburg 和 San Joaquin 地图基本匿名分区的道路数目. 如图 6 所示，80% 以上的匿名区域内道路分布都小于等于 1，根据 3.1 节的分析，遭受推理攻击时成功的概率非常大.

3.2.2　攻击耐受力　针对文中提出的恶意攻击模型可知，算法主要目标就是提高匿名空间内查询分布的复杂性来增加攻击的难度，因此，在评估算法攻击耐受力时，使用查询在匿名区域内各个道路上分

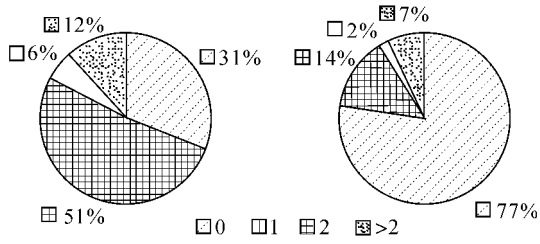

图 6 Oldenburg 和 San Joaquin 基本匿名区道路分布比较

Fig. 6 Roads in Oldenburg and San Joaquin anonymous regions

布的信息熵来表示. 较大的信息熵表示攻击者分析某个用户位于某条道路较强的不确定性. 使用基于 cloak box 的 K-anonymity 算法作为比较,其基本匿名区域的大小与 KD-anonymity 中蜂窝的平均大小近似,在 Oldenburg 中取 200×200,San Joaquin 中为 2 500×2 500.

实验结果如图 7、8 所示,其横坐标为匿名组的 k 值,纵坐标 e 为信息熵. 从图中可以看到在 Oldenburg 和 San Joaquin 两个实验中获得的结果显示了相同的变化趋势,KD-Anonymity 中 NRF 成组最终得到的匿名组平均信息熵值大于 NNF,而 NNF 和 NRF 都明显大于 K-anonymity 算法. 而随着用户密度的降低,San Joaquin 实验中显示两者区别明显加大.

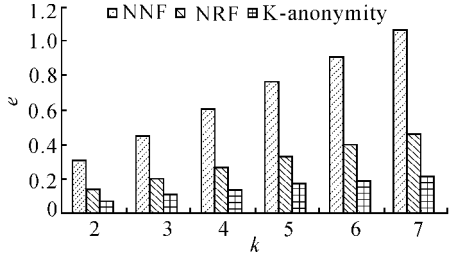

图 7 Oldenburg 匿名组平均信息熵

Fig. 7 Entropy in Oldenburg experiment

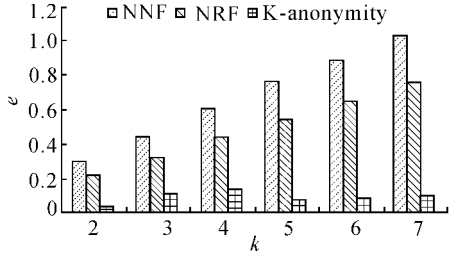

图 8 San Joaquin 匿名组平均信息熵

Fig. 8 Entropy in San Joaquin experiment

3.2.3 **一次匿名成功率** 论文使用一次匿名成功率来分析不同算法的效率. 一次匿名成功率表示匿名时没有经过匿名区域的合并就达到匿名要求的比率,显然,较大的匿名成功率代表更短的用户响应时间及算法性能. 在计算成功率时,论文没有区分 NNF 和 NRF.

图 9 和图 10 展示了不同 k 值时一次匿名成功率的变化. 可以观察到随着 k 值的变化,不管是 Oldenburg 和 San Joaquin,一次匿名成功率都逐渐降低,但是基于 cloak box 的 K-anonymity 算法降幅明显大于 KD-anonymity.

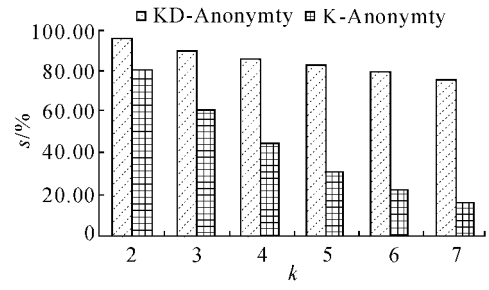

图 9 Oldenburg 一次匿名成功率

Fig. 9 First round success rate in Oldenburg experiment

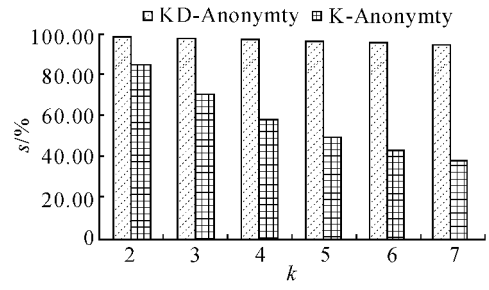

图 10 San Joaquin 一次匿名成功率

Fig. 10 First round success rate in San Joaquin experiment

4 结 语

在道路网络限制移动的环境中,移动用户向不可信的 LBS 提供自身位置信息获取服务的时候面临推理攻击而导致位置隐私泄漏的风险. 一般的位置隐私保护算法仅考虑在欧几里德空间内用户的移动问题,不完全适用路网限制的上下文环境. 本文提出了一个基于匿名蜂窝的位置隐私保护方法,根据路网环境的点、线网络拓扑结构,先将路网网络处理成匿名蜂窝网络,基于匿名蜂窝对用户的查询进行处理,从而增强应对推理攻击的鲁棒性.

在今后的研究中,将对算法进一步完善,并对其在实际应用中的其他问题进行研究,例如较长道路和单行线道路中匿名蜂窝适应性移动的可行性分析;路网限制环境中连续查询攻击问题等.

(下转第 439 页)

1574.

[12] 王国瑾，蒋素荣. 两类新的广义 Ball 曲线曲面的求值算法及其应用[J]. 应用数学学报，2004，27(1)：52－63.
WANG Guo-jin, JIANG Su-rong. The algorithms for evaluating two new types of generalized Ball curves/surfaces and their applications [J]. **Acta Mathematicae Applicatae Sinica**, 2004, 27(1): 52－63.

[13] 檀结庆，方中海. 区间 Wang-Said 型广义 Ball 曲线的降阶[J]. 计算机辅助设计与图形学学报，2008，20(11)：1483－1493.
TAN Jie-qing, FANG Zhong-hai. Degree reduction of interval generalized ball curves of Wang-Said type [J]. **Journal of Computer-Aided Design and Computer Graphics**, 2008, 20(11): 1483－1493.

[14] 葛传丰，朱晓临. G2 保形的分段 4 次广义 Ball 插值曲线[J]. 合肥工业大学大学学报：自然科学版，2008，31(11)：1875－1877.
GE Chuan-feng, ZHU Xiao-lin. G2 shape preserving segmented quartic generalized Ball curve [J]. **Journal of Hefei University of Technology: Natural Science**, 2008, 31(11): 1875－1877.

[15] 田奕丰，王国瑾. 有理三次圆弧的标准正交基与广义 Ball 基表示[J]. 浙江大学学报：工学版，2009，43(1)：1－7.
TIAN Yi-feng, WANG Guo-jin. Representing rational cubic circular arc by normalized totally positive or generalized Ball methods [J]. **Journal of Zhejiang University: Engineering Science**, 2009, 43(1): 1－7.

[16] 王国瑾，汪国昭，郑建民. 计算机辅助几何设计[M]. 北京：高等教育出版社；海德堡：施普林格出版社，2001：7－8.

[17] DELGADO J, PENA J M. A shape preserving representation with an evaluation algorithm of linear complexity [J]. **Computer Aided Geometric Design**, 2003, 20(1): 1－10.

(上接第 434 页)

参考文献(References)：

[1] KARIM W. Privacy implications of personal locators: why you should think twice before voluntarily availing yourself to GPS monitoring[J]. **Journal of Law & Policy**, 2004, 14: 485－515.

[2] BARKHUUS L, DEY A. Location-based services for mobile telephony: a study of users's privacy concerns [C] // **Proceedings of Interact**. Zurich: IOS, 2003: 709－712.

[3] SWEENEY L. K-anonymity: a model for protecting privacy[J]. **Journal on Uncertainty, Fuzziness and Knowledge-based Systems**, 2002, 10(5): 557－570.

[4] GRUTESER M, GRUNWALD D. Anonymous usage of location based services through spatial and temporal cloaking[C] // **Proceedings of Mobile Systems, Applications and Services**. San Francisco: ACM, 2003:31－42 .

[5] KIDO, YANAGISAWA H. A anonymous communication technique using dummies for location-based services[C] // **Proceedings of ICPS**. Santorini: IEEE, 2005:88－97.

[6] GHINITA G, KALNIS P, SKIADOPOULOS S. PRIVE: anonymous location-based queries in distributed mobile systems[C] // **Proceedings of World Wide Web**. Banff: ACM, 2007:371－380.

[7] MOHAMED M, CHOW Chi-Yin, WALID A. The new Casper: query processing for location services without compromising privacy[C] // **Proceedings of Very Large Data Bases**. Korea: VLDB, 2006:763－774.

[8] BAMBA B, LIU L, PESTI P. Supporting anonymous location queries in mobile environments with privacy grid[C] // **Proceedings of World Wide Web**. Beijing: IEEE, 2008: 237－246.

[9] GEDIK B, LIU L. Protecting location privacy with personalized k-anonymity: architecture and algorithms[J]. **IEEE Transactions on Mobile Computing**, 2008, 7(1):1－17.

[10] XU J, TANG X, HU H, et a1. Privacy-conscious location-based queries in mobile environments[J]. **IEEE Transactions on Parallel and Distributed Systems**, 2010, 21(3):313－326.

[11] KYRIAKOS M, YIU Man-Lung. Anonymous query processing in road networks[J]. **IEEE Transactions on Knowledge and Data Engineering**, 2010, 22(1):2－15.

[12] MOON B, JAGADISH H V, FALLOUTSOS C, et a1. Analysis of the clustering properties of the Hilbert space-filling curve [J]. **IEEE Transaction on Knowledge and Data Engineering**, 2001, 13(1):124－141.

[13] THOMAS B. Generating network-based moving objects[C] // **Proceedings of International Conference on Scientific and Statistical Database Management**. Berlin: IEEE, 2000:253－256.