DOI：10.7652/xjtuxb201402002

位置服务中的查询隐私度量框架研究

张学军[1,2]，桂小林[1]，冯志超[1]，田丰[1]，余思[1]，赵建强[1]
（1.西安交通大学电子与信息工程学院，710049，西安；2.兰州交通大学电子与信息工程学院，730070，兰州）

摘要：针对查询隐私度量机制存在过高评价用户隐私保护水平的问题，提出一个泛化的查询隐私度量框架。该框架将影响用户查询隐私的各种因素放在一起统一考虑，形式化定义用户、攻击者、隐私保护机制和隐私度量指标，提供了一种融合攻击者背景知识和推理能力的系统的隐私度量方法，可在相同条件下正确度量多种查询隐私保护机制的有效性，帮助用户选取合适的隐私需求以获得隐私保护和服务质量之间的平衡。在由 Thomas Brinkhoff 路网数据生成器生成的数据集上进行的模拟实验验证了该框架的有效性和准确性。

关键词：基于位置的服务；查询隐私；隐私度量；隐私保护机制
中图分类号：TP309 **文献标志码**：A **文章编号**：0253-987X(2014)02-0008-06

A Quantifying Framework of Query Privacy in Location-Based Service

ZHANG Xuejun[1,2], GUI Xiaolin[1], FENG Zhichao[1], TIAN Feng[1], YU Si[1], ZHAO Jianqiang[1]
(1. School of Electronics and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China;
2. School of Electronics and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China)

Abstract: A generalized quantification framework for evaluating user query privacy is proposed to address the overestimation of user's privacy protection level offered by query privacy protection mechanism in location-based services. The framework is to comprehensively take the various elements and relations that influence the query privacy of mobile users into account together and to formally define the mobile user, adversary, privacy protection mechanism and privacy metrics which reflect the user's privacy and service quality requirements. Moreover, the framework provides a systematic approach with integrating adversary's background knowledge available to and reasoning abilities in privacy quantification. The method can correctly evaluate the effectiveness of various query privacy protection mechanisms under the same conditions, and can help user to select appropriate privacy requirements to achieve a right tradeoff between privacy protection and service quality. The effectiveness and accuracy of the framework are verified through a set of experiments on datasets generated by network-based generator of moving objects proposed by Thomas Brinkhoff.

Keywords: location-based service; query privacy; privacy metric; privacy protection mechanisms

智能移动设备(例如智能手机、PDA 等)的广泛使用促进了位置服务(location-based service, LBS)的快速发展。据美国 ABI Research 最新预测，LBS 全球总收入将由 2009 年的 26 亿美元上升到 2014

收稿日期：2013-05-18。 作者简介：张学军(1977—)，男，博士生；桂小林(通信作者)，男，教授。 基金项目：国家自然科学基金资助项目(61172090,61163010)；国家高技术研究发展计划重大专项资助项目(2012ZX03002001)；教育部高等学校博士学科点专项科研基金资助项目(20120201110013)；陕西省科技发展计划"十五"攻关资助项目(2012K06-30)。
网络出版时间：2013-11-13 网络出版地址：http://www.cnki.net/kcms/detail/61.1069.T.20131113.1651.001.html

年的 140 多亿美元[1]。虽然 LBS 为用户提供了极大的方便，但也造成了严重的隐私关注。恶意攻击者可以将包含位置信息的 LBS 查询与用户关联起来，推断出其生活习惯、兴趣爱好、健康状况等个人敏感信息。研究者就如何允许用户在享受 LBS 的同时限制其隐私泄露提出了许多查询隐私保护机制[1-5]，但是，对这些保护机制的评价通常忽略了攻击者可能具有的背景知识和推理能力，而这些背景知识和推理能力可以帮助攻击者减少其对用户真实位置的不确定性[4]。因此，已有的查询隐私度量方法过高地评价了给定保护系统提供的隐私保护水平。而且，目前的 LBS 隐私保护研究主要集中在保护技术研发上，对隐私保护系统可信性及隐私保护机制有效性的评估还不够成熟和完善，明显缺乏一个形式化的框架来说明和度量各种查询隐私保护机制。

本文针对上述问题，设计了一个融合攻击者背景知识和推理能力的形式化查询隐私度量框架。该框架综合考虑 LBS 系统中影响用户查询隐私的各种因素，能正确度量各种查询隐私保护机制的有效性。

1　相关工作

在查询隐私中，位置 k 匿名[5]是最常使用的度量指标，k 是匿名集的大小，k 越大，查询的隐私性越高。林欣等指出，当匿名集中各个用户发送查询的概率不相等时，匿名集的大小就不能反映每个用户的真正匿名性[6]。Shokri 等分析了位置 k 匿名在攻击者具有用户实时位置信息、统计信息和无信息时保护隐私的有效性，得出 k 匿名对查询隐私有效，对位置隐私无效，并指出攻击者可以利用 k 匿名的缺陷推断出用户的当前位置[7]。为此，他们提出了一种基于扭曲的隐私度量指标，通过比较攻击者跟踪用户获得的轨迹和用户真实轨迹之间的差异来反映用户的隐私保护水平[8]。最近，Shokri 等设计了一个位置隐私量化工具，假定攻击者可以利用从用户轨迹采样中抽取的用户移动模型、LBS 访问模式等推断其截获的轨迹的真正制造者，并使用攻击者的期望估计误差作为隐私度量指标[9]，它在思想上和我们的工作很接近。通过对以上工作的研究发现：①已有的查询隐私度量方法大都忽略了攻击者的背景知识和推理能力，而这些信息和隐私度量密切相关。②已有的查询隐私度量方法大都针对某个特定的隐私保护技术，缺乏一个泛化的度量框架来量化各种隐私保护机制，因此建立融合攻击者背景知识和推理能力的查询隐私框架就显得非常必要。

2　查询隐私度量框架

为了正确度量各种查询隐私保护机制在不同攻击模型下的有效性，需要综合考虑 LBS 系统中影响用户查询隐私的各种因素和它们之间的关系，例如用户的时空状态和需求、攻击者的背景知识和推理能力、隐私保护机制实现算法、隐私指标等。本文将这些高度相关的因素放在一起，针对匿名和泛隐私保护机制，设计了一个形式化查询隐私度量框架，如图 1 所示。

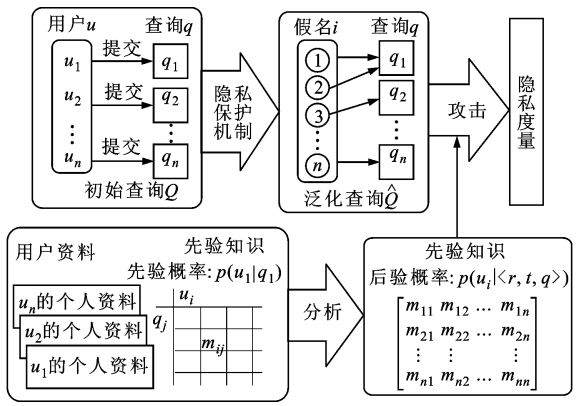

图 1　查询隐私度量框架

定义该框架为六元组 $<U, Q, P_{\mathrm{PM}}, \hat{Q}, A_{\mathrm{attacker}}, M_{\mathrm{metric}}>$，其中 U 是用户 u 的集合。Q 为用户 u 在 t 时刻 l 位置上发出的 LBS 查询集合。P_{PM} 是隐私保护机制，它对查询 $q(q\in Q)$ 进行扭曲处理并产生泛化查询 $\hat{q}(\hat{q}\in\hat{Q})$。$\hat{Q}$ 是攻击者可以观察的泛化查询 $\hat{q}$ 的集合。A_{attacker} 是攻击者，指能够通过 LBS 访问用户位置和查询的任何实体。A_{attacker} 能够利用他对 P_{PM}、用户资料的背景知识和 $\hat{Q}$，通过执行推理攻击推断出 q 的真正请求者。M_{metric} 是查询隐私度量指标，用于度量攻击者的攻击性能。攻击者的推理是统计意义上的，他利用获得的用户资料抽取每个用户发送 q 的先验概率，形成概率矩阵 $\boldsymbol{M}=(m_{ij})$，进而利用 $\boldsymbol{M}$ 尝试重构泛化区域内每个用户发送 $\hat{q}$ 的后验概率。

2.1　移动用户

在移动无线网络环境中，用户是在一定的时空状态下执行 LBS 查询，时空特性不仅包含它的实际状态，也包含来自观察者角度的状态。

记 $U=\{u_1, u_2, \cdots, u_N\}$ 为在区域 L 内移动的 N

个用户的集合，$L=\{l_1,l_2,\cdots,l_M\}$ 为 M 个不同位置的集合，时间 $T=\{1,2,3,\cdots,T\}$ 是离散的，为可记录时间实例的集合。由于用户是移动的，其位置随时间发生变化，所以记函数 whereis：$U\times T\rightarrow L$ 为用户在任意时刻的真实位置。用户 u 在时刻 t 的位置分布定义为集合 $S(t)=\{(u,\text{whereis}(u,t))\mid u\in U\}$。

定义 1（初始查询 Q） 设 Q_u 为 LBS 支持的所有查询 q_u 的集合，则定义用户 u 的初始查询 q 为一个四元组$<u,t,\text{whereis}(u,t),q_u>$，其中 $u\in U$ 为用户标识，$t\in T$ 为查询 q 发生的时间戳，$\text{whereis}(u,t)\in L$ 是用户 u 在 t 时刻与查询 q 相关联的位置，$q_u\in Q_u$ 是查询体。用户 U 在时刻 T 所有可能的初始查询的集合为 $Q\subseteq U\times T\times L\times Q_u$。初始查询 Q 代表了用户在真实环境中的实际状态。

定义 2（泛化查询 $\hat{Q}$） 设 Q_u 为 LBS 支持的所有查询 q_u 的集合，I 为用户所有假名的集合，$P(L)$ 是 L 的幂集，$R\subset P(L)$ 为所有可能的包含多个用户的位置区域集合，则定义用户的泛化查询 $\hat{q}$ 为一个四元组$<i,t,r,q_u>$，其中 $i\in I$ 是用户的假名且可为空，$r\in R$ 是与查询相关联的位置区域且 $\text{whereis}(i,t)\in r$，$q_u\in Q_u$ 是查询体。用户 U 在时刻 T 的所有可能的泛化查询集合为 $\hat{Q}\subseteq I\times T\times R\times Q_u$。泛化查询 $\hat{Q}$ 反映了来自观察者角度的用户时空状态。

2.2 隐私保护机制

LBS 系统中，攻击者能够利用其可用的背景知识从初始查询 Q 中推理出用户的隐私信息或定位其位置。为了防止这种威胁，需要在 LBS 服务器得到用户查询之前，对其进行扭曲处理产生泛化查询 $\hat{Q}$。隐私保护机制就是实现这种扭曲处理的。

定义 3（隐私保护机制 P_{PM}） 设 Q 是初始查询，$\hat{Q}$ 泛化查询，则定义隐私保护机制为映射函数 $f:Q\rightarrow\hat{Q}$。例如，有 $f(q)=\hat{q}$。转换函数 f 实现初始查询到泛化查询的映射，攻击者的目标就是通过观察到的泛化查询 $\hat{Q}$ 的子集，尽力重构出查询 Q。

2.3 攻击者

为了正确评估隐私保护机制的有效性，必须对攻击模型进行建模。攻击模型可通过攻击者的背景知识和推理能力来刻画。对每种隐私保护机制，攻击者所拥有的背景知识和推理能力直接影响攻击这种保护技术的难易程度。攻击者得到的有用背景知识越多，推理能力越强，就越有可能推断出用户的隐私。所以，只有将攻击者的背景知识和推理能力融入到隐私度量中，才能正确度量用户的隐私保护水平。

2.3.1 *攻击者背景知识和推理能力假设* 在隐私度量中融入攻击者的背景知识，就需要知道攻击者可能拥有多少或哪些背景知识，这在实际中是不可行的。因此，一般会对攻击者所拥有的背景知识和推理能力做出各种假设。本文做如下假设[10]：①攻击者具有用户实时位置分布的全局知识；②攻击者可获取匿名服务器转发给 LBS 服务器的任何泛化查询，匿名服务器可信且用户和匿名服务器之间的通信是安全的；③匿名服务器使用的隐私保护方法是公开的，即对匿名集中的每个用户，攻击者通过隐私保护方法得到的匿名区域和其他人进行计算得到的匿名区域一样；④攻击者可利用获得的用户资料得到关于用户的先验知识；⑤攻击者不具备用户隐私需求的决策过程，但在获得匿名服务器转发的泛化查询后，可以了解用户的隐私需求；⑥攻击者不能关联同一用户的任意两个查询。

2.3.2 *攻击者背景知识的表达与量化* 在查询隐私中，攻击者的目标是通过观察到的 $\hat{Q}$ 的子集，利用相关的背景知识尽可能正确地重新识别出 q 的真正发送者 u。攻击者观察到泛化查询子集的范围取决于它的能力。对全局攻击者，它观察到泛化查询的子集等于 $\hat{Q}$。攻击者识别出 $\hat{Q}$ 的真正发送者越正确，用户隐私泄露的可能性就越大。因此，可以将 u 和初始查询 q 以及 u 和泛化查询 $\hat{q}$ 之间的关联描述为条件概率 $p(u|q)$ 和 $p(u|\hat{q})$。将所有 $p(u|q)$ 和 $p(u|\hat{q})$ 都看作变量，而将背景知识表达为这些变量的约束。

本文中，假定攻击者具有用户资料的背景知识。用户资料可描述为与用户相关的一组属性的集合，分为描述信息（例如国家、种族等）、联系信息（例如姓名，地址等）和偏好信息（例如移动模型等）[11]。攻击者还可以通过对用户的观察截取到一些服务请求的背景知识，例如用户的假名、查询内容、位置等。攻击者将具有的背景知识和观察到的信息进行关联攻击，得到 $p(u|q)$ 和 $p(u|\hat{q})$，这就将背景知识表达成了变量 $p(u|q)$ 和 $p(u|\hat{q})$ 的约束。

2.3.3 *先验知识的获取* 用户资料的每个属性都有一定的取值范围，其取值可离散化为明确的形式。例如，性别的取值可表示为“男”、“女”等。这样，每个属性的取值都是有限的。设 $A=\{a_1,a_2,\cdots,a_n\}$ 是要考虑的 n 个属性，则每个用户的资料可表示为 $\Phi_u=\{a_1:v_1,a_2:v_2,\cdots,a_n:v_n\}$，其中 v_i 是用户 u 的属性 a_i 的取值，a_i 的取值可以为空。

给定一个属性 a_i，先将其值离散化。如果是数字型，则把连续数据空间划分为不相交且互斥的间隔区域，离散后的每个属性用二进制串表示。于是用户的资料可以用一个向量 $\boldsymbol{P}_u=<l_1,l_2,\cdots,l_n>$ 表示，其中 l_i 是一个二进制数字序列，该序列的长度等于 a_i 的所有可能离散值的个数。如果 v_i 满足相应离散值，则相应二进制位取为 1，否则为 0。如性别所有可能的值是“男”和“女”，可用两位二进制 男|女 表示：“男”对应的位取 1，为“10”，“女”对应的位取 1，为“01”；

对每个查询 q，都有一个相关属性的子集用于推理该查询的真正请求者。由于用户资料中的每个属性对攻击者识别查询 q 的重要性不同，所以相关属性的每个值都应有一个权重以确定该属性值识别查询 q 的重要程度。例如奢侈品的查询，相关的属性应包括工作、月薪和年龄。在这 3 个属性中，月薪比年龄重要，而且月薪超过 10 000 元比月薪低于 1 000元对该查询更重要。因此，引入相关向量 $\boldsymbol{W}(q)=<w_1,w_2,\cdots,w_n>$ 表示属性值和查询 q 之间的关系，其中 n 为属性的个数。对于任意用户 $u\in U$ 和 $q\in Q$，$\phi(u,q)=\sum_{i\leqslant n}w_i\cdot\boldsymbol{P}_u$ 表示用户 u 的资料和查询 q 的相关程度。于是，用户 u 发送查询 q 的概率为

$$p(u\mid q)=\phi(u,q)/\sum_{u'\in U}\phi(u',q) \tag{1}$$

$p(u|q)$满足约束条件 $\sum_{u_i\in U}p(u_i\mid q_j)=1$。

2.3.4　*后验概率的获取*　设变量 U 为泛化查询 $\hat{q}$ 的请求者，$u_l:R\times T\rightarrow P(U)$ 为把一个区域映射到 t 时刻位于该区域的用户集合的函数 $u_l(r,t)=\{u\in U\mid \text{whereis}(u,t)\in r\}$。于是，给定一个泛化查询 $\hat{q}$，匿名区域 r 内用户 u 发送查询 $\hat{q}$ 的后验概率为

$$p(u\mid\hat{q})=p(u\mid q)/\sum\nolimits_{u'\in u_l(r,t)}p(u'\mid q) \tag{2}$$

2.4　隐私度量

为了客观评价查询隐私保护机制的实际效果，需要建立一些评价指标来度量用户的查询隐私。

2.4.1　*位置 k 匿名*

定义 4（位置 k 匿名）　设 $q=<u,t,\text{whereis}(u,t),q_u>\in Q$ 是一个初始查询，$\hat{q}=<i,r,t,q_u>\in\hat{Q}$ 是其对应的泛化查询。如果匿名集中的所有用户 u 满足 $|\{u\in U\mid\text{whereis}(u,t)\in r\wedge f(q)=\hat{q}\}|\geqslant k$，则称用户 u 是位置 k 匿名。

因为匿名集中的所有用户都把 r 作为 t 时刻泛化查询 $\hat{q}$ 的匿名区域，所以匿名集中所有用户都是位置 k 匿名的。当攻击者具有相关背景知识时，匿名集的大小就不能正确评价用户的查询隐私，特别是后验概率很高的发送用户，很容易被选为候选发送者。

2.4.2　*β 熵匿名*　在查询隐私保护中，如果用户和查询关联的不确定性越大，攻击者越难推断出查询的真正发送者，因此可以借鉴信息论中的熵理论，用攻击者识别泛化查询的不确定性来度量用户的查询隐私。

设变量 U 表示查询 $\hat{q}$ 的请求者，则攻击者对查询的不确定性可表示为

$$E(U\mid\hat{q})=-\sum_{u_i\in u_l(r,t)}p(u_i\mid\hat{q})\log p(u_i\mid\hat{q}) \tag{3}$$

从攻击者的角度看，当它没有任何背景知识时，匿名区域内每个用户发送查询的可能性相同，匿名区域的熵最大，用户的查询隐私保护水平最高。

给定泛化查询 $\hat{q}$ 和 β，如果区域 r 内所有用户提交相同查询后得到的匿名区域都是 r，且 $E(U|\hat{q})\geqslant\beta$，则称查询发送者 u 是 β 熵匿名。

定义 5（β 熵匿名）　设 $\beta>0$，$q=<u,t,\text{whereis}(u,t),q_u>\in Q$ 是初始查询，$\hat{q}=<i,r,t,q_u>\in\hat{Q}$ 是其对应的泛化查询，如果对于所有的用户 $u\in u_l(r,t)$，满足 $E(U\mid\hat{q})\geqslant\beta\wedge f(q)=\hat{q}$，则称用户 u 是 β 熵匿名。

2.4.3　*δ 互信息匿名*　互信息可用于度量两个随机变量之间的相关性。在查询隐私中，可以用互信息度量泛化查询被攻击者截取后查询不确定性的减少程度。攻击者在截取到泛化查询之前，仅知道 q 是由用户 U 以概率 $p(U|q)$ 发送，其不确定性可描述为 $E(U|q)$。攻击者截取到泛化查询 $\hat{q}$ 后，其不确定性可以表示为 $E(U|\hat{q})$。于是，攻击者截取到泛化查询前后不确定性的增益为

$$\begin{aligned}I(U\mid q;\hat{q})&=E(U\mid q)-E(U\mid\hat{q})=\\&-\sum_{u_i\in u_l(r,t)}p(u_i\mid q)\log p(u_i\mid q)+\\&\sum_{u_i\in u_l(r,t)}p(u_i\mid\hat{q})\log p(u_i\mid\hat{q})\end{aligned} \tag{4}$$

$I(U|q;\hat{q})$的值反映了攻击者对查询不确定性的变化情况，$I(U|q;\hat{q})$值越大，攻击者的不确定性就越大，用户的隐私保护水平越高。

给定泛化查询 $\hat{q}$ 和 δ 值，如果区域 r 内所有用户提交相同查询后得到的泛化区域都是 r，且 $I(U|q;\hat{q})\leqslant\delta$，则称查询发送者是 δ 互信息匿名。

定义 6(δ 互信息匿名) 设 $\delta>0$,$q=<u,t,\text{whereis}(u,t),q_u>\in Q$ 是初始查询,$\hat{q}=<i,r,t,q_u>\in\hat{Q}$ 是其对应的泛化查询,如果对所有用户 $u\in u_l(r,t)$,满足 $I(U|q;\hat{q})\leqslant\delta\wedge f(q)=\hat{q}$,则称查询发送者 u 是 δ 互信息匿名。

3 实验分析

利用框架对 dichotomicPoints[10] 隐私保护算法的有效性进行度量。实验在产生用户位置坐标和 LBS 请求时,采用业界认可的 Thomas Brinkhoff 路网数据生成器[12] 模拟生成用户在 Oldenburg 市区(面积为 23.57 km×26.92 km)交通路网内的运动轨迹,随机产生10 000个用户,用户的移动速度采用默认设置。攻击者关于用户资料的背景知识可用 UCI 机器学习中的 Adult 数据集,利用 2.3 节的方法求出。因为本文主要关注框架对隐私保护机制有效性的度量,所以在实验中采用随机产生用户先验知识的方法。模拟算法在 Windows7 平台下使用 Java 实现,测试机器的基本参数为 3.51 GHz Intel Core(TM)i5 处理器、4 GB 内存。实验结果取算法执行 100 次仿真的平均值。

3.1 位置 k 匿名度量

图 2 给出了攻击者有、无背景知识时,利用位置 k 匿名度量查询隐私保护机制的情况。无背景知识时,在攻击者看来,由 dichotomicPoints 算法产生的匿名区域内每个用户发送查询的可能性相同。当观察到泛化查询 $\hat{q}$ 时,攻击者只能推断出匿名区域内用户发送 $\hat{q}$ 的概率为该匿名区域内用户数目的倒数,即图中"无背景知识"对应的曲线。有背景知识时,从攻击者的角度看,匿名区域内每个用户发送查询的概率不再相等,攻击者使用攻击算法找出最有可能发送查询 $\hat{q}$ 的用户,即图中"有背景知识"对应的曲线。首先用户的后验概率随着 k 值的增加而减小,这是因为 k 越大,隐私保护机制产生匿名区域的用户数就越多。其次,给定一个 k 值,攻击者没有背景知识时,发送用户的后验概率均小于 $1/k$,即攻击者不能以大于 $1/k$ 的概率识别出查询发送者,k 的大小能反映用户的隐私水平;攻击者有背景知识时,发送用户的后验概率大于 $1/k$,即攻击者能以大于 $1/k$ 的概率识别出查询发送者,用户的隐私有可能暴露,匿名集的大小已不能正确反映用户的隐私水平。因此,在查询隐私度量中考虑攻击者背景知识是必不可少的,同时也表明查询隐私的度量结果应该包括用户查询隐私水平值和攻击者背景知识的假设,这可帮助用户理解他们隐私受保护的程度以及攻击者背景知识对用户隐私保护的重要性。

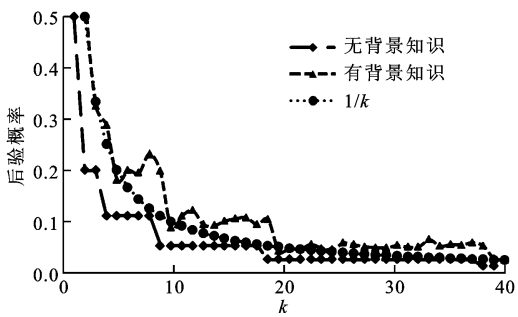

图 2 位置 k 匿名度量

3.2 β 熵匿名和 δ 互信息匿名度量

β 熵匿名和 δ 互信息匿名指标用于度量攻击者对查询的不确定性,反映了用户的查询隐私水平。

一个满足 β 熵匿名的匿名区域可以确保该匿名区域的熵值大于 β,而满足 δ 互信息匿名的匿名区域可确保攻击者对查询不确定性的减少小于 δ。图 3 和图 5 描述了不同 β 和 δ 对应的熵和互信息,可以看出,由 dichotomicPoints 算法产生匿名区域的所有用户都满足 β 熵匿名和 δ 互信息匿名的定义,且当 β 和 δ 分别接近整数时,熵值和互信息值的变化比较急速,这是由熵的本质所确定的。图 4 和图 6 给出了匿名区域平均面积占数据空间的比率随 β 和

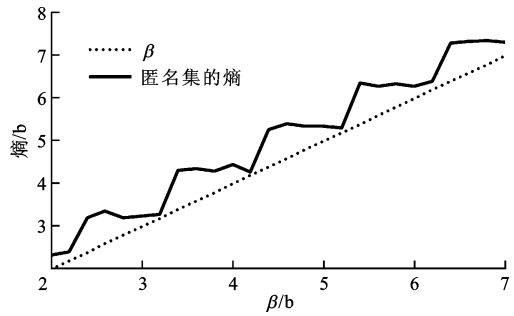

图 3 β 熵匿名度量

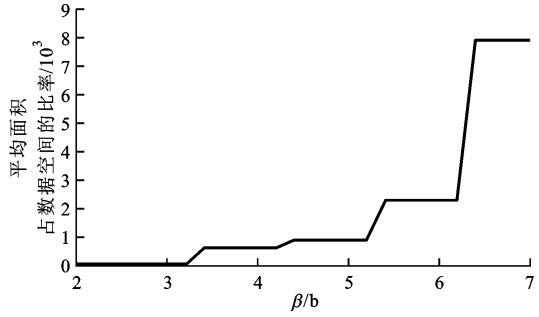

图 4 β 熵匿名区域面积

δ 分别变化的情况，图 4 中匿名区域的面积随 β 增大而增大，这是因为 β 越大，匿名区域的熵越大，匿名区域内包含的用户越多，匿名区域的面积就越大，用户的隐私保护水平也越高，但服务质量越低。如图 6 所示，匿名区域的面积随着 δ 的增大而减小，因为 δ 越大，攻击者的不确性越小，匿名区域内包含的用户数越少，满足隐私需求的匿名区域就越小，用户的隐私保护水平也越低，但服务质量越高。因此，β 熵匿名和 δ 互信息匿名指标能够正确反映攻击者具有背景知识时，查询隐私保护机制提供的用户隐私保护水平，可以帮助用户确定隐私需求和服务质量之间的平衡。

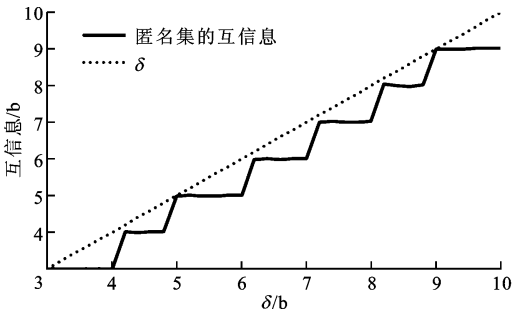

图 5　δ 互信息匿名度量

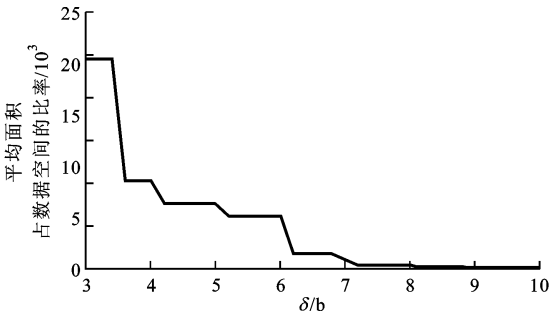

图 6　δ 互信息匿名区域面积

4　结　论

本文提出了一个形式化的查询隐私度量框架。该框架综合考虑了 LBS 系统中影响用户查询隐私的各种因素，定义了攻击者可用背景知识和推理能力的假设及隐私度量指标。实验结果验证了该框架的有效性。同时，对实验结果的分析表明，考虑攻击者的背景知识和推理能力对 LBS 查询隐私保护的度量是必不可少的，在设计隐私保护机制时需要对攻击者的知识和其目标的一致性进行建模，才能设计出更适合实际的查询隐私保护机制。后续工作中，将考虑把位置服务应用集成到框架中，分析隐私保护机制对于这些应用的有效性。另外，对攻击者具有用户移动模型和查询历史等背景知识进行合理的建模也是将要进行的重点工作之一。

参考文献：

[1] PAN Xiao, XU Jianliang, MENG Xiaofeng. Protection location privacy against location-dependent attacks in mobile services [J]. IEEE Transactions on Knowledge and Data Engineering, 2012, 24(8): 1506-1519.

[2] XU T, CAI Ying. Exploring historical location data for anonymity preservation in location-based services [C]// Proceedings of the 27th IEEE International Conference on Computer Communications. Piscataway, NJ, USA: IEEE, 2008: 1220-1228.

[3] PING A, ZHANG Nan, FU Xinwen, et al. Protection of query privacy for continuous location based services [C]// Proceedings of the 30th IEEE International Conference on Computer Communications. Piscataway, NJ, USA: IEEE, 2011: 1710-1718.

[4] SHOKRI R, THEODORAKOPOULOS G, TRONCOSO C, et al. Protecting location privacy: optimal strategy against localization attacks [C]// Proceedings of the 19th ACM Conference on Computer and Communication Security. New York, USA: ACM, 2012: 617-626.

[5] GRUTESER M, GRUNWALD D. Anonymous usage of location-based services through spatial and temporal cloaking [C]// Proceedings of the First International Conference on Mobile Systems, Application, and Services. New York, USA: ACM, 2003: 31-42.

[6] 林欣，李善平，杨朝会. LBS 中连续查询攻击算法及匿名性度量 [J]. 软件学报，2009，20(4)：1058-1068.
LIN Xin, LI Shanping, YANG Zhaohui. Attacking algorithms against continuous queries in LBS and anonymity measurement [J]. Journal of Software, 2009, 20(4): 1058-1068.

[7] SHOKRI R, TRONCOSO C, DIAZ C. Unraveling an old cloak: k-anonymity for location privacy [C]// Proceedings of the 2010 ACM Workshop on Privacy in the Electronic Society. New York, USA: ACM, 2012: 115-118.

[8] SHOKRI R, FREUDIGER J, JADLIWALA M, et al. A distortion-based metric for location privacy [C]// Proceedings of the 8th ACM Workshop on Privacy in the Electronic Society. New York, USA: ACM, 2009: 21-30.

（下转第 37 页）