doi:10.11959/j.issn.1000-436x.2017163

关联概率不可区分的位置隐私保护方法

张磊[1,2]，马春光[1]，杨松涛[1,2]，李增鹏[1]
（1. 哈尔滨工程大学计算机科学与技术学院，黑龙江 哈尔滨 150001；
2. 佳木斯大学信息电子技术学院，黑龙江 佳木斯 154007）

摘 要：首先量化了在快照查询服务和连续查询服务中攻击者可能通过关联关系建立的关联概率攻击方法。然后，针对这些攻击方法提出了与之对应的基于广义差分隐私的隐私保护模型。基于建立的隐私保护模型设计了基于位置偏移产生关联概率不可区分的隐私保护方法，并证明了这种方法的隐私保护效力。最后，通过实验进一步验证所提模型和方法的隐私保护效力和算法执行效率。

关键词：基于位置服务；关联概率；差分隐私；隐私保护
中图分类号：TP311 **文献标识码**：A

Correlation probability indistinguishable location privacy protection algorithm

ZHANG Lei[1,2], MA Chun-guang[1], YANG Song-tao[1,2], LI Zeng-peng[1]
(1.College of Computer Science and Technology, Harbin Engineering University, Harbin 150001,China;
2. College of Information and Electronic Technology, Jiamusi University, Jiamusi 154007, China)

Abstract: The attack of correlation probability used by the adversary in snapshot query and continuous query services were measured. Then, on account of these attacks, a privacy protection framework based on the differential privacy and a location-shift scheme to achieve the indistinguishable of correlation probabilities was provided, and the protection effectiveness of this method was proved. At last, security analysis and the experiment results further verify the protection effectiveness and guarantee the execution efficiency of the proposed scheme.

Key words: location-based service, correlation probability, differential privacy, privacy protection

1 引言

随着无线通信技术和定位技术的逐渐成熟，基于位置服务（LBS, location based service）得到了广泛的应用。这种通过用户向服务器提交位置相关信息的服务（如查询最近的餐厅或在半小时内每5 min查询最近的加油站等）为使用者带来极大的便利。然而，为了享受这种便利，用户需要向位置服务器提交位置相关信息，这些信息可能会造成用户隐私的泄露。随着对隐私问题的逐渐关注，研究者基于位置服务的服务方式分别提出了基于快照查询服务的隐私保护方法[1~11]和基于连续查询服务的隐私保护方法[12~18]。这些方法大多针对当前用户的真实位置，采用泛化、扰乱以及模糊的方式，降低真实位置被攻击者识别的概率，以此来保护用户的位置隐私。但是，随着位置服务数据的大量产生以及数据挖掘技术和位置预测技术的快速发展，攻击者可利用其掌握的大量背景知识，使用更多的方法去获取用户的隐私信息，进而导致已有的隐私保护方法很难抵抗基于背景知识统计预测攻击[19~22]。例如，

收稿日期：2016-12-19；**修回日期**：2017-06-14
通信作者：马春光，machunguang@hrbeu.edu.cn
基金项目：国家自然科学基金资助项目（No.61472097）；高等学校博士学科点专项科研基金资助项目（No.20132304110017）；黑龙江省自然科学基金资助项目（No.F2015022）
Foundation Items: The National Natural Science Foundation of China (No.61472097), The Specialized Research Fund for the Doctoral Program of Higher Education (No.20132304110017), The Natural Science Foundation of Heilongjiang Province (No.F2015022)

当用户需要查询最近的海滨浴场时，传统的泛化方法将会生成至少 k–1 个相似查询来隐藏用户的真实位置，但是在泛化的位置集合中 k 个位置针对内容为海滨浴场的查询，显然存在用户位于海边的概率大于用户位于购物中心概率的情况。这使攻击者能够有效地剔除匿名位置，进而获得用户的位置隐私。同样，在连续查询服务中，相同查询的移动用户位置转移概率也存在差异，攻击者同样可以基于以上转移概率之间存在的差异识别出用户的真实轨迹。

对于这些基于统计预测攻击，研究者们提出了文献[20, 23, 24]等方法分别保护用户在快照查询服务和连续查询服务中的位置隐私。然而，这些方法多是基于 k-匿名模型[1]提出的，即假设存在至少 k–1 个位置具有与真实用户相似的查询或转移概率，使攻击者无法在这些提交的位置中通过概率差异识别出用户的真实位置。但是，当攻击者掌握更多的背景知识（如 k–1 个匿名位置的真实情况）时，攻击者可使用匿名集合与背景集合做差的方法获得真实位置信息，该攻击被视为差分攻击（differential attack）。面对这种差分攻击，Dwork[25]提出了差分隐私概念和隐私保护模型，该模型一经提出就引起了广泛的研究与关注，并在数据挖掘[26, 27]以及数据发布[28, 29]等方面被大量使用。同时，在位置隐私保护方面，Dewri 等[30]基于差分隐私提出了位置不可区分性；Andrés 等[31]提出了地理不可区分性；随后，Bordenabe 等[32]对这种地理不可区分性进行了优化；Primault 等[33]对这种地理不可区分性加以实现并在实际环境中加以应用；Chatzikokolakis 等[34]将这种地理不可区分性应用到位置轨迹的隐私保护中；Perazzo 等[35]提出了查询一致的隐私保护方法。这些方法基于差分隐私模型，在一定程度上保护了用户的位置隐私。

但是，当前已有的方法多数针对的都是位置概率的差分隐私问题，而基于位置服务并不仅是单纯的随机位置，还存在着特定查询对位置的关联关系。攻击者可以通过这种关联关系，在大量的查询与位置数据之间建立关联概率，通过这种关联概率识别出用户的真实位置，进而获得用户的位置隐私。针对这种情况，本文首先定义了在快照查询服务和连续查询服务中存在的这种关联概率，并基于这些关联概率提出了潜在的关联概率攻击模型。同时，利用这些模型对已有的隐私保护方法进行攻击测试，获得了较好的攻击效果。其次，针对这些攻击模型，本文建立了广义的基于差分隐私的概率不可区分性定义，并根据该定义提出了基于位置偏移的隐私保护方法，该方法能够有效地满足差分隐私模型的基本要求，抵抗多种潜在的基于概率分析、预测的攻击方法。最后，在安全性分析中通过信息论中的熵和互信息，进一步证明了该方法对以上攻击的抵抗能力，并通过实验进一步验证了所提算法的隐私保护效力和执行效率。

本文的主要贡献可概括如下。

1) 提出了基于关联概率和关联转移概率的攻击模型，并对当前存在的部分隐私保护方法进行了攻击测试。

2) 提出了广义的基于差分隐私的概率不可区分性定义，并根据该定义提出了基于位置偏移的隐私保护方法。

3) 通过信息论中的熵和互信息，证明了所提出算法的安全性。

4) 模拟实验验证了所提算法的有效性和算法执行效率。

2 相关工作

随着人们开始越来越多地关注基于位置服务中的隐私问题，近年来，研究者提出了大量的隐私保护方法。基于服务方式，这些隐私保护方法可以分为面向快照查询服务的隐私保护方法和面向连续查询服务的隐私保护方法 2 种。

在面向快照查询服务的隐私保护方法中，k-匿名方法[1]最早被 Gruteser 等从数据发布的隐私保护引入到位置隐私保护，该方法通过寻找至少 k–1 个用户与真实用户同时提交查询，以此泛化真实查询达到保护用户位置隐私的目的。Gedik 等[2]通过建立匿名区域，最大程度地模糊了用户的真实位置。Liu 等[3]针对存在的同质攻击，提出了查询多样性的概念。Rebollo 等[4]针对中心服务器的不可信问题，提出了用户协作的查询信息交换方法。Rebollo 等[5]又对这种方法加以改进，提出在协作用户交换查询之后，由协作用户集合中具有最大熵值的用户提交查询的方法。Niu 等[7]针对用户协作的隐私保护方法中协作用户一般与真实用户之间的距离和最小的情况，提出了基于方差的攻击方法。同时，针对这种方法提出了随机行走的隐私保护方法，进一步提高用户协作隐私保护方法的隐私保护效力。在众多基于快照查询服务的隐私保护方法中，基于加密技术

的隐私保护方法能够提供最大的隐私保护效力。在这种类型的隐私保护方法中，Khoshgozaran 等[8]提出的可计算 PIR（privacy information retrival）是最为典型的应用。在该方法中，用户提交的信息不会泄露给包括位置服务器在内的任何用户以外的实体，实现了零隐私泄露。之后，Khoshgozaran 等[9]又提出硬件 PIR 方法，进一步提高了这种方法的执行效率。Lien 等[10]基于加法同态加密的主要思想，建立了安全的 *k*-nearest neighbor (*k*NN)查询方法。Paulet 等[11]基于不经意传输和 PIR 技术，同时保护了用户的隐私和位置服务器中保存的数据。

在面向连续查询服务的隐私保护方法中，Niu 等[24]通过可信第三方服务器(TTP)缓存历史查询的方法，最大限度地使用缓存中的数据完成用户的查询申请，降低了与位置服务器之间的交互。Ma 等[13]使用 Voronoi 优化了锚点的部署位置，用大量的锚点位置代替用户的真实位置，在连续的位置服务中隐藏用户的真实轨迹。Schlegel 等[14]将查询区域进行网格划分，并通过选择不同网格进行兴趣点查询，以此扰乱攻击者所能获得的用户位置。Wang 等[15]利用路网的区段特点，将用户的真实位置扩展为相应路段，以此模糊用户的连续位置。Palanisamy 等[16]使用 mix-zone 来切断连续位置之间的关联关系，并通过不规则的 mix-zone 形状变换以及时间等待，降低进出 mix-zone 之间的用户关联概率。Gao 等[17]假设协作用户可共享真实用户的行进方向，进而产生相同方向的行进路径，实现位置轨迹匿名。Hwang 等[18]通过选择历史轨迹进行匿名，实现了轨迹的时间间隔、位置和轨迹分段的共同匿名。

然而，以上这些方法主要是基于 *k*-匿名隐私保护发展而来，这种模型的最大问题是无法提供一种有效且严格的方法来证明其隐私保护水平，Dwork[25]提出的差分隐私保护模型很好地解决了这一问题。基于差分隐私保护模型，Dewri 等[30]提出了位置不可区分性，即在基于位置扰乱机制的基础上，要求扰乱后的位置之间彼此满足差分隐私要求，实现位置集合的不可区分。同样，基于差分隐私保护模型，Andrés 等[31]提出了地理不可区分性，通过在用户所在的服务申请范围内大量添加噪声，使噪声位置在地理上与真实位置不可区分来保护用户的位置隐私。之后，基于这种地理不可区分性，Bordenabe 等[32]降低了这种方法的约束条件，优化了这种机制对位置集合的计算复杂性。Primault 等[33]通过在真实环境中应用这种地理不可区分方法，并尝试对这种方法加以攻击。最后，通过攻击结果展示了这种方法存在的不足。Chatzikokolakis 等[34]对这种地理不可区分性进一步扩展，并将其应用在连续查询服务的隐私保护中。针对位置服务的二元特性，Perazzo 等[35]基于差分隐私保护模型提出了查询一致的隐私保护方法。

但是，这些当前存在基于差分隐私模型的隐私保护方法主要针对位置服务中的单一随机变量，而位置服务的二元特性使基于特定查询的关联概率更容易获得用户的位置隐私。针对这一问题，本文从位置和查询之间的关联关系入手，根据差分隐私保护模型的基本特性，提出了概率不可区分模型，并依据该模型制定了位置偏移算法，以此来保护在位置和查询关联概率攻击下的用户位置隐私。

3　预备知识

3.1　相关概念

在 2 种不同的服务中，攻击者可通过掌握的大量背景知识获得 2 种不同服务之间的位置查询相关概率，这 2 种概率可被称为关联概率和关联转移概率。

定义 1　（关联概率）关联概率是指对于指定查询 q 与位置集合 $L=\{l_1,l_2,\cdots,l_n\}$ 中任意位置 l_i $(1\leqslant i\leqslant n)$，其对应的百分比可表示为 $cor(l_i,q)=\dfrac{I_t(l_i,q)}{\sum_{i=1}^{n}I_t(l_i,q)}$，其中，$I_t$ 表示 l_i 与 q 同时出现的次数。

定义 2　（关联转移概率）对于指定查询 q 和非敏感区域中的位置 l，以及当前不确定性位置集合 $L=\{l_1,l_2,\cdots,l_n\}$，存在条件概率 $cor_t(l\to l'|q)=\dfrac{I_t(l\to l'\,|\,q)}{\sum_{i=1}^{n}I_t(l\to L\,|\,q)}$ 使用户从位置 l 移动到位置 l'，其中，I_t 表示历史数据中的转移次数。

3.2　攻击模型和攻击效果

基于以上 2 种不同服务方式下计算得到的位置查询关联概率，攻击者可以发起 4 种不同的攻击来获取用户位置。

定义 3　（极大关联攻击）对于指定查询 q，存在不确定性位置集合 $L=\{l_1,l_2,\cdots,l_n\}$ 中的任意位置 l_i，满足 $cor(l_i,q)=\max(cor(L,q))$，则攻击者可根据 $cor(l_i,q)$ 猜测 l_i 是用户提交的真实当前位置。

定义 4　（关联差分攻击）针对指定的查询 q，

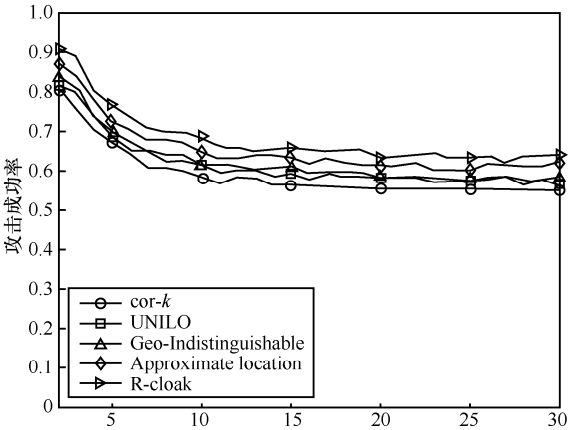
存在不确定性位置集合 $L=\{l_1,l_2,\cdots,l_n\}$ 中的任意子集 L'，满足 $cor(L,q)-cor(L',q)=cor(l,q)$，其中，$l$ 是用户当前的真实查询位置，则攻击者在掌握 $cor(L',q)$ 的情况下可获得用户的真实位置。

定义 5 （极大关联转移攻击）对于指定查询 q 和非敏感区域中的位置 l，以及当前不确定性位置集合 $L=\{l_1,l_2,\cdots,l_n\}$，存在 $cor_t(l\to l'|q)=\max(cor_t(l\to L|q))$，则攻击者可根据 $cor_t(l\to l'|q)$ 猜测 l'是用户转移后的真实位置。

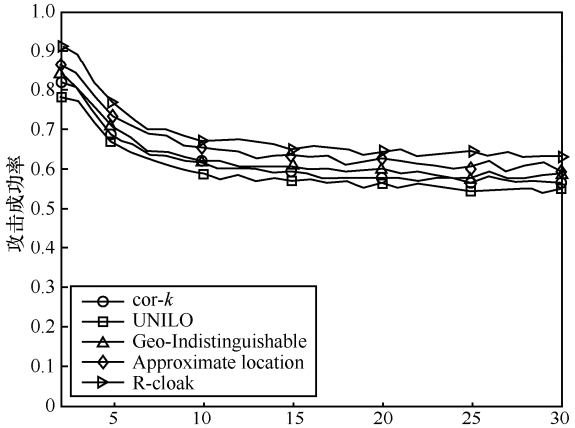
定义 6 （关联转移差分攻击）对于指定查询 q 和非敏感区域中的位置 l，存在不确定性位置集合 $L=\{l_1,l_2,\cdots,l_n\}$ 中的任意子集 L'，满足 $cor_t(l\to L|q)-cor_t(l\to L'|q)=cor_t(l\to l'|q)$，则攻击者可根据 $cor_t(l\to l'|q)$ 猜测 l'是用户转移后的真实位置。

为验证以上潜在攻击方法的攻击效果，本文针对基于快照查询服务隐私保护的 cor-k[23]、UNILO[35]、Geo-Indistinguishable[32]、Approximate location[30]和 R-cloak[7]等 5 种算法，以及针对连续查询服务隐私保护的 Enhanced-CaDSA[12]、Cooperative[13]、Snet[15]、DGS[14]和 MobiMix[16]等 5 种方法，分别采用上面提出的针对 2 种不同服务的隐私攻击方法进行攻击测试，所产生的攻击结果如图 1 和图 2 所示。

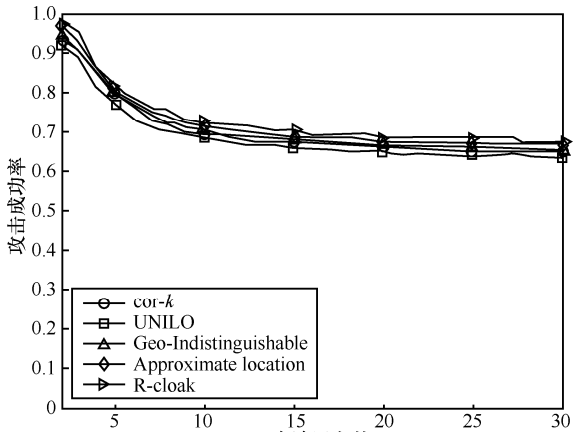
从图 1 中可以看到，将极大关联攻击和关联差分攻击相结合的方法其攻击效果要远好于单独使用任一种攻击方法。极大关联攻击对以上隐私保护方法的攻击效果如图 1(a)所示，从图中可以看出，cor-k 能够较好地抵抗极大关联攻击，这是由于这种方法主要是基于对关联概率的泛化建立的。而查询一致性的方法 UNILO 和地理不可区分的方法 Geo-Indistinguishable 也在匿名人数增加的情况下相对较好地抵抗这种攻击。位置相似的方法 Approximate location 与前 2 种方法相差不大，主要是因为相似位置与地理不可区分性存在一定的位置差异，进而导致关联概率之间的差异使该方法在抵抗此类攻击方面存在一定不足。最后，尽管 R-cloak 考虑到某一位置是否存在查询，但是这种方法并不能很好地模糊关联概率之间的差异，因此，造成在极大关联攻击下其隐私保护效力最低。从图 1(b)中可以看到在关联差分攻击下，查询一致性的方法 UNILO 在保证查询一致的情况下，存在部分位置关联概率差与真实位置之间的差异，进而使其有较好的隐私效力表现。cor-k 和 Geo-Indistinguishable 在相关概率泛化和地理不可区分的基础上，同样可产生少部分的概率差的模糊效果使这 2 种方法具备一定的隐私保护效力。最后，Approximate location 由于和 Geo-Indistinguishable 稍有不同，其隐私保护效力稍差，而 R-cloak 在关联概率差分攻击下的隐私保护效力最差。从图 1(c)中可以看到，将 2 种攻击方法结合使用，极大地提高了攻击能力，造成以上攻击效果的原因与前述分析相同。



(a) 极大关联攻击的攻击结果

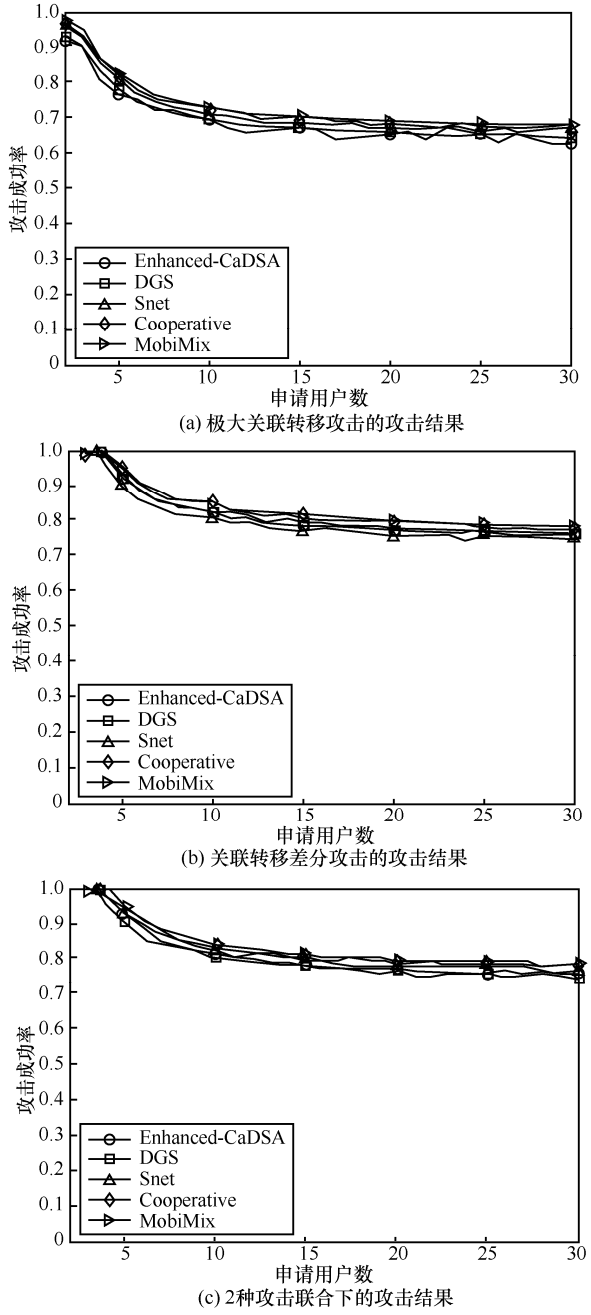


(b) 关联差分攻击的攻击结果



(c) 2种攻击联合下的攻击结果

图 1 针对部分快照查询下隐私保护方法的攻击结果



(a) 极大关联转移攻击的攻击结果

(b) 关联转移差分攻击的攻击结果

(c) 2种攻击联合下的攻击结果

图 2　针对部分连续查询下隐私保护方法的攻击结果

从图 2 中可以看到，极大关联转移攻击和关联转移差分攻击对连续查询服务进行保护的方法存在较高的攻击成功率，这主要是由于这些算法多数隐藏真实位置，而未能对转移概率加以泛化或隐藏，使基于概率的攻击方法取得较好的攻击效果。从图 2(a)中可以看到，随着匿名人数的增加，极大关联转移攻击的成功率始终高于 0.6。其中，Enhanced-CaDSA 由于大量的提供缓存结果使其攻击成功率最低。DGS 由于采用了泛化兴趣点的方法，其对极大关联转移攻击的抵抗能力与泛化当前位置成为路段的 Snet 和将真实位置转换成锚点位置的 Cooperative 相似，这些方法的位置隐藏能力在一定程度上阻止了此类攻击。而 MobiMix 采用的 mix-zone 方法显然由于其在非 mix-zone 区域中的位置提交，使关联转移攻击成功率最高。从图 2(b)中可以看到，在连续查询服务中，通过关联转移差分攻击对以上保护方法的攻击效果极好，这是由于在连续查询服务过程中，在识别转移概率且已知查询的情况下，攻击者能够最大限度地获知用户的当前位置，进而获得用户隐私，此时，以上方法均无法抵抗这种攻击，产生的攻击成功差异主要取决于暴露位置的数量以及转移概率矩阵预测的成功率。从图 2(c)中可以看到，在 2 种方法结合的情况下，产生作用的主要是关联转移差分攻击，这是由于大量的位置被关联转移差分攻击所识别，剩下的位置一般是极大关联转移攻击所无法识别的位置。

综上，可以认为所提出的攻击方法对当前较常见的隐私保护方法有较好的攻击效果，因此，需要提出新的隐私保护模型以保护用户位置隐私。

3.3　隐私保护模型和基本思想

由于以上攻击方式均基于位置与查询之间的相关概率产生，因此，可以基于这种概率关系提出一种概率不可区分性。基于该不可区分性，本文将概率不可区分性计算及位置偏移算法部署在中心服务器上，由可信的中心服务器提供隐私保护服务，并将 LBS 服务器反馈的查询结果提纯后发送给用户，最终完成查询服务。

对于指定查询 q 和不确定位置集合 $L=\{l_1,l_2,\cdots,l_n\}$，其关联概率满足概率函数 $K(L,q)\to P(Z)$，其中，Z 是记录值集合，因此，概率分布的相似性可以定义为

$$M_P(\mu_1,\mu_2)=Sup_{z\in Z}\,|\mu_1(z)-\mu_2(z)|$$

当 $\mu_1(z)$ 和 $\mu_2(z)$ 同时为 0 或∞时，$|\mu_1(z)-\mu_2(z)|=0$，即 $M_P(\mu_1,\mu_2)$ 在 μ_1、μ_2 对每个值 z 有相似概率。

M_P 表示 p 和 p'之间的不可区分级别，该值越小，表示 2 个位置针对某一指定查询的不可区分性越大，而当该值超过某一限定时，可认为攻击者能够区分 2 个提交位置。由此，可以得到在 $M_p(p,p')$ 度量下的广义差分隐私（ε—位置—查询关联隐私的定义）。

定义 7　机制 $P(K(L),q)\to P(Z)$ 满足 ε—位置—

查询关联隐私，当且仅当对于不确定性位置集合中的任意位置 l 和 l'，存在

$$M_P(p(K(l),q),p(K(l'),q)) \leqslant \varepsilon M_P(p(l,q),p(l',q)) \quad (1)$$

式(1)可等价表示为 $K(l,q)(z) \leqslant \mathrm{e}^{\varepsilon M_P(p(l,q),p(l',q))} \cdot K(l',q)(z)$，对于所有 $z \in Z$，参数 ε 可看作是对 M_P 的缩放。

针对广义差分隐私，传统的解决方法添加满足拉普拉斯分布的随机噪声。使用随机算法 $f:l \to l'$ 作用在概率 p 上，有 $f(L,q)=p(L,q)+Y,\ Y \sim Lap\left(\frac{\Delta}{\varepsilon}\right)$，其中，$Lap\left(\frac{\Delta}{\varepsilon}\right)$ 的概率密度函数为 $P(x)=\frac{\varepsilon}{2\Delta}\mathrm{e}^{-\frac{|x|\varepsilon}{\Delta}}$。但这种方法更适合于发布数据的隐私保护，而在位置服务的隐私保护中，由于位置与查询之间关联概率保存在位置服务器中，而位置服务器又被视为半可信实体，即该实体能够完成基于位置的服务请求，但对用户的位置信息存在非恶意的好奇，所以，简单地改变具体位置的相关概率是不能更改位置服务器中通过历史数据计算得到的这种概率值。因此，在相关概率中添加拉普拉斯噪声的方法无效。

基于概率泛化的思想，可以采用一种相同查询下的位置偏移方法来实现随机算法 $f:l \to l'$，使对不确定性位置集合 $L=\{l_1,l_2,\cdots,l_n\}$ 中的任意子集 L'，有 $M_P(K(l,q),K(l',q)) \leqslant \varepsilon M_P(p(l,q),p(l',q))$。在真实位置关联概率过高，即偏移后的位置仍无法满足隐私保护条件时，算法执行失败。

4 隐私保护方法

4.1 面向快照查询服务的隐私保护方法

基于位置偏移的基本思想，可实现各种关联概率泛化。因此，可设计位置随机算法 $f:l \to l'$，对不确定性位置集合 $L=\{l_1,l_2,\cdots,l_n\}$ 中的位置进行偏移，使每个位置对应的位置查询关联概率与当前位置集合 L 中的平均概率不可区分，由此可得到如下定理。

定理 1 随机算法 $f:l \to l'$ 满足 ε—位置—查询关联隐私。

证明 对于不确定位置集合 L 中的任意 2 个位置 l 和 l'，存在随机算法 $f:l \to l'$，使 $|cor(f(l),q)-avg(cor(L,q))| \leqslant \varepsilon$，则该随机算法满足

$$\begin{aligned} &M_P(cor(f(l),q),cor(f(l'),q)) \\ &\quad = M_P(avg(cor(L,q)) \pm \varepsilon, avg(cor(L,q)) \pm \varepsilon) \\ &\quad \leqslant \varepsilon M_P(cor(l,q),cor(l',q)) \end{aligned} \quad (2)$$

其中，$avg()$ 为取平均值。由此可知随机算法 $f:l \to l'$ 满足 ε—位置—查询关联隐私。

基于定理 1，可以得到不确定位置集合 L 的转移过程，其执行过程如算法 1 所示。

算法 1 快照查询下的位置偏移

输入 $S, cor(S,q), L \in S, \varepsilon$

输出 L'

1) p_a=avg(cor (L, q));
2) for (i=1; i<=n; ++i)
3) if($|cor(l'_i,q)-p_a| <= \varepsilon$)
4) L'=l'_i+ L';
5) $T(i)$= l_i–l'_i;
6) end if
7) end for

在算法 1 中，通过步骤 2)~步骤 7)迭代对不确定位置集合 L 中的每个位置进行转移，并保存在集合 L'中。同时，将每步的转移距离保存在 $T(i)$中，以便于中心服务器对查询反馈结果进行处理。

4.2 面向连续查询服务的隐私保护方法

同样，基于位置偏移的思想，可设计位置随机算法 $g:l \to l'$，对不确定性位置集合 $L=\{l_1,l_2,\cdots,l_n\}$ 中的位置进行偏移，使非敏感区域中的位置 l 与 L 中每个位置之间产生的关联转移概率与平均关联转移概率不可区分，由此可得到如下定理。

定理 2 随机算法 $g:l \to l'$ 满足 ε—位置—查询关联隐私。

证明 对于非敏感区域中位置 l 和不确定位置集合 L 中的任意 2 个位置 l'和 l''，存在随机算法 $g:l \to l'$，满足 $|cor_t(l \to g(l')|q)-avg(cor_t(l \to L|q))| \leqslant \varepsilon$，则该随机算法满足

$$\begin{aligned} &M_P(cor_t(l \to g(l')|q),cor_t(l \to g(l'')|q)) \\ &\quad = M_P(avg(cor_t(l \to L|q)) \pm \varepsilon, \\ &\qquad avg(cor_t(l \to L|q)) \pm \varepsilon) \\ &\quad \leqslant \varepsilon M_P(cor_t(l \to l'|q),cor_t(l \to l''|q)) \end{aligned} \quad (3)$$

由此可知随机算法 $g:l \to l'$ 满足 ε—位置—查询关联隐私。算法 2 是基于定理 2 产生的位置偏移算法，以此实现连续查询服务下的隐私保护。

算法 2 连续查询下的位置偏移

输入 $S, cor_t(l \to S|q), L \in S, \varepsilon$

输出 L'

1) p_a=avg($cor_t(l \to L|q)$);
2) for (i=1; i<=n; ++i)

3)　if($|cor_i(l \to l'_i|q) - p_a| <= \varepsilon$)

4)　　$L'=l'_i+L'$;

5)　　$T(i)= l_i-l'_i$;

6)　end if

7) end for

算法 2 与算法 1 极为相似，其中，通过步骤 2)~步骤 7)迭代，对不确定位置集合 L 中的每个位置进行转移，并保存在集合 L'中，同时将每步的转移距离保存在 $T(i)$中，以便于中心服务器对查询反馈结果进行处理。差别在于对不同关联概率的比较上。

4.3　通用隐私保护方法

对 2 种不同的位置服务类型所提出的随机算法 f 和 g，可以得到新的针对 2 种服务共同的随机算法 $F: l \to l'$，其中，$F = f \cup g$，且 F 满足 2ε—位置—查询关联隐私。

证明　对于不确定位置集合 L，存在随机算法 $f: l \to l'$ 和 $g: l \to l'$，满足 ε—位置—查询关联隐私，由差分隐私的序列组合性可知，$F = f \cup g$ 满足 2ε—位置—查询关联隐私。

该隐私保护方法的算法实现过程可视为算法 1 和算法 2 的组合，即首先找到满足关联概率不可区分的位置集合，然后在该集合中找到满足关联转移概率的位置集合，同时记录真实位置的转移距离，以便在返回结果中进行提纯处理。

5　安全性分析

为验证所提出的随机算法能够有效地保护用户在使用位置服务过程中的个人隐私，本文引入信息论中的信息熵和互信息作为度量标准，以此评估所提出算法的安全性。其中，信息熵可视为经过位置偏移后各种概率的不确定性，而互信息可视为一个集合导致的另一个集合变化不确定性的缩减量。

查询 q 和位置 l 之间的信息熵可表示为

$$H(L) = \sum_{l \in L} p(l,q) \mathrm{lb} p(l,q) \tag{4}$$

位置集合 L 和偏移位置集合 L'之间的互信息可表示为

$$I(L;L') = \sum_{l \in L, l' \in L'} p(l,l') \mathrm{lb} \frac{p(l,l')}{p(l)p(l')} \tag{5}$$

本文使用互信息 $I(L;L')$ 来表示位置集合 L 和偏移位置集合 L'之间的相互独立程度，$I(L;L')$=0 表示 L 和 L'相互独立不存在关联关系，即此时攻击者无法通过 $cor(l',q)$ 获得用户的真实位置关联概率 $cor(l,q)$，进而推测出潜在的真实位置。

针对基于快照的位置服务方式，随机算法 $f: l \to l'$存在如下定理。

定理 3　随机算法 $f: l \to l'$ 可抵抗极大概率攻击。

证明　对于指定查询 q 以及不确定位置集合 $L = \{l_1, l_2, \cdots, l_n\}$，存在位置集合 L 中的任意 2 个位置 l 和 l'，有查询 q 和位置 l 之间关联关系的信息熵

$$H(L) = \sum_{l,l' \in L} cor(L,q) \mathrm{lb}\, cor(L,q) \tag{6}$$

对于位置集合 L 有随机算法 f 使 L 中的任意位置 l 满足 $|cor(f(l),q) - avg(cor(L,q))| \leqslant \varepsilon$，由此，可获得 $|cor(f(l),q) - cor(f(l'),q)| \leqslant 2\varepsilon$。当 ε 足够小时，对于 L 中的任意 2 个位置 l 和 l'有 $cor(f(l),q) \approx cor(f(l'),q)$。因此，可认为偏移后的 2 个位置之间的关联概率彼此相等且独立，进而获得最大熵 $H(L)$。根据 Jaynes 最大熵理论，可知最大熵表示最大不确定性，因此，可认为攻击者可获得的位置 $f(l)$ 和 $f(l')$ 之间无法通过关联概率加以区分，即攻击者无法通过 $cor(f(l),q) = \max(cor(L,q))$ 识别出用户的偏移位置，更无法识别出用户的真实位置。

定理 4　随机算法 $f: l \to l'$ 可抵抗关联差分攻击。

证明　对于指定查询 q 以及不确定位置集合 $L = \{l_1, l_2, \cdots, l_n\}$，存在随机算法 f 使 L 偏移获得偏移后位置集合 L'，其中，L 和 L'之间的联合概率可表示为 $p(L,L')$，其边际概率分别为 $p(L)$ 和 $p(L')$，则不确定位置集合 L 和 L'之间的互信息可表示为

$$I(L;L') = \sum_{l \in L, l' \in L'} p(l,l') \mathrm{lb} \frac{p(l,l')}{p(l)p(l')} \tag{7}$$

由位置集合 L 和偏移位置集合 L'中的位置数量可知 $p(L)$=$p(L')=\frac{1}{n}$，而随机算法 f 为使 L 中的任意位置 l 满足 $|cor(f(l),q) - avg(cor(L,q))| \leqslant \varepsilon$，对 L 中的位置进行位置偏移。当 ε 足够小时，攻击者获得各位置的关联概率近似等于 $avg(cor(L,q))$，使攻击者无法获知 L 中的任意位置 l 在 f 的作用下的偏移结果，有联合概率 $p(L,L')=\frac{1}{n^2}$。因此，$I(L;L')$=0，可知 L 和 L'之间彼此独立，即攻击者通过差分关联攻击在获得到偏移后的相关概率 $cor(f(l),q)$ 的情况下，无法推测出真实的关联概率 $cor(l,q)$，此时攻击者无法通过真实关联概率识别出用户的真实位置。

针对基于连续查询的位置服务方式，随机算法 $g:l\to l'$ 存在如下定理。

定理 5 随机算法 $g:l\to l'$ 可抵抗关联转移概率攻击。

证明 对于非敏感区域中位置 l、指定查询 q 以及不确定位置集合 $L=\{l_1,l_2,\cdots,l_n\}$，存在位置集合 L 中的任意 2 个位置 l'和 l''，在查询 q 条件下的位置 l 与不确定位置 l'和 l''之间的关联转移关系信息熵为

$$H(L)=\sum_{l',l''\in L} cor_t(l\to g(L)|q)\mathrm{lb}\,cor_t(l\to g(L)|q) \tag{8}$$

对于位置集合 L 有随机算法 g 使 L 中的任意位置 l'满足 $|cor_t(l\to g(l')|q)-avg(cor_t(l\to L|q))|\leqslant\varepsilon$，由此，可获得 $|cor_t(l\to g(l')|q)-cor_t(l\to g(l'')|q)|\leqslant 2\varepsilon$。当 ε 足够小时，对于 L 中的任意 2 个位置 l'和 l''有 $cor_t(l\to g(l')|q)\approx cor_t(l\to g(l'')|q)$。因此，可认为 2 个关联转移概率相等且独立，进而获得最大熵 $H(L)$。同样根据 Jaynes 最大熵理论，可知最大熵表示最大不确定性，可认为位置 l'和 l''之间无法通过关联转移概率加以区分，即攻击者无法通过 $cor_t(l\to g(l')|q)=\max(cor_t(l\to L|q))$ 识别出用户的真实偏移位置，真实的转移位置更难识别。

定理 6 随机算法 $g:l\to l'$ 可抵抗关联转移差分攻击。

证明 对于非敏感区域中位置 l、指定查询 q 以及不确定位置集合 $L=\{l_1,l_2,\cdots,l_n\}$，存在随机算法 g 使 L 偏移获得偏移后位置集合 L'，其中，L 和 L'之间在条件 q 下的联合概率可表示为 $p(L,L'|q)$，其边际概率分别为 $p(L)$和 $p(L')$，则不确定位置集合 L 和 L'之间的条件互信息可表示为

$$\begin{aligned}I(L;L'|q)&=D(p(L,L'|q)\,\|\,p(L|q)p(L'|q))\\&=E_{p(L,L',q)}\mathrm{lb}\frac{p(L,L'|q)}{p(L|q)p(L'|q)}\end{aligned} \tag{9}$$

其中，D 为相对熵。由位置集合 L 和偏移位置集合 L'中的位置均与查询 q 相关，可知 $p(L|q)=p(L'|q)=\frac{1}{n}$，而随机算法 g 为使 L 中的任意位置 l' 满足 $|cor_t(l\to g(l')|q)-avg(cor_t(l\to L|q))|\leqslant\varepsilon$，对 L 中的位置进行位置偏移。当 ε 足够小时，攻击者获得各位置的关联转移概率近似等于 $avg(cor_t(l\to L|q))$，使攻击者无法获知 L 中的任意位置 l'在 g 的作用下的偏移结果，有联合概率 $p(L,L'|q)=\frac{1}{n^2}$。因此，$I(L;L'|q)=0$，此时可知 L 和 L'之间彼此独立，即攻击者在通过差分关联攻击获得偏移后的关联转移概率 $cor_t(l\to g(l')|q)$ 的情况下，无法推测出真实的关联转移概率 $cor_t(l\to l'|q)$ 以及 $cor(l,q)$，此时攻击者无法通过真实关联转移概率识别出用户的真实位置。

综上，可得到满足随机算法 $f:l\to l'$ 和 $g:l\to l'$ 的 $F:l\to l'$，其中，$F=f\cup g$，可同时抵抗以上提出的 4 种攻击方式，其证明过程如上所示。

6 实验验证

6.1 实验准备

为验证以上算法的隐私保护效力与执行效率，本文将所涉及的算法在 Windows 7 上使用 Matlab 7 加以模拟。其运行环境为 1.70 GHz Intel Core i5，内存 4 GB。实验数据集采用 BerlinMOD Data Set 真实数据集中的城市中心区域，以获取更多的具有相似关联概率的位置，并假设存在足够的可信第三方提供保护服务。

在算法的执行效率方面，主要从用户位置的偏移距离、相似概率的查找范围和执行时间 3 个方面加以验证。通常情况下，偏移距离越大表明获取的兴趣点位置与真实位置之间的差异越大，服务的精确度越低；算法执行时间越长，则基于位置服务的响应时间越长，用户获得服务的时效性越差；同样，相似概率的查找范围越大，则直接导致算法的执行时间过长，也会间接影响服务质量。对于算法的隐私保护效力，主要通过偏移后不同概率的不确定性和每对概率的相似程度 2 个方面利用信息熵加以验证。当不确定性越高时，算法的隐私保护效果越好；当每对概率的相似程度增大时，不确定性相应提高，进而增强隐私保护效果。

6.2 实验结果

从图 3 中可以看到，3 种不同的随机算法随着申请用户数的增加都需要扩大寻找范围以获得满足概率不可取分的用户位置。同时，随着 ε 值逐渐降低，3 种算法的寻找范围逐渐扩大，其扩大程度随着小数点位数的减小产生近似指数倍的增长。这主要是由于隐私级别的提高造成对用户位置要求的提升，使算法需要在更广阔的范围内寻找满足概率不可区分的用户位置，且 ε 取值成倍降低导致概率不可区分性的成倍提高，进而造成如图 3 所示的寻找范围的增长。其次，从图 3 中可以看出，3 种不同的随机算法存在寻找范围差异，具体表现为

$F>g>f$。这是由于随机算法 F 是随机算法 g 和 f 的功能组合，该算法需要寻找到同时满足关联概率和关联转移概率不可区分的用户位置，这种较为严格的要求使该算法需要在更为广阔的范围内寻找合适的位置。另外，随机算法 g 的偏移距离高于 f，这是由于很多满足关联概率不可区分的位置并不存在关联转移概率，导致 g 的寻找范围随申请人数的变化高于随机算法 f。这种寻找范围变化例外的情况可以在图 3(a)中可以看到，f 与 g 的寻找范围曲线几近重合，这主要是在较小隐私级别要求的情况下，寻找到的满足关联转移概率的位置一般能够同时满足关联概率不可区分，使随机算法 f 与 g 仅寻找到满足关联转移概率不可区分的位置即完成算法。

从图 4 中可以看到，随着申请用户数的增加，3 种随机算法产生的偏移距离逐渐增大。另外，随着 ε 取值的逐渐降低，3 种随机算法产生的偏移距离逐渐增大，且当 ε 从 0.01 减小到 0.001 时，偏移距离变化较大。产生这一现象的主要原因是，在隐私级别增大的情况下，能够满足概率不可区分的用户位置大量减少，需要在更大的范围内进行位置寻找。而从 0.01 减小到 0.001 将会大幅度提高隐私级别，导致需要大幅度扩张寻找范围，以便寻找到满足要求的用户位置。因此，为满足与平均值之间的不可区分，需要用户的真实位置产生较大的偏移距离增长。最后，从图 4 中还可以看出，3 种随机算法产生的偏移距离满足 $F>g>f$。首先，随机算法 F 是 g 和 f 的功能组合，需要寻找到同时满足关联概率和关联转移概率不可区分的位置，这需要在更为广阔的范围内进行寻找，进而产生较大的偏移距离；其次，随机算法 g 的偏移距离高于 f，因为很多满足关联概率不可区分的位置并不存在关联转移概率，在这些位置上并不存在相同查询导致的位置转移；再次，在图 4(a)中可以看到，F 与 g 的偏移距离曲线几近重合，这是由于在较小隐私级别要求的情况下，寻找到的满足关联转移概率的位置一般满足关联概率不可区分，因此，存在随机算法 F 与 g 偏移距离相等的情况。

从图 5 中可以看到，3 种随机算法随着申请用户数的增加，其算法的执行时间逐渐上升，但即使图 5(d)所示的 ε=0.001 申请人数为 30 的情况下，各种算法的执行时间都控制在毫秒级别，因此，算法具有较好的执行效率。其次，随着 ε 取值的逐渐降低，3 种随机算法的执行时间存在差异，其差

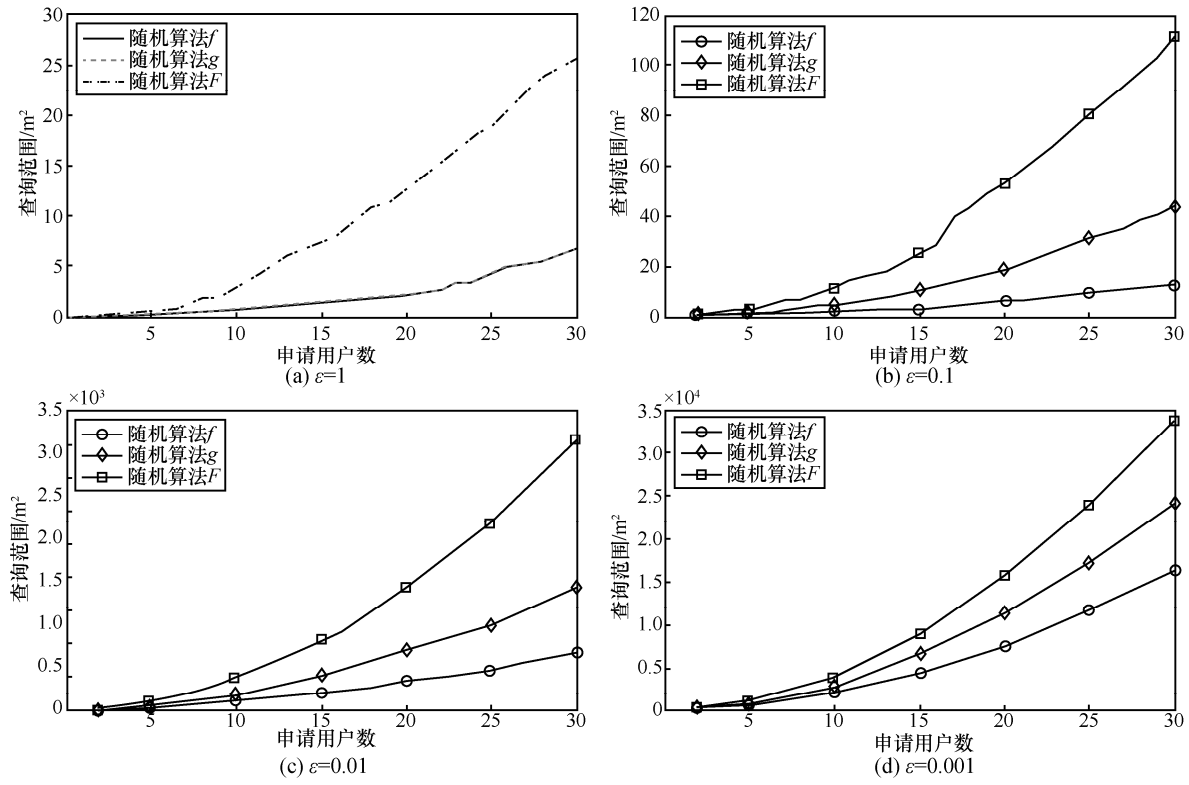


图 3　不同 ε 取值下 3 种算法的查询范围

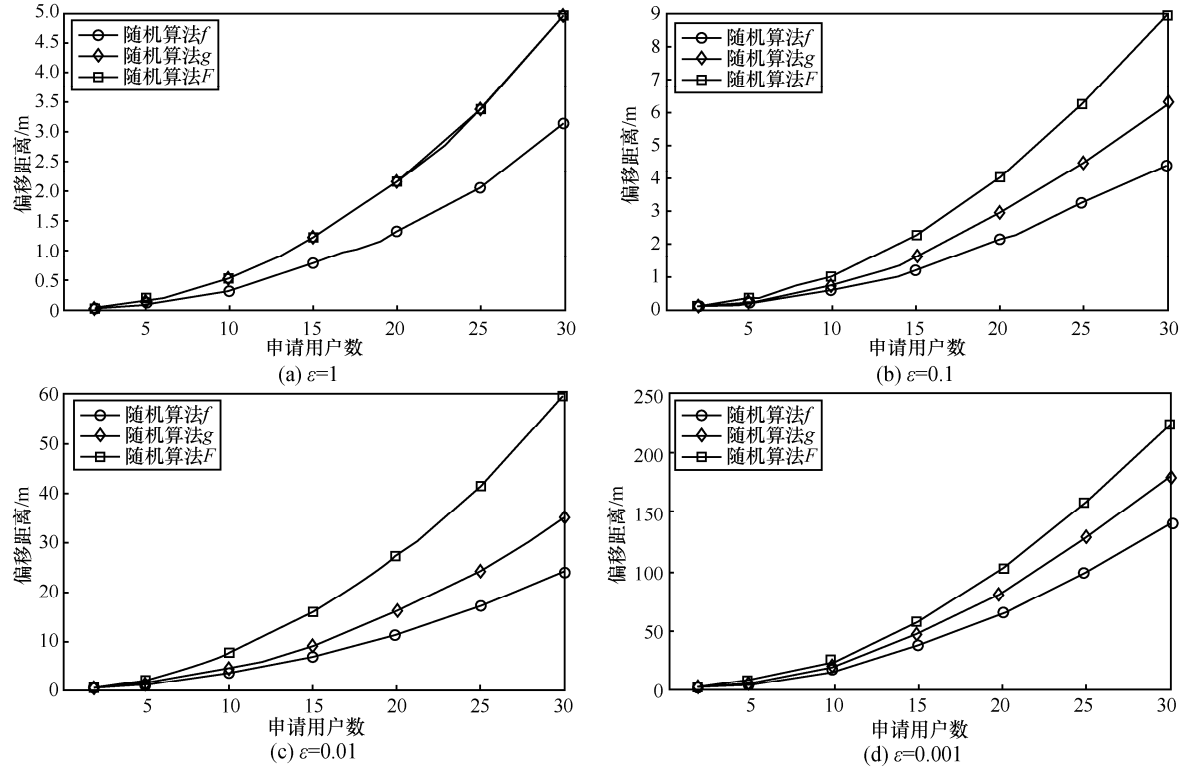


图 4　不同 ε 取值下 3 种算法的偏移距离

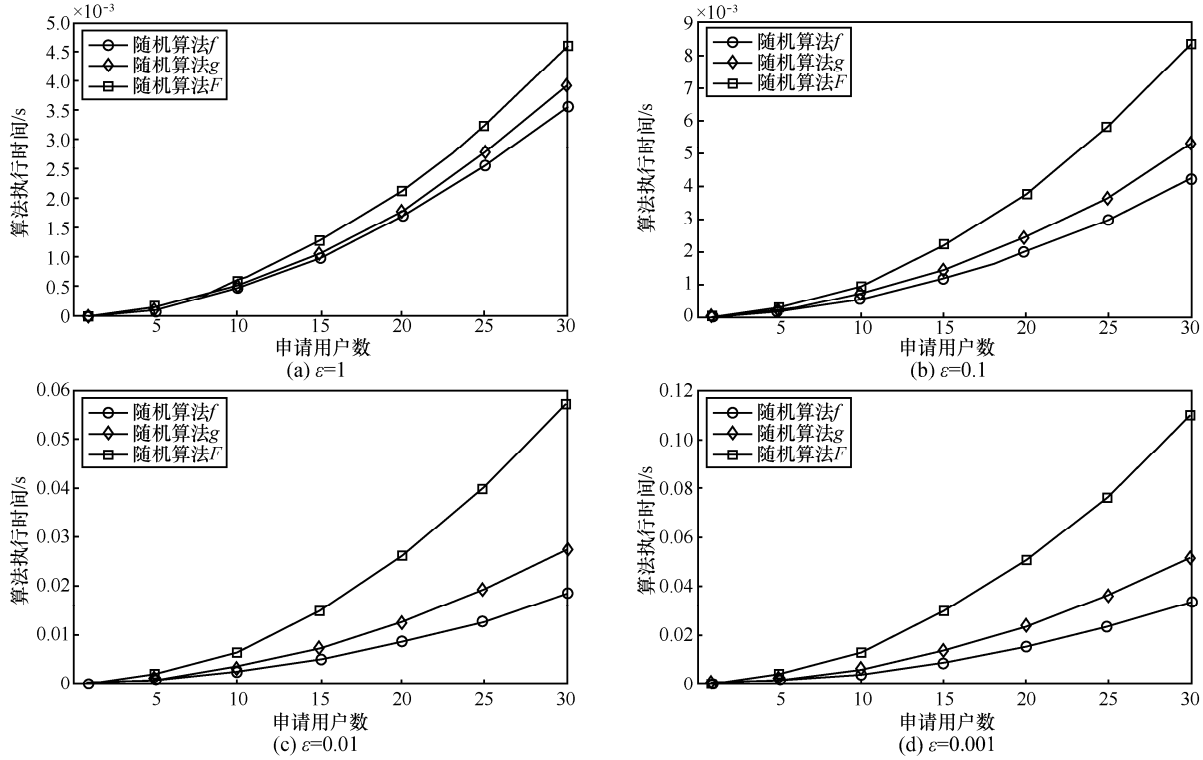


图 5　不同 ε 取值下 3 种算法的运行时间

异具体表现为 $F>g>f$，造成这种现象的原因与寻找范围扩大有直接的关系。最后，与寻找范围和偏移距离不同，在 ε 值成倍减少时执行时间并没有成倍增加。这是因为不可区分概率的寻找主要基于概率数据的查询比较，其执行时间主要受中心服务器性能的影响。

从图 6 中可以看到 3 种不同的随机算法产生的用于表示位置集合中位置不确定性的熵值变化 3 种算法在隐私保护级别较高，即 ε 取值较大的情况下均可以达到最大熵值。而在 ε 取 1 和 0.1 的情况下，由于用户申请位置各种概率的不确定性，使概率之间的差值与原始概率相差不大，存在被攻击者识别的情况，其熵值低于最大熵。同时，由于用户申请位置概率的随机性，使产生的熵值存在不规则性的波动，其位置隐私的安全性受到一定影响。

从图 7 中可以看到，3 种不同的随机算法中每对位置之间的不可区分性。在 ε 取值较小的情况下，偏移后的位置存在较强的不可区分性，攻击者很难分辨各概率之间的差异，且随着 ε 取值的逐渐升高，这种成对熵逐渐达到最大熵，并且与最大熵之间很难区分。而在 ε=1 的情况下，由于偏移后的位置概率对真实概率的影响较小，且用户提交位置存在概率差异性，导致计算获得的熵值低于其他取值，并且存在熵值波动。

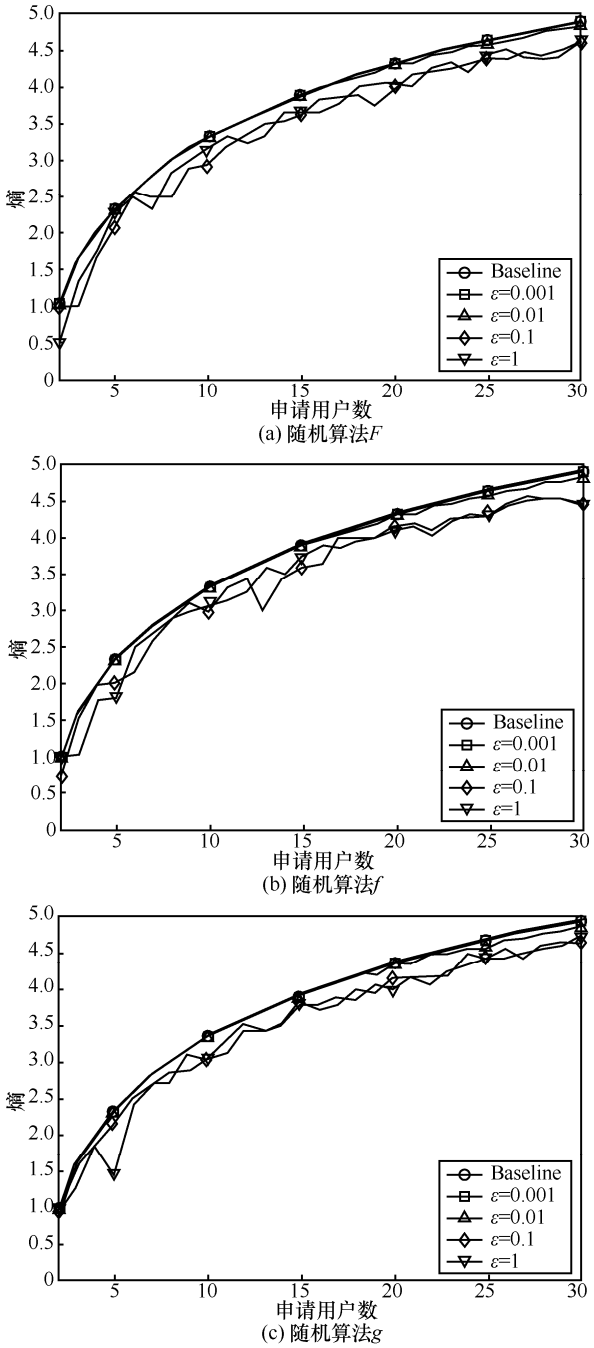


图 6　不同运算机制下的熵变化

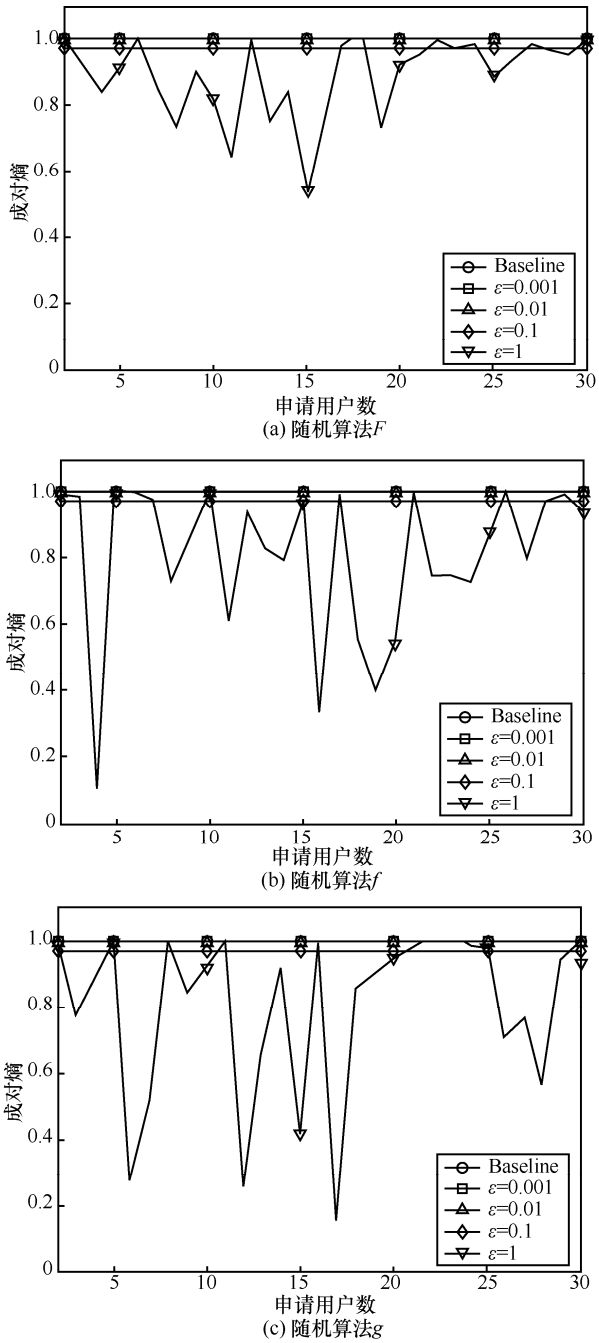


图 7　不同运算机制下的成对熵变化

通过以上实验验证可以得出，本文所提出的方法在选定合适 ε 值的基础上，能够提供较好的隐私保护级别和服务质量。

7 结束语

针对大量位置相关数据可作为背景知识被攻击者使用的情况，本文首先通过建立基于位置相关数据统计概率攻击，描述了潜在的攻击方法以及攻击效果。然后，针对潜在的各种概率攻击，提出了位置概率不可区分性机制，并在这种机制上建立了 3 种基于位置偏移的隐私保护方法。最后，通过安全性分析证明了这 3 种方法对潜在攻击的有效抵抗性，并通过实验进一步验证了 3 种方法的算法执行效率和隐私保护效力。尽管所提出的方法能够较好地解决基于位置相关数据的统计概率攻击，且具有较好的算法执行效率，但仍存在部分问题尚未解决。例如，ε 的取值直接影响算法的寻找范围、偏移距离和执行时间，如何选择合适的 ε 成为首要问题；偏移距离一方面提高了用户位置的隐私保护程度，但同时又影响了用户服务质量，如何选择合适的偏移位置以便平衡隐私和服务质量。

参考文献：

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

[2] GEDIK B, LING L. Location privacy in mobile systems: a personalized anonymization model[C]//Distributed Computing Systems, ICDCS 2005. 2005: 620-629.

[3] FUYU L, HUA K A, YING C. Query *l*-diversity in location-based services[C]//Mobile Data Management: Systems, Services and Middleware, 2009 MDM'09 Tenth International Conference. 2009: 436-442.

[4] REBOLLO-MONEDERO D, FORNE J, SOLANAS A, et al. Private location-based information retrieval through user collaboration[J]. Computer Communications, 2010, 33(6): 762-774.

[5] REBOLLO-MONEDERO D, FORNE J, DOMINGO-FERRER J. Query profile obfuscation by means of optimal query exchange between users[J]. IEEE Transactions on Dependable and Secure Computing, 2012, 9(5): 641-654.

[6] SHOKRI R, THEODORAKOPOULOS G, PAPADIMITRATOS P, et al. Hiding in the mobile crowd: location privacy through collaboration[J]. IEEE Transactions on Dependable and Secure Computing, 2014, 11(3): 266-279.

[7] NIU B, ZHU X Y, LI Q H, et al. A novel attack to spatial cloaking schemes in location-based services[J]. Future Generation Computer Systems, 2015, 49(C): 125-132.

[8] KHOSHGOZARAN A, SHIRANI-MEHR H, SHAHABI C. SPIRAL: a scalable private information retrieval approach to location privacy[C]//2008 Ninth International Conference on Mobile Data Management Workshops. 2008: 49-56.

[9] KHOSHGOZARAN A, SHAHABI C, SHIRANI-MEHR H. Location privacy: going beyond *K*-anonymity, cloaking and anonymizers[J]. Knowledge and Information Systems, 2011, 26(3): 435-465.

[10] LIEN I T, LIN Y H, SHIEH J R, et al. A novel privacy preserving location-based service protocol with secret circular shift for *k*-NN search[J]. IEEE Transactions on Information Forensics and Security, 2013, 8(6): 863-873.

[11] PAULET R, KAOSAR M G, YI X, et al. Privacy-preserving and content-protecting location based queries[J]. IEEE Transactions on Knowledge and Data Engineering, 2014, 26(5): 1200-1210.

[12] BEN N, QINGHUA L, XIAOYAN Z, et al. Enhancing privacy through caching in location-based services[C]//Computer Communications (INFOCOM), 2015 IEEE Conference. 2015: 1017-1025.

[13] MA C G, ZHOU C L, YANG S T. A Voronoi-based location privacy-preserving method for continuous query in LBS[J]. International Journal of Distributed Sensor Networks, 2015, 11(3): 1-17.

[14] 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.

[15] WANG Y, XIA Y, HOU J, et al. A fast privacy-preserving framework for continuous location-based queries in road networks[J]. Journal of Network and Computer Applications, 2015, 53: 57-73.

[16] PALANISAMY B, LIU L. Attack-resilient mix-zones over road networks: architecture and algorithms[J]. IEEE Transactions on Mobile Computing, 2015, 14(3): 495-508.

[17] GAO S, MA J F, SHI W S, et al. LTPPM: a location and trajectory privacy protection mechanism in participatory sensing[J]. Wireless Communications & Mobile Computing, 2015, 15(1): 155-169.

[18] HWANG R H, HSUEH Y L, CHUNG H W. A novel time-obfuscated algorithm for trajectory privacy protection[J]. IEEE Transactions on Services Computing, 2014, 7(2): 126-139.

[19] OZER M, KELES I, TOROSLU, et al. Predicting the next location change and time of change for mobile phone users[C]//Third ACM SIGSPATIAL International Workshop on Mobile Geographic Information Systems. Dallas, Texas, ACM. 2014: 51-59.

[20] XUE A Y, ZHANG R, ZHENG Y, et al. Destination prediction by sub-trajectory synthesis and privacy protection against such prediction[C]//29th IEEE International Conference on Data Engineering (ICDE). Brisbane, AUSTRALIA.2013: 254-265.

[21] SU H, ZHENG K, WANG H, et al. Calibrating trajectory data for similarity-based analysis[C]//2013 ACM SIGMOD International Conference on Management of Data. ACM, 2013: 833-844.

[22] CHEN X, PANG J, XUE R. Constructing and comparing user mobility profiles[J]. ACM Transactions on the Web, 2014, 8(4): 21-25.

[23] LEI Z, CHUNGUANG M, SONGTAO Y. Location association similar based anonymus algorithm[J]. China Sciencepaper, 2016, 11(2): 197-201, 213.

[24] NIU B, QINGHUA L, XIAOYAN Z, et al. Achieving *k*-anonymity in privacy-aware location-based services[C]//INFOCOM, 2014 Proceedings

IEEE. 2014: 754-762.

[25] DWORK C. DIfferential privacy[C]//Lecture Notes in Computer Science. 2006: 1-12.

[26] XIONG L. Adaptive differentially private data release for data sharing and data mining[J]. 2013 IEEE 13th International Conference on Data Mining Workshops (ICDMW). 2013: 891-891.

[27] CHENG X, SU S, XU S Z, et al. DP-Apriori: a differentially private frequent itemset mining algorithm based on transaction splitting[J]. Computers & Security, 2015, 50:74-90.

[28] XU J, ZHANG Z J, XIAO X K, et al. Differentially private histogram publication[J]. VLDB Journal, 2013, 22(6): 797-822.

[29] RIBONI D, BETTINI C. Incremental release of differentially-private check-in data[J]. Pervasive and Mobile Computing, 2015, 16:220-238.

[30] DEWRI R. Local differential perturbations: location privacy under approximate knowledge attackers[J]. IEEE Transactions on Mobile Computing, 2013, 12(12): 2360-2372.

[31] ANDRÉS M E, BORDENABE N E, CHATZIKOKOLAKIS K, et al. Geo-indistinguishability: differential privacy for location-based systems[C]//2013 ACM SIGSAC Conference on Computer & Communications Security. ACM, 2013: 901-914.

[32] BORDENABE N E, CHATZIKOKOLAKIS K, PALAMIDESSI C. Optimal geo-indistinguishable mechanisms for location privacy[C]// 2014 ACM SIGSAC Conference on Computer and Communications Security. ACM, 2014: 251-262.

[33] PRIMAULT V, MOKHTAR S B, LAURADOUX C, et al. Differentially private location privacy in practice[J]. arxiv: 1410.7744.

[34] CHATZIKOKOLAKIS K, PALAMIDESSI C, STRONATI M. Geo-indistinguishability: a principled approach to location privacy[M]//Distributed Computing and Internet Technology. 2015: 49-72.

[35] PERAZZO P, DINI G. A uniformity-based approach to location privacy[J]. Computer Communications, 2015, 64: 21-32.

作者简介：

张磊（1982-），男，黑龙江绥化人，哈尔滨工程大学博士生，佳木斯大学讲师，主要研究方向为信息安全、隐私保护。

马春光（1974-），男，黑龙江双城人，博士，哈尔滨工程大学教授、博士生导师，主要研究方向为密码学、数据安全与隐私保护、无线自组织网络及安全。

杨松涛（1972-），男，黑龙江鹤岗人，博士，佳木斯大学教授，主要研究方向为信息安全、隐私保护。

李增鹏（1989-），男，山东青岛人，哈尔滨工程大学博士生，主要研究方向为密码学、密码协议。