doi:10.11959/j.issn.1000-436x.2017184

基于属性基加密的用户协作连续查询隐私保护策略

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：在基于位置服务（LBS, location-based service）中，不可信的服务器可利用掌握的用户属性信息作为背景知识，在快照或连续查询过程中关联不确定位置集合中的位置，进而识别出潜在的用户真实位置造成用户位置隐私的泄露。针对这种攻击行为，当前主要的隐私保护方法是进行属性泛化。然而，已有的属性泛化方法一般假设存在一个可信的第三方服务器对参与匿名用户的属性进行寻找与比较，而这个第三方服务器很可能因为攻击焦点或服务瓶颈问题变得不再可信。针对攻击者可能使用的用户属性进行分析攻击以及第三方服务器潜在的不可信问题，提出一种基于属性基加密（CP-ABE, ciphertext policy attribute based encryption）的方法，通过用户协作完成对用户属性的匿名操作，并且整个过程中第三方服务器以及协作用户无法获知该用户任何隐私信息。最后，安全性分析和实验验证进一步证明了所提方法具有较高的隐私保护效力和算法执行效率。

关键词：基于位置服务；隐私保护；属性基加密；属性匿名
中图分类号：TP311 **文献标识码**：A

CP-ABE based users collaborative privacy protection scheme for continuous query

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: In location-based services (LBS), as the untrusted LBS server can be seen as an adversary, and it can utilize the attribute as background knowledge to correlate the real location of the user in the set of uncertain locations. Then the adversary can gain the location privacy when the user enjoys the snapshot and continuous query through the correlation inference attack. In order to cope with this attack, the main scheme in privacy protection is to generalize the attribute and achieve attribute anonymity. However, algorithms of this type usually assumes a trusted third party (TTP) which provides the service of similarity attribute finding and comparing, and it is unpractical in the real environment, as the TTP may become the point of attack or the bottleneck of service and it cannot be considered as the trusted one all the time. Thus, to cope with the correlation inference attack as well as the semi-trusted third party, ciphertext policy attribute based encryption (CP-ABE) and users collaboration based attribute anonymous scheme was proposed. In this scheme, the user coupled achieve location and attribute anonymity. Furthermore, this scheme could also provide security for attacks from the semi-trusted third party as well as semi-trusted collaborative users. At last, security analysis and the experiment results further verify the effectiveness of privacy protection and the efficiency of algorithm execution.

Key words: location-based services, privacy protection, attribute encryption, attribute anonymity

收稿日期：2016-12-19；**修回日期**：2017-07-17
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), Specialized Research Fund for the Doctoral Program of Higher Education (No.20132304110017), The Natural Science Foundation of Heilongjiang Province (No.F2015022)

1　引言

无线通信技术以及卫星定位技术的出现为基于位置服务带来了无限的商业发展空间。这种通过用户移动客户端提供当前位置并反馈所需的查询信息、导航、广告推送等便捷服务的商业模式创造了巨大的社会财富。然而，用户在使用这些服务的过程中，不可避免地需要面对个人隐私泄露的问题。并且随着使用位置服务人数的逐渐增加，导致攻击者可能掌握的背景知识也越来越多，移动用户的隐私问题越来越严峻。

当前，快照查询服务和连续查询服务成为基于位置服务中的 2 个主要服务类型。快照查询服务是指当前用户仅申请一次基于位置的查询服务在获得查询结果后不再进行查询申请；连续查询服务指用户在一段时间内，连续地向位置服务器提供自身位置，并获得以连续位置为基础的多组查询结果的服务形式[1]。快照查询服务由于仅进行单次查询其隐私保护相对较易实现。而在基于位置的连续查询服务中，攻击者可通过用户表现出的属性关联获得用户轨迹信息，进而通过该信息分析获得用户的个人隐私。这种分析攻击行为使当前已有的基于泛化[2,3]、假位置[4,5]以及轨迹匿名[6]等隐私保护方法均表现出其隐私保护能力的局限性。当前，已有研究者开始关注这种基于属性关联分析攻击，并提出了时空属性关联泛化[7]、时空属性模糊[8]等相关隐私保护方法，更有基于可信中心服务器建立属性泛化模型[9]提供多属性泛化隐私保护服务的相关技术。然而，这些已有的基于属性泛化或模糊的方法大多依靠可信的第三方服务的系统架构，无法保障在第三方因攻击焦点以及服务瓶颈等问题下导致的用户隐私安全。

针对这种情况，本文基于属性基加密的基本思想，通过采用用户协作的方式实现匿名用户的相似属性泛化，提出了一种属性相似的协作用户选择（SACU, similar attribute collaborative users）算法。该算法使中心服务器无法获得用户相关信息，同时，最大限度地将中心服务器的计算开销分散到协作用户当中，不仅提高了中心服务器结构的安全性，而且降低了中心服务器的计算复杂性。实验结果表明，该方法不仅能最大限度地增加对用户属性的泛化程度提高隐私保护能力，而且在提供服务的过程中降低了对中心服务器计算能力的依赖，保证了基于位置服务的时效性与服务质量。

2　相关工作及预备知识

2.1　相关工作

早在 2003 年，基于位置服务的隐私保护问题就被研究者所关注。基于可信第三方的 *k*-匿名方法[10]为用户提供了随机的 *k* 个匿名位置，泛化了用户真实位置信息。之后，针对互质性问题又分别提出了语义多样性[11]和查询多样性[12]的位置隐私保护模型。出于对第三方可信的质疑，Chow 等[13]提出了点对点寻找匿名用户的位置隐私保护方案；Ma 等[14]通过信息交换扩大了协作用户的查找范围；Gabriel 等[15]则基于零隐私泄露的思想，通过对使用可计算 PIR 的方法实现对查询信息的秘密检索。同样，对于连续位置服务中的隐私保护问题，Ryo 等[16]针对连续查询过程中可能出现的不规则暂停，提出了一种更加人性化的假轨迹生成方法。Moein 等[17]通过轨迹相似性计算寻找参与匿名的轨迹。Palanisamy 等[18]则通过 mix-zone 截断了不同轨迹段之间的关联，降低了攻击者获得移动用户完整轨迹的概率。

随着对基于位置服务隐私保护研究的深入，简单地对位置或轨迹进行匿名显然无法保障用户的个人隐私，攻击者特别是由半可信 LBS 以自身掌握的查询信息作为背景知识，可以更为方便地关联用户提供的查询位置，较明显的就是通过连续查询过程中提供的各种属性信息进行关联，进而获得用户轨迹的情况。半可信 LBS 指可提供 LBS 服务但同时对用户隐私信息具有非恶意好奇驱使下的 LBS 服务器，而半可信第三方指参加或提供隐私保护，但同样对用户隐私信息好奇的在中心服务器或协作用户。针对这种属性关联行为，Zhang 等[19]针对查询与位置间的关联概率提出了关联概率相似的属性匿名方案。马春光等[20]通过将真实位置转换为锚点的方法，使所有查询过程中的用户均表现为锚点位置所具有的属性，实现了属性泛化。Alicia 等[21]进一步通过信息熵来评估用户的移动属性，通过移动属性的扰乱降低用户的关联程度。

然而，这些方法在很大程度上都依赖于可信的第三方中心服务器来完成匿名或属性信息泛化，而不依赖第三方的方法又由于移动用户自身通信计算能力的局限，无法寻找最大空间范围内相同属性的可匿名用户。基于属性基加密的方法为相同属性协作用户的寻找带来了一个很好的解决方向。不同

于使用属性基加密方法进行访问控制，本文主要通过属性基加密来寻找具有相似属性的协作用户，出于这种目的，对于属性基加密方法的安全以及计算要求较低，降低了对密码协议计算复杂性的要求。同时，通过使用混合式系统架构，利用半可信第三方进行查询信息广播，极大地提高了协作用户的寻找范围，增加了属性匿名计算的成功率。另外，由于将解密计算放到协作用户客户端进行，也在一定程度上降低了中心服务器的计算负载，提高了中心服务器的服务质量。

2.2　预备知识

为方便对本文的理解，首先将本文所涉及的各种变量符号加以详细说明。所有符号及其解释说明如表 1 所示。

表 1　　文中涉及的各种变量符号

符号	描述
p	用户关联概率
A	所有属性集合
S	不确定位置集合
k_{u}	用户提供的对称密钥
pk_{L}	LBS 提供的加密公钥
pk	用户使用的属性基加密公钥
A_{u}	用户指定的部分属性
E	密文信息
E_{c}	CS 广播的密文信息
E_{cu}	CU 反馈给 CS 的密文信息
q_t	兴趣点类型集合
k	匿名值

本文所使用到的属性基加密方案为文献[22]所提供的 CP-ABE，该方案通过分发对应的解密密钥使仅具有相同属性的用户能够对密文进行解密，而属性不同的用户无法解密获得对应的明文信息。其处理流程如下。

设置阶段：输入隐式的安全参数，输出公钥 pk 以及主密钥 M_K。

加密阶段：输入待加密信息 M、访问策略 A 以及公钥 pk，输出对应的加密密文 E。

密钥生成阶段：输入属性集合 $A = \{A_1, A_2, \cdots, A_n\}$、主密钥 M_K 以及公钥 pk，输出针对指定属性的解密密钥 D。

解密阶段：输入在访问策略 A_{u} 加密下的密文信息 E，在属性集合 γ 下产生的解密密钥 D 以及公钥 pk，当且仅当属性 A 满足访问策略 A_{u} 时，可输出明文信息 M。

本文假设所有的访问策略在进行基于位置的查询服务前已经协商完毕，每个用户已经获得根据自身属性建立的私钥。因此，在以下隐私保护技术的描述中不再叙述 CP-ABE 的加解密处理过程。

2.2.1　系统架构

在基于位置服务的隐私保护系统架构中，一般存在双层、三层或混合式 3 种不同的架构模型。本文采用的是最能够保护用户个人隐私的混合式系统架构，其架构模型如图 1 所示。

架构包含 4 个实体：用户（user）、中心服务器（CS）、协作用户（CU）、位置服务器（LBS）。其中，CS 为半可信实体，一方面该实体能够按照协议执行匿名任务，另一方面对用户提交的查询具有非恶意的好奇，希望从中获得用户个人隐私；CU 同样是半可信实体，可看作是好奇用户集合，该集合中包含可参与协作的用户和非协作用户；LBS 同样是半可信的，该实体能够完成查询结果的反馈，但可以利用用户的属性信息间存在的差异，关联连续的位置以获得用户轨迹，进而分析出潜在的个人隐私。本文假设所有参与实体与 LBS 不存在共谋。

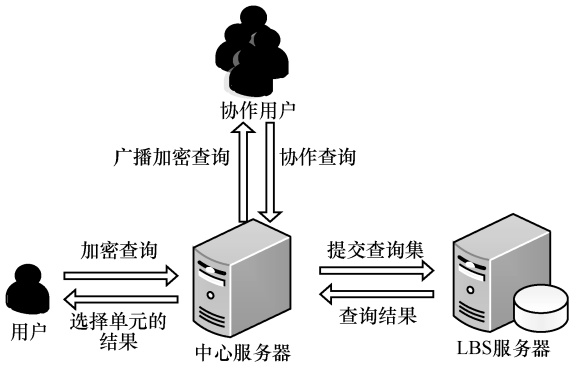


图 1　混合式系统架构模型

2.2.2　攻击手段

本文针对的攻击手段包括：CS、CU 在传递用户查询信息时对用户隐私的窥探；半可信 LBS 根据背景信息通过属性关联对连续查询过程中的用户隐私信息采用的关联分析攻击。其中，本文所提出的隐私保护方法主要针对关联分析攻击。这种关联分析攻击采用属性关联方式如文献[9]。对不确定位置集合 S 中的任一位置 l'，其用户属性的关联概率 p 可表示为已知子轨迹 L 中的位置 l 与不确定性位置 l'之间的相似程度，于是有

$$p(l')=\frac{1}{n}\sum_{i=1}^{n}sim(A_i^{l},A_i^{l'}),0\leqslant p\leqslant 1 \tag{1}$$

其中，A_i^{l}和 $A_i^{l'}$分别表示攻击者已知位置 l 和不确定位置 l'所对应的属性信息，相似度 sim 表示为

$$sim(A_i^{l},A_i^{l'})=\frac{\min(v_{A_i^{l}},v_{A_i^{l'}})}{\max(v_{A_i^{l}},v_{A_i^{l'}})},0<sim\leqslant 1 \tag{2}$$

其中，v 表示属性信息数值，min 和 max 分别表示最小和最大值。当 p=1 时表示 2 个位置属于同一轨迹 T。

在实际环境中攻击者计算得到的 p 可能低于预期值，此时，还可以通过排除不确定位置集合 S 中其他位置与 l 之间的关联概率较低的位置来识别真实位置。

3　基于属性基加密的 SACU 方法

本文所采用的隐私保护的思想是，针对用户在申请基于位置连续查询服务下可能遭遇的基于属性的关联分析攻击，利用属性进行泛化的思想，通过具有相似属性的用户参与匿名，泛化用户真实属性，使攻击者无法获得攻击预期，进而保护用户个人隐私。由此建立了 SACU 方法。

SACU 方法的主要执行过程如下。

1) 用户将信息 $M=\{l,q_t,k_{\mathrm{u}},A_{\mathrm{u}},k\}$（其中，$l$ 为用户当前位置）分别使用 LBS 提供的公钥以及属性加密的访问策略 A 加密获得密文信息 $E=\{E_{pk_{\mathrm{L}}}(l),E_{pk}(q_t,A_{\mathrm{u}}),\ E_{pk_{\mathrm{L}}}(k_{\mathrm{u}}),k\}$，并将 E 发送给 CS。

2) CS 在获得密文信息 E 后，将 $E_{\mathrm{c}}=\{E_{pk_{\mathrm{L}}}(l),E_{pk}(q_t,A_{\mathrm{u}}),E_{pk_L}(k_{\mathrm{u}})\}$广播给当前查询范围内的可参与匿名的协作用户集。

3) CU 在接收到密文信息 E_{c}后，首先确定是否参与匿名，然后尝试进行解密，并发送匿名查询 $E_{\mathrm{cu}}=\{E_{pk_{\mathrm{L}}}(l),E_{pk}(q_t,A_{\mathrm{u}}),E_{pk_{\mathrm{L}}}(k_{\mathrm{u}})\}$给 CS 参与到匿名过程中，否则该协作用户放弃当前密文信息。

4) CS 随机选择至少 k–1 个协作用户，并将用户与 CU 的查询信息发送给 LBS 服务器。

5) LBS 服务器在获得中心服务器提交的查询后，完成查询并分别使用各用户提供的对称密钥对结果进行加密后，将结果发送给 CS。

6) 在获得 LBS 服务器反馈的查询结果后，CS 将查询结果发送给用户，完成整个查询过程。

在整个查询过程中，CS 先后获得加密后的查询和查询结果，但在该过程中并没有获得任何该用户的信息；而 CU 并不能解密用户提供的位置信息，而且 CU 根据属性基加密的原则只能在具有相同属性的条件下能够完成解密，保证了参与匿名的用户具有相同属性；LBS 获得具有相同属性不同位置的查询集合，无法通过属性将多次匿名的真实位置关联成为轨迹。

CS 在整个服务过程中起到信息广播、结果反馈和 k 匿名检测的作用，其处理过程如算法 1 所示。

算法 1　CS 的信息处理过程

输入　E

1) procedure 在当前区域中广播 E；
2) 从 CU 获得 $E_{\mathrm{cu}}(i)$；
3) if count($E_{\mathrm{cu}}(i)$)$\leqslant k$
4)　　在更大的区域广播 E；
5) else
6)　　将 E 和 $E_{\mathrm{cu}}(i)$发送给 LBS；
7) end if
8) 从 LBS 获取 R；
9) 将 R 发送给用户和 CU；
10) end procedure

3.1　CU 的信息处理过程

传统的中心服务器结构一般通过 CS 选择匿名用户，完成对指定查询的匿名操作。但随着攻击者掌握的背景知识的逐渐增加，由 CS 完成对相同属性用户的选择势必会造成 CS 的计算负载过高，而采用 CS 生成假轨迹的方法在生成仿真性较高假轨迹过程中同样需要较大的计算负载。本文通过相同属性协作用户解密的方式降低 CS 计算复杂度，因此，部分由 CS 处理的信息转交给 CU 处理。CU 的处理过程如算法 2 所示。

1) 在获得由 CS 广播的加密信息 E 后，CU 确认是否希望参与匿名，如果希望则尝试解密 E，否则抛弃该信息。

2) CU 尝试解密 E，若成功则根据解密后获得的兴趣点类型建立自身查询，否则放弃信息 E。

3) CU 根据自身查询或协作查询建立新的加密信息 $E_{\mathrm{cu}}=\{E_{pk_{\mathrm{L}}}(l),E_{pk}(q_t,A_{\mathrm{u}}),E_{pk_{\mathrm{L}}}(k_{\mathrm{u}})\}$，并将该信息发送给 CS，完成协作。

算法 2　CU 的处理过程

输入　E_{c}

1) procedure 从CS收到 E_c
2) if 想要成为协作用户 CU
3) if 可以解密 E_c
4) 利用 q_t 生成查询;
5) 发送 E_{cu} 给 CS;
6) else
7) 删除 E_c;
8) end if
9) else
10) 删除 E_c;
11) end if
12) End procedure

3.2 基于 cache 的连续结果反馈

在基于位置的连续查询过程中，用户每次查询 CS 都会获得大量的查询结果，在这些结果中可能会存在用户在后续移动过程中所需要的兴趣点，若每次都将这些结果丢弃不仅造成查询资源的浪费，也间接增加了 LBS 的查询负载。针对这一问题，可以利用 CS 服务器的 cache 功能，将获得的查询结果保存在 CS 的 cache 中，用户通过从 CS 的 cache 中获得查询结果，一方面减少了与 LBS 可能存在的信息交互，另一方面降低了 LBS 的查询负载。

为了使 CS 获知是否需要对获得的反馈结果保存在 cache 中，用户提交的密文信息需要更改为 $E=\{E_{pk_L}(l),E_{pk}(q_t,A_u),E_{pk_L}(k_u),k,c\}$，其中，$c$={0,1}，当 c 为 0 时，表示不需对查询结果信息加以缓存；当 c 为 1 时，表示需要对查询结果信息缓存并在后续查询过程中将结果集合 R 再次发送给用户。将查询结果保存在cache后，每次用户首先从CS的cache中获得查询结果，若该结果能够满足当前查询需求，则用户使用当前结果；否则用户再次提交查询给 CS，由 CS 重新发起新的一次查询。

4 安全性分析

在本文中，由于采用混合式系统架构，需要分析在不同实体间的信息传递可能导致的隐私信息泄露情况。因此，本文分别讨论在 LBS、CS 以及协作用户之间的用户信息安全情况。

4.1 用户与 CS 和 CU 之间的安全性

在提交查询过程中，CS 收到加密后的查询信息，其可获知的明文信息包括属性集合、匿名值和是否使用 cache，而对于加密信息的安全性取决于属性加密方案[22]，本文不再讨论该方案的安全性。可知在查询处理过程中，CS 无法获得用户所希望得到的兴趣点类型、当前所处位置以及用户提供的对称密钥，因此在查询提交给 CS 的过程中，用户自身的敏感信息并未泄露给 CS。在收到从 LBS 获得查询结果并反馈给用户的过程中，整个查询结果使用用户提交的对称密钥进行加密，CS 在不掌握解密密钥的情况下很难对该结果解密并获得用户隐私。同样，保存在 CS 的 cache 中的反馈结果也使用对称密钥进行加密，使在连续使用 cache 的情况下 CS 仍无法获得用户的个人隐私。

在密文信息传递到 CU 的过程中，若 CU 具有与用户提供不同的属性，则 CU 不具备解密用户信息的能力无法获知任何用户信息；当 CU 能够解密该信息时，CU 仅获得该用户提供的兴趣点类型集合，仍无法获知该用户包括位置和查询信息等任何相关隐私信息。

在最坏的情况下，CU 与 CS 之间共谋，此时，CS 可获得除用户外所有匿名用户的位置信息，以及查询的兴趣点类型集合。但是在整个过程中，用户的位置信息和反馈回来的兴趣点都保持在密文状态，即使这 2 个中心实体共谋也不会获得更多的用户信息。因此，本文所提供的方法更能保障在中心实体共谋情况下的用户隐私。

4.2 用户与 LBS 之间的安全性

在基于位置查询中，LBS 可掌握所有用户提交的位置和查询信息，使 LBS 可通过用户属性信息分析不确定位置集合 S 中与用户属性相似的位置，进而关联获得用户的位置轨迹。这种分析关联的方法可表现为第 3 节所述的用户属性相似性比较。在本文中，任何一次独立查询 CS 所提供的匿名集合都是能够解密由用户提供的属性基加密兴趣点的协作用户，该解密过程保障 CU 具有与用户相同的属性信息，这使 LBS 无法通过相似程度的差异识别出潜在的用户位置。另外，由 CS 发送给 LBS 的位置包含用户真实位置以及至少 k−1 个具有相同属性的匿名位置，这使 LBS 在 k 个位置中准确识别真实用户的概率为 $\frac{1}{k}$。同样，在连续查询过程中，每个连续的匿名集合都表现出相似的属性，使 LBS 将当前集合中的用户与下一集合中特定用户关联的概率等同于随机猜测。当用户设定的匿名值为 k，连续

查询的次数为 n 时，LBS 利用属性关联获得用户的真实轨迹的概率等于 $\frac{1}{nk}$。当用户通过使用 CS 中 cache 的缓存数据时，进一步降低了用户与 LBS 之间的信息交互概率， LBS 在该过程中无法获知任何与用户相关信息，使 LBS 收到的位置信息出现信息断层的情况。这种情况进一步增加了 LBS 对位置轨迹判断的不确定性，进而无法判断当前轨迹是否结束，或当前轨迹存在新的位置转换。因此，在使用 cache 的情况下，LBS 对真实轨迹的判断概率不再是 $\frac{1}{nk}$，而是 $\frac{1}{nmk}$，其中，m 是使用 cache 导致的连续轨迹出现的断层次数。这更增加了攻击者获取用户轨迹的难度。

整个算法的复杂度取决于用户设定的 k 值，当 k 值越大时算法需要搜索较大范围寻找到希望成为协作用户并且能够解密的 CU，因此，在最坏的情况下寻找不到可用的协作用户，此时算法失败；在最好的情况下，即每次均能找到合适的协作用户，此时算法 2 被执行 k 次，且算法 2 中未进行循环，因此其时间复杂度为 $O(k)$。在通常情况下，用户在广播几次之后仍未能找到协作用户即可认为算法执行失败，所以在一般情况下该算法的时间复杂度为 $O(k)$。CP-ABE 的算法复杂度计算参考文献[22]。

5　实验验证

5.1　评估标准

对于本文所提出的隐私保护方法，主要从隐私保护效力和算法的执行效率 2 个方面加以评估。其中，隐私保护效力从单个匿名集合攻击者识别用户信息熵和连续查询过程中的平均熵[5]变化 2 个方面来验证；算法的执行效率从加解密算法的执行时间，以及寻找到足够协作用户完成匿名的成功率 2 个方面加以评估。

由于本文方法主要是通过对用户属性的泛化来实现用户匿名的，所以为比较算法优势，本文主要与文献中的 IRDA 算法、部分属性泛化的 TOA 算法[7]、基于假位置的锚点 HINN 算法[20]、基于假轨迹的 TPDA 算法[16]等算法进行比较。所有评估实验均在 Windows 7 操作系统上使用 Matlab 7 加以模拟。其运行环境为 1.70 GHz Intel Core i5，内存大小为 4 GB。实验数据集采用位置隐私保护公认的 BerlinMOD Data Set 真实数据集，并选择城市中心区域，以获取更多的申请用户，以便获得较好的实验结果。为验证匿名效果本文假设分别有 30%、50%、80%的用户同意参与匿名。

通常，攻击者对于用户位置的最大不确定程度可由信息熵加以度量。假设对于不确定位置集合 S 中的任意位置，攻击者准确识别的概率为 $p(i)$，则在每个单次查询建立匿名集合中，攻击者对用户真实位置的不确定程度可表示为

$$H(i)=-\sum_{i=1}^{k}p(i)\mathrm{lb}p(i) \tag{3}$$

其中，k 为 CS 提交的属性相同的匿名用户数量。

由此，可以计算在连续位置匿名下的用户平均熵为

$$\overline{H_c}=\frac{\sum_{i=1}^{n}H_c(i)}{n} \tag{4}$$

其中，n 表示连续查询中的快照查询次数。最后，可获得每次快照查询的平均熵方差

$$\sigma^2=\mathrm{E}[(H_c-\overline{H_c})^2]=\frac{\sum_{i=1}^{n}(H_c-\overline{H_c})^2}{n} \tag{5}$$

显然，由于在不同快照查询下的不同位置熵的不确定性，使 σ^2 的取值越小，表明当前的隐私保护级别越高。

加解密算法的执行时间取决于用户所需要泛化的属性数量，本文主要检测随着用户属性量增加造成的加解密时间的增长，以便检测是否具有较好的执行效率。匿名成功率主要指在用户提出的匿名值基础上，每次快照查询过程中在协作用户比例变化的情况下成功完成匿名服务和提出匿名要求用户之间的比值，该值可表示为

$$p_s=\frac{\mathrm{sum}(U_s)}{U} \tag{6}$$

其中，U_s 表示成功匿名用户数，U 表示当前申请匿名用户总数。根据式(6)可得在连续位置服务过程中的连续匿名成功率

$$p_c=\frac{\sum_{i=1}^{n}\mathrm{sum}(U_s(i))}{nU} \tag{7}$$

根据以上设定的评估标准本文进行相应实验

验证，下一节将对所取得的实验验证结果以及可能造成该结果的具体原因加以详细说明。

5.2 实验结果

从图 2 可以看到不同算法的隐私保护能力，熵值按照 TOA<TPDA<HINN<IRDA<SACU 的顺序排列。这是由于在 HINN 中，用户将自身的真实位置使用指定的假位置锚点来代替真实位置。同一的锚点位置在一定程度上对用户的属性信息起到了泛化的作用，但是诸如兴趣点等特殊属性并没有得到较好的匿名处理，攻击者仍可根据部分未能有效处理的用户属性识别用户。TPDA 由于使用生成假轨迹的方法，在一定程度上实现了属性泛化，但是每次查询中产生的假位置在不可到达等方面仍与真实位置之间存在差异，这使攻击者在一定程度上可利用假位置与真实位置之间的不可到达性进行筛选，进而识别出用户的真实位置。TOA 由于主要关注连续位置服务中的时间泛化，因此，在用户属性的泛化要求较低，这使很多用户自身属性表现出与匿名用户之间的差异，进而攻击者对真实位置的识别概率要高于其他方法。SACU 和 IRDA 均能达到最大熵值，也就是在当前不确定位置集合中，攻击者对用户的真实位置存在最大不确定性，即使攻击者通过使用用户属性加以关联仍无法准确识别出用户真实位置。可见在快照查询下，相比于其他隐私保护算法 SACU 算法能够得到较好的隐私保护效果。

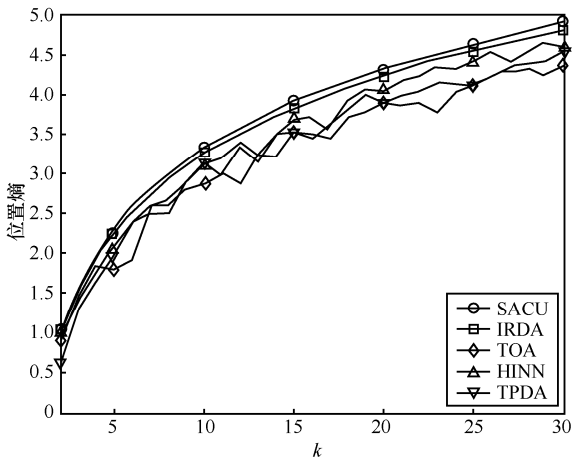


图 2 快照查询下的位置熵

图 3 给出了在连续 20 次查询下，随用户匿名值增加情况下的各方法平均熵方差变化，用以验证连续查询时不同算法的隐私保护能力。从图 3 可以看出，各算法的曲线范围按照 SACU<IRDA<HINN<TPDA<TOA 的顺序，且在连续查询下表现出的平均熵方差之间的差异较小，正如上文所说，该值越小表示连续位置服务过程中每次快照服务产生的熵差异越小，进而用户的隐私保护程度越高。在各种算法中，SACU 由于在每次快照查询中均可以保证用户可以获得最大熵，即攻击者对用户位置存在最大不确定性，因此，在连续查询下，其平均熵方差取值最低，随着匿名值变化始终保持为零。IRDA 由于在连续匿名过程中需要有中心服务器筛选具有相同属性的匿名用户，其中心服务器计算量较大，存在一定的匿名失败概率，因此，在连续匿名时可能会因连续匿名失败造成攻击者识别真实用户的情况，使其平均熵方差取值要高于 SACU。HINN 在连续查询时使用连续锚点代替真实位置，存在某些查询反复使用同一锚点的情况，这使在连续查询的过程中攻击者可利用锚点之间变换的时间差分析出用户潜在的移动速度，进而通过速度这一特殊属性识别该用户。TPDA 由于更关注整个轨迹的相似程度，对用户属性尤其是连续表现出的用户属性关注较少，使该方法在攻击者通过连续属性关联情况下泄露用户隐私的情况较为突出，因此其平均熵方差取值较大。TOA 方法由于本身在快照查询过程中对用户属性匿名存在一定缺陷，使该方法在连续查询过程中，其单次快照查询产生的熵值与平均熵值之间的差异较大，进而导致连续查询中所计算的平均熵方差取值较高，可以断定在连续查询下攻击者通过属性关联攻击可获得用户位置隐私的概率要远高于其他方法。

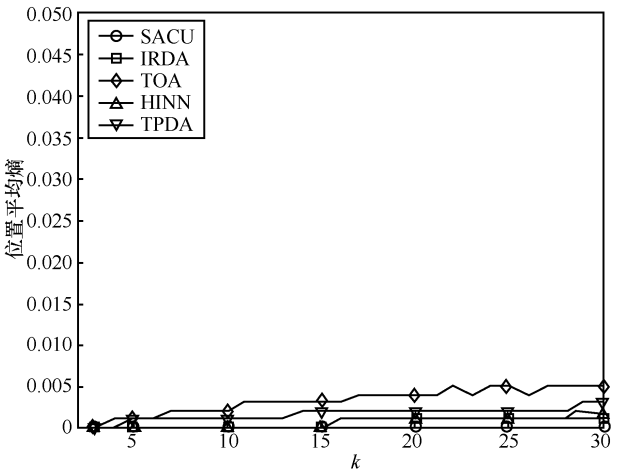


图 3 连续查询下的位置平均熵

从图 4 可以看到使用基于属性的加密方法在进行加解密过程中的时间消耗，以此检测 SACU 算法是否能够满足基于位置服务的实时特性。通常情况下，用户加解密的时间取决于用户提供属性的数量，当属性数量越大其加解密的时间越长。从图 4 可以看出，无论是在加密过程中还是在解密过程中，属性基加密都可在 0.6 s 以内完成操作，而且是在用户所提出的属性达到 30 个不同类型情况下的最长处理时间。实际上用户在进行连续查询时能够被攻击者识别的属性不会过多，主要表现在用户的查询时间间隔、移动速度、兴趣点类型等几个方面，其真实的属性数量要远小于 30。由此可以断定，该方法能够在较短的时间内收集到足够的具有相同属性的匿名用户，其算法时间效率能够得到较为有效的保障。由于加解密算法均在较短的时间范围内完成，因此 SACU 算法不会对基于位置服务的服务质量造成较大影响。

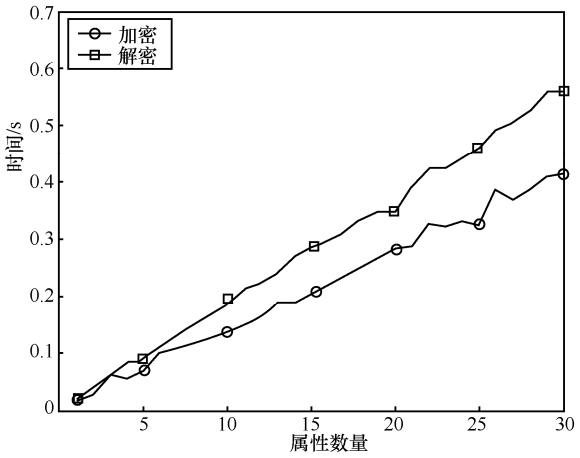


图 4　属性基加密的加解密执行时间

图 5 显示的是各算法在连续查询中每次快照查询下的匿名成功率，用以检测快照查询时各算法执行时的执行效率。在图 5 中可以看到快照查询下的成功率按照 IRDA<TOA<HINN<SACU<TPDA 排列，这是由于在存在匿名用户不足的情况使随匿名值增长各算法的隐私保护成功率逐渐降低。其中，IRDA 算法所体现出的隐私保护成功率最低，这是由于该算法通过中心服务器选取当前不确定用户集合中存在的具有相同属性的匿名用户，整个选择过程需要对所有用户的属性进行计算比较，其计算量和检测量十分巨大，使匿名过程可能超过用户的时间忍耐程度，导致用户放弃匿名要求。TOA 方法通过历史数据或当前中心服务器收到的其他用户数据进行匿名，其匿名过程需要考虑到一定的用户属性泛化，对参与匿名的用户要求较高，因此存在无法找到足够匿名用户的风险。HINN 方法需要在每次查询时选择最邻近的锚点，这也正是该方法使用 Voronoi 进行预划分，并部署锚点的主要原因。这种锚点部署存在部分区域无法有效使用锚点的问题，同时在同一时间区段内，使用同一锚点的用户数量也间接影响着用户的匿名成功率。这使 HINN 算法在快照查询中的匿名成功率相对不高。SACU 算法由于使用广播的方式，使较大范围内的用户可收到广播的协作请求，因而可以获得更大范围内的匿名响应用户，同时将属性相似性计算分布到协作用户客户端进行计算，降低了中心服务器的计算复杂性，提高了算法的执行时间，其快照查询下的隐私保护成功率相对较高。TPDA 方法由于使用生成的假位置参与匿名，无论在任何时间以及任何地点范围都可以生成大量的虚假位置参与匿名，因此，其匿名成功率最高，但是由于该方法需要针对生成假位置进行真实性筛选，使该方法也存在一定的匿名失败情况。在快照查询匿名的成功率方面 SACU 算法效果稍好于其他几种算法，但低于 TPDA。

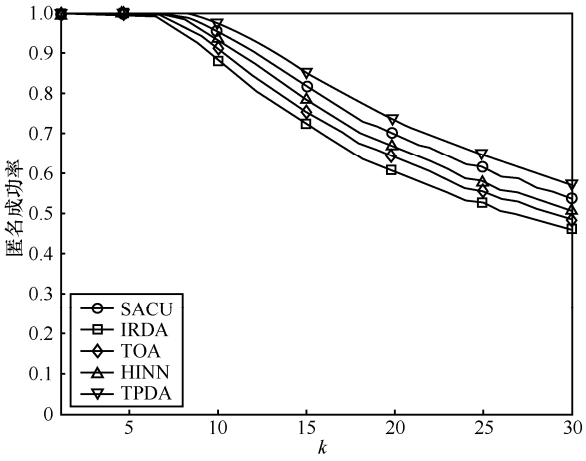


图 5　快照查询下的匿名成功率

图 6 给出了在 20 次连续查询情况下的隐私保护成功率，用以检测连续查询时算法的隐私保护执行效率。与快照查询的匿名成功率不同之处在于，一旦又一次单独的快照查询出现隐私保护失败的情况则表示连续查询下隐私保护失败。这是由于单次失败导致连续查询中的某一阶段被攻击者识别，根据这种位置阶段的暴露，攻击者可以通过轨迹校正的方式分析获得真实轨迹。从图 6

可以看出，各算法虽然在整体匿名成功率上与快照查询极为相似，但是由于将连续查询过程作为一个整体进行比较，使所有算法均表现出远低于快照查询的成功率。其中的主要原因与快照查询中的匿名成功率相似，不同的就是在于每一次快照查询的匿名失败则标志着当前连续查询中的匿名失败。

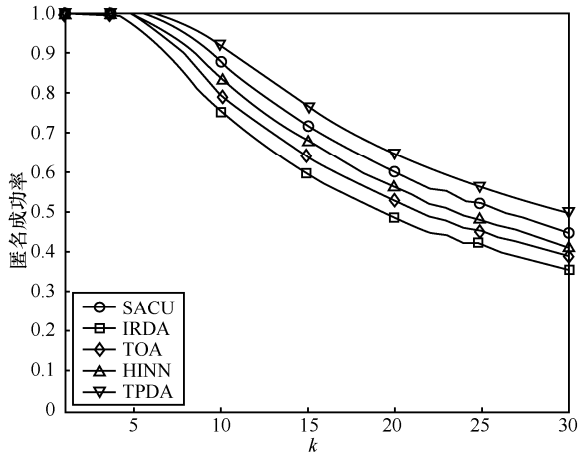


图 6 连续查询下的匿名成功率

综上，可以看到，SACU 算法能够提供当前同类算法无法提供的更好的基于用户属性的隐私保护效力，同时该算法可保障在中间传输过程中，半可信第三方对传递的查询信息的非恶意的好奇分析。另外，将属性比较过程发布到协作用户客户端通过解密的方式比较，一方面降低了中心服务器的计算负载，另一方面也提高了对相似属性匿名协作用户的查找概率。最后，与同类方法的匿名成功率比较进一步说明了 SACU 算法具有较好的算法执行效率。

6 结束语

本文通过使用属性基加密的方法，解决了中心服务器在半可信情况下的用户属性匿名问题，并且通过协作用户解密的方式选择具有相同属性信息的匿名用户，实现了在连续位置查询过程中的用户属性信息匿名，降低了攻击者通过用户属性信息分析将用户位置关联成用户轨迹的攻击成功率。同时，基于协作用户的解密方式，也在一定程度上降低了中心服务器的计算负载，提高了基于位置服务的服务质量。但是，本文在取得较好隐私保护效果的同时仍存在部分尚未解决的问题，如在共谋情况下的潜在隐私泄露度量问题、协作用户解密的效率保障问题等，这些问题将作为今后的研究方向，以便于对位置隐私保护进行更加深入的研究。

参考文献：

[1] 马春光, 张磊, 杨松涛. 位置轨迹隐私保护综述[J]. 信息网络安全, 2015(10): 24-31.
MA C G, ZHANG L, YANG S T. Review on location trajectory privacy protection[J]. Netinfo Security, 2015(10): 24-31.

[2] 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.

[3] LIN C, WU G W, YU C W. Protecting location privacy and query privacy: a combined clustering approach[J]. Concurrency and Computation-Practice & Experience, 2015, 27(12): 3021-3043.

[4] 周长利, 马春光, 杨松涛. 基于敏感位置多样性的 LBS 位置隐私保护方法研究[J]. 通信学报, 2015, 36(4): 129-140.
ZHOU C L, MA C G, YANG S T. Research of LBS privacy preserving based on sensitive location diversity[J]. Journal on Communications, 2015, 36(4): 129-140.

[5] NIU B, GAO S, LI F H, et al. Protection of location privacy in continuous LBSs against adversaries with background information[C]// 2016 International Conference on Computing, Networking and Communications (ICNC). 2016.

[6] 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.

[7] 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.

[8] NI W, GU M, CHEN X. Location privacy-preserving *k* nearest neighbor query under user's preference[J]. Knowledge-Based Systems, 2016, 103: 19-27.

[9] 张磊, 马春光, 杨松涛, 等. 基于轮廓泛化的位置隐私保护模型及方法[J]. 系统工程与电子技术, 2016, 38(12): 2894-900.
ZHANG L, MA C G, YANG S T, et al. Location privacy protection model and algorithm based on profiles generalization[J]. Systems Engineering and Electronics, 2016, 38(12): 2894-2900.

[10] 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, F, 2003: 1189037.

[11] XIAO Z, XU J, MENG X. P-sensitivity: a semantic privacy-protection model for location-based services[C]//International Conference on Mobile Data Management Workshops. 2008.

[12] FUYU L, HUA K A, YING C. Query l-diversity in location-based services[C]//Mobile Data Management: Systems, Services and Middleware. 2009.

[13] CHOW C Y, MOKBEL M F, LIU X. A peer-to-peer spatial cloaking algorithm for anonymous location-based service[C]//14th Annual ACM International Symposium on Advances in Geographic Information Systems. 2006: 171-178.

[14] MA C, ZHANG L, YANG S, et al. Achieve personalized anonymity through query blocks exchanging[J]. China Communications, 2016, 13(11): 106-18.

[15] GHINITA G, KALNIS P, KHOSHGOZARAN A, et al. Private queries in location based services: anonymizers are not necessary[C]//2008 ACM SIGMOD International Conference on Management of data. 2008: 121-32.

[16] KATO R, IWATA M, HARA T, et al. A dummy-based anonymization method based on user trajectory with pauses[C]//20th International Conference on Advances in Geographic Information Systems. 2012.

[17] GHASEMZADEH M, FUNG B C M, CHEN R, et al. Anonymizing trajectory data for passenger flow analysis[J]. Transportation Research Part C-Emerging Technologies, 2014, 39(2014): 63-79.

[18] PALANISAMY B, LIU L, LEE K, et al. Anonymizing continuous queries with delay-tolerant mix-zones over road networks[J]. Distributed and Parallel Databases, 2014, 32(1): 91-118.

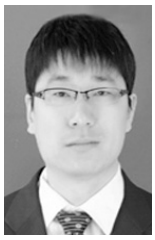
[19] 张磊, 马春光, 杨松涛. 基于位置关联相似性的匿名算法[J]. 中国科技论文, 2016, 11(2): 197-201, 13.
ZHANG L, MA C G, YANG S T. Location association similar based anonymus algorithm[J]. China Sciencepaper, 2016, 11(2): 197-201, 13.

[20] 马春光, 周长利, 杨松涛, 等. 基于 Voronoi 图预划分的 LBS 位置隐私保护方法[J]. 通信学报, 2015, 36(5): 5-16.
MA C G, ZHOU C L, YANG S T, et al. Location privacy-preserving method in LBS based on Voronoi division[J]. Journal on Communications, 2015, 36(5): 5-16.

[21] RODRIGUEZ-CARRION A, REBOLLO-MONEDERO D, FORNE J, et al. Entropy-based privacy against profiling of user mobility[J]. Entropy, 2015, 17(6): 3913-3946.

[22] BETHENCOURT J, SAHAI A, WATERS B. Ciphertext-policy attribute-based encryption[C]//IEEE Symposium on Security and Privacy. 2007.

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。