基于多方安全计算的属性泛化 mix-zone

王斌[1,2]，张磊[2]，张国印[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：针对路网环境下 mix-zone 无法有效地实现属性进行隐藏或泛化和抵御伪装攻击的问题，基于属性泛化和同态加密，提出了一种秘态属性泛化的隐私保护方法。该方法通过同态加密，实现了秘密出价选择计算代理、秘密计算相似属性，并以相似属性完成属性泛化的整体处理。通过属性泛化，解决了 mix-zone 可被攻击者利用属性追踪的问题，同时秘密计算的属性处理不会泄露任何信息给参与者，也防止伪装攻击者获得 mix-zone 中各用户的隐私信息。最后，通过安全性分析和实验验证分别在理论和实践这 2 个方面对所提算法的优势加以分析和比较。

关键词：路网环境；隐私保护；属性泛化；同态加密；安全计算

中图分类号：TP311

文献标识码：A

doi: 10.11959/j.issn.1000−436x.2019077

Attribute generalization mix-zone based on multiple secure computation

WANG Bin[1,2], ZHANG Lei[2], ZHANG Guoyin[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 order to cope with the problems of resist the attack of user tracking with attributes and resist the disguised attack in road networks, based on the conception of attribute generalization and the methods of homomorphic encryption, a privacy protection method to complete the calculation of attribute generalization in private state and achieve attribute generalization was proposed. The proposed method utilized the homomorphic encryption to achieve secret bidding selection with an agent, and then utilize secure multi-party computation to calculate the similar attributes and achieve the attribute generalization in the end. With the help of fully attribute generalization, this method can solve the problem of users can be tracked by the potential attributes, and at the same time the privacy calculation also does not reveal any information to participants, so this method can prevent the disguised attacker that obtains user's information in mix-zone. At last, in order to demonstrate the superiority of the proposed method in both of academically and practicality, the security analysis and experimental verification are given, and the procedure of formulation verification and the result of experiment further substantiate the superiorities of the proposed method.

Key words: road network, privacy protection, attribute generalization, homomorphic encryption, secure computation

---

收稿日期：2018–12–12；**修回日期**：2019–02–04

通信作者：张磊，8213662@163.com

基金项目：黑龙江省自然科学基金资助项目（No.F2015022）；黑龙江省普通本科高等学校青年创新人才培养计划基金资助项目（No.UNPYSCT-2017149）；国家级大学生创新创业训练基金资助项目（No.201810222033）；黑龙江省省属本科高校基本科研业务费基金资助项目

Foundation Items: The Natural Science Foundation of Heilongjiang Province (No.F2015022), The University Nursing Program for Young Scholars with Creative Talents in Heilongjiang Province (No.UNPYSCT-2017149), The National Undergraduate Innovation and Entrepreneurship Training Program of China (No.201810222033), Basic Scientific Research Service Fee Project of Heilongjiang Provincial Undergraduate Universities

1 引言

随着定位技术和无线通信技术的不断发展，基于这类技术的位置服务已经广泛地深入人们的日常生活当中，这种以提交用户当前或一段时间真实位置并获得如导航、兴趣点查询、广告推送等相关信息的服务方式为广大用户的日常生活带来了极大的便利。但是，随着越来越多地使用这种服务，人们逐渐发现基于位置的服务存在着获取及威胁用户个人隐私的情况，这使很多人逐渐减少对其的使用，并很大程度上制约了这种服务方式和服务技术的发展。

针对隐私泄露问题，很多研究者都提出了自己的观点。其中，提出最早并影响最深的是 Gruteser 等[1]提出的 *k*-匿名观点，该观点认为把用户的真实位置与其他至少 *k*−1 个位置共同提交给基于位置服务的服务提供商，利用真实位置与其他用户位置之间的不确定性来降低攻击者准确识别指定用户的概率，以此保护用户的个人隐私。随着该观点的提出，区域泛化[2]、位置多样性[3]、查询多样性[4]等相关观点被分别提出，并逐渐演化为欧氏空间的快照查询隐私保护[5-7]和欧氏空间的连续查询隐私保护[8-9]。但这类方法较难应用到真实环境当中，因为人们真实活动的空间被道路组成的网络结构所覆盖，按照欧氏空间得到的匿名位置可能位于一些不可到达或者易被攻击者识别的区域，进而无法有效地起到泛化用户真实位置的目的。因此，研究者又专门针对路网环境，提出了在该环境下的连续查询隐私保护方法[10-12]。不过，这些方法或者无法有效地应对攻击者使用的追踪攻击，或者在隐私保护过程中过于追求隐私保护的效果而对服务质量影响较大。这种情况导致另外一种被称作 mix-zone[13-18]的可抵抗追踪攻击并能有效减少对服务质量影响的方法被广泛研究和应用。通常情况下，mix-zone 可看作是一个外部用户无法获知任何消息的黑盒区域，在该区域中，内部用户可以进行信息交互、信息处理等一系列操作，而外部用户无法通过各种方式获取任何信息传递情况。同时，由于 mix-zone 采用的是按照关键位置节点部署的方式，使 2 个连续节点之间的用户未受到各种算法的影响，因此对服务质量的影响相对较弱。综上，mix-zone 的这一系列优点使其更适合在路网环境部署应用。

不过，用户在连续移动并查询的过程中并非仅仅去掉假名、查询时间等有限属性便可有效地抵抗攻击者的攻击行为。按照张磊等[19]的观点，攻击者可以根据连续用户所表现出的多种属性进行关联，进而猜测出用户的连续位置，分析获得用户的隐私信息。尽管在文献[20-22]中已经利用属性泛化实现隐私保护，但由于所针对环境的差异，这些方法并不适合直接应用到 mix-zone 中。针对这个问题，本文以属性泛化为基本出发点，从 mix-zone 构成的基本特性考虑，基于同态加密的基本思想，设计并提出了一种基于 mix-zone 半可信的用户代理、且用户属性信息全程不可获知的隐私保护方法，简称 AG mix-zone 算法。基于该方法，mix-zone 中用户可秘密计算获知泛化后的属性集合信息，利用泛化后的属性集合信息在离开 mix-zone 后表现为用户属性，以此令攻击者无法在非 mix-zone 区域对用户属性加以关联。最后，本文通过安全性分析和实验验证，证明了本文所提算法的隐私保护能力及其在实际部署中的执行效率。

2 预备知识

2.1 系统架构

mix-zone 主要设计用来防止攻击者实施追踪攻击，即通过 mix-zone 切断攻击者对用户的连续暴露位置所表现出的用户信息。但是，由于用户的属性彼此之间存在差异，使假名、查询间隔等有限属性互换很难隐藏可被攻击者利用的全部属性信息，进而表现为如图 1 所示的可被攻击者利用的属性关联模型。图 1 中，进出 mix-zone 的移动用户均表现出自身特定的用户属性，如用户 ID、查询频率、查询内容等，并用三角形、五角形、五边形等不同图形来表示。攻击者可利用这些属性的相似性对进出 mix-zone 的用户加以关联，使 mix-zone 的作用失效。

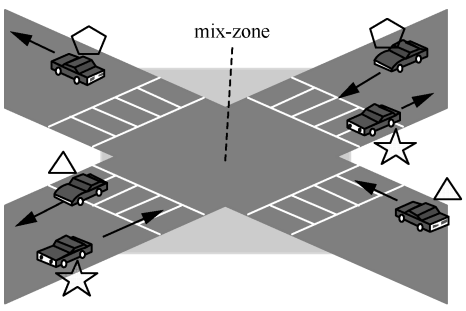


图 1 未经属性泛化导致的用户属性表现

基于上述分析及图 1 的观测比较发现，这种关联的攻击方式可以表示为用户可被观测的全部属性的相似性比较，因此使用 $\mathrm{sim}(A_s, A'_s)$ 来量化这种相似性，其中，$\mathrm{sim}(\cdot)$ 表示相似性。设某一指定被追踪用户的属性为 $A_s=(a_{s,1},a_{s,2},\cdots,a_{s,m})$，该用户离开 mix-zone 之后表现出的属性为 $A'_s=(a'_{s,1},a'_{s,2},\cdots,a'_{s,m})$，若存在 $\mathrm{sim}(A_s, A'_s)<\lambda$，则可认为这 2 种属性为同一用户所表现，即进出 mix-zone 的用户为同一用户，进而攻击者可对该用户进行持续追踪。其中，m 为可被识别的属性数量，λ 为攻击者可使用的属性相似最小阈值。

2.2　算法基本思想

针对利用用户表现出的属性相似性进行攻击的方法，最有效的手段就是进行属性泛化。属性泛化的基本思想是令所有离开 mix-zone 的用户表现出相似属性，使攻击者无法利用进出 mix-zone 的用户属性关联来识别出指定的追踪用户。图 2 给出了这种思想处理后表现出的属性泛化结果。

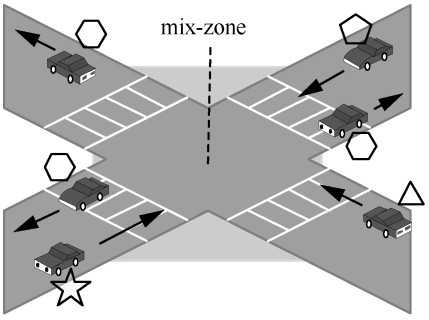


图 2　经过属性泛化后的用户属性表现

从图 2 中可以看出，进入 mix-zone 时表现出不同属性类型的移动用户，在经过 mix-zone 的处理之后，均表现出五边形属性，这使攻击者无法利用进出 mix-zone 的用户之间的属性相似来识别用户。形式化表示为 $\mathrm{sim}(A_s, A'_n)<\lambda, s\in n$，其中，$n$ 为当前 mix-zone 中的用户数量。但是，简单地通过 mix-zone 中的用户彼此之间交换并建立这种相似属性是不安全的，因为攻击者可以伪装成为参与者并被包含在 mix-zone 区域内，进而造成交换数据被攻击者获得从而泄露用户隐私，因此，需要一种既能够保障 mix-zone 中的用户彼此之间不能获得相关信息，同时又能通过选择的半可信代理用户完成相同属性信息处理的隐私保护方法。基于这一思路，本文利用同态加密的基本特性，分别利用秘密数据比较和秘密相似属性值计算对属性泛化的观点进行处理。在整个属性泛化的处理过程中，用户的信息既不会被代理获知，也不会被 mix-zone 中参与者所了解，所有用户仅能获得最后泛化的属性均值，并利用该均值实现离开 mix-zone 后的用户属性泛化。

3　半可信代理的属性相似 mix-zone

3.1　基于秘密数据比较的代理选择

完成属性相似计算的首要问题是在当前 mix-zone 区域内由所有用户共同选择一个半可信代理来完成属性相似计算。在各用户均无法获知真实出价的情况下，由出价最高的用户被选为代理。首先代理应是如图 3 所示的 mix-zone 内成员，以便与用户进行信息交互；其次，代理是在连续的竞标过程中胜出的一方，这样能够获得由其他用户反馈给该用户的消耗补偿。竞标过程可表现为如图 4 所示的分组竞标价格比较过程。

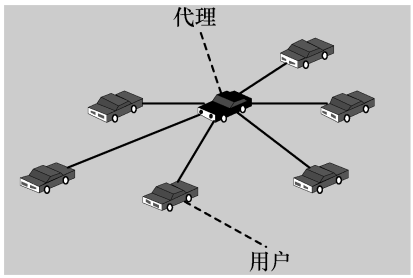


图 3　mix-zone 中代理与用户关系

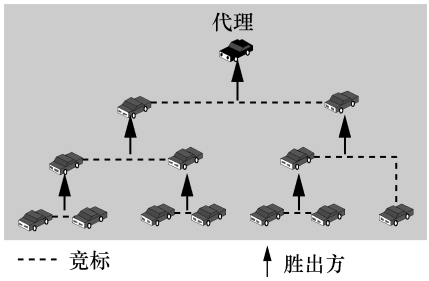


图 4　mix-zone 中分组竞标价格比较过程

因此竞标比较问题可转换为多方的秘密数据比较问题，借鉴文献[23]给出的秘密比较方法，可获得如下所述的秘密比较处理过程。

假设当前 mix-zone 中存在的参与者人数为 n，则参与者 $P_i(1\leqslant i\leqslant n)$ 每人拥有的长度为 l bit 的代理参选投标出价为 $b_1,b_2,\cdots,b_n$，转换为二进制表示为 $\vec{b_i}=(b_{i,l-1},b_{i,l-2},\cdots,b_{i,0})$，$0\leqslant l\leqslant n$，则出价 $b_1,b_2,\cdots,b_n$ 可按照以下步骤进行比较。

步骤 1　$P_i(1\leqslant i\leqslant n)$ 分别基于全同态加密方案生成公私钥对 $(\mathrm{pk}_i,\mathrm{sk}_i)$，并将公钥 pk_i 公开发布给 mix-zone 内所有参与者。

步骤 2 P_1 利用自己的公钥 pk_1 对自身的二进制出价 $\vec{b}_1=(b_{1,l-1},b_{1,l-2},\cdots,b_{1,0})$ 进行逐比特加密，得到加密后的信息 $E(b_1)=(E(b_{1,l-1}),E(b_{1,l-2}),\cdots,E(b_{1,0}))$，并向其他 $n-1$ 个用户 $P_k(2\leqslant k\leqslant n)$ 公开加密后的出价信息 $E(b_1)$。

步骤 3 当 P_k 收到加密后的出价 $E(b_1)$ 时，首先利用 pk_k 对自身的出价 $\vec{b}_k=(b_{k,l-1},b_{k,l-2},\cdots,b_{k,0})$ 进行加密，获得 $E(b_k)=(E(b_{k,l-1}),E(b_{k,l-2}),\cdots,E(b_{k,0}))$，然后计算

$$\overline{e}\triangleq E(c_k)=(E(c_{k,l-1}),E(c_{k,l-2}),\cdots,E(c_{k,0}))=(E(c_{k,l-1})\odot E(l),E(c_{k,l-2})\odot E(l),\cdots,E(c_{k,0})\odot E(l))$$

其中，$E(l)$ 是 l 在经过全同态加密后获得的密文，$E(c_k)\odot E(l)$ 表示 $E(c_k)+E(l)\bmod d$，即将加密后的 l 位与出价 $E(l)$ 的模之间进行 d 进位加法，如果其结果超出 l 位整数的表示范围，则只需去除低于 l 位的结果。

在完成上述计算之后，P_k 计算 $\mathrm{ans}(E(b_1),E(b_k))=E(b_1)+\overline{e}\triangleq E(g_k)$，其中，$g_k$ 表示第 k 位计算结果，$E(b_1)+\overline{e}$ 表示 $E(b_1)$ 和 $E(c_k)$ 对应的 2 个 λ 位整数的进位加法，$E(g_0)=E(d_{1,0})\odot E(c_{k,0})$；$E(t_0)=E(d_{1,0})E(c_{k,0})$；$E(g_i)=E(d_{1,i})\odot E(c_{k,i})\odot E(t_{i-1})$；$E(t_i)=(E(d_{1,i})\ E(c_{k,0}))\odot(E(d_{1,i})\ E(t_{i-1}))\odot(E(c_{k,0})\ E(t_{i-1}))$，$1\leqslant i\leqslant l-1$。$E(g_i)$和 $E(t_i)(0\leqslant i\leqslant l-1)$ 分别表示第 i 位的计算结果和第 i 位向第 i+1 位的进位；$E(d_{1,i})\ E(c_{k,0})$ 表示 $E(d_{1,i})$ 与 $E(c_{k,0})$ 的模 d 进位乘法，模 d 进位加法和模 d 进位乘法均按照全同态加密方案中同态运算步骤加以执行，且每次对 2 个密文完成加法或乘法运算后，执行同态解密操作以降低密文中存在的噪声。在完成以上操作后，得到 $\mathrm{ans}(E(b_1),E(b_k))\triangleq(E(g_{k,l-1}),\cdots,E(g_{k,0}))$ 将 $E(g_{k,l-1})$ 发送给用户 P_1。

步骤 4 P_1 对 $E(g_{k,l-1})(2\leqslant k\leqslant n)$，解密后得到 $g_{k,l-1}$，若 $g_{k,l-1}$=0，则 $b_1\geqslant b_k$；若 $g_{k,l-1}$=1，则 $b_1<b_k$。P_1 将结果发送给 P_k。

同理，经过 $\dfrac{n(n-1)}{2}$ 次两方安全计算后，可得到参与者中的最大竞标价格，由竞标价格的发起者作为代理完成后续的属性相似操作。

3.2 基于同态加密的属性模糊计算

通常情况下，用户在道路环境下的连续移动过程中可被攻击者获得并被识别的 m 个用户属性可表示为 $A=(a_1,a_2,\cdots,a_m)$，针对 mix-zone 中的 n 个用户，可将这种属性表示为 $A_i=(a_{i,1},a_{i,2},\cdots,a_{i,m})$，$1\leqslant i\leqslant n$。属性模糊的目的就是使 mix-zone 中的多个用户表现出的属性集合之间存在 $A_1\approx A_2,\cdots,A_{n-1}\approx A_n$，即针对任意选择的 2 个用户 u_i 和 u_j $(i\in n,j\in n)$，其表现出的属性之间满足 $A_i-A_j\leqslant\lambda$，其中，λ 为攻击者不可分辨出的最小阈值。为实现这种属性相似计算，且 mix-zone 中各用户（包括所选择的代理用户）在计算过程中无法获知任何用户属性信息，本文修改文献[24]提供的单轮集合地点计算的方式，对不同用户提供的属性进行相似属性计算。该计算过程可表示如下。

代理用户利用 ElGamal 生成公钥 $\mathrm{pk}_s=\{p,g,Z\}$，私钥为 $\mathrm{sk}_s=\alpha$。其中，p 是一个大素数，g 是 $\mathbb{Z}_p$ 的发生器，满足 $Z=g^{\alpha}\bmod p$。

步骤 1 对于 mix-zone 中的任意 2 个用户 u_i 和 $u_j(i<j)$，用户 u_j 分发 $r_{ij,1}\in\mathbb{Z}_p^*$、$r_{ij,2}\in\mathbb{Z}_p^*$、$r_{ij,3}\in\mathbb{Z}_p^*$ 和 $r_{ij,4}\in\mathbb{Z}_p^*$ 这 4 个私密随机数；用户 u_i 分发 $r_{ij,1}^{'}\in\mathbb{Z}_p^*$、$r_{ij,2}^{'}\in\mathbb{Z}_p^*$、$r_{ij,3}^{'}\in\mathbb{Z}_p^*$ 和 $r_{ij,4}^{'}\in\mathbb{Z}_p^*$ 这 4 个私密随机数，其中，$r_{ij,t}+r_{ij,t}^{'}=r_{ij}(t=1,2,3,4)$。用户 u_i 随机选择一个正整数 $s_i\in\mathbb{Z}_p^*$ 满足 $\|s_i\|\leqslant\|p\|-1$，并将 s_i 保存，其中，$\|\cdot\|$表示比特长度。所有 mix-zone 中的用户共享一个通用的正整数 $r\in\mathbb{Z}_p^*$，且 r 满足 $\|r\|\leqslant\|p\|-35$。

步骤 2 将属性值平均分成两部分，用户 u_i 的属性为 $A_i=(x_i,y_i)$，$(x_i,y_i)\in(\mathbb{Z}_p^*)^2$，用户 u_j 的属性为 $A_j=(x_j,y_j)$，$(x_i,y_i)\in(\mathbb{Z}_p^*)^2$。首先，用户 u_j 计算密文，具体如下。

$$C_{j,1}=r(x_j^2+y_j^2)Z^{r_{ij,1}}\bmod p$$
$$C_{j,2}=Z^{r_{ij,2}}\bmod p$$
$$C_{j,3}=2rx_jZ^{r_{ij,3}}\bmod p$$
$$C_{j,4}=2ry_jZ^{r_{ij,3}}\bmod p$$
$$C_{j,5}=g^{r_{ij,4}}\bmod p$$

然后，用户 u_j 将密文 $(C_{j,1},C_{j,2},C_{j,3},C_{j,4},C_{j,5})$ 发送给用户 u_i。当收到密文 $(C_{j,1},C_{j,2},C_{j,3},C_{j,4},C_{j,5})$ 后，用户 u_i 计算如下密文。

$$C_{ij,1}=C_{j,1}Z^{r'_{ij,1}}+C_{j,2}r(x_i^2+y_i^2)Z^{r'_{ij,2}} \bmod p=$$
$$r(x_i^2+y_i^2+x_j^2+y_j^2)Z^{r_{ij}} \bmod p$$
$$C_{ij,2}=(C_{j,3}x_i+C_{j,4}y_i)Z^{\tilde{r}_{ij,3}} \bmod p=$$
$$2r(x_ix_j+y_iy_j)Z^{r_{ij}} \bmod p$$
$$C_{ij,3}=C_{j,2}s_iZ^{\tilde{r}_{ij,2}} \bmod p=s_iZ^{r_{ij}} \bmod p$$
$$U_{ij}=C_{ij,1}-C_{ij,2}+C_{ij,3}=$$
$$[r((x_i-x_j)^2+(y_i-y_j)^2)+s_i]Z^{r_{ij}} \bmod p$$
$$V_{ij}=C_{j,5}g^{r'_{ij,4}}=g^{r_{ij}} \bmod p$$

最后，用户 u_i 将密文 (U_{ij},V_{ij}) 发送给代理。

步骤 3　代理在从用户 u_i 处获得密文 (U_{ij},V_{ij}) 后，使用自身的私钥 $\mathrm{sk}_s=\alpha$ 对其进行解密，并获得明文 $m_{ij}=r((x_i-x_j)^2+(y_i-y_j)^2)+s_i \bmod p=rd_{ij}^2+s_i \bmod p$，其中，$d_{ij}=\sqrt{(x_i-x_j)^2+(y_i-y_j)^2}$ 表示 2 个用户之间的属性差异距离。由于 $\|r\|\leqslant\|p\|-35$，$\|s_i\|\leqslant\|p\|-1$ 且 $0<rd_{ij}^2+s_i<p$，有 $0\leqslant m_{ij}=rd_{ij}^2+s_i<p$。

步骤 4　对于 mix-zone 中的每一个用户，代理计算 m_{ij} 的平均值 $\bar{m}_{ij}=\dfrac{\sum_{j=1,j\neq i}^{n}m_{ij}}{n-1}=r\dfrac{\sum_{j=1,j\neq i}^{n}d_{ij}^2}{n-1}+s_i=r\bar{d}_i^2+s_i$，然后计算差异 $r\Delta_{ij}=|m_{ij}-\bar{m}_{ij}|=|rd_{ij}^2-r\bar{d}_i^2|=r|d_{ij}^2-\bar{d}_i^2|$。对于每个差异 $r\Delta_{ij},(j\in[1,n],j\neq i)$，代理需要计算 r 次，获得对每个用户之间的属性差异 σ_i，即 $r\sigma_i=\dfrac{\sum_{j=1,j\neq i}^{n}r\Delta_{ij}}{n-1}$。在获得 $r\sigma_i(i\in[1,n])$ 后，代理在其中选择最小值 $r\sigma_k(k\in[1,n])$ 及其索引 k。由于 r 是正整数，可知 σ_k 是在集合 $\{\sigma_1,\sigma_2,\cdots,\sigma_n\}$ 之中的最小值。即计算后获得的 A_k 的属性值 σ_k 是与 mix-zone 中所有用户属性值最相近的，可以使用该值作为模糊后所有用户可使用的泛化属性。在整个计算处理过程中，每个用户均不知道其他用户的真实属性，而代理仅知道模糊后所有用户共同使用的泛化属性值，即在整个处理过程中没有用户属性信息被其他实体所获知。

3.3　安全性分析

由于本文所提算法是在属性相似的思想基础上，基于同态加密的处理方法实现的，该方法既可针对全程追踪的攻击者，又可针对参与 mix-zone 中的伪装攻击者，因此，其安全性需要从 2 个方面考虑：1) 攻击者对泛化后的属性信息准确猜测的不确定性；2) 算法中所使用的加密算法对参与到 mix-zone 中的 2 种用户（代理和一般用户）的信息泄露情况。

攻击者对泛化后属性猜测的不确定性使用信息熵进行度量，即设攻击者在离开 mix-zone 的众多用户中，根据表现出的属性相似性准确识别出追踪用户的概率为 $p_i(1\leqslant i\leqslant n)$，则攻击者在众多离开 mix-zone 中的用户中，利用相似性猜测的情况可表示为 $H_i=-\sum_{i=1}^{n}p_i\mathrm{lb}p_i$。按照 Jaynes 最大熵定理可知，当 H_i 最大时，攻击者具有最大不确定性，即无法在离开的 n 个用户中准确地识别出追踪用户。为证明这一点，假设攻击者掌握的追踪用户属性为 $A=(a_1,a_2,\cdots,a_m)$，在离开 mix-zone 的 n 个用户中的任意 2 个用户 u_i 和 u_j 的属性为 $A_i=(a_{i,1},a_{i,2},\cdots,a_{i,m})$ 和 $A_j=(a_{j,1},a_{j,2},\cdots,a_{j,m})$，根据属性相似，攻击者可通过计算 $\mathrm{sim}(A,A_i)$ 和 $\mathrm{sim}(A,A_j)$ 来判断这 2 个用户哪一个是被追踪用户。假设 $p_{u,i}$ 和 $p_{u,j}$ 分别是攻击者将这 2 个用户与追踪用户进行相似属性关联获得的猜测概率，则有 $H=-p_{u,i}\mathrm{lb}p_{u,j}\ -p_{u,j}\mathrm{lb}p_{u,i}$。而经过属性泛化，使 $\mathrm{sim}(A,A_i)\cong\ \mathrm{sim}(A,A_j)<\lambda$。攻击者无法通过相似属性对这 2 个用户进行区分，进而有 $p_{u,i}\approx p_{u,j}$，此时根据 Jaynes 最大熵定理可知 H 取最大值，即攻击者在 n 个用户中准确地识别出追踪用户的不确定性最大。

加密算法对用户信息的披露可以从代理竞选和属性私密计算两方面加以分析。在代理竞选方面，任意用户之间只存在竞选代理所需要的计算补偿，只有在计算补偿最小的情况下，该用户才能被选为代理。整个竞选过程中，所有出价均在加密状态下完成，任何用户均无法获知其他用户的出价。由于代理竞选是在 n 个用户中选择，可能会出现参与竞选用户共谋获取其中某一用户出价的情况，但仅在 $P_i(1\leqslant i\leqslant n-1)$ 的情况下，可猜测获得出价排位，但并不能获知具体出价。同时，由于 mix-zone 中的用户是一种随机建立的用户群组，一方面很难出现全部参与用户中仅余一个用户为待攻击对象的情况；另一方面，由于代理在属性泛化的过程中仍不能准确地获得待攻击对象的属性信息，因此其

攻击投入与收获不符，即攻击者很难通过这种方式对待攻击对象加以攻击并获得其隐私信息。

在属性私密计算方面，用户信息的披露程度可以从相似属性计算的私密性方面加以分析，即在相似属性的计算过程中没有隐私信息被 mix-zone 中的其他参与者获知。对于代理用户，由于代理需获得密文 (U_{ij},V_{ij}) 并通过自己的私钥 sk_s 解密计算泛化后的属性信息，而 $U_{ij}=C_{ij,1}-C_{ij,2}+C_{ij,3}=[r((x_i-x_j)^2+(y_i-y_j)^2)+s_i]Z^{r_{ij}} \bmod p$ 和 $V_{ij}=C_{j,5}g^{r'_{ij,4}}=g^{r_{ij}} \bmod p$ 分别是通过用户 u_i 和用户 u_j 的属性信息经过一系列模操作计算获取的。在模值（即 mod 计算后所取得的结果）为确定数的情况下，原始值（即加密前的明文信息）是不确定的，因而代理是不能利用获得的密文将用户的属性信息还原的，即使恶意用户与代理共谋攻击也无法还原该信息。

对于除代理外的参与用户，设 ε 为隐私参数，$\mathcal{A}$ 为攻击者，$\mathcal{B}$ 为待攻击对象，攻击者对属性信息的攻击可转换为 $\mathcal{A}$ 和 $\mathcal{B}$ 之间的博弈 $\mathrm{game}_{\mathcal{A}}^{\mathcal{B}}$。在整个属性信息泛化处理过程中，假设 $\mathcal{A}$ 可获得代理用户的私钥 sk_s 及 $\mathcal{B}$ 的公开参数。所有 mix-zone 中的参与者的属性信息可表示为 $A_1,A_2,\cdots,A_n$。$\mathcal{B}$ 执行秘密属性计算可得到泛化后的属性信息 $f(A_{\mathrm{gen}})=g(f(A_1),\cdots,f(A_n))$，并将密文 $f(A_1),\cdots,f(A_n)$ 及对 A_{gen} 的索引发送给 $\mathcal{A}$。$\mathcal{B}$ 再选择随机属性 $A_r=(x_r,y_r)$ 并随机选择 $\varepsilon\in\{0,1\}$。当 ε=0 时，将 A_{gen} 发送给 $\mathcal{A}$；当 ε=1 时，则将 A_r 发送给 $\mathcal{A}$。$\mathcal{A}$ 猜测针对隐私参数 ε 的猜测值 $\varepsilon'\in\{0,1\}$，当且仅当 $\varepsilon'=\varepsilon$ 时，$\mathcal{A}$ 获胜。此时攻击者在此博弈中的优势可表示为 $\mathrm{Adv}_{\mathcal{A}}=2\mathrm{pr}[\varepsilon'=\varepsilon]-1$。由于在整个过程中，$\mathcal{A}$ 无法通过有效的手段准确猜测 ε' 的取值，进而使 $\varepsilon'=\varepsilon$ 的概率为随机猜测概率，由此可获得攻击者的博弈优势 $\mathrm{Adv}_{\mathcal{A}}=2\mathrm{pr}[\varepsilon'=\varepsilon]-1=2\times\frac{1}{2}-1=0$，即攻击者不具备优势。因此，可认为除代理外的参与用户无法获知任意用户的属性信息。

基于上述分析，可认为本文所提的基于多方安全计算的属性泛化的 mix-zone 能够有效地提供隐私保护，且针对不同类型的攻击者均具有较好的防护能力。

4 实验验证

4.1 实验设定

为了验证本文所提 AG mix-zone 算法在隐私保护能力和算法执行效率这 2 个方面的优势，本文利用 BerlinMOD Data Set 提供的路网数据，并随机选择位置生成 mix-zone 进行测试。实验使用笔记本电脑的处理器为 Intel core I7，内存为 4 GB，操作系统为 Windows10，并采用 Matlab R2017a 作为测试工具进行模拟实验。同时，为进一步验证 AG mix-zone 算法的优势，在测试的过程中与当前的一些同类算法进行了比较，参与比较的算法有通过移动耽搁泛化查询间隔时间的等待忍耐 mix-zone（delay-tolerant mix-zone）[25]、利用 mix-zone 变形降低关联程度的偏移 mix-zone（shifted mix-zone）[26]、多维 mix-zone 授权的多维 mix-zone（multiple mix-zone）[17]和基于身份验证加密的加密 mix-zone（cryptographic mix-zone）[18]。实验将从隐私保护能力和算法执行效率这 2 个方面展开，并且同时针对规则 mix-zone 和不规则 mix-zone 进行测试比较，以便获得部署 mix-zone 的最佳形状。其中，隐私保护能力将在属性泛化后的信息熵度量、泛化后的属性差异率、进出 mix-zone 用户的可关联性这 3 个方面展开；算法执行效率将在规则 mix-zone 和不规则 mix-zone 状态影响下的算法执行时间和隐私保护成功率等这 2 个方面展开。在验证属性数量对算法的影响时，将用户限定在 10 个，而在验证用户数量对算法的影响时则将属性设定为 15 个。为简化算法验证过程，整个实验中所使用的用户和属性数量均为 1~30。

4.2 隐私保护能力分析

从图 5 中可以看出，除本文所提 AG mix-zone 算法外，在用户数量一定的前提下，其他算法的信息熵均随属性数量的增加而减小。这是由于本文算法主要针对 mix-zone 中的用户利用量化的多属性相似计算完成属性泛化，该方法处理的属性数量远超过其他算法。相比之下，加密 mix-zone 由于同样使用加密手段隐藏属性，表现要好于其他 3 种算法。而多维 mix-zone 表现要好于 mix-zone 和等待忍耐 mix-zone 算法，这主要是由于该算法使用多重的 mix-zone 进行重复覆盖，在一定程度上提升了属性隐藏的效率。偏移 mix-zone 和等待忍耐 mix-zone 由于针对的属性有限，且未能通过多重覆盖的方式

降低属性暴露的风险，随属性增加导致的信息熵取值相对较低，但偏移 mix-zone 表现要好于等待忍耐 mix-zone，这是由于偏移后的 mix-zone 在一定程度上起到了覆盖的作用，虽然低于多维 mix-zone 但要好于等待忍耐 mix-zone。

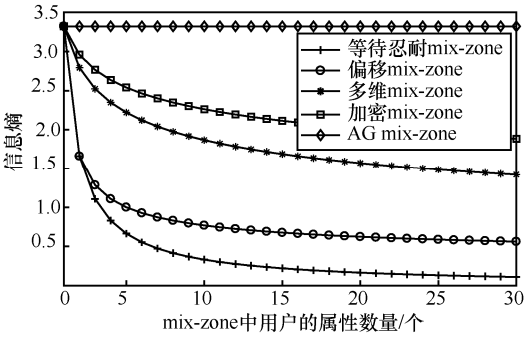


图 5　属性变化导致的泛化后信息熵曲线

从表 1 可以看出，5 种算法在属性数量确定的前提下，信息熵均随着参与用户数量的提升而增长，即攻击者对属性泛化后用户的不确定性在持续增加。这是由于用户数量提升了攻击者的不确定性，令攻击者猜测的基数不断增长。在 5 种算法中，AG mix-zone 算法能够取得信息熵最大值，这是由于该算法将所有显露的用户属性均加以泛化，最大程度地减少了攻击者可利用的属性。而其他算法中，多维 mix-zone 使用了多重覆盖的原理，在多个不同的 mix-zone 中使用大量用户进行多种属性的泛化处理，其信息熵取值比 AG mix-zone 算法稍低。加密 mix-zone 虽然采用加密手段进行属性隐藏，但其所针对的属性有限，其信息熵低于多维 mix-zone。另 2 种算法中，偏移 mix-zone 由于偏移性与覆盖相似，但偏移距离有限，限制了其对属性的泛化效果。最后，等待忍耐 mix-zone 由于主要应对时间间隔这一属性，其对于其他类型的属性泛化能力有限，导致该算法表现不佳。

从图 6 中可以看出，随着属性数量的增加，除本 AG mix-zone 算法外，其他算法在经过处理后用户表现出的属性之间仍存在较大差异，且属性数量越大这种差异越明显。由于 AG mix-zone 算法是针对所有潜在属性进行整体性泛化，因此泛化后的属性差异并不随着属性数量的增加而产生变化。其他算法则针对有限的属性加以处理，其处理能力有限，无法对所有属性展开泛化处理，属性差异随着属性数量的增加而逐渐变大。

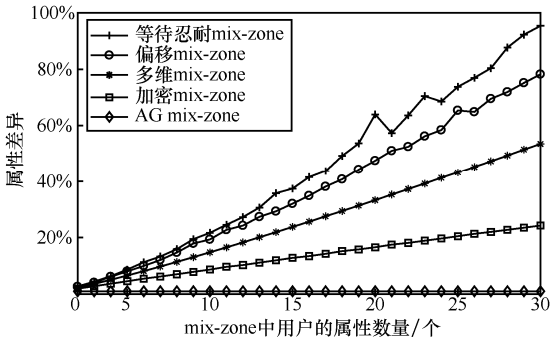


图 6　属性泛化后的差异

对于进出 mix-zone 用户的可关联性，本文采用成对熵加以度量，即成对熵取值越高，可关联性越低。从表 2 可以看出，AG mix-zone 算法可取得最高成对熵取值，因为该算法将用户间的属性加以泛化，使离开 mix-zone 的用户均表现出相似属性，攻击者无法将这些用户中的任意一个与进入 mix-zone 之前的用户相关联。而其他 4 种算法均无法对所有属性加以处理，进而造成攻击者可通过属性加以关联，致使成对熵取值随属性数量的增加而逐渐降低，即攻击者可将进出 mix-zone 用户加以关联的概率增加，用户存在被攻击者识别的风险。

从图 7 中可以看出，随着用户数量的变化，成对熵表示的用户可关联性均无变化，呈直线状态。

表 1　　**5 种算法在不同用户数量下的信息熵**

算法	用户数量为 5	用户数量为 10	用户数量为 15	用户数量为 20	用户数量为 25	用户数量为 30
等待忍耐 mix-zone	2.089 7	2.989 7	3.516 2	3.889 7	4.179 5	4.416 2
偏移 mix-zone	2.136 2	3.056 2	3.594 3	3.976 2	4.272 3	4.514 3
多维 mix-zone	2.275 5	3.255 5	3.828 8	4.235 5	4.551 0	4.808 8
加密 mix-zone	2.205 8	3.155 8	3.711 5	4.105 8	4.411 7	4.661 5
AG mix-zone	2.321 9	3.321 9	3.906 9	4.321 9	4.643 9	4.906 9

表 2 5 种算法在不同用户数量下的成对信息熵

算法	用户数量为 5	用户数量为 10	用户数量为 15	用户数量为 20	用户数量为 25	用户数量为 30
等待忍耐 mix-zone	0.244 5	0.133 1	0.084 4	0.067 0	0.053 9	0.050 3
偏移 mix-zone	0.253 9	0.153 5	0.110 3	0.086 5	0.073 1	0.063 0
多维 mix-zone	0.365 6	0.237 1	0.184 0	0.153 7	0.133 7	0.119 3
加密 mix-zone	0.446 2	0.311 2	0.257 4	0.223 6	0.199 6	0.182 6
AG mix-zone	1.000	1.000	1.000	1.000	1.000	1.000

这是由于在实验测试过程中，针对的是已确定的进出用户，成对熵表现的是同一用户在进入和离开 mix-zone 中的可关联性变化，这种变化不随其他用户的加入而受影响。但是在这些平行的直线中，AG mix-zone 算法的成对熵取值要高于其他算法，这是由于 AG mix-zone 算法对表现出来的所有属性展开了泛化，其可关联特性降到最低，即使在确定进出用户是同一用户的前提下，仍不可用属性相似的原理来确定该用户。而其他 4 种算法由于针对的属性数量有限，且相互间存在属性处理差异，使它们均无法取得成对熵的最高值。

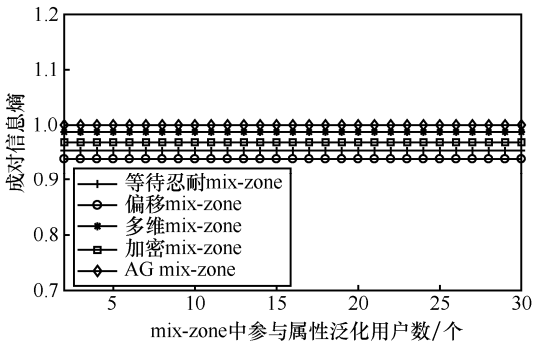


图 7 随用户数量变化导致的进出 mix-zone 用户可关联性差异

4.3 算法执行效率分析

从图 8 中可以看出，除 AG mix-zone 算法外，其他算法均随着属性数量的增加，算法执行时间显著增长。这是由于 AG mix-zone 算法采用的是 mix-zone 中的用户共同处理相似属性的方法，这种方法是对所有属性量化后由代理和用户之间共同完成的，整个处理过程不随着属性数量的增加而改变。而其他算法由于针对的属性种类等方面的限制，使在属性增加的情况下，算法不得不重复执行或者多次处理，进而导致随着属性数量的变化，算法的整体执行时间不断增加。其中，加密 mix-zone 和多维 mix-zone 由于需要多次执行算法以处理更多的属性，因此执行时间受属性变化的影响最高，且在超过某一值后，这种变化加剧。而等待忍耐 mix-zone 和偏移 mix-zone 算法的执行时间未随属性的增加而加剧并不是因为这 2 种算法很好地解决了重复执行的问题，而是因为这 2 种算法针对的属性有限，即使多次重复也无法有效地处理过多的属性，因而其执行时间的变化并未表现出急剧上升的情况。

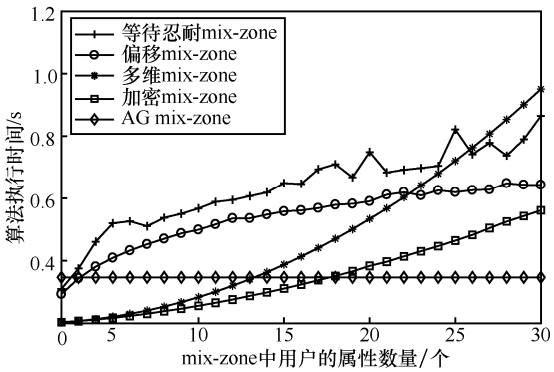


图 8 随属性变化的执行时间（规则的 mix-zone）

与规则 mix-zone 下算法随属性变化而导致的时间差异不同，从图 9 中可以看出，AG mix-zone 算法在不规则 mix-zone 下其执行时间要高于规则 mix-zone，这是因为在不规则 mix-zone 下，AG mix-zone 算法需要更多的时间来选择足够的且通信便利的用户来进行多方安全计算，通信距离的限制在不规则 mix-zone 中表现得更为明显，因而在一定程度上影响了 AG mix-zone 算法的表现。其他算法中，偏移 mix-zone 算法的表现要好于另外 3 种算法，这是由于该算法主要设计为不规则 mix-zone 下的区域偏移扩展，更适于在不规则 mix-zone 的条件下部署。

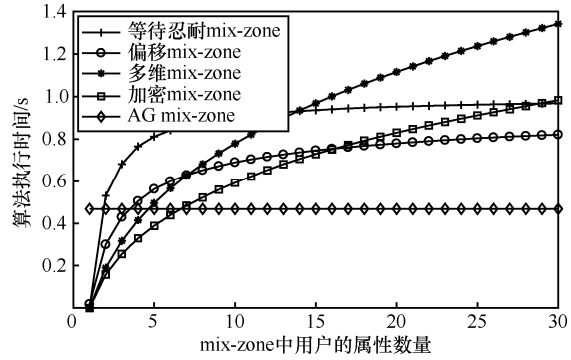


图 9　随属性变化的执行时间（不规则 mix-zone）

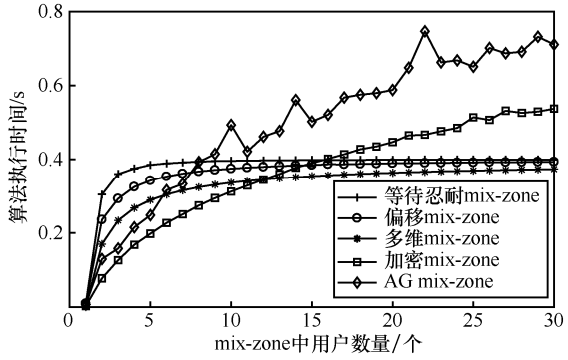


图 10　随用户数量变化的算法执行时间（不规则 mix-zone）

从表 3 中可以看出，随用户数量的变化导致的算法执行时间与随属性变化导致的执行时间存在较大差异。首先，AG mix-zone 算法的执行时间不再保持固定不变，而是随着人数的增加逐渐延长。这是由于该算法需要 mix-zone 内用户进行多方安全计算，而增加的用户数量势必会导致计算的轮次或计算的基数增大，进而导致算法执行时间的延长。以加密为手段的加密 mix-zone 同样受到这一影响，其算法执行时间随 mix-zone 中用户数量的增长而延长。其他 4 种算法由于仅需建立 mix-zone，并在区域中简单地对属性进行隐藏或者泛化，处理过程不随着用户数量的增加而更加复杂，因此在完成基本处理后无论 mix-zone 中的用户增加多少，处理时间保持稳定。

与规则 mix-zone 下的用户数量变化导致的算法执行时间差异相似，从图 10 中可以看出，AG mix-zone 算法的执行时间依旧受用户数量的影响较大，且同类以加密为手段的加密 mix-zone 表现依旧相似。所不同的是多维 mix-zone 在不规则 mix-zone 的条件下，受影响的程度有所降低，这是由于该算法在不规则 mix-zone 的条件下不再需要部署更多重复性的 mix-zone 来实现属性隐藏，因此降低了其部署 mix-zone 的数量，间接地导致算法执行时间的下降，但是这种降低较为有限。

从图 11 中可以看出，在规则的 mix-zone 中随着用户数量的增加，各算法的执行成功率逐渐降低，这是由于所有算法均需要在 mix-zone 中找到足够的用户以满足当前属性泛化所要求的用户数量，当不能找到足够数量的用户情况下，算法执行失败。在这些算法中，AG mix-zone 算法的执行成功率受用户数量的影响较小，这是由于该算法通过 mix-zone 中用户彼此间的多方安全计算来完成属性泛化，算法的执行仅需要找到足够数量的用户即可，并不需要对用户情况加以限制。其他算法中，多维 mix-zone 算法的执行成功率仅次于 AG mix-zone 算法，这是由于该算法仅需要在用户数量不足的情况下重复或者延伸部署 mix-zone 通过多重 mix-zone 来提升算法的执行成功率。相比之下，同样基于加密算法的加密 mix-zone 的受用户数量的影响相对较大，这是由于该算法需要通过寻找具有相似属性的用户来完成属性泛化，因此在很大程度上相似属性用户的寻找限制了其执行成功率。偏移 mix-zone 由于通过移动方式提升了相似属性用户的寻找能力，但是由于对查询时间的属性泛化，使该算法需要设定等待时间，而这种等待时间在很大程度上影响了用户参与的积极性，造成随着用户数量增加而

表 3　　**随用户数量变化的算法执行时间对比（规则 mix-zone）**

算法	用户数量为 5	用户数量为 10	用户数量为 15	用户数量为 20	用户数量为 25	用户数量为 30
等待忍耐 mix-zone	0.395 3	0.399 1	0.399 6	0.399 8	0.3998	0.399 9
偏移 mix-zone	0.379 2	0.392 1	0.395 5	0.397 0	0.3978	0.398 3
多维 mix-zone	0.371 0	0.387 4	0.392 2	0.394 5	0.3958	0.396 6
加密 mix-zone	0.315 8	0.381 2	0.511 3	0.551 2	0.5874	0.617 1
AG mix-zone	0.339 3	0.478 1	0.541 8	0.586 1	0.720 6	0.741 1

导致的算法执行成功率降低。最后，等待忍耐 mix-zone 既没有偏移包含更多用户的能力，又受到等待时间的影响，成功率最低。

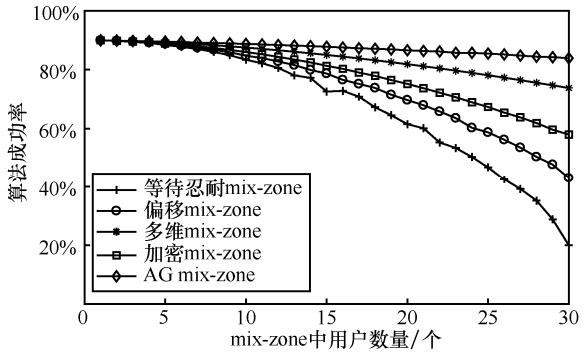


图 11　随用户数变化的算法执行成功率（规则 mix-zone）

不规则 mix-zone 中随用户数量变化导致的算法成功率差异与规则 mix-zone 中算法成功率的差异大体相同，所不同的是不规则 mix-zone 的算法成功率更高，进而表现为如表 4 所示的算法成功率整体上升的状态。这主要是不规则 mix-zone 由于设计特性导致其能够包含更多的用户在 mix-zone 当中，相比于规则 mix-zone 更多的用户为属性泛化提供了更多的选择，进而提升了属性泛化的算法执行成功效率。

从图 12 中可以看出，在规则 mix-zone 下随属性变化导致的算法执行产生的成功率差异。AG mix-zone 算法的执行成功率受属性变化的影响较小，仅当属性数量超过某一阈值时才逐渐表现出成功率的降低，这是由于该算法是通过属性量化后展开的相似属性泛化实现的隐私保护，算法所处理的是属性共有值而不是单独某一个值，而表现为不受属性数量影响的处理过程。而其他算法中，由于是直接对属性展开泛化，因此在属性增加的情况下，需要寻找满足属性相似的大量用户，在一定程度上，因属性数量增加而导致的相似属性用户寻找的困难造成了算法执行成功率的降低。另外，由于多维 mix-zone、等待忍耐 mix-zone 和偏移 mix-zone 针对的属性数量有限，在成功率下降的同时，这 3 种算法反而表现出平稳的下降趋势，这是由于这 3 种算法仅针对某些特定属性，在无法提供所有属性泛化的情况下，这 3 种算法仅需完成对所针对属性的泛化便可完成算法执行，因而其下降趋势有所缓解。

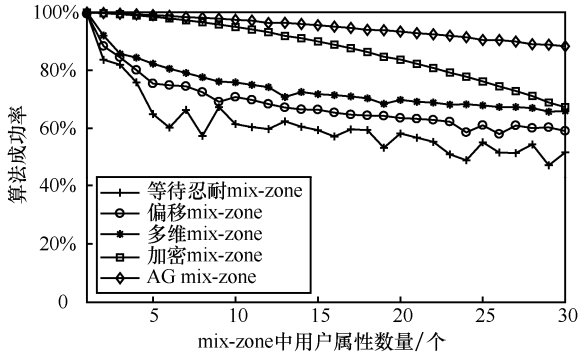


图 12　随属性变化的算法执行成功率（规则 mix-zone）

与在规则 mix-zone 下随属性变化的算法执行成功率相似，在不规则 mix-zone 下，各算法表现出的成功率差异可从如图 13 所示的曲线差异中看出。所不同的是在不规则 mix-zone 下，各算法的成功率要高于规则 mix-zone，这是由于不规则的 mix-zone 能够包含更多的用户，进而为算法执行提供了便利，充足的用户数量提升了算法的执行成功率。

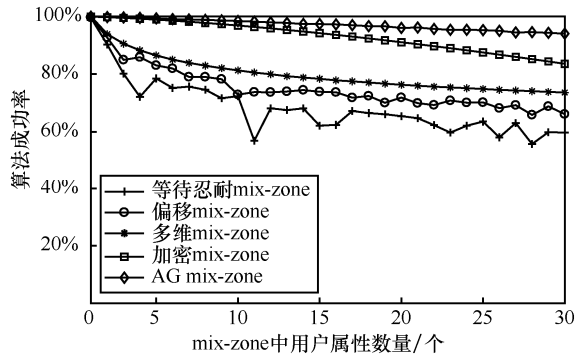


图 13　随属性变化的算法执行成功率（不规则 mix-zone）

表 4　　**随用户数变化的算法执行成功率对比（不规则 mix-zone）**

算法	用户数量为 5	用户数量为 10	用户数量为 15	用户数量为 20	用户数量为 25	用户数量为 30
等待忍耐 mix-zone	87.87%	81.97%	76.62%	70.33%	64.36%	52.78%
偏移 mix-zone	88.19%	83.68%	81.85%	78.00%	71.78%	64.39%
多维 mix-zone	88.88%	86.84%	84.19%	81.05%	82.50%	78.56%
加密 mix-zone	88.69%	86.00%	82.33%	82.88%	77.73%	76.91%
AG mix-zone	89.30%	88.31%	87.42%	86.33%	85.20%	83.98%

通过上述实验验证结果可以看出，本文 AG mix-zone 算法比其他同类算法具有更好地隐私保护能力和算法执行效率，因而该算法可更好地应用在实际路网环境的部署当中，有效地保护了用户的个人隐私。

5　结束语

在路网环境中，mix-zone 作为一种有效地防治攻击者对用户展开追踪攻击的手段被广泛的研究和应用。然而，在现实使用的过程中，一方面用户在离开 mix-zone 区域后所展现的属性可被关联；另一方面一些潜在的能够伪装并加入 mix-zone 中的伪装攻击者可获得用户彼此间传递的信息。针对这 2 种存在的问题，本文基于属性泛化和同态加密，提出了一种能够实现用户属性不可关联且 mix-zone 中用户无法获知彼此真实信息的隐私保护方法。该方法通过秘密选择代理，并由代理完成私密相似属性计算，最终实现离开 mix-zone 的用户彼此间的属性不可关联。为证明本文所提 AG mix-zone 算法的有效性和算法执行效率，在安全性分析中使用了博弈论的观点对加密手段和泛化效果进行了理论证明，同时在实验验证中给出了算法在实际部署中取得的各种效果，并通过与其他 4 种算法的比较进一步说明本文所提 AG mix-zone 算法的优势。尽管本文所提 AG mix-zone 算法能够解决 mix-zone 中存在的固有问题，但是由于所采用的多方安全计算需要较大的计算量和计算处理时间，AG mix-zone 算法在执行效率上仍存在较大缺陷，今后的工作将在如何提升执行效率方面展开。

参考文献：

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

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

[3] BAMBA B, LIU L, PESTI P, et al. Supporting anonymous location queries in mobile environments with privacygrid[C]//International Conference on World Wide Web. ACM, 2008: 237-246.

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

[5] 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-the International Journal of Grid Computing and Escience, 2015, 49(2015): 125-132.

[6] FEI F, LI S, DAI H, et al. A *K*-anonymity based schema for location privacy preservation[J]. 2017, PP(99): 1.

[7] SUN Y M, CHEN M, HU L, et al. ASA: against statistical attacks for privacy-aware users in Location Based Service[J]. Future Generation Computer Systems-the International Journal of Escience, 2017, 70: 48-58.

[8] YE A Y, LI Y, XU L. A novel location privacy-preserving scheme based on l-queries for continuous LBS[J]. Computer Communications, 2017, 2017(98): 1-10.

[9] PENG T, LIU Q, MENG D C, et al. Collaborative trajectory privacy preserving scheme in location-based services[J]. Information Sciences, 2017, 387: 165-179.

[10] 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(2015): 57-73.

[11] ZEBERGA K, JIN R, CHO H J, et al. A safe-region approach to a moving *k*-RNN queries in a directed road network[J]. Journal of Circuits Systems and Computers, 2017, 26(5): 1-18.

[12] SEPAHKAR M, KHAYYAMBASHI M. A novel collaborative approach for location prediction in mobile networks[J]. Wireless Networks, 2018, 24(1): 283-294.

[13] YING B, MAKRAKIS D, MOUFTAH H T. Dynamic mix-zone for location privacy in vehicular networks[J]. IEEE Communications Letters, 2013, 17(8): 1524-1527.

[14] PALANISAMY B, RAVICHANDRAN S, LIU L, et al. Road network mix-zones for anonymous location based services[C]// International Conference on Data Engineering. IEEE, 2013: 1300-1303.

[15] PALANISAMY B, LIU L. Effective mix-zone anonymization techniques for mobile travelers[J]. Geoinformatica, 2014, 18(1): 135-164.

[16] KANG J W, YU R, HUANG X M, et al. Location privacy attacks and defenses in cloud-enabled internet of vehicles[J]. IEEE Wireless Communications, 2016, 23(5): 52-59.

[17] ARAIN Q A, DENG Z L, MEMON I, et al. Privacy preserving dynamic pseudonym-based multiple mix-zones authentication protocol over road networks[J]. Wireless Personal Communications, 2017, 95(2): 505-521.

[18] ZHANG L. OTIBAAGKA: a new security tool for cryptographic mix-zone establishment in vehicular Ad Hoc networks[J]. IEEE Transactions on Information Forensics and Security, 2017, 12(12): 2998- 3010.

[19] 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): 2894-2900.
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.

[20] 张磊, 马春光, 杨松涛, 等. 基于属性基加密的用户协作连续查询

限私保护策略[J]. 通信学报, 2017, 38(9): 76-85.
ZHANG L, MA C G, YANG S T, et al. CP-ABE based users collaborative privacy protection scheme for continuous query[J]. Journal on Communications, 2017, 38(9): 76-85.

[21] ZHANG L, YANG S, LI J, et al. A particle swarm optimization clustering-based attribute generalization privacy protection scheme[J]. Journal of Circuits, Systems and Computers, 2018, 27(11): 171-121.

[22] DARGAHI T, AMBROSIN M, CONTI M, et al. ABAKA: a novel attribute-based k-anonymous collaborative solution for LBSs[J]. Computer Communications, 2016, 85(2016): 1-13.

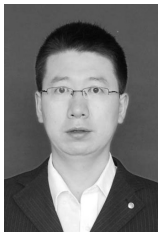
[23] 汤全有, 马传贵, 光焱. 基于全同态加密的秘密数据比较方案[J]. 信息工程大学学报, 2012, 13(6): 654-657.
TANG Q Y, MA C G, GUANG Y. Comparing private numbers based on fully homomorphic encryption[J]. Journal of Information Engineering University, 2012, 13 (6): 654-663.

[24] WANG X F, MU Y, CHEN R M. One-round privacy-preserving meeting location determination for smartphone applications[J]. IEEE Transactions on Information Forensics and Security, 2016, 11(8): 1723-1732.

[25] 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.

[26] 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.

[作者简介]



王斌（1979−），男，黑龙江安达人，哈尔滨工程大学博士生，佳木斯大学副教授，主要研究方向为机器学习、隐私保护。

张磊（1982−），男，黑龙江绥化人，博士，佳木斯大学副教授，主要研究方向为信息安全、隐私保护。

张国印（1962−），男，博士，山东黄县人，哈尔滨工程大学教授、博士生导师，主要研究方向为嵌入式系统与体系结构、网络技术与信息安全。