第48卷第6期 中国科学技术大学学报 Vol.48,No.6
2018年6月 JOURNAL OF UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA Jun. 2018

文章编号:0253-2778(2018)06-0512-07

k-匿名机制下查询隐私的一种度量方法

陈家明,王 丽,肖亚飞,方贤进

(安徽理工大学计算机科学与工程学院,安徽淮南 232001)

摘要:在k-匿名机制下,提出一种以信息熵和对数函数为基础的查询隐私度量方法.首先,建立k-匿名机制下的查询隐私的度量模型,包含4种角色和4种操作,为隐私保护的度量提供形式化的描述.然后,介绍两种背景知识的量化方式.针对第二种方式,由于用户属性离散化后的数值会被计算入背景知识概率表达中,造成背景知识概率表达的不准确,为此提出将离散化后的用户属性值作为特定查询和用户属性相关量的下标来求得相关量,从而进一步得到用户发出此特定查询的概率,摆脱了用户属性离散化后的数值对量化结果的影响.最后,提出查询隐私的度量方法.实验结果表明,该隐私度量方法能够较为有效地度量k-匿名机制下查询隐私算法的保护水平.

关键词:位置服务;k-匿名;隐私度量;背景知识

中图分类号:TP391 **文献标识码**:A doi:10.3969/j.issn.0253-2778.2018.06.010

引用格式:陈家明,王丽,肖亚飞,等.k-匿名机制下查询隐私的一种度量方法[J].中国科学技术大学学报,2018,48(6):512-518.

CHEN Jiaming, WANG Li, XIAO Yafei, et al. A privacy preserving measurement for query under k-anonymity mechanism[J]. Journal of University of Science and Technology of China, 2018,48(6):512-518.

A privacy preserving measurement for query under k-anonymity mechanism

CHEN Jiaming, WANG Li, XIAO Yafei, FANG Xianjing

(*College of Computer Science and Engineering, Anhui University of Science and Technology, Huainan* 232001, *China*)

Abstract: A query privacy measurement was proposed under the k-anonymity mechanism. The method was based on information entropy and logarithmic function. First, a framework for query privacy under the k-anonymity mechanism is established, which contains four roles and four operations provides a formal description for privacy measurement. Then, two quantitative methods of background knowledge are introduced. For the second step, user attribute discretization values will be calculated as a probability expression of background knowledge, affects the accuracy of the probability expression. The value of each user attribute after discretization was proposed as the index of the array to calculate the relevant quantities, the index of the array being generated by the relevance of the particular query and the attributes of the user, so as to further obtain the probability of the user issuing the particular query, thus avoiding the influence of discretized values of user attributes on the quantification results. Finally, a query privacy measurement is proposed. The experimental results show that the method can effectively measure the level

收稿日期:2017-09-05;**修回日期**:2018-04-10

基金项目:国家自然科学基金项目(61572034)资助.

作者简介:陈家明,男,1996年生,硕士生.研究方向:信息安全.E-mail:1732827657@qq.com

通讯作者:王丽,博士/讲师.E-mail:liwang@aust.edu.cn

of protection of the query privacy protection algorithm under k-anonymity mechanism.

Key words: location-based services; k-anonymity; privacy measure; background knowledge

0　引言

近年来，基于位置的服务（location based service，LBS）越来越多地融入人们的生活.在导航、交通调度、兴趣点查找、救援任务中，广泛地使用LBS.瑞典市场研究公司 Berg Insight 发布的报告预测，全球 LBS 市场规模将以 22.5%的复合年增长率（CAGR）从 2014 年的 103 亿欧元，增加至 2020 年的 348 亿欧元[1].LBS 在给人们的生活带来巨大方便的同时，也会带来隐私信息的泄露.如查询者所在位置及查询请求可能关联到家庭住址、兴趣习惯、健康状况、政治立场等隐私信息.

针对以上问题，研究者提出一些隐私保护方法，如假名技术、加密技术、位置 k-匿名、位置模糊等[2-5].这些方法能否保护个人隐私，需要隐私度量的结果来判断.隐私度量的结果既可以为研究者在设计隐私保护算法时提供参考，又可以让用户了解其隐私保护水平.现有的隐私保护的度量机制大都忽略了攻击者的背景知识和推理能力，直接影响到度量结果的可用性，也会造成过高评价隐私保护水平的问题.同时，LBS 隐私保护的研究主要集中在构造隐私保护算法上，对隐私保护机制的评估缺乏度量手段[6].本文在信息熵度量[7]的基础上，以 GLKA 算法[8]（Casper 算法[9]的改进算法）为例，提出一种融入攻击者背景知识的查询隐私度量方法.

1　相关工作

隐私保护方法的度量可以从服务质量度量和隐私信息度量两个方面来考虑，隐私信息的度量又包含轨迹隐私的度量和查询隐私的度量.现有的文献，大多是对隐私信息度量的研究[10].

关于服务质量度量，Shoki 等[11]提到服务质量的丢失与用户的真实位置及其扰动位置的距离、保护机制产生用户扰动位置的方法、攻击者关于用户历史位置记录的背景知识有关.Theodorakopoulos 等[12]提出不同形式的服务质量度量方法.

关于隐私信息度量，k-匿名是最常用的度量指标.k 是匿名集的大小，k 越大，查询的隐私性就越高[13].林欣等[14]指出，在连续查询的情况下，匿名集中各用户发送查询请求的概率不再相等，攻击者将概率最大的用户作为查询的发出者，此时匿名集的大小就不能反映用户的真正匿名性.基于信息熵的度量方式也常被用在隐私信息的度量上.Huang 等[15]用信息熵度量通信系统中用户的匿名性，即用来度量攻击者识别系统中信息的发出者或接收者的隐私水平.Hoh 等[16]将信息熵的度量方式用在轨迹跟踪的不确定程度上，随机变量表现为每个位置实例包含在当前跟踪车辆轨迹的概率.Chen 等[17]说明了在攻击者有背景知识和无背景知识的情况下，可以用互信息来度量系统的隐私水平.张学军等[6]建立 β 熵匿名和 δ 互信息匿名等指标来度量查询隐私.其中 β 熵匿名是指攻击者判断当前系统的信息熵不大于 β，即判断某查询请求是匿名区域内某用户发送出来的系统信息熵不大于 β.δ 互信息匿名可以度量两个随机变量的相关性，如用来度量泛化查询被攻击者截取后查询不确定性的减少程度.彭长根等[7]提出几种信息熵模型，随后引入隐私信息熵、平均互信息量、条件熵及条件互信息等来描述隐私保护系统信息源的隐私度量、隐私泄露度量、含背景知识的隐私度量及泄露度量.另外，基于错误的度量方式也被用在隐私信息的度量上，即将攻击者计算出的用户位置或用户信息出错的概率作为度量隐私信息的指标.Freudiger 等[18]用错误概率法衡量混合区域中的轨迹隐私水平.Zhang 等[19]提出基于贝叶斯条件概率的隐私度量模型，并采用扭曲度法度量隐私保护算法的隐私水平.

GLKA 算法[8]是史敏仪等在 Casper 算法[9]的基础上提出的一种 k-匿名机制下的隐私保护算法.在 Casper 算法中，用户的隐私需求包括匿名度 k 和匿名区域最小面积 A.Casper 算法基于四叉树模糊算法，递归地将空间分成 4 个象限空间，然后根据用户的隐私需求将用户准确位置泛化成匿名区域，当不能满足隐私需求时，先把匿名区域扩大到相邻象限，若仍不能满足要求，再将区域扩大到父类象限中，直到满足隐私需求为止.这种算法在合并单元格时，没有充分考虑目标用户相邻区域中用户的分布，使形成的匿名区域存在空间冗余.GLKA 算法根据目标用户相邻区域的用户分布情况来合并网格区域，确保目标用户与相邻用户的混淆，减小冗余空间.

本文以 GLKA 算法为例，在信息熵的基础上，提出一种融入攻击者背景知识的查询隐私度量方

法.攻击者的背景知识表现为用户发出某特定查询的概率.将攻击者背景知识融入度量方法后,度量结果更加可信.

2 查询隐私度量方法

2.1 查询隐私度量模型

为了更好地说明查询隐私度量方法,本文建立了一种 k-匿名保护机制下查询隐私的度量模型,该模型考虑攻击者具有背景知识,能够更好地贴近实际.具体的模型如图 1 所示.

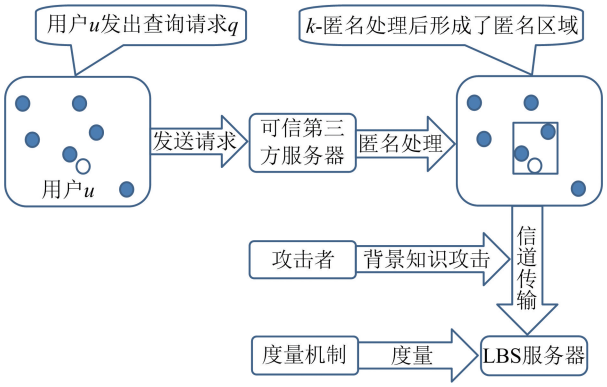

图 1 k-匿名保护机制下的查询隐私的度量模型

Fig.1 A privacy preserving measure model for query under k-anonymity mechanism

在某一区域中,用户 u 在 l 位置,t 时刻发起查询请求 q,组成初始查询元组 $Q = <u,l,q,t>$.经过可信任匿名服务器的匿名机制处理后,Q 形成泛化查询 $\hat{Q} = <r,\hat{q},\hat{t}>$,用户 u 的唯一标识符被删除,l 经过位置 k-匿名机制处理后,被包含 k' 个用户的区域 r 所代替,用户集合为 U,$\hat{q}$ 可以是经过扭曲处理的 q,而 $\hat{t}$ 可以是 t 时刻转化的一个时间段.在没有任何背景知识的情况下,攻击者识别 q 的发出者的概率等于 $1/k'$.假设攻击者已经从信道上获取了 $\hat{Q}$ 以及匿名区域 r 中所有用户的背景知识,那么攻击者就可以根据每个用户的背景知识推测出每个用户发送出查询请求 $\hat{q}$ 的概率.概率最大的用户即攻击者认为是发出查询请求 $\hat{q}$ 的用户.

模型中的 4 种角色:

(Ⅰ)发出请求的用户.移动用户在一定区域中,当其有需求时便会发出查询请求,即请求 LBS 服务器提供服务.

(Ⅱ)可信的第 3 方服务器.在用户和 LBS 之间的服务器.用来收集用户的身份、位置、查询请求、提交查询请求的时刻等信息,并将这些信息泛化成满足用户隐私需求的匿名信息.

(Ⅲ)LBS 服务器.LBS 提供商提供的服务器,以响应用户的各种查询请求.

(Ⅳ)攻击者.带有善意或恶意目的,试图从用户查询请求中获得更多信息的人.

模型中的 4 种操作:

(Ⅰ)发送查询请求.用户将查询请求发送给 LBS 服务器,以获得服务.在此之前需要经过可信的第 3 方匿名服务器进行匿名化处理.

(Ⅱ)匿名机制处理.用户的位置信息经过 k-匿名机制处理,身份、查询请求、提交查询请求的时刻等信息可通过其他方式匿名化.

(Ⅲ)LBS 服务器查询.LBS 服务器根据用户的查询请求,运行算法,进行有关查询运算,并将结果反馈用户.

(Ⅳ)背景知识攻击.攻击者从传输的信道上获取匿名处理后的信息,利用自身的背景知识推测查询请求的发出者.

2.2 查询隐私的度量方法

本文考虑了攻击者具有背景知识情况后,在信息熵的基础上,提出 k-匿名机制下查询隐私度量方法.现有的度量机制大都忽略了攻击者具有背景知识的情况,而这些背景知识可以帮助攻击者减小其推测用户隐私信息的不确定性,因此将攻击者的背景知识融入度量中是十分必要和必须的.同时,信息熵是隐私信息度量的有效工具,在隐私信息的度量上有着广泛应用,度量效果也较可靠.

由于用户属性离散化后的数值会被计算到背景知识概率表达中,造成背景知识概率表达的不准确,为此我们做出改进,即首先将离散化后的用户属性值作为特定查询和用户属性相关量的下标来求得相关量,从而进一步得到用户发出此特定查询的概率;然后提出 k-匿名机制下查询隐私的具体度量方法.

2.2.1 两种背景知识的量化方式

在现实中攻击者通常都会拥有一定的背景知识,因此将背景知识融入度量中更能贴近实际.将背景知识融入度量方法中又是一项非常具有挑战性的工作,一方面,是由于背景知识比较复杂,如关于数据本身的认知可以分为:医疗知识(如统计出某个地区患某种病的概率很高)、地图信息等;关于个人的背景知识,由于个体的不同又会有很多差异.另一方面,由于攻击者的不同,其素质、知识水平、攻击能力都会有很大的不同.通常情况下,我们不是预测攻击者具有怎样的背景知识,而是假设在怎样的背景知识情况下,隐私具有怎样的泄露风险.

k-匿名机制下，背景知识需要表达成条件概率的形式，从而将其应用到攻击者推测 q 的真正发出者中.文献[20]提出用 Me method 方法来量化背景知识，主要思想是首先将背景知识表达为正关联规则($A \Rightarrow B$)和负关联规则(如 $A \Rightarrow \neg B$、$\neg A \Rightarrow \neg B$ 等)，再把这种关联规则表达为条件概率，而条件概率的大小则表示关联规则的强度，如 $P(B \mid A)=0.5$，就是正关联规则 $A \Rightarrow B$ 且规则强度为 0.5 的表示.之后分别将两类规则按照其强度排序，并从中选出最强的 K_+ 个正关联规则和最强的 K_- 个负关联规则，使用(K_+，K_-)最强关联规则来量化背景知识.王彩梅等[21]指出，(K_+，K_-)最强关联规则量化背景知识，仅使用了关联规则的数量来衡量背景知识，忽略了关联规则的强弱，因此她们将不同关联规则按照强弱分成 n 类，用不同区间关联规则的数量来描述所假设的攻击者拥有的背景知识的强弱，概率大小总区间为[0,1].这样既考虑到关联规则的数量，又考虑到其概率分布.

文献[17]提出了一种量化背景知识的方法.给定一个查询，其通常与用户的某些属性相关联.例如，化妆品的查询可能与用户的性别、年龄、收入等属性相关，而且这个查询与“女性”的相关性要比与“男性”的相关性大，与“20＜年龄≤50”的相关性比与“年龄≤20 岁”和“年龄＞50 岁”的相关性要大.这时对于一个给定的查询 q，我们可以给这个 q 分配一个与用户的 n 个属性相关的相关量 $W(q)=(w_1, w_2, \cdots, w_n)$，$w_i$ 为查询 q 和用户属性 a_i 具体值的相关性；对于某个用户 u 的表示，我们首先需要将用户 u 的属性离散化，每个属性都有有限个不同的值，每个属性用这个属性中不同值个数的二进制位表示，当属性值对应某个二进制位时，就将这个二进制位改为 1，其余都用 0 表示.如性别属性中有“男”和“女”两个不同的值，我们可以用两个二进制位男|女表示，当用户为女时，对应的“女”位修改为“1”，“男”位还是“0”，此时用户性别属性用“01”表示.如果有用户 u 的属性为“女性”且“50 岁≥年龄＞20 岁”，那么有：$\Phi_u=<01,010>$，Φ_u 是用户 u 离散化后的表示.发出查询的用户 u 和特定查询 q 的相关值为 $V(u,q)$，可以表示为

$$V(u,q)=\sum w_i \cdot \phi_u[i].$$

用户 u 形成的匿名区域中的用户集合为 U，u 发出查询请求 q 的概率 $p(u|q)$，可以表示为

$$p(u \mid q)=\frac{V(u,q)}{\sum_{u' \in U} V(u',q)}(u \in U).$$

在某些情况下，求得的概率可能不太准确.比如“离我最近的餐馆”的查询中，由于“男”是用“10”表示，而“女”是用“01”表示，通过这样的方式计算出来的概率，男性用户会比女性用户的概率高，但是男性用户和女性用户在其他属性相同时，此查询的概率应该一样.本文做出改进，主要思想如下：

(Ⅰ)属性相关量 $W(q_j,a_i)$ 的表达

用户的每个属性都有一个或多个不同的值，对于一个给定的查询请求 q_j，它与用户属性 a_i 的不同值有不同的相关性，根据相关性的不同，给每个 a_i 的不同值都赋上一个表示与 q_j 的相关值，将 q_j 与 a_i 每个不同值的相关值表达成一个数组 $W(q_j,a_i)$，即查询请求 q_j 与属性 a_i 的相关量.具体的 $W(q_j,a_i)$ 赋值描述如表 1 所示.

表 1　查询请求与用户属性相关量赋值描述表

Tab.1　Correlative evaluation description of query request and user attribute

查询请求	用户属性相关量						
	a_1		…	a_2		…	…
q_1	$W(q_1,a_1)[1]$	$W(q_1,a_1)[2]$	…	$W(q_1,a_2)[1]$	$W(q_1,a_2)[2]$	…	…
q_2	$W(q_2,a_1)[1]$	$W(q_2,a_1)[2]$	…	$W(q_2,a_2)[1]$	$W(q_2,a_2)[2]$	…	…
…	…	…	…	…	…	…	…

表 1 中，$W(q_1,a_1)[1]$ 和 $W(q_1,a_1)[2]$ 表示查询请求 q_1 与第一个属性 a_1 前两个不同值的相关值.$W(q_j,u)$ 即查询请求 q_j 与具体用户 u 的相关量.

(Ⅱ)背景知识的概率表达

我们首先将具体用户的属性 a_i 离散化，然后将离散化后的数值作为数组 $W(q_j,a_i)$ 的下标，找出 q_j 与属性 a_i 的相关值.表 1 中，a_1 如果表示“性别”属性，$W(q_1,a_1)$ 表示查询请求 q_1 与“性别”的相关量，“女性”离散化后的值可以是“01”，在 $W(q_1,a_1)$ 中找到下标是“01”的值即 q_1 与 a_1 为“女性”的相关值.按此方式，找出 q_1 和具体用户 u 的所有属性值的相关值得到 q_1 与 u 的相关量 $W(q_1,u)$，将 $W(q_1,u)$

中所有值求和得到 q_1 与 u 的相关值 $V(q_1,u)$.在包含 k' 用户的匿名区域中找出 q_1 与每个用户的相关值 $V(q_1,u_i)$,并求和:$\sum_{i\leqslant k'}V(q_1,u_i)$,再将 $V(q_1,u)$ 与此值相比,得到的值作为匿名区域中用户 u 发出查询请求 q_1 的概率值 $p(u|q_1)$,表达式如下:

$$p(u\mid q_1)=\frac{V(q_1,u)}{\sum_{i\leqslant k'}V(q_1,u_i)}.$$

这里得到的是用户 u 发出查询请求 q_1 的概率值,同理,我们可以按照此方式得到 u 与 q_j 的概率值.我们将用户属性离散化后的值作为下标得到用户和查询请求的相关量,再求和得到两者的相关值,而不是将其带入相关值得计算中,排除了离散化后的值对计算相关值的影响,从而使得到的概率更准确.

2.2.2 查询隐私度量方法

Shannon 等提出的信息论是通信学科的理论基础,用于确定一个通信系统中消息的真正发出者.后来,有学者将信息熵应用于隐私信息的度量.本文提出的位置 k-匿名机制下的查询隐私度量方法也是基于信息熵的.

信息熵可以用来描述一个系统的混乱程度,熵值越大,系统越混乱.用户 u 发出查询请求 q_j 后,经过位置 k-匿名机制下的保护算法形成的匿名区域中包含了 k' 个用户,其集合为 U.

u 发出查询请求 q_j 的系统信息熵可通过如下公式描述:

$$H(u)=-\sum_{u\in U}p(u\mid q_j)\log_2 p(u\mid q_j)$$

如果攻击者没有任何背景知识,他在匿名区域中识别出 q_j 的真正发出者的概率为匿名区域中所有用户数量的倒数,即等于 $1/k'$,此时,系统的信息熵最大,即

$$H_{\max}(u)=-\sum_{u\in U}\frac{1}{k'}\log_2\left(\frac{1}{k'}\right)=\log_2 k'.$$

由于对数函数具有单调性的特点,所以我们用这个性质作为基础,来改进文献[21]中轨迹隐私度量方法,即分别在分子(当前信息熵)和分母(当前系统最大熵)前加上对数(以‘2’为底).若用 PL(u)来表示,则有

$$\mathrm{PL}(u)=\left(\frac{\log_2 H(u)}{\log_2 H_{\max}(u)}\right)\times 100\%.$$

式中,PL(u)的值反映了当 k 值大于 2、给定某一查询 q_j 的情况下,用户 u 在当前 k-匿名算法下的隐私保护强度,PL(u)越大意味着当前 k-匿名保护算法的隐私保护水平越高;反之,则越低.

3 实验分析与验证

实验以 k-匿名机制下的 GLKA 算法为例,进行有效性度量.实验机器的主要参数为 Core i3 CPU、4G 运行内存,在 Windows 7 系统下采用 Matlab 仿真.采用 Thomas Brinkhoff 路网生成器模拟用户在市区网络中的位置信息,将德国奥登伯格市区图作为输入(图 2),生成了市区内 10 000 个随机用户,生成参数采用默认值.由于本文关注的是保护机制下保护水平的度量问题,我们随机生成背景知识的概率表达,并从产生的 10 000 个用户中随机选取 100 个用户作为查询用户,将地图划分为 32、64、128、256、512、1 024 大小的网格,利用 GLKA 算法分别计算不同隐私需求(即 k 值)情况下的数据,通过本文提出的度量方法来度量 100 个查询用户的隐私保护水平,并取平均值作为输出结果.

图 2 奥登伯格市区图

Fig.2 Oldenburg city map

3.1 有无背景知识对隐私保护水平的影响

我们首先将地图划分成不同小大的网格后,通过 GLKA 算法生成 6-匿名的区域,分别统计出在有无背景知识的情况下,匿名区域中所包含的用户个数,然后生成用户的背景知识概率表达;最后计算出两种情况下的隐私保护水平.如图 3 所示,横坐标表示地图的网格划分大小,纵坐标是通过本文提出的隐私度量方法度量的隐私保护水平.L_1 是有背景知识的情况下绘制的折线.在不同大小网格的情况下,隐私保护水平有所不同.当攻击者具有背景知识时,匿名区域中的每个用户以不相等的概率被认为是查询请求的发出者,攻击者认为概率最大的用户是查询请求的发出者.隐私信息遭到了泄露,隐私保护水平也有所不同;地图网格增大,生成的匿名区域也变

大，区域中的用户数量也会增多，所以得到的隐私保护程度也随之升高.L_2 是在没有背景知识的情况下绘制的折线.从图 3 可以看出，这条折线一直保持着“100%”的隐私保护水平.在没有背景知识的情况下，攻击者认为匿名区域中每个用户发出查询的可能性相等.如果我们假设这种情况下用户的查询隐私得到了完全的保护，那么得到的结果也应该是一直保持着隐私保护水平的最大值.

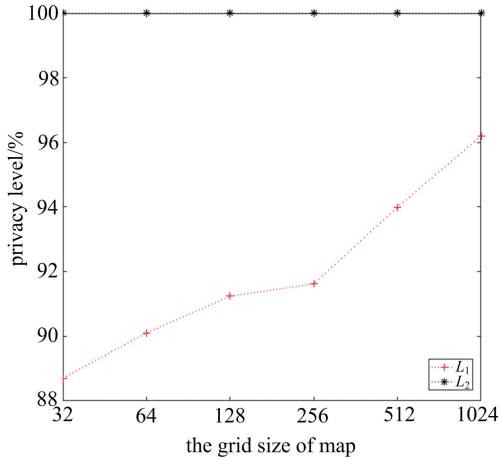

图 3　背景知识对隐私保护水平的影响

Fig.3　Influence of background knowledge on the level of privacy protection

3.2　不同隐私需求和不同地图网格大小对隐私保护水平的影响

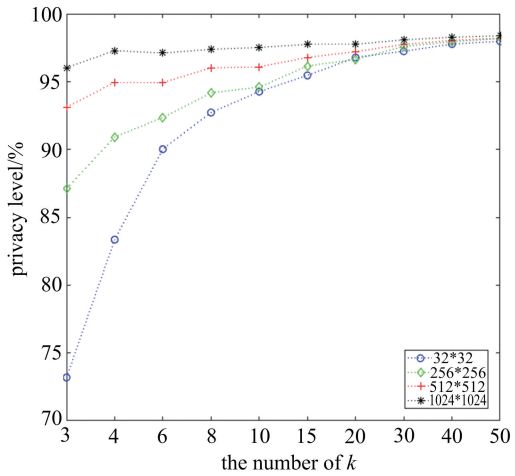

图 4　隐私需求对隐私保护水平的影响

Fig.4　Influence of privacy demand on the level of privacy protection

我们将地图划分成不同小大的网格后，在每种不同大小的网格下，利用 GLKA 算法得到隐私需求为 3、4、6、8、10、15、20、30、40、50 情况下的数据，再计算出不同隐私需求情况下的隐私保护水平.如图 4 所示，横坐标表示隐私需求 k，纵坐标表示隐私水平.同一网格大小的情况下，随着隐私需求的不断增大，隐私保护水平也随之升高.隐私需求增大，匿名区域中的用户变多，系统的不稳定性升高，隐私保护水平也逐渐升高，所以在网格大小一定的情况下，隐私需求越大，隐私保护水平也越高.图 5 中，横坐标表示地图网格大小，纵坐标表示隐私水平.从图 5 可看出，当隐私需求不变时，随着网格增大，隐私保护水平也随之升高.网格越大，即使 k 值不变，通过 GLKA 算法得到的匿名区域也会越大，匿名区域中的用户数量可能会越多，用户的查询隐私信息就可能得到更好的保护，所以在隐私需求一定的情况下，网格大小越大，隐私保护水平就越高.

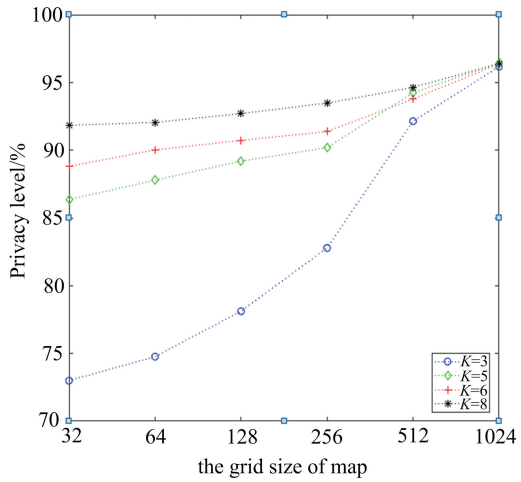

图 5　网格大小对隐私保护水平的影响

Fig.5　Influence of grid size on the level of privacy protection

4　结论

本文提出一种 k-匿名机制下查询隐私算法的度量方法，该方法考虑攻击者具有匿名区域中所有用户的背景知识，相比已有方法，度量结果更接近实际情况，为研究者设计隐私保护算法提供参考.实验以 GLKA 算法为例，证明了本文的查询隐私度量方法的有效性.下一步我们将着重背景知识的量化工作，查询请求的多样性问题以及用户属性和查询请求相关性赋值问题仍然还值得进一步研究.

参考文献(References)

[1] 崔宁宁，杨晓春，王斌，等. 移动 k-支配最近邻查询验证研究[J/OL]. 计算机学报，2017，40：113[2017-08-05]. http://kns.cnki.net/kcms/detail/11.1826.TP.20170728.1258.032.html.
CUI Ningning，YANG Xiaochun，WANG Bin，et al. Research on authentication of moving k-dominant NN

queries[J/OL]. Chinese Journal of Computers, 2017, 40: 113 [2017-08-05]. http://kns.cnki.net/kcms/detail/11.1826.TP.20170728.1258.032.html.

[2] ARTAIL H, ABBANI N. A pseudonym management system to achieve anonymity in vehicular Ad Hoc networks[J]. IEEE Transactions on Dependable and Secure Computing, 2016, 13(1):106-119.

[3] CHEN D, LI H, ZHOU S. CSEP: Circular shifting encryption protocols for location privacy protection [C]// IEEE/ACIS International Conference on Computer and Information Science. Piscataway, NY, USA: IEEE Press, 2017:45-50.

[4] BASERI Y, HAFID A, CHERKAOUI S. K-anonymous location-based fine-grained access control for mobile cloud[C]// Consumer Communications and NETWORKING Conference. Piscataway, NY, USA: IEEE Press, 2016:720-725.

[5] 万盛, 李凤华, 牛犇, 等. 位置隐私保护技术研究进展[J]. 通信学报, 2016, 37(12):124-141.
WAN Sheng, LI Fenghua, NIU Ben, et al. Research progress on location privacy-preserving techniques[J]. Journal on Communications, 2016, 37(12):124-141.

[6] 张学军, 桂小林, 冯志超, 等. 位置服务中的查询隐私度量框架研究[J]. 西安交通大学学报, 2014, 48(2): 8-13.
ZHANG Xuejun, GUI Xiaolin, FENG Zhichao, et al. A quantifying framework of query privacy in location-based service[J]. Journal of Xi'an Jiaotong University, 2014, 48(2):8-13.

[7] 彭长根, 丁红发, 朱义杰, 等. 隐私保护的信息熵模型及其度量方法[J]. 软件学报, 2016, 27(8):1891-1903.
PENG Changgen, DING Hongfa, ZHU Yijie, et al. Information Entropy Models and Privacy Metrics Methods for Privacy Protection [J]. Journal of Software, 2016, 27(8):1891-1903.

[8] 史敏仪. 面向位置服务的轨迹隐私保护技术研究[D]. 南京:南京邮电大学, 2014.

[9] MOKBEL M F, CHOW C Y, AREF W G. The new casper: A privacy-aware location-based database server [C]// IEEE International Conference on Data Engineering. Turkey: IEEE Computer Society, 2007:1499-1500.

[10] 王玲玲, 马春光, 刘国柱. 基于位置服务的隐私保护机制度量研究综述[J]. 计算机应用研究, 2017, 34(3): 647-652.
WANG Lingling, MA Chunguang, LIU Guozhu. Survey on metrics for location-based privacy protection mechanisms[J]. Application Research of Computers, 2017, 34(3):647-652.

[11] SHOKRI R, THEODORAKOPOULOS G, TRONCOSO C, et al. Protecting location privacy: Optimal strategy against localization attacks [C]// Proceedings of the 2012 ACM Conference on Computer and Communications Security. New York, NY, USA: ACM, 2012:617-627.

[12] THEODORAKOPOULOS G, SHOKRI R, TRONCOSO C, et al. Prolonging the hide-and-seek game: Optimal trajectory privacy for location-based services[C]// Proceedings of the 13th Workshop on Privacy in the Electronic Society. New York, NY, USA: ACM, 2014:73-82.

[13] KELLY D J, RAINES R A, GRIMAILA M R, et al. A survey of state-of-the-art in anonymity metrics[C]// ACM Workshop on Network Data Anonymization. New York, NY, USA: ACM, 2008:31-40.

[14] 林欣, 李善平, 杨朝晖. 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.

[15] HUANG L P, YAMANE H, MATSUURA K. Silent cascade: Enhancing location privacy without communication Qos degradation [C]// International Conference on Security in Pervasive Computing. Berlin: Springer-Verlag, 2006: 165-180.

[16] HOH B, GRUTESER M, XIONG H, et al. Preserving privacy in GPS traces via uncertainty-aware path cloaking[C]// ACM Conference on Computer and Communications Security. New York, NY, USA: ACM, 2007: 161-171

[17] CHEN X, PANG J. Measuring query privacy in location-based services[C]// Proceedings of the second ACM conference on Data and Application Security and Privacy. New York, NY, USA: ACM, 2012: 49-60.

[18] FREUDIGER J, SHOKRI R, HUBAUX J P. On the optimal placement of mix zones[C]//Proceedings of the 9th International Symposium on Privacy Enhancing Technologies. Berlin: Springer-Verlag, 2009: 216-234.

[19] ZHANG X J, GUI X L, TIAN F. Privacy quantification model based on the bayes conditional risk in location-based services[J]. TsingHua Science and Technology, 2014, 19(5): 452-462.

[20] DU W, TENG Z, ZHU Z. Privacy-maxENT: Integrating background knowledge in privacy quantification[C]// Proceedings of the 2008 ACM SIGMOD international conference on Management of data. New York, NY, USA: ACM, 2008:459-472.

[21] 王彩梅, 郭亚军, 郭艳华. 位置服务中用户轨迹的隐私度量[J]. 软件学报, 2012, 23(2):352-360.
WANG Caimei, GUO Yajun, GUO Yanhua. Privacy metric for user's trajectory in location-based services [J]. Journal of Software, 2012, 23(2):352-360.