路网环境下敏感位置匿名区域的生成方法

戴佳筑 华 亮
(上海大学计算机工程与科学学院 上海 200444)

摘 要 用户的位置信息涉及个人隐私,用户精确的位置信息可能会暴露其爱好、行为等敏感信息,因此,位置信息的匿名显得非常重要。现有的位置隐私保护方法大多是在欧氏空间下基于k-匿名算法生成位置匿名区域。欧氏空间下的k-匿名算法虽然可以在一定程度上解决用户位置信息的匿名保护问题,但是在现实生活中,用户的位置受路网环境影响较大,同时,欧氏空间下的k-匿名算法在生成匿名区域后对该区域是否还处于敏感范围内未做考虑。因此提出了一种路网环境下敏感位置匿名区域的生成方法。该方法基于空间划分,首先按照路网L-差异性要求对路网交叉点生成维诺图单元;接着考虑用户所处位置的敏感度,对用户位置生成匿名区域。实验结果表明,与一般的k-匿名算法生成的匿名区域相比,提出的算法能较好地解决一般k-匿名算法生成的匿名区域仍然处于敏感范围内的问题,从而更好地保护用户的位置隐私。

关键词 维诺图,匿名区域,路网结构,位置敏感度,位置隐私
中图法分类号 TP393 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2016.3.027

Method of Anonymous Area Generation for Sensitive Location Protection under Road Networks

DAI Jia-zhu HUA Liang
(Department of Computer Engineering and Science, Shanghai University, Shanghai 200444, China)

Abstract Location information contains personal privacy, and the precise location information may disclose users' hobby, behavior and other sensitive information. Therefore, the protection of location information privacy is very important. The existing methods of generating anonymous area for location privacy protection are mostly based on the k-anonymous algorithms under Euclidean space. Though they can anonymize user's location in a certain extent, in real life, the location is influenced by the road networks environment seriously, at the same time, the k-anonymous algorithms under Euclidean space do not consider whether the generated anonymous area is still within sensitive region. This paper proposed a method of generating anonymous area for sensitive location protection under road networks. The method is based on space partition. Firstly, it generates the Voronoi cell for the node of road networks according to the demands of road networks L-diversity. Then combined with the sensitivity of the user's location, it generates the anonymous area for user's location. The experiment results show that compared with the existing k-anonymous algorithms, this algorithm achieves the goal that the generated anonymous area is not within sensitive region, therefore, it provides better protection for location privacy.

Keywords Voronoi, Anonymous area, Road networks structure, Location sensitivity, Location privacy

1 引言

近些年来,普适计算研究得到长足发展,基于位置感知的移动社交网络应用也得到了很大的发展。移动社交网络应用将用户的位置信息同实际应用结合在了一起,如微信、微博等,都为用户提供了位置服务。通过位置服务,用户可以进行位置签到或服务评价。此外,近年来出现的打车软件,正是利用了移动社交网络应用的位置服务功能,用户通过分享自己的位置给服务器,服务器广播用户的位置给周围的出租车司机,司机获取到周围乘客的位置后,根据距离远近,选择要服务的乘客。正因为位置服务带来的便利和乐趣,使得其用户数量数以亿计,并且还在快速增长中。

然而,由于位置信息的敏感性,攻击者可以通过伪基站截获、窃听位置信息传输通道等方式访问位置数据,并由此推理获取到与位置信息相关的用户个人隐私信息,从而导致用户行为模式、兴趣爱好、健康状况等个人隐私信息的泄露,使得用户的隐私安全受到严重的威胁。因此,采取一些方法对用户的精确位置信息进行匿名化处理,使其真实位置信息不被其他用户直接获取,是位置服务中亟待解决的问题。

传统的位置匿名大多采用欧氏空间下的k-匿名算法。用户发送给位置服务器的服务请求中,以一块空间区域表示其自身位置,该区域中不仅包含该用户本身,还包含其他至少

到稿日期:2015-01-06 返修日期:2015-03-17
戴佳筑(1970—),男,博士,副教授,主要研究方向为网络安全、云计算安全、信息系统取证与审计、物联网和应用密码学,E-mail:daijz@shu.edu.cn;
华 亮(1990—),男,硕士,主要研究方向为信息安全。

$k-1$ 个用户，即使攻击者截获此位置信息，也无法将该用户的位置与其他 $k-1$ 个用户的位置区别开来，这时待匿名用户满足位置 k-匿名，这 k 个用户共同组成匿名集，同时，匿名集用户所在的模糊区域称为位置匿名区域，以此位置匿名区域来完成对用户精确位置信息的保护。但是，传统的 k-匿名算法并没有考虑用户所处的环境。在位置 k-匿名中，当寻找到 k 个用户满足匿名集对用户数量的要求时，匿名区域即可根据这个匿名集生成，这样生成的匿名区域很可能仍然在敏感区域范围内。例如，用户在“医院”请求位置服务，按照 k-匿名算法，寻找到的其他 $k-1$ 个用户可能也处于“医院”，当对这 k 个用户所在的位置生成匿名区域时，该匿名区域可能还在医院范围内，对于用户来说，发送匿名区域给位置服务器请求服务时，自己“在医院”的隐私还是暴露了。

同时，在欧氏空间下，移动用户的位置签到可以是任意的，只要在空间范围内，其可以不受约束。而在现实生活中，移动用户的运动往往受到路网的限制，且路网的结构对于查询效果和通信效率都有着重要的影响。因此，将欧氏空间下位置隐私保护算法直接应用于真实路网时，可能因为用户的签到位置和移动方向等原因导致用户位置隐私的泄漏。

基于此，本文提出一种路网环境下敏感位置匿名区域的生成方法，在生成匿名区域的过程中考虑了用户所处的路网环境。本方法根据位置敏感度生成相应的用户位置匿名区域，使得生成的匿名区域能够较好地满足用户的隐私需求。通过相关实验验证，与传统的 k-匿名算法相比，本方法匿名效果较好，匿名成功率更高。

本文第 2 节简要回顾相关工作；第 3 节介绍提出的系统模型；第 4 节详细描述位置信息匿名算法；第 5 节是实验分析；最后总结全文。

2 相关工作

近年来，对于用户位置信息隐私保护的研究相对较多，并且取得了一些成果。目前的位置信息保护方法大多基于 k-匿名算法。k-匿名模型最早是由 Sweeney L 等在文献[2]中提出的，它要求发布的数据中存在一定数量（至少为 k）的不可区分的记录，使攻击者不能判别出隐私信息所属的具体个体，从而保护了个人隐私。一般来说，k 值越大则隐私保护效果越好，然而丢失的信息也越多。

Gruteser[3] 最早将 k-匿名技术应用到位置隐私保护中来，使在某个时间段内的某个位置区域中至少存在 k 个用户，攻击者无法通过 ID 相互区别这些用户，即便攻击者取得了某个具体的位置信息，也无法从 k 个用户中找到确定用户。该方法的位置匿名区域依据用户的隐私度 k 的值确定，若存在一个用户隐私度需求 k 较大，则整个系统最后的匿名区域很大。

对于位置匿名区域的生成方法，很多研究者也进行了深入的研究。文献[4]提出了一种圆形区域划分的 k-匿名位置隐私保护方法。借助位置 k-匿名模型，提出了圆形区域划分匿名方法。将整个区域划分为相切圆及相邻的 4 个相切圆的顶点组成的曲边菱形所形成的组合区域，当用户位置区域含有的用户数量不满足隐私保护安全系数要求时，利用区域扩充公式得到合适的匿名区域。文献[5]提出利用网格划分平面，通过不断翻倍扩增网格宽度寻找满足用户隐私需求的匿名区，最终完成对用户位置的匿名，同时算法在运行中能够与邻近节点分享计算所得中间结果，并对其进行缓存，可显著降低网络带宽的消耗，减少位置匿名耗时，同时能够避免匿名区中心攻击，抗查询采样攻击的能力也得到较大提升。文献[6]则证明了在匿名区域面积相等的情况下，查询结果的精确度由匿名区域周长的大小确定，这个结论可以指导我们在保证查询质量的情况下生成合适的位置匿名区域。

以上研究均基于欧氏空间，没有考虑实际路网环境。在现实生活中，用户的移动路线大都受限于道路网络。根据获取到的路段信息，攻击者利用路网背景知识可以很容易地将用户锁定于路网范围之内，并通过推断攻击进一步锁定用户所处路段。因此，以上研究成果虽然在理论上具有很好的效果，但是在路网环境下并不具有很好的可行性和应用性，位置隐私保护方法必须将路网结构考虑在内。

2006 年，Machanvajjhala A[8] 等人提出了“L-多样性”的概念，2009 年，Ting Wang 等[9]将 L-多样性应用到路网环境中，提出了“路段多样性”的概念，要求匿名框不仅要满足用户的位置 k-匿名要求，还必须包含至少 L 条路段，以增加恶意攻击者在路网环境下进行推断攻击的难度。文献[10]根据公路网络的结构特点，提出了隐匿环和隐匿树这两种子图结构，利用隐匿环和隐匿树模糊移动用户在公路网络中的位置信息，可以有效地保护位置隐私。文献[11]针对公路网络下用户分布不均可能导致的推断攻击，设计出一种面向路网限制的位置隐私保护算法。算法通过对公路网络的边权进行排序，并结合路段地理位置分布，进行隐匿边集的构造，以降低边权不均引起推断攻击的风险。文献[12]提出在路网环境下利用维诺图对位置进行隐私保护。它首先根据路网模型形成路网结构的维诺图，然后根据隐私模型和匿名算法生成匿名集和匿名区域。但是，这些方法都未考虑用户的位置敏感度，对于生成的匿名区域依然处于敏感区域这一会暴露用户隐私的弊端尚未提出很好的解决方案，因此需要进一步研究。

本文在以上文献基础上，提出一种路网环境下敏感位置匿名区域生成方法，对用户的位置隐私进行保护。本方法在兼顾服务质量的前提下，根据位置敏感度生成相应的用户位置匿名区域，使得生成的匿名区域能够较好地满足用户的隐私需求。

3 系统模型

3.1 威胁模型

如图 1 所示，假设一所医院位于图中三角形 V1V2V3 组成的区域内，待匿名用户（以方形表示）在此区域内签到后请求服务，为了避免自己“在医院”这个敏感信息被别人获取，用户需要请求位置匿名服务器对其精确位置信息进行匿名。匿名服务器收到位置匿名服务请求后，按照传统 k-匿名算法，选取周围其他用户（以三角形表示）构成匿名集后，将生成的匿名区域（以圆形表示）和匿名集一起发送给位置服务器请求位置服务。但是，传统 k-匿名算法仅仅满足了匿名集对用户数量的要求，对用户处于敏感位置时的匿名未做过多考虑，对生成的匿名区域仍然位于敏感范围内的情况也未做处理。在这种情况下，如果匿名服务器生成的匿名区域仍然位于医院这个范围内，虽然满足了匿名度要求，但是对于用户来说，其隐私还是泄露了。例如，图 1 中产生的匿名区域仍然位于三

角形 V1V2V3 所示的敏感范围内，还是泄露了用户“在医院”的隐私。

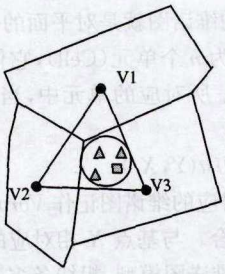

图 1　敏感位置区域示意图

同时，匿名集用户的隐私度也可能会威胁到用户的隐私安全。在选取匿名集用户的过程中，如果除了待匿名用户外，匿名集中其余用户的隐私度需求均大于待匿名用户的隐私度需求，则以待匿名用户的隐私度为匿名度生成的匿名区域虽然可以满足待匿名用户的隐私需求，但不能满足匿名集中其他用户的隐私需求，通过排除法，还是可能泄露用户的隐私。如图 1 所示，待匿名用户（方形）的隐私度需求为 $k=4$，因此在周围寻找到其他 $k-1$ 个用户共同生成匿名区域。此时，如果这 $k-1$ 个匿名集用户（三角形）的隐私度需求均大于待匿名用户的隐私度需求，则生成的匿名区域（圆形）无法完成对匿名集内用户的完全匿名，通过排除法，可以轻易地找到待匿名用户。

本文将研究解决以上威胁模型下的用户位置隐私保护问题。

3.2　位置匿名系统结构

本文采取中心服务器结构完成对位置信息的匿名。相关术语说明如表 1 所列，中心服务器结构主要由 3 部分组成：移动用户（User）、位置服务器（Location Server，LS）和可信第三方匿名服务器（Anonymous Server，AS）。移动用户只需要进行位置信息的生成和发布，不需要对位置信息做任何处理，减少了移动端的处理需求。本文中的位置匿名由 AS 来完成。引入 AS，是为了更好地减轻移动用户客户端设备的计算压力，AS 将收集移动用户的位置信息，根据匿名算法对用户的位置进行匿名，生成匿名区域，匿名结束后需要将匿名结果发送至 LS 请求服务，LS 将把查询结果返回给 AS，AS 再对 LS 返回的结果集进行求精，选出合适的结果返回给移动用户，如图 2 所示。

表 1　相关术语说明

缩写	描述
User	移动用户
AS	位置匿名服务器
LS	位置服务器
k	用户匿名度，匿名区域包含的最少用户数量
L	路网 L-多样性，即匿名区域中需要包含的最少路段数
S	路网交叉点的敏感度
λ	路段敏感度
ε	匿名区域隐私门限值
R_{min}	匿名集用户的最小查找半径
R_{max}	匿名集用户的最大查找半径
T	位置匿名的最大有效时间
Pro	用户隐私需求（k，R_{min}，R_{max}，T）
Q	位置服务请求，Q={ID，Loc，λ，Pro，q}，q 为要查询的位置
C	匿名用户集
K-Area	用户匿名区域
ASQ	Q 经过 AS 处理后的匿名查询请求，ASQ={C，K-Area，q}

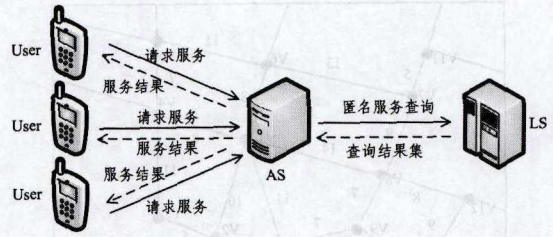

图 2　中心服务器结构

在发送位置信息到 AS 前，移动用户需要在 AS 注册一个特定的隐私协议，该隐私协议被定义为一个四元组 $Pro=(k, R_{min}, R_{max}, T)$。$k$ 表示用户的隐私度，即该用户位置的匿名区域中至少包含 k 个用户的位置信息。若用户处于密集区域，k-匿名方法生成的匿名区域将会很小，很小的空间区域虽然满足用户 k-匿名的需求，但在极端情况下，k 个用户的位置信息集中在某个很小的敏感范围内，会引起用户位置隐私的泄露。为了防止选取的匿名集用户过度集中，使得生成的匿名区域过小，我们设定参数 R_{min} 表示匿名集用户的最小查找半径。同时，在真实路网环境中，满足匿名需求的用户所在的位置可能较为分散，导致寻找到的匿名区域会很大，过大的匿名区域会严重影响位置请求的服务质量，给服务器负载造成影响，因此，设定参数 R_{max} 表示匿名集用户的最大查找半径。若查找范围超过 R_{max}，则认为此次匿名失效。设定 T 表示用户位置信息匿名的最大时间，如果匿名时间大于 T，则本次匿名失效。根据隐私协议和位置匿名算法，可由 AS 生成用户位置信息的匿名区域，然后将匿名区域发送到 LS 请求位置服务。

3.3　路网抽象模型

现实生活中，用户的移动路线大多受限于道路网络，真实的路网环境是由许多道路交叉后形成的，因此，考虑对路网结构进行抽象。

定义 1（抽象路网模型）　以无向图 $G=(V,E,S,e)$ 来表示路网模型，含义如下：

1）V 是图的顶点集合，表示路网交叉点的集合；

2）$E\subseteq V\times V$ 是图中非空边的集合，表示两个路网交叉点之间的路段；

3）S 是路网交叉点的敏感度，反映出本路段周围敏感位置点的密度；

4）两个路网交叉点之间的距离记为 e，称为路段权重。

图 3 是上海市杨浦区五角场区域某路网真实环境，图 4 是对此路网生成的抽象模型。在模型中，地图由各个区域组成，一个区域是一个连接子图，表示为 $G'=(V',E',S',e')$，其中 $V'\subseteq V, E'\subseteq E, S'\subseteq S, e'\subseteq e$。

图 3　真实路网环境

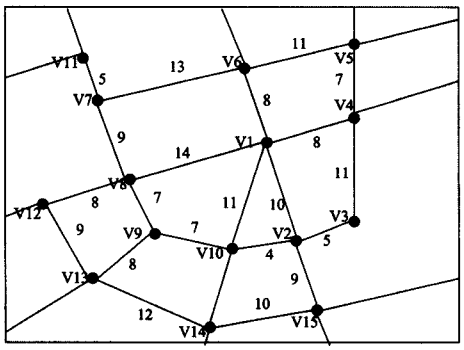

图 4　抽象路网模型

3.4　路网 L-差异性

图 5 中的加粗矩形部分为基于欧氏空间生成的用户位置匿名区域，由于中间包含多个用户位置，因此攻击者无法将原始用户的位置与该区域中的其他用户的位置区分，原始用户的位置信息受到了保护，如图中匿名区域中包含了 A、B 两个用户，降低了攻击者区分出真实用户的概率。图 6 考虑了路网环境，图中的直线代表路网中可以行走的路段，生成的位置匿名区域中包含与图 5 相同的用户 A 和用户 B，但是这两个用户所处的路段相同，一旦攻击者截获了位置匿名区域的信息，依然可以通过相关攻击手段推断出用户所处的某一具体路段位置，用户的位置隐私就会被泄漏。

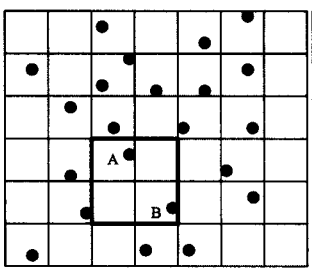

图 5　欧氏空间下的匿名区域

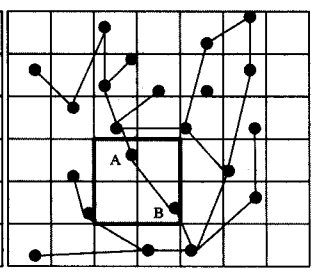

图 6　路网空间下的匿名区域

定义 2(路网 L-差异性)　路网 L-差异性是指在位置匿名区域生成过程中，为了防止生成的匿名区域中包含的路段太少而提高攻击者推断出位置隐私的成功率，要求生成的匿名区域中至少包含 L 条不同的路段[8]。

3.5　维诺图单元模型

一般的匿名区域生成方法大多数基于空间区域的固定网格划分和网格合并，若多个路网节点和用户聚集在一个网格或多个网格中，则 k-匿名算法在查找匿名集 C 时，需要进行频繁的距离计算和距离比较，花费的时间过多，造成位置服务查询质量的下降。维诺图单元是一个多边形区域，每一个多边形区域只对应一个点，我们将这个点称为“基点”。若某个要查询最近邻的用户被包含于某个多边形区域，则该用户的最近邻一定是这个多边形区域的基点。因此，通过维诺图的最近邻特性来寻找匿名集用户，会减少查找时间，进而减少匿名区域生成的时间。同时，为了满足路网 L-差异性的需求，在选取路网交叉点构造维诺图单元时，根据实际路网结构，我们为 L 设置取值范围，选择那些连接的路段数在 L 范围内的路网交叉点，从而保证生成的维诺图单元中的路段数符合要求，避免为了满足路段 L-差异性而频繁地进行区域扩展的操作，减少匿名时间。

定义 3(欧氏距离)　任意两点 $X(X_A,X_B)$ 和 $Y(Y_A,Y_B)$ 之间的欧氏距离，记作 $dist(X,Y)$。

$$dist(X,Y)=\sqrt{(X_A-X_B)^2+(Y_A-Y_B)^2} \tag{1}$$

定义 4(维诺图)　设 $X=\{X_1,\cdots,X_n\}$ 为平面上任意 n 个互异的点，X 对应的维诺图就是对平面的一个子区域划分，整个平面因此被划分为 n 个单元(Cell)，它们具有这样的性质：任一点 Y 位于点 X_i 所对应的单元中，当且仅当对于任何的 $X_j\in X, j\neq i$，都有

$$dist(Y,X_i)<dist(Y,X_j) \tag{2}$$

我们将与 X_i 对应的维诺图记作 $Voronoi(X_i)$。

X 为基点的集合。与基点 X_i 相对应的单元记作 X_i 的维诺图单元。图 7 是维诺图模型，粗线条多边形为点 X_i 对应的维诺图单元。

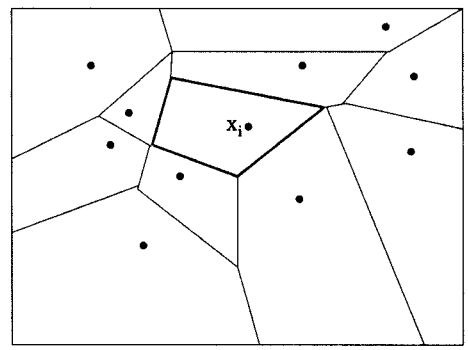

图 7　维诺图模型

3.6　路网交叉点的维诺图单元

路网中的每个交叉路口都可以抽象为一个点 V_i，它们共同构成一个离散点集合 $V_T=\{V_1,V_2,\cdots,V_n\}$，考虑到路网 L-差异性，根据实际路网环境，一般情况下，一个路网交叉点连接的路段在 2 条到 5 条之间，因此设定 L 的取值范围为 $2\leqslant L\leqslant 5$。根据 L 的具体取值，从集合 V_T 中筛选出连接的路段数大于等于 L 的离散点集合 $V=\{V_1,V_2,\cdots,V_i\}, i=1,2,\cdots,n, V\subseteq V_T$。将集合 V 中的每一个点与其周围的 n 个点做连线，对每一条连线做中垂线，这 n 条中垂线将会相交围成一个有 n 条边的多边形，这个多边形称为维诺图单元。由于生成的多边形只包含一个路网交叉点，我们也将此多边形称为路网交叉点 V_i 的维诺图单元。图 8 是对应图 4 中连接的路段数大于 2 的路网交叉点生成的维诺图单元。

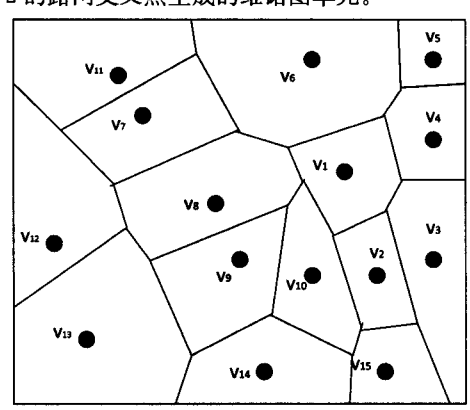

图 8　路网交叉点对应的维诺图

生成的路网节点的维诺图单元具有如下性质：

(1)各个路网节点的维诺图单元互不重合且相邻的维诺图单元共享一条边；

(2)各维诺图单元均满足路网 L-差异性；

(3)用户签到的位置均位于路网交叉点的某个维诺图单元内；

(4)每个维诺图单元内的点到本维诺图单元基点的距离小于到其他基点的距离。

4 敏感位置匿名区域的生成方法

本文提出一种路网环境下敏感位置匿名区域的生成算法来对用户的位置隐私进行保护。

4.1 相关定义

当用户向 LS 请求位置服务之前，需要对其当前所在的位置进行匿名，此工作由 AS 来完成。当用户进行位置签到后，其精确的位置坐标 Loc 会位于某路网交叉点所在的维诺图单元内，用户发送一个位置查询请求 $Q=\{ID,Loc,\lambda,Pro,q\}$ 给 AS，其中 Pro 为 AS 根据用户的需求生成的用户隐私协议，$\lambda(\lambda\in(0,1))$ 为路段敏感度，由 AS 根据路网状况赋予相关路段 λ 值。用户在某路段签到后，其位置敏感度就是路段敏感度。q 为用户要查询的位置。根据隐私协议中的隐私度 k，AS 将在用户所在位置的周围选取其他 $k-1$ 个请求位置服务的用户，组成匿名集 C，同时生成匿名区域 $K\text{-}Area$，完成位置信息的匿名。当 AS 完成位置匿名后，会将生成的匿名位置请求信息 $ASQ=\{C,K\text{-}Area,q\}$ 发送给 LS 请求服务。

对于待匿名用户本身来说，匿名集的选取过程中需要遵循以下规则。

定义 5(匿名集用户选取规则) 假设有用户组 $U=\{u_1,u_2,u_3,\cdots,u_n\}$，各个用户对应的隐私度的集合为 $K=\{k_{u1},k_{u2},k_{u3},\cdots,k_{un}\}$，如果位置服务请求者是 u_l，则设定匿名区域的匿名度为 k_{ul}。对于其他用户的隐私度集合 $K'=\{k_{u2},k_{u3},\cdots,k_{ui}\}$，$i=2,\cdots,n$，选择隐私度 $k_{ui}\leqslant k_{u1}$ 的用户作为匿名集用户。当满足条件的匿名集用户个数达到 $k_{u1}-1$ 时，匿名集用户选取结束。

在描述具体算法之前，需要说明几个参数。

定义 6(路段敏感度 λ) 路网环境中，AS 为路段赋予一个参数 λ，该参数称为路段敏感度。λ 值越大，表示用户位置信息的敏感度越高，也就需要进行更加严格的位置匿名。

定义 7(路网交叉点敏感度 S) 每个路网交叉点会连接若干条路段，对于每条路段 r，都会有路段敏感度 $r.\lambda$。将每个路网交叉点所连接路段的敏感度的平均值称为路网交叉点敏感度，用 S 来表示。

$$S=\frac{\sum_{i=1}^{n}r_i.\lambda}{n} \tag{3}$$

其中，n 代表路网交叉点连接的路段总数，S 的取值范围为 $(0,1)$。

定义 8(匿名区域隐私度 μ) 匿名区域的隐私度用来衡量生成的匿名区域是否能够满足用户的隐私需求，使该区域不包含在敏感区域范围内。对于路网环境下的用户匿名区域隐私度 μ，其与匿名区域覆盖的路网交叉点的个数和其敏感度 S 有关。μ 可形式化表示为：

$$\mu=\frac{\sum_{i=1}^{m}V_i.S}{m} \tag{4}$$

其中，V_i 为当前位置匿名区域覆盖的路网交叉点，m 为其个数的总和，$V_i.S$ 为各个路网交叉点的敏感度。

定义 9(敏感区域隐私门限 ε) 敏感区域隐私门限 ε 是使用户位置的匿名区域范围大于敏感区域范围的最小边界值。只有匿名区域隐私度满足 $\mu\leqslant\varepsilon$ 时，才能保证生成的匿名区域满足用户的位置隐私需求。对于某个敏感区域，根据其包含的路网交叉点，该区域的隐私门限 ε 可由式(4)求出。

4.2 敏感位置信息匿名区域的生成算法

敏感位置匿名区域的生成算法主要分为 5 个阶段。

(1)确定匿名区域的匿名度

为了满足待匿名用户的匿名需求，我们将待匿名用户的隐私度作为匿名区域的匿名度，假设该用户的隐私度为 k_u，则令匿名区域的匿名度 $K_A=k_u$，以此匿名度为标准来选取匿名集用户。

(2)选取匿名集成员

以待匿名用户所在的路网交叉点为基准，以 $R_{\min}$ 为最小查找半径，以路网交叉点的维诺图单元为初始查找范围，在当前路网交叉点周围进行广度优先搜索。当搜索到用户 u_i 时，根据匿名集用户选取规则，判断 u_i 的隐私度 k_{ui} 是否不大于匿名区域的匿名度 K_A，如果满足条件，则将用户 u_i 加入到匿名集中，并记录其所在的路段到 $Line$ 集合，记录此用户与待匿名用户之间的距离 $dist_{ui}$ 到集合 $dist$；否则对该用户不予处理，等待下一轮匿名。当匿名集内的用户数达到 K_A，且路段集合 $Line$ 中的元素个数达到 L 时，停止搜索。

之所以选择维诺图单元作为初始查找范围，是基于维诺图的基点与用户最邻近原则。根据最近邻原则，在选取匿名集用户时，不需要对空间中所有的用户进行逐个距离计算后再选取符合距离要求的用户作为匿名集用户，只需要根据匿名集用户的选取规则选取 K_A 个匿名集用户后，计算这 K_A 个匿名集用户与待匿名用户之间的距离以便在生成匿名区域时使用，这样可以减少服务器的计算量，进而减少匿名时间。

(3)匿名集用户的扩展搜索

若待匿名用户所在的维诺图单元内未能找到足够的匿名集用户或路网 L-差异性尚未满足，则根据路网权值 e，在 $R_{\max}$ 限制的查找半径内，选择离待匿名用户所在的维诺图单元最近的相邻维诺图单元，重复步骤(2)，搜索合适的匿名集用户，直至匿名集用户选取成功，且满足路网 L-差异性为止。

(4)初始匿名区域生成

以待匿名用户所在的位置为圆心，在 $dist$ 集合中找到离待匿名用户最远的匿名集用户 u_i，以 $dist_{ui}+\phi$(ϕ 是一个极小值)为半径，生成待匿名用户的初始匿名区域 C-K-Area，同时根据式(4)计算此时的匿名区域隐私度 μ，查看 $\mu\leqslant\varepsilon$ 是否成立，如果成立，将此初始匿名区域作为最终的位置匿名区域发送给 LS 请求位置服务，否则，转步骤(5)进行匿名区域扩展。

(5)匿名区域的扩展

若生成的初始匿名区域还在敏感范围内，则需要对初始匿名区域进行扩展。在 $R_{\max}$ 限制的查找半径内，选择不在当前初始匿名区域内且离待匿名用户距离最近的路网交叉点，以待匿名用户所在位置为圆心，以选取的路网交叉点和用户的距离为半径生成新的匿名区域，并计算新生成的匿名区域的隐私度 μ，查看 $\mu\leqslant\varepsilon$ 是否成立，如果成立，将此时的匿名区域作为最终的位置匿名区域发送给 LS 请求位置服务，否则，重复本步骤直至 $\mu\leqslant\varepsilon$。如果在 $R_{\max}$ 限制的查找半径内都未达到 $\mu\leqslant\varepsilon$，则认为匿名失败。

算法伪代码如下。

算法 1　敏感位置的匿名区域生成算法

输入：路网 Voronoi(V_i)，路段集合 Line，用户位置签到集合 U，待匿名用户 u，位置服务请求 Q={ID，Loc，λ，Pro，q}

输出：ASQ={C，K-Area，q}

1)Function K-Area(Q)
2)Send(Q) to AS；//发送位置服务请求到 AS 完成位置匿名
3)$K_A=k_u$；//初始化匿名区域的匿名度为待匿名用户的隐私度 k
4)C←u；n=1；//将用户 u 加入到匿名集 C 中，同时记录匿名集中用户的个数
5)Line←$line_u$；m=1；//记录用户所在的路段号，同时记录匿名区域中的路段数
6)While(R_{min}≤搜索半径≤R_{max})
7)If($loc_u\in$ Voronoi(V_i))
8)If(n!=K_A&&$loc_{ui}\in$ Voronoi(V_i)&&$k_{ui}\leq K_A$)//用户 u_i 在 Voronoi(V_i)中且隐私度 $k_{ui}\leq K_A$
9)　C←u_i；n++；
10)　Line←$line_{ui}$；
11)　If($line_{ui}\notin$ Line)
12)　　m++；//若新找到的匿名集用户所在的路段不在 Line 集合中，则路段数加 1
13)　　dist←$dist_{ui}$；//将选取的用户与待匿名用户之间的距离存入 dist
14)End if
15)End if
16)If (n!=K_A ‖ m!=L)
17)　Select(离 Voronoi(V_i)最近的相邻维诺图单元)；
18)Continue；
19)End if
20)End if
21)End while
22)If (n==K_A&&m==L)
23)　Select max(dist)；//选取与待匿名用户距离最远的用户
24)　$R_{C\text{-}k\text{-}Area}$=max(dist)+φ；
25)　C-K-Area=Circle(Loc_u，$R_{C\text{-}k\text{-}Are}$)；//生成初始匿名区域
26)End if
27)While(C-K-Area.μ≥ε)
28)　Select($V_j\notin$C-K-Area&&Min(dis(V_j，Loc_u)//选择不在当前初始匿名区域内且离待匿名用户距离最近的路网交叉点
29)C-K-Area=Circle(Loc_u，dis(V_j，Loc_u))
30)　If(C-K-Area.μ≤ε)
31)　　K-Area←C-K-Area；//匿名完成
32)Else
33)　Continue；
34)End if
35)End while
36)Send ASQ={C，K-Area，q} to LS；
37)End Function

5　实验结果与分析

5.1　系统开发环境

本文实验在 PC 机和手机上完成。采用的开发环境：位置匿名服务器为 HL-PC，CPU 为 Intel Core(TM) 2 Duo E7500 2.93GHz，内存空间为 4GB，硬盘空间是 500GB，安装的操作系统是 Microsoft Windows 7 Professional Service Pack 3 的 32 位系统，编程语言是 Java 语言。开发工具是 Eclipse 3.7.2，服务器数据库是 MySQL 6.5。移动终端开发平台是 Android SDK 2.3.3，终端数据库是 Sqlite3.65，系统开发环境是 JDK1.6.0，测试手机是 MI 2。

5.2　实验设计与参数设置

本节通过实验模拟分析匿名算法性能，将本文提出的敏感位置匿名区域生成方法(4.2 节)与传统的匿名区域生成方法做比较。实验采用 ThomasBrinkhoff 基于路网的移动对象生成器[18]，以城市 Oldenburg 的交通路网作为输入，生成城市路线图上的移动对象，目标对象在空间均匀分布。为了更好地验证位置匿名效果而不受其他因素影响，本实验在密集网络环境下进行，生成 5000 个注册移动用户进行验证。用户的隐私度 k 取值范围设置为[3，10]。根据实际路网环境，一般情况下，一个路网交叉点在连接的路段数为 2～5 条之间，因此设定 L 的取值范围为[2，5]。R_{min} 取值为 20m，R_{max} 取值为 2km，匿名时间 T 的上限设置为 1000ms，路段敏感度 λ 取值范围为(0，1)。实验对算法的匿名成功率和匿名时间两个指标进行评测。

5.3　实验结果分析

根据以上实验设计和参数设置，实验分为两部分：(1)检验算法的匿名成功率；(2)检验算法的匿名时间。

(1)匿名成功率

匿名成功率是指成功匿名的消息数在所有移动用户提出的匿名请求的消息数中所占的比例，本文中所指的匿名成功是指生成的用户位置匿名区域超过敏感区域范围，否则，即使在 R_{max} 范围内可以成功生成匿名区域，只要该区域还在敏感范围内，就认为是匿名失败。

图 9 显示了 k 值和匿名成功率之间的关系。如图 9 所示，在位置敏感度不变的情况下，随着 k 值的增大，传统 k-匿名算法和本文提出的敏感位置匿名区域生成算法的匿名成功率都会降低，这是因为 AS 需要寻找更多满足待匿名用户隐私度的用户来构成匿名集。与传统的 k-匿名策略相比，本文提出的算法考虑生成的匿名区域是否还在敏感范围内，如果在，需要进行区域扩展，使新的匿名区域范围大于敏感区域范围，因此匿名成功率更高，能更好地保护用户位置隐私。在位置敏感度不变的情况下，实验结果表明，随着 k 值的增大，本文所提算法的平均匿名成功率在 80%左右，高出传统 k-匿名算法近 10%。

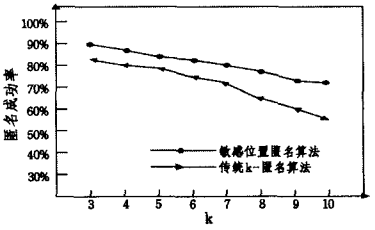

图 9　随着 k 值的变化两种位置匿名算法的匿名成功率比较

图 10 显示了路段敏感度 λ 值和匿名成功率之间的关系。如图 10 所示，在用户隐私度 k 值不变的情况下，随着路段敏感度的增大，对路段上的位置使用传统 k-匿名算法和敏感位置匿名算法的匿名成功率都会下降。其中，由于传统的 k-匿名算法在生成匿名区域时不考虑位置所在路段的敏感度，当用户在敏感度较低的路段进行位置匿名时，其成功率较高，与本文提出的匿名算法基本无差别。但是当用户处于敏感路段位置时，传统 k-匿名算法的匿名成功率较低，这是由于传统

k-匿名算法生成的匿名区域还在敏感范围内。

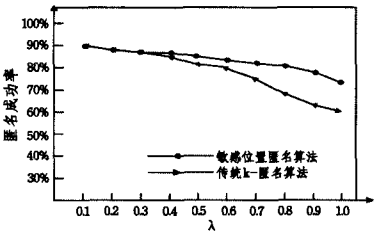

图 10 随着 λ 值的变化两种位置匿名算法的匿名成功率比较

同时，随着路段敏感度 λ 的增大，传统 k-匿名算法匿名成功率的下降幅度也比敏感位置匿名算法的匿名成功率下降幅度大。在用户隐私度 k 不变的情况下，实验结果表明，随着 λ 值的增大，本文所提算法的平均匿名成功率在 80%左右，传统 k-匿名算法的平均匿名成功率在 65%左右。

(2)匿名时间

匿名时间是指查询请求提出到匿名成功或匿名失败的时间。

图 11 显示了 k 值和匿名时间之间的关系。如图 11 所示，在位置敏感度不变的情况下，随着 k 值的增大，传统的 k-匿名算法和敏感位置的匿名区域生成算法的匿名时间都会相应增多，这是因为 AS 需要花时间来寻找匿名集用户。同时，与传统的 k-匿名算法相比，由于本文的算法考虑位置信息的敏感度，需要检查生成的匿名区域是否还处于敏感范围内，如果还处于敏感区域，需要进行匿名区域扩展，这就造成匿名时间多于传统 k-匿名算法的现象，但是匿名时间依然低于规定的匿名时间上限 T，与传统 k-匿名算法的匿名时间相差不多。

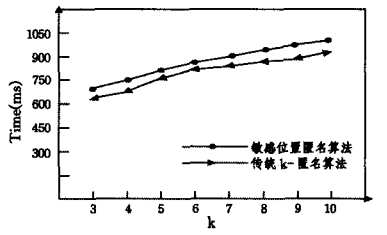

图 11 随着 k 值的变化两种位置匿名算法的匿名时间比较

图 12 显示了路段敏感度 λ 值和匿名时间之间的关系。如图 12 所示，在用户隐私度 k 值不变的情况下，随着路段敏感度 λ 的增大，对路段上的位置使用传统 k-匿名算法和敏感位置匿名算法的匿名时间都会增多。传统的 k-匿名算法只要匿名集用户的数量达到要求即停止匿名，不考虑生成的区域是否还在敏感范围内，因此当用户在敏感度较低的路段进行位置匿名时，所需的匿名时间与敏感位置匿名算法差别不大。但是当用户处于敏感路段的位置时，需要进行匿名区域扩展以保证用户的匿名区域不包含在敏感范围内，因此匿名时间会有所增加。

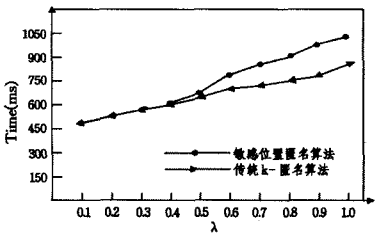

图 12 随着 λ 值的变化两种位置匿名算法的匿名时间比较

由以上实验可以看出，本文提出的算法能够在匿名时间上限 T 以内很好地提高匿名成功率，从而为用户的位置查询提供保障。

结束语 针对移动社交网络中用户位置隐私保护问题，本文提出一种敏感位置的匿名区域生成方法，能很好地提高敏感位置信息的匿名成功率。实验证明该方法具有很好的位置匿名效果。今后的研究将考虑如下 3 个方面：(1)本文依旧采取了第三方匿名服务器对用户的位置信息进行匿名，因此，服务器的可信度对匿名结果影响很大，未来考虑去掉第三方服务器的位置隐私保护方法；(2)在移动社交网络中更复杂的攻击模型(例如结合用户的社会关系等背景知识产生的隐私位置推理攻击)下的隐私保护也是值得深入研究的问题；(3)稀疏网络环境下的位置匿名会因为用户稀疏而失败，因此，如何提高稀疏网络环境下的位置匿名成功率也是未来需要考虑的问题。

参 考 文 献

[1] Huo Zheng, Meng Xiao-feng. A Survey of Trajectory Privacy-Preserving Techniques[J]. Chinese Journal of Computers, 2011, 34(10): 1820-1830(in Chinese)
霍峥，孟小峰. 轨迹隐私保护技术研究[J]. 计算机学报，2011，34(10)：1820-1830

[2] Sweeney L. K-anonymity: a model for protecting privacy[J]. International Journal on Uncertainty, Fuzziness and Knowledge based Systems, 2002, 10(5): 557-570

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

[4] Zhao Ze-mao, Hu Hui-dong, Zhang Fan, et al. A k-anonymous algorithm in location privacy protection based on circular zoning [J]. Journal of Beijing Jiaotong University, 2013, 37(5): 13-18 (in Chinese)
赵泽茂，胡慧东，张帆，等. 圆形区域划分的 k-匿名位置隐私保护方法[J]. 北京交通大学学报，2013，37(5)：13-18

[5] Wang Jia-hui, Cheng Jiu-jun. The location of the P2P mode based on grid amplification anonymous algorithm[J]. Computer Science, 2014, 41(4): 90-94(in Chinese)
王嘉慧，程久军. P2P 模式下基于网格扩增的位置匿名算法[J]. 计算机科学，2014，41(4)：90-94

[6] Xu J, Tang X, Hu H, et al. Privacy-conscious location based queries in mobile environments[J]. IEEE Transactions on Parallel and Distributed Systems, 2010, 21(3): 313-326

[7] Huang Yi, Huo Zheng, Meng Xiao-feng. CoPrivacy: A Collaborative Location Privacy-Preserving Method without Cloaking Region[J]. Chinese Journal of Computers, 2011, 34(10): 1976-198(in Chinese)
黄毅，霍峥，孟小峰. CoPrivacy：一种用户协作无匿名区域的位置隐私保护方法[J]. 计算机学报，2011，34(10)：1976-198

[8] Machanvajjhala A, Kifer D, Gehrke J, et al. L-diversity: Privacy beyond k-anonymity[J]. ACM Transactions on Konwledge Discovery from Data, 2007, 1(1): 3

[9] Wang Ting, Liu Ling. Privacy-aware mobile services over road networks[J]. VLDB Endowment, 2009, 2(1): 1042-1053

[10] Xue Jiao, Liu Xiang-yu, Yang Xiao-chun, et al. A Location Priva-

cy Preserving Approach on Road Network[J]. Chinese Journal of Computers,2011,34(5):865-878(in Chinese)
薛姣,刘向宇,杨晓春,等.一种面向公路网络的位置隐私保护方法[J].计算机学报,2011,34(5):865-878

[11] Sun Lan,Luo Zhao,Wu Ying-jie,et al. An Algorithm for protection location privacy in road network[J]. Journal of Shandong University(Engineering Science),2012,42(5):96-101(in Chinese)
孙岚,罗钊,吴英杰,等.面向路网限制的位置隐私保护算法[J].山东大学学报(工学版),2012,42(5):96-101

[12] Zhao Ping,Ma Chun-guang,Gao Xun-bing,et al. Protecting Location Privacy with Voronoi Diagram over Road Networks[J]. Computer Science,2013,40(7):116-120(in Chinese)
赵平,马春光,高训兵,等.路网环境下基于 Voronoi 图的位置隐私保护方法[J].计算机科学,2013,40(7):116-120

[13] Shun K G,Ju X,Chen Z,et al. Privacy protection for users of location-based services[J]. IEEE Communications Society,2012,19(1):30-39

[14] Chow C Y,Mokbel M F,Liu X. A peer-to-peer spatial cloaking algorithm for anonymous location-based service[C]// Proceedings of the 14th Annual ACM International Symposium on Advances in Geographic Information Systems. ACM,2006:171-178

[15] Hu H,Xu J. 2PASS:Bandwidth-optimized location cloaking for anonymous location-based services[J]. IEEE Transactions on Parallel and Distributed Systems,2010,21(10):1458-1472

[16] Wang Lu,Meng Xiao-feng. Location Privacy Preservation in Big Data Era:A Survey[J]. Journal of Software,2014,25(4):693-712(in Chinese)
王璐,孟小峰.位置大数据隐私保护研究综述[J].软件学报,2014,25(4):693-712

[17] Brinkhoff T. A framework for generating network based moving objects[J]. Geo Informatic,2002,6(2):153-180

[18] Sousa M, Techmer A, Steinhage A, et al. Human tracking and identification using a sensitive floor and wearable accelerometers [C]//2013 IEEE International Conference on Pervasive Computing and Communications (PerCom). IEEE,2013:166-171

[19] Gedik B, Liu L. Protecting location privacy with personalized k-anonymity:Architecture and algorithms[J]. IEEE Transactions on Mobile Computing,2008,7(1):1-18

[20] Xu J,Tang X,Hu H,et al. Privacy-conscious location-based queries in mobile environments[J]. IEEE Transactions on Parallel and Distributed Systems,2010,21(3):313-326

[21] Pan X,Xu J,Meng X. Protecting location privacy against location-dependent attacks in mobile services[J]. IEEE Transactions on Knowledge and Data Engineering,2012,24(8):1506-1519

[22] Zhu Qing,Zhao Tong,Wang Shan. Privacy Preservation Algorithm for Service-Oriented Information Search[J]. Chinese Journal of Computers,2010,33(8):1315-1323(in Chinese)
朱青,赵桐,王珊.面向查询服务的数据隐私保护算法[J].计算机学报,2010,33(8):1315-1323

(上接第 126 页)

结束语 本文提出了一种基于多样化历史信息的协商策略,解决了传统协商策略在重复协商时出现的协商效率低下的问题。根据资源的使用情况分别提出了历史票证和信任票证,根据两种票证的具体用途及其对协商安全性的影响程度,分别设计了票证的格式及验证方法,并将票证的工作过程与协商过程融合,提高了同一用户多次访问相关资源的协商效率。最后进行了实验仿真,结果表明,当协商双方围绕多个资源进行协商时,相对于传统的协商策略,基于多样化历史信息的协商策略在协商效率上有明显的提高;随着资源之间访问控制策略相似度的提升,基于多样化历史信息的协商策略在协商效率上的优势更加突出。当信任票证对证书的覆盖率达到100%时,双方只需一轮协商即可建立针对具体资源的信任。

参 考 文 献

[1] Winsborough H,Seamons K E,Jones V E. Automated trust negotiation[C]// DARPA Information Survivability Conference and Exposition. 2000:88-102

[2] Liao Zhen-song,Jin Hai,Li Chi-song,et al. Automated trust negotiation and its development trend [J]. Journal of Software,2006,17(9):1933-1948(in Chinese)
廖振松,金海,李赤松,等.自动信任协商及其发展趋势[J].软件学报,2006,17(9):1933-1948

[3] Shen Hai-bo,Hong Fan. Survey of research on access control model [J]. Application Research of Computers,2005,6(3):9-11(in Chinese)
沈海波,洪帆.访问控制模型研究综述[J].计算机应用研究,2005,6(3):9-11

[4] Yu T, Ma X, Winslett M. PRUNES: an efficient and complete strategy for automated trust negotiation over the Internet[C]// Proceedings of the 7th ACM Conference on Computer and Communications Security. ACM,2000:210-219

[5] Yao Hui,Gao Cheng-shi,Dai Qing,et al. Dynamic programming-based strategy for automated trust negotiation [J]. Computer Application,2009,28(4):892-895(in Chinese)
姚慧,高承实,戴青,等.一种基于动态规划的自动信任协商策略[J].计算机应用,2009,28(4):892-895

[6] Lu H,Liu B. DFANS:A highly efficient strategy for automated trust negotiation [J]. Computers & Security,2009,28(7):557-565

[7] Pikulkaew T, Kikuchi H. Improving Efficiency in Privacy-Preserving Automated Trust Negotiation with Conjunctive Policies [C]//2011 14th International Conference on Network-Based Information Systems (NBiS). IEEE,2011:679-684

[8] Yu D,Chen N. An automated trust negotiation model based on improved coloured timed Petri net [J]. International Journal of Sensor Networks,2014,16(2):61-69

[9] Squicciarini A,Bertino E,Ferrari E,et al. PP-trust-X:A system for privacy preserving trust negotiations [J]. ACM Transactions on Information and System Security (TISSEC),2007,10(3):12

[10] Liu B. Efficient trust negotiation based on trust evaluations and adaptive policies [J]. Journal of Computers,2011,6(2):240-245

[11] Li Jian-li,Gao Yong,Huo Guang-lei,et al. Reputation-based P2P trust system [J]. Computer Application,2011,31(1):147-150(in Chinese)
李健利,高勇,霍光磊,等.基于声誉的 P2P 信任系统[J].计算机应用,2011,31(1):147-150