位置服务（LBS）中的隐私保护

位 置 服 务（LBS）

位置服务

简言之，就是基于地理位置数据而展开的服务。是指通过移动终端和无线通讯网络（或卫星定位系统）的相互配合，以空间数据库为基础，确定移动用户的实际地理位置或坐标，并将其与其他信息整合起来，为用户提供所需要的与位置相关的增值服务。

LBS

LBS：Location Based Service

Services that intergrate a mobile device's location or position with other information so as to provide added value to a user

紧急救援服务

定位最近的救援者位置

对航空物流进行追踪

跟踪服务

GPS

RFID

SENSOR

查找到达目的地的最优路径

交通导航

信息娱乐服务

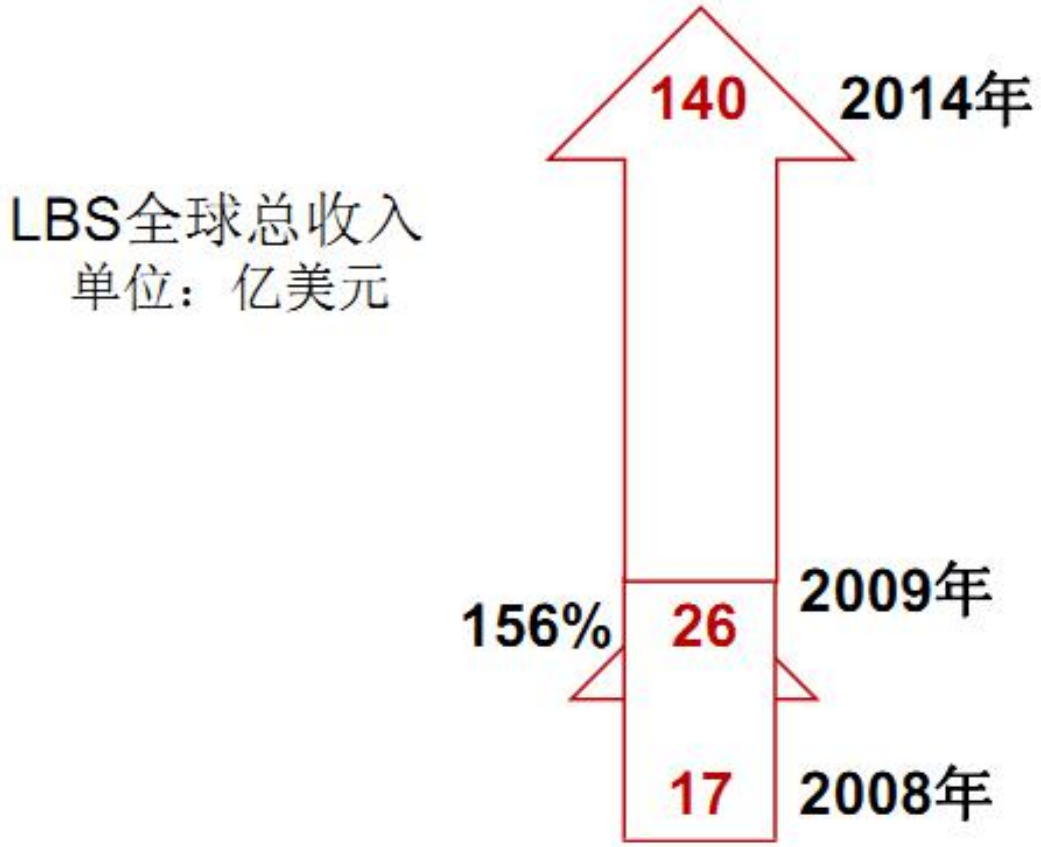
查找距离我百米内最近的银行

务 分发服

向我餐馆500米范围内的用户发送电子优惠券

位置服务前景广阔……

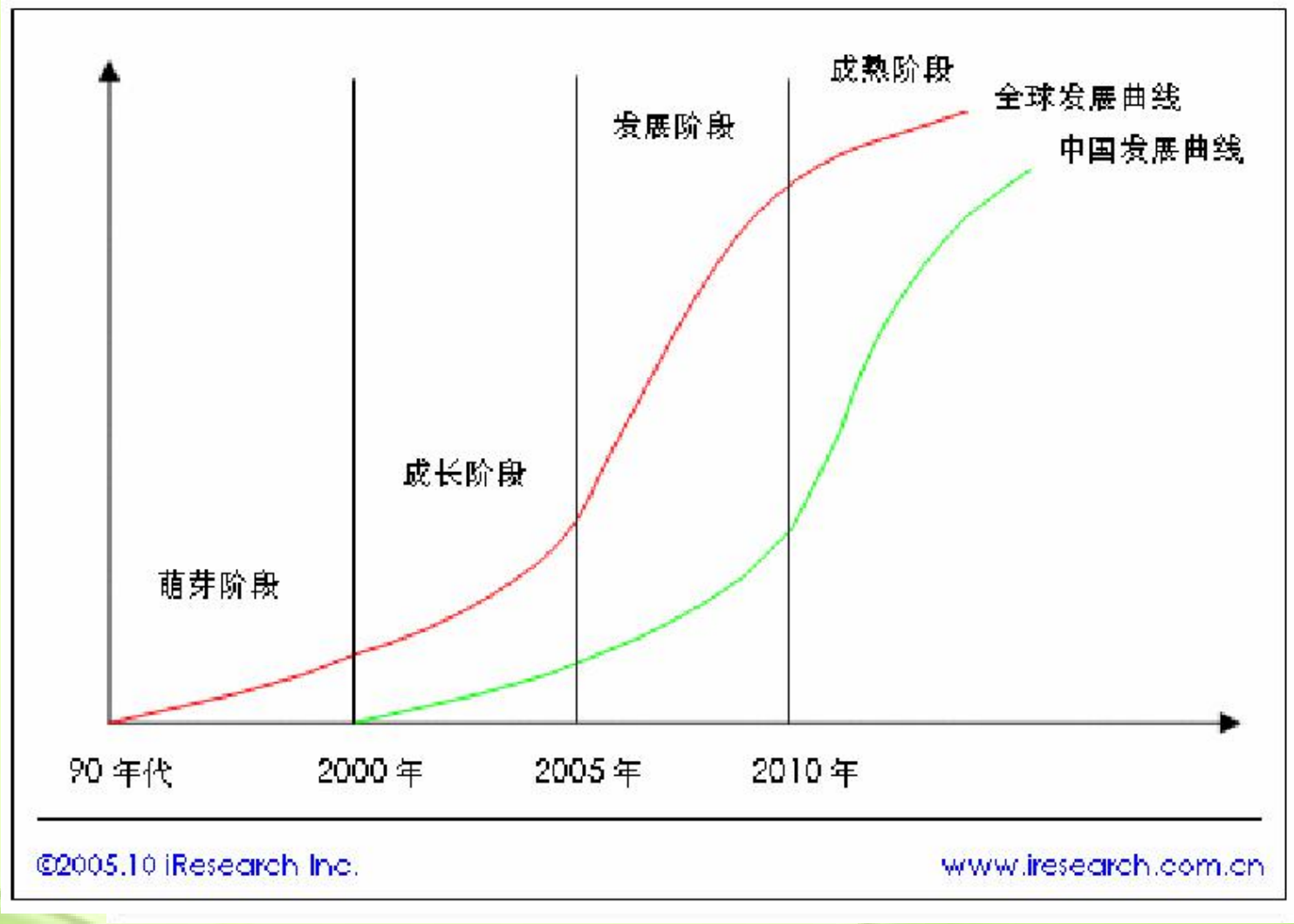
- 美国著名市场研究公司ABI research日前发布预测



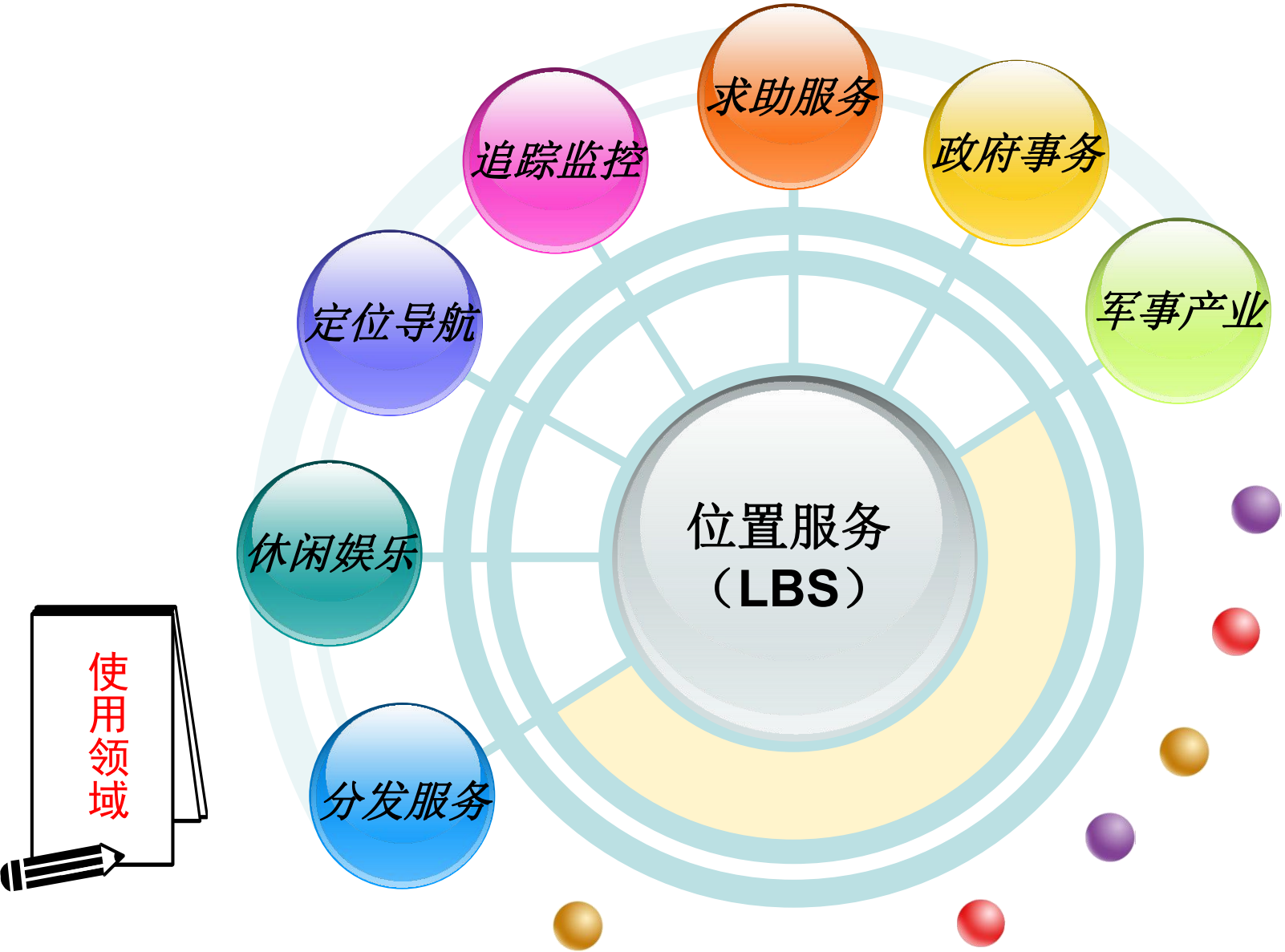
市场规模

位置服务前景广阔……

全球及国内移动位置服务市场发展阶段对比



位置服务前景广阔……



Form oracle

也许某天，当你经过熟悉的星巴克，打开手机记录你的位置，却惊喜地发现一位好友也正在这里享用咖啡。而在不远处的商场内，某件你关注了很久的名牌衬衫正在打折……

又或者你去一个遥远的城市旅行，打开手机，却发现自己瞬间对这个陌生的地方变得如此熟悉，想去哪儿就能随心所欲，想玩什么立马就能朝着目标出发……此时，你肯定会为位置服务所带来的便捷生活方式而兴奋！

隐私问题！！

可是你又是否设想过，未来，无论你走到哪里，上班，下班，旅行亦或是散步，都会有漫天飞舞的各种广告铺天盖地向你席卷而来，你无处躲藏，只能被淹没在信息的垃圾中烦恼叹息……

甚至有一天，对你而言，无论是熟悉的亲朋还是陌生人，都没有了秘密，每天无论你走到哪儿，干了什么，将要去哪，甚至你最近喜欢什么，想要计划什么，大家都能知道的时候，你会不会又对这一新兴的服务产生恐惧呢。。。

财经 股票 港股 基金 体育 NBA 彩票 奥运 科技 数码 手机 下载 时尚 购物 旅游 小说
中国平安 车险计算器
专享非事故道路救援 免费无限次
车辆所在省* 车辆所在市* 购车时间*
车牌号* 车价* 万元 姓名*

手机定位广告兴起或引发隐私保护问题

2010年06月14日14:01 腾讯科技 明轩 我要评论 (0)

腾讯科技讯（明轩）北京时间6月14日消息，据国外媒体报道，《华尔街日报》周日撰文指出，营销和软件公司正逐步使用智能手机和其它移动设备中的高端技术来追踪消费者，从而根据消费者所在的位置向他们提供广告。不过这种通过手机用户的地理位置来提供广告的业务，正在引起美国立法者和隐私保护组织的担忧。

最近的一份调查显示，许多使用定位科技的公司并不会明确告知消费者，他们正在收集哪类数据、如何使用这些数据及与哪些公司分享。受高速无线网络增长、智能手机日渐流行等因素的推动，在过去的一年当中，类似的营销活动获得了迅速的发展。当用户访问特定移动网络或打开应用时，这些地理位置公司就能够得知用户的位置。这些公司能够利用这些信息，向用户发送附近商店的相关广告。

内 容 导 览

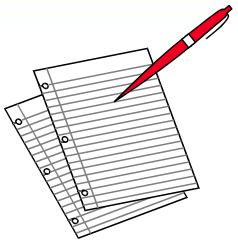
隐私问题举例

生活中这样的例子很常见：张某利用自己带有GPS的手机提出了“寻找距离我（现在所在位置）最近的肿瘤医院”的查询。这是导航系统中一个普通的最近邻查询（即NN查询）。此查询被提交给服务提供商，比如Google Maps，再由服务提供商通过查询空间数据库返回给用户查询结果。

此时涉及到的隐私问题包含两个方面：一是用户不想让人知道他此刻所在的位置，比如银行。二是用户不想让任何人知道自己提出了某方面的查询，如张某不想让人知道自己将去肿瘤医院。

LOGO

位置服务中的隐私泄露



位置隐私泄露

其他人以任何方式获知对象过去或现在的位置。

查询隐私泄露

查询请求的内容，可能透露用户的行为模式（如经常走的道路及经过的频率）、兴趣爱好（如喜欢去哪个商店、哪种俱乐部）、健康状况（如经常去的心理咨询诊所）以及政治倾向等个人隐私信息。

隐私泄露的三种途径

直接交流（Direct Communication）

攻击者从位置设备或者位置服务器中直接获取用户的位置信息。（如利用手机内置GPS定位导致的位置信息泄露）

观察（Observation）

攻击者通过观察被攻击者的行为直接获取位置信息。（如攻击者跟踪被攻击者）

连接泄露（Link Attack）

攻击者通过“位置”连接外部的数据源（或者背景知识）从而确定在该位置或者发送该消息的用户。（如查询信息A是从别墅B发出的，当攻击者知道别墅B的主人是C这一背景时，那么就能确定A是由C发出的）

LOGO

位置服务中的隐私保护

位置隐私保护

避免用户与某一精确位置匹配。。

查询隐私保护

避免用户与某一敏感查询匹配。

www.themegallery.com LOGO

隐私保护的工作

位置匿名

寻找一种状态，在这种状态下很多对象组成一个集合（匿名集），从集合的外面向里看，组成集合的各个对象无法被区分，使得某一个位置的信息无法与确切的个人/组织/机构相匹配。

查询处理

位置信息经过匿名处理后不再是用户的真实位置，可能是多个位置的集合或者一个模糊化的位置，查询系统应该对应地返回包含真实查询信息的模糊信息集

www.themegallery.com LOGO

位置服务VS隐私保护

不能兼得！！只能寻找平衡点！

鱼与熊掌可否兼得？

在最小的位置暴露下获得最佳的服务，使隐私暴露处于“可控状态”！

位置服务

提供精确位置

隐私保护

隐藏用户位置

隐私保护的特点

保护位置隐私与享受服务是一对矛盾

位置信息的多维性特点

位置匿名的即时性特点

基于位置匿名的查询处理

位置隐私需求个性化

www.themegallery.com

隐私保护的系统结构

独立结构

Non-cooperative Architecture

仅有客户端与位置数据库服务器的C/S结构。该结构需要移动用户拥有具有强大定位、计算、存储能力的设备，自己完成位置匿名和查询结果求精。

中心服务器结构

Centralized Architecture

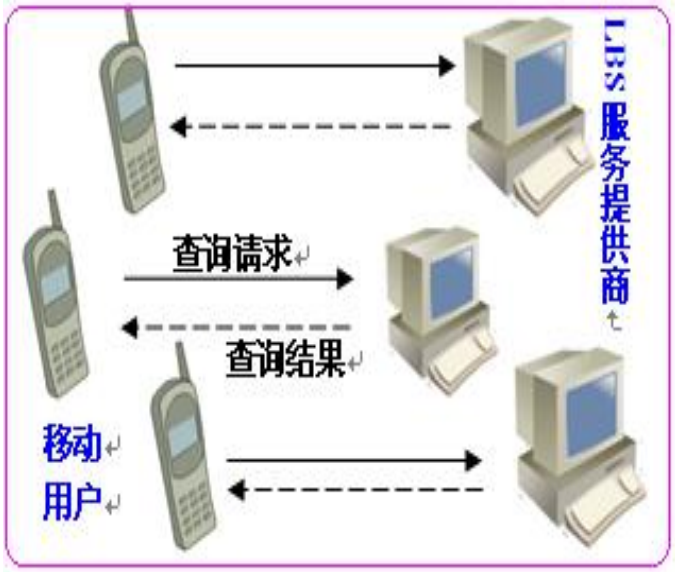
在用户和位置数据库服务器之间加入第三方可信中间件——位置匿名服务器，并由位置匿名服务器完成匿名处理和查询结果求精。

分布式点对点结构

Peer-to-peer Architecture

由客户端和位置数据库服务器组成，移动客户端根据匿名算法找到其他一些移动用户组成一个匿名组，以此进行位置匿名。同时组内的客户端之间的节点具有平等性，共同协调完成查询结果求精。

独立结构



独立结构示意图

处理流程：将匿名后的位置连带查询一起发送给位置数据库服务器；位置服务器根据匿名的位置进行查询处理并将候选结果集返回给用户；用户根据真实位置挑选出真正的结果。

优点：结构简单，易于配置

缺点：对客户端要求高；
缺乏全局信息，隐蔽性弱

中心服务器结构

具有全局信息，
隐私保护效果好。

优 点

位置匿名服务器成为系统的瓶颈；

位置匿名服务器还可以成为隐私泄露的唯一攻击点。

缺 点

处理过程：（1）用户发送包含精确位置的查询请求给位置匿名服务器。（2）匿名服务器使用某种匿名算法完成位置匿名后，将匿名后的请求发送给位置数据库服务器。（3）位置数据库服务器根据匿名区域进行查询处理，并将查询结果的候选集返回给位置匿名服务器。（4）位置匿名服务器从候选结果集中挑出真正的结果返回给移动用户。

分布式点对点结构

头结点平衡网络负载！！！

处理过程：（1）移动客户端根据匿名算法建立一个匿名组，进行位置匿名。（2）根据实际的网络通讯情况选择一个合适的头结点，头结点将匿名后的查询请求发送给位置数据库服务器。（3）位置数据库服务器根据匿名区域进行查询处理，并将查询结果的候选集返回给头结点。（4）头结点可以选择出真实结果发送给提出查询的用户，也可以将查询结果的候选集发送给用户，由用户自己挑选出真实的结果。

优点：拥有全局信息，隐私效果好；消除了系统瓶颈；消除了唯一攻击点。

缺点：网络通讯代价高。

隐私保护模型

位置K-匿名模型（Location K-Anonymity Model）

模型描述

把某一用户的真实位置点扩大为一个**模糊的位置范围**，使得该范围覆盖**k个**用户的位置，此时无法将该用户与其他（k-1）个用户相区别，称此位置满足位置k-匿名。

模型定义

将任意用户的位置以一个三元组 ($[x_1, x_2]$, $[y_1, y_2]$, $[t_1, t_2]$)表示，其中($[x_1, x_2]$, $[y_1, y_2]$)描述了对象所在的二维空间区域，$[t_1, t_2]$表示一个时间段。($[x_1, x_2]$, $[y_1, y_2]$, $[t_1, t_2]$)表示用户在这个时间段的某一个时间点出现在($[x_1, x_2]$, $[y_1, y_2]$)所表示的二维空间中的某一点。除此用户外，还有其他至少（k-1）个用户也在此时间段内的某个时间出现在($[x_1, x_2]$, $[y_1, y_2]$)的二维空间的某一点， 这样的用户集合满足位置k-匿名。

隐私保护需求的四个参数：

k：即 k-匿名，用户要求返回的匿名集中至少包含的用户数。

A_{min}：匿名空间的最小值，返回的匿名空间必须要超过此值，可以是面积或半径等。

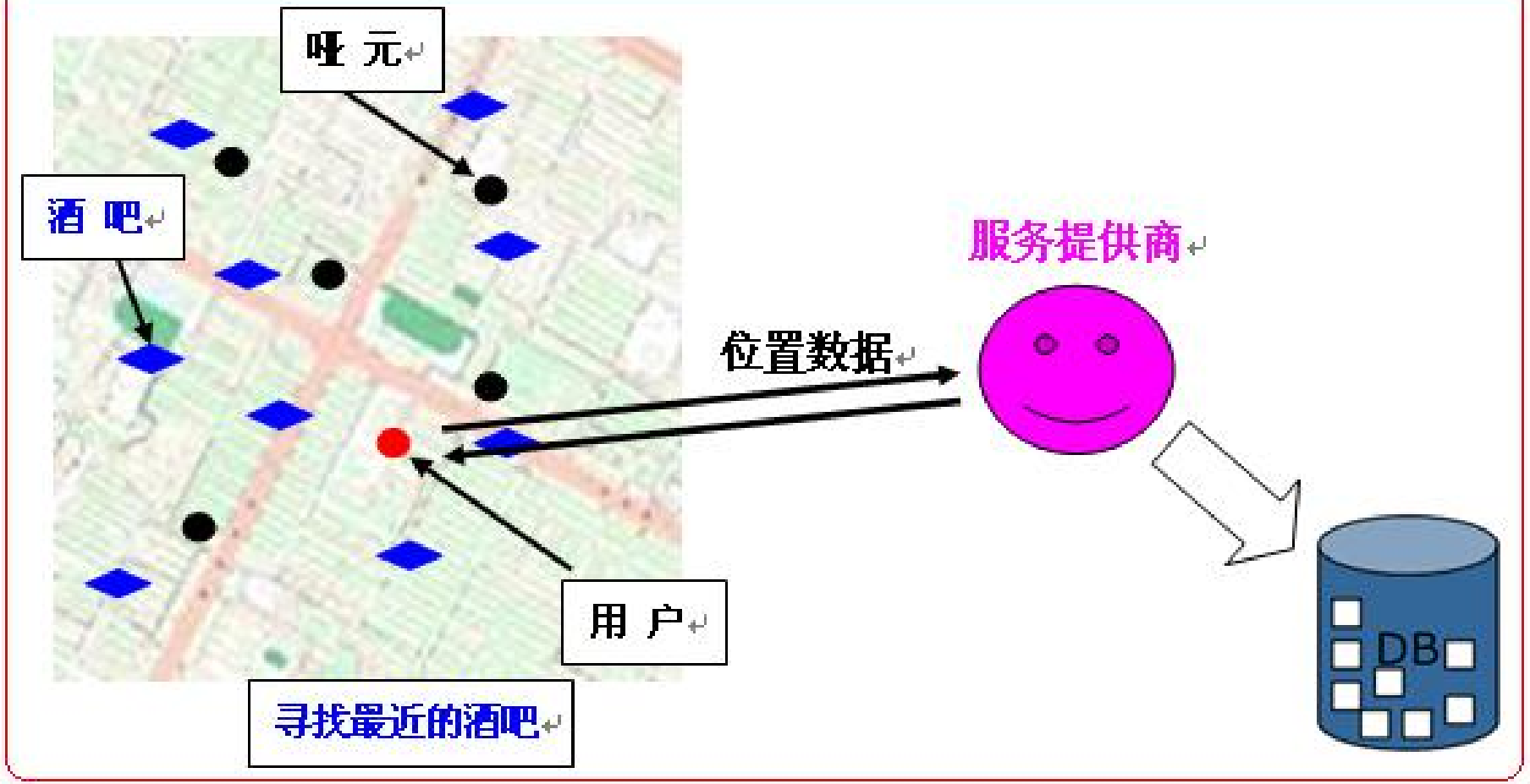
A_{max}：匿名空间的最大值，返回的匿名空间必须不能超过此值，也可以是面积或半径等。

T_{max}：可容忍的最长匿名延迟时间。即从用户提出请求的时刻起需要在T_{max}的时间范围内完成用户的匿名。

隐私保护技术


假位置
空间匿名
时空匿名
Dummy
Spatial Cloaking
Spatio-Temporal Cloaking

一、假位置

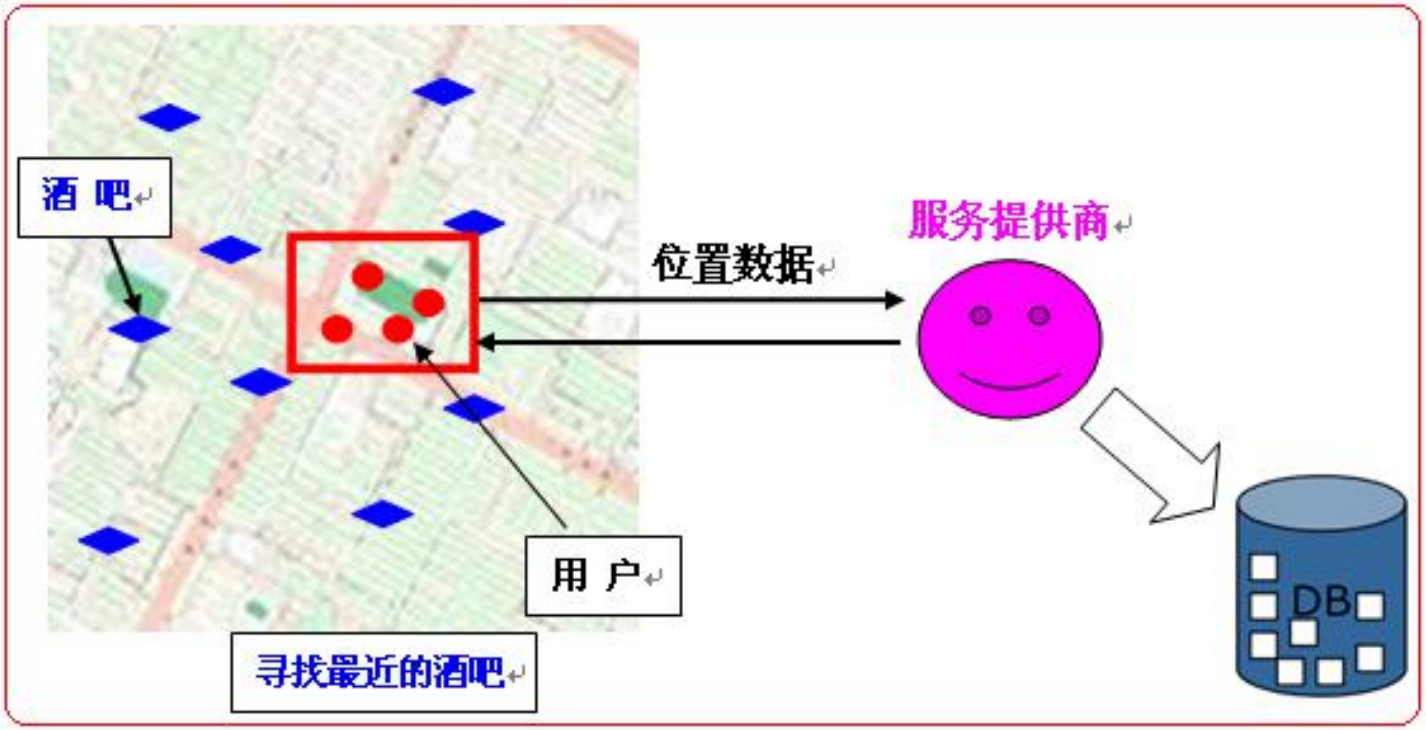


位置匿名

查询处理

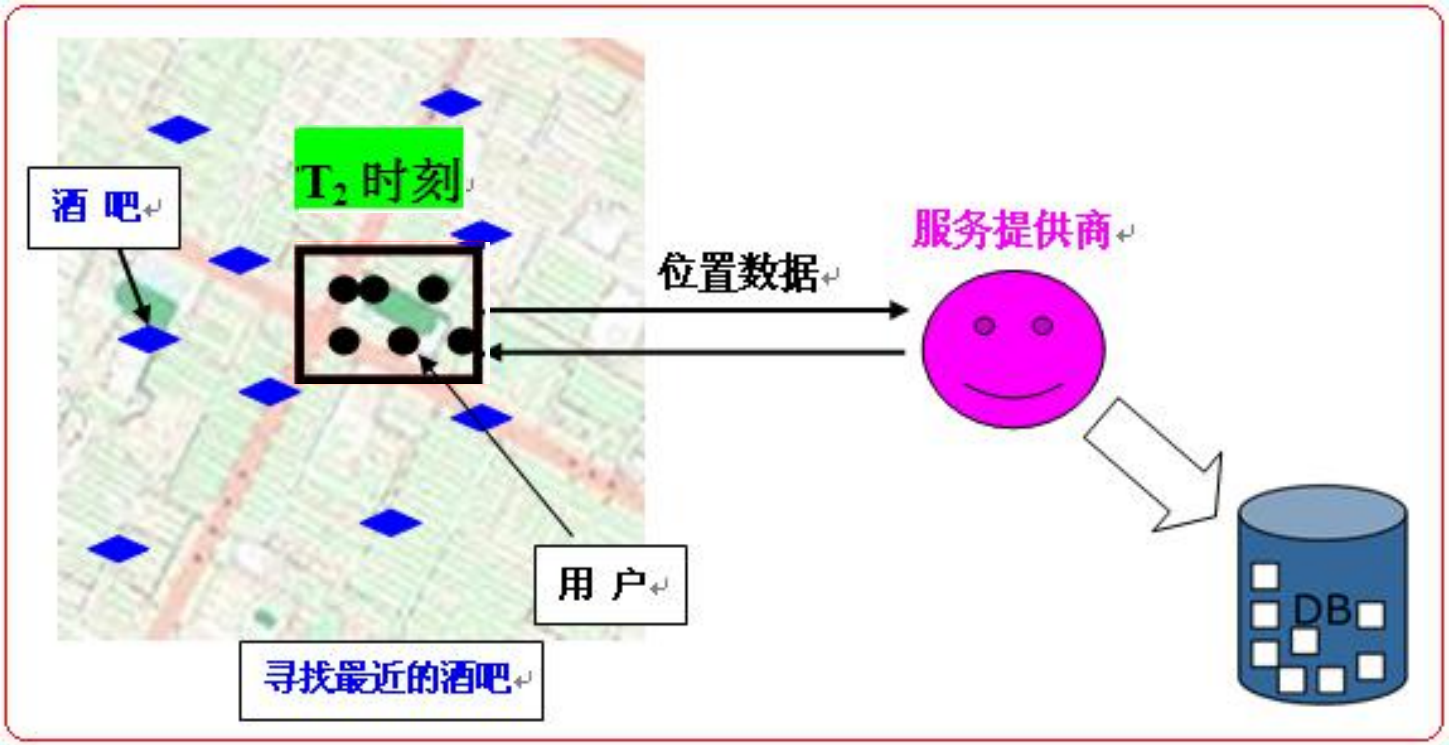
二、空间匿名

降低对象的空间粒度，用一个空间区域来表示用户的真实精确位置，区域的形状不限，普遍使用圆和矩形。同时结合 k-匿名的思想来构造，即用户提供给服务器的位置区域不仅需要包含该用户的准确位置，并且还需要包含其他 k-1 个移动用户。这样用区域表示位置之后，用户在此区域内每一个位置上出现的概率相同。攻击者仅能知道用户在这个空间区域内，却无法确定具体的位置。



三、时空匿名

时空匿名就是在空间匿名的基础上，增加一个时间轴。在扩大位置区域的同时，延迟响应时间，可以在这段时间中出现更多的用户、提出更多的查询，隐私匿名度更高。



基于空间匿名与时空匿名的查询处理

查询处理器

用户发送给位置服务器的不再是精确的位置点，而是一个匿名框。用户在框内的每一个位置出现的概率相同，查询处理器无法获知移动用户的确切位置。传统的查询方式不再适用。

查询的类别

基于位置数据库服务器中的数据可以分为两种：公开数据和隐私数据。相应的，根据查询点和被查询点是否隐私，可以将查询分为四种。

查 询 结 果

一、在候选结果集中随机挑选一个对象作为结果返回给用户；二、返回整个候选集，用户或第三方可信件求精；三、以概率查询处理技术处理查询结果。

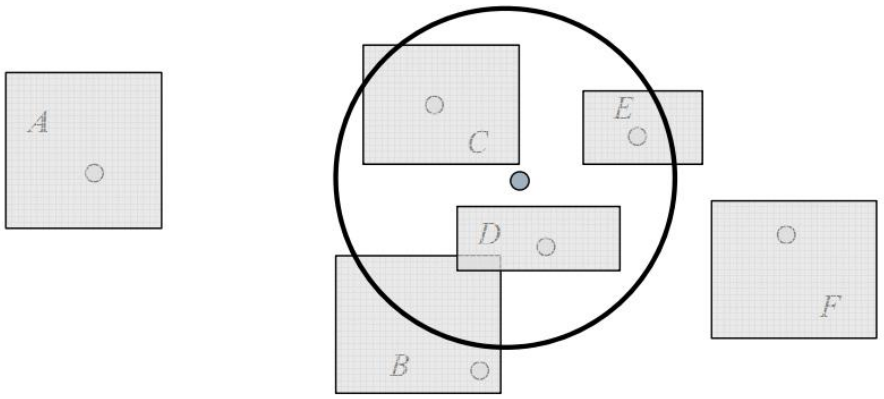
示　例

某加油站500米内的所有出租车

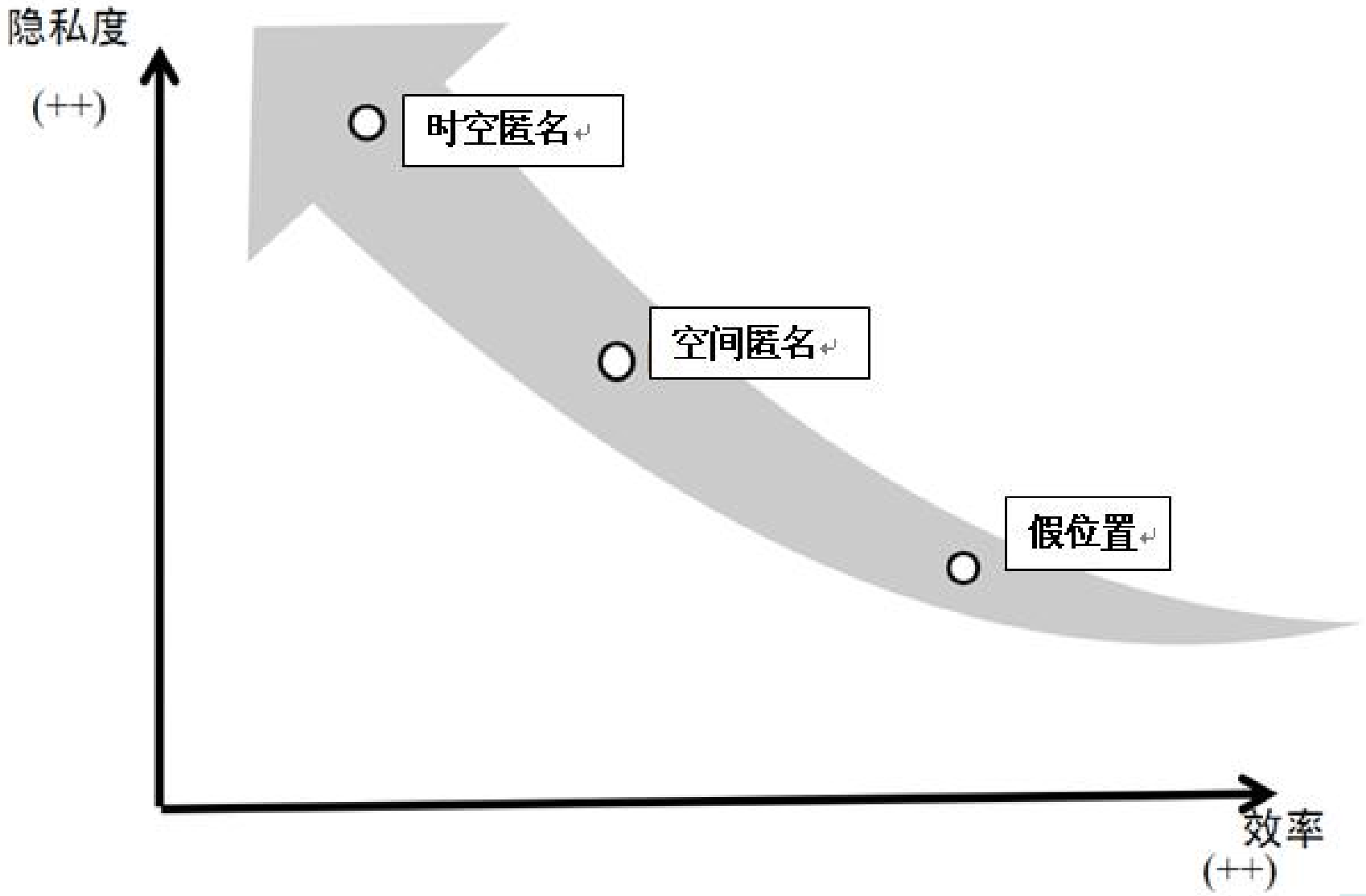
※将所有与查询范围相交的匿名区域都作为候选集。

※将匿名框与圆形框的重叠区域面积大小表示查询结果是真正结果的概率。

查询结果：（B，50%）（C，90%）（D，1）（E，60%）

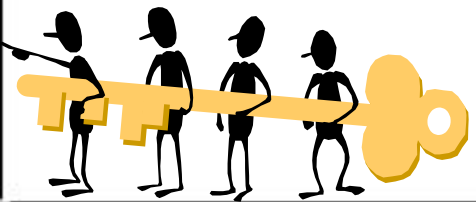


三种匿名技术的对比



基于P2P的位置匿名算法实现

现有系统区域被划分成网格的形式，要求移动用户通过与其他用户的通信查找到 k-1 个用户，将自己真实的位置点扩大为一个模糊的匿名空间，使得该范围覆盖了 k 个用户的位置，同时还必须满足A_{min}。假定匿名空间是一个矩形，用一个左下-右上的二元（$[x_l,y_b][x_r,y_u]$）表示。如何实现这一过程呢？？



基于P2P的位置匿名算法实现

算法一：Function P2P Anonymity-Requestor(k,d)

1、初始化：令已找到用户的集合S′ ={Ø}，已经找到用户的个数k′=|s′|=0，跳数的初始值h=1，d为用户的度，k为匿名度，h_{max}为最大跳数，flag为标志位。

2、if(d≥k-1)then

{

以单跳形式广播查找请求，寻找k-1用户；

}

else

{

h_{max}=最大跳数；

flag=0；

do

{

以多跳形式开始广播查找请求；

S 是经过多跳形式返回的已找到的用户集合，则 k′=|s|；

移动用户 m 如果想要获得服务，首先将已找到用户的集合S′ 置为空集，已找到用户的个数k′ 赋为 0。然后，用户 m 通过比较隐私档案中k 值和自身度d 的大小，选择相应的方式（单跳或多跳）向其周围的用户发送一个消息序号 ID 来广播查找请求，用户 m 监听网络，等待周围用户的响应。

度是指与该用户在一个通讯范围内用户的数目。

单跳指由发站直接到收站的一次性传输；多跳指由发站到收站经过多次转发的方式。

基于P2P的位置匿名算法实现

算法二: Function P2P Anonymity-Receiver (h)

1、**if**（一个请求是重复的）**then**

通过一个 **ACK** 消息回应服务的请求者；

return；

2、**if (h =1) then**

发送一个元组**(ID，** $[x_r,y_r]$**，** v_{max}**，** t_r**，h)** 回应服务的请求者；

else

{

h = h-1；

广播查找请求；

S 是查找到的用户的集合；

for $S_i \in S$ **do**

$S_i.h = S_i.h+1$;

$S=S \cup$ **(ID，** $[x_r,y_r]$**，** v_{max}**，** t_r**，h)**；

返回**S** 到服务的请求者；

}

（1）当请求的接收者 r 接收到来自服务请求者m（可能是请求的发起者，也可能是请求的转发者）的查找请求后：首先基于消息序号 ID 请求的接收者 r检查该请求是否为一个重复的请求。若是，则返回一个 ACK 消息。否则，基于 h 的大小处理请求。

（2）当 h=1（单跳） 时，r 返回一个包含 ID、当前位置、最大运动速度、时间戳和跳数的元组(ID，$[x_r, y_r]$，v_{max}，t_r，h)给请求者；

（3）当 h>1 （多跳）时，r 递减 h 的值，再向周围用户发送原始的消息序号 ID 来广播查找请求。r 监听整个网络，等待周围用户的响应，直到收集到所有其周围用户的回复。最后，r 分别递增收集到的元组中的 h 值，同时将自己的元组信息一起添加到已收集的元组中构成集合 S，返回给请求的发起者或转发者 r。

基于P2P的位置匿名算法实现

```
do
    {
     以多跳形式开始广播查找请求;
     S 是经过多跳形式返回的已找到的用户集合，则 k′=|s′|;
    if (flag = = 0) then
      {
       服务请求者中心位置的调整;
       flag=1;
       }
     if ( k ′<k-1) then
       if (S = S′|| h=h_max) then
         请求挂起;
          h = h+1;
    S ′ =S;
  } while (k′= k-1);
}
```

- 当请求的发起者 m 经过一次查找请求后得到的用户数小于 k-1 时，它将继续递增 h的值，并且以一个新的消息序号 ID 重新广播请求，重复此过程，直到找到 k-1 个用户。
- 值得注意的是，如果在两次连续的广播后得到的结果集相同，或者，当跳数 h 的值超过规定的最大跳数时，m 则需要释放他的隐私档案，选择一个更小的 k 值，或是将请求挂起一段时间，随即再次发出查找请求服务。

当用户只是通过单跳的形式就可以查找到满足匿名要求的用户数时，即 $d \geq k-1$，这说明用户处于一个相对密集的区域，用户的移动对其位置匿名影响不大。但当用户采用多跳的形式查找用户时，即用户所处的区域相对稀疏时，必须考虑用户位置的调整。

基于P2P的位置匿名算法实现

3、**for all $S_i' \in S'$do**
调整集合中每个用户的位置范围；
4、形成一个包含这 **k** 个用户的最小匿名空间**A**；
5、**if ($A<A_{min}$) then**
扩大区域 **A** 到 A_{min}；

算法结束！！！

如果所形成的匿名空间 $A<A_{min}$，则扩大空间 A 到 A_{min}。最后，服务的请求者 m 将包含匿名空间 A 的查询发送到基于位置的社会性网络数据库服务器。经过处理之后，返回给用户一个查询结果的候选集，用户根据自己的精确位置对候选结果集求精。


谢
谢