

基于信号衰减和孤立点检测的移动定位算法

朱江^{1,2}, 戚正伟²

(1. 杭州市公安局, 杭州 310002; 2. 上海交通大学软件学院, 上海 200240)

摘要: 提出一种基于信号衰减的 GSM 移动定位算法, 该算法从移动台上行信令中提取相关网络信号衰减参数, 通过改进的无线信道数学模型计算出移动台与附近基站的距离, 通过孤立点检测算法降低偶然误差, 给出移动台的估计位置。实验表明, 该算法有效提高定位精度, 且对现有网络和移动台无任何要求, 实用性强。

关键词: 定位算法; 移动定位; 信号衰减; 孤立点检测

Mobile Location Algorithm Based on Signal Attenuation and Outliers Detection

ZHU Jiang^{1,2}, QI Zheng-wei²

(1. Public Security Bureau of Hangzhou, Hangzhou 310002; 2. School of Software, Shanghai Jiaotong University, Shanghai 200240)

【Abstract】 This paper presents a mobile location algorithm based on signal attenuation and outliers detection. It takes advantages of original parameter of signal attenuation from mobile station's uplink channel. It can translate the distance information from mobile station to base station through the improved wireless channel mathematics model, and can decrease occasional error through the new outliers' detection algorithms. Experiment result indicates that the algorithm is efficient. The position accuracy is improved without modification on both network and mobile stations.

【Key words】 position algorithm; mobile location; signal attenuation; outliers detection

1 概述

蜂窝移动定位技术在公共安全服务等多方面都有广泛的应用。蜂窝无线定位所采用的定位方法有很多种, 根据所检测的特征测量值的不同, 可以将基本的定位方法分成以下几种: CELL ID 定位方法, 基于信号衰减的方法, 基于时间测量的方法, 基于信号到达角度的方法以及 GPS 定位方法。其中, CELL ID 定位方法精度较低; 基于时间测量的方法、基于信号到达角度的方法以及 GPS 定位方法需要对网络和终端进行改造; 基于信号衰减(SA)的定位技术既有精度优势又降低了定位实施所需的成本。

这项技术利用手机靠近基站或远离基站所带来的信号衰减变化估计手机的方位。假定基站发出的信号功率已知, 那么在另一点测量信号功率时, 就可以利用一定的传播模型估计出基站与该点的距离。以前, 这项技术被认为是定位技术中最不可靠的一种, 主要原因是: 信号并不只因为传输距离而产生衰减, 其他的因素如穿越墙、植物、金属、玻璃、车辆等都会对信号功率产生影响; 在进行第三方定位时, 基站信号到达终端的强度很难获取。

为克服信号衰减定位技术的这些缺点, 研究者提出了通过检测接收信号的场强值, 利用事先测量出的网络覆盖范围内不同位置上的场强分布的数据库来估算出移动台的大致位置, 但该方法对环境变化敏感。为了获取基站信号到达终端的强度, 文献[1]提出使用拨打的方式, 在手机振铃前终止, 同时使用其设计的单站定位系统从空中截取手机上行的测量报告, 从中解析出基站信号达到目标手机的强度, 但该方法若控制不好容易造成目标手机振铃, 而且单站定位系统对无线测量报告的截获率也是一个问题。

本文提出遵循 GSM 相关规范, 通过发送一种特殊的短信给目标移动台, 对该移动台上行测量报告中的 GSM 网络参数进行解析, 可获得基站信号到达终端的强度, 并进一步根据孤立点检测(outliers detection)算法获取该移动台的位置信息。该方法可使被定位移动台没有任何察觉, 对现有移动网络无大改动, 且对移动台无任何要求, 此外, 它还具有较高的精度以满足不同用户的需求。

2 算法模型

本文的算法是从测量报告的网络参数中提取出当前及相邻小区到达被定位移动台的场强值和它们的识别码, 从这些参数可以计算出移动台到达基站的距离, 如图 1 所示。

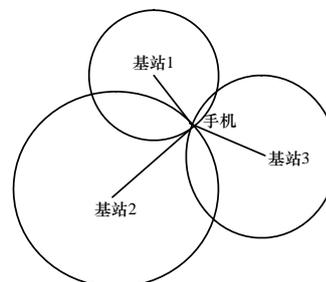


图1 定位算法原理示意图

从理论上说, 移动台应该位于以基站为圆心, 计算距离

基金项目: 国家自然科学基金资助项目(60873209)

作者简介: 朱江(1978-), 男, 硕士研究生, 主研方向: 软件工程, 移动定位; 戚正伟, 讲师、博士

收稿日期: 2009-10-20 **E-mail:** ok6688@hotmail.com

为半径的圆周上,若没有误差,3个基站的圆应该交于一点,这就是移动台的位置。由于基站的发射功率从建站后一般不会变化,因此可以通过无线信道的数学模型较准确地计算出移动台接收到的场强值距离即圆的半径。

2.1 无线信道数学模型

在无线定位过程中,电波传播的特性是一个不可回避的问题,认识移动通信无线传播信道本身的特性和电波传播规律,对于理解和评价蜂窝网移动台定位误差和提高定位精度具有重要意义。而要保证比较高的场强定位精度,建立合适的电波传播数学模型是十分重要的,然后才能在传播数学模型的基础上利用场强定位算法进行定位计算。

到目前为止,人们已经根据各种不同地形环境中的场强实测数据总结出一些场强损耗的传播模型,主要包括:Okumura Hata 模型, COST-231Hata 模型, Okumura 模型, Egli 模型, Walfisch 和 Bertoni 模型, WIM 模型, Lee 模型等。其中, Okumura Hata 模型是根据 Okumura 实际数据曲线图建立的经验公式,已广泛应用于实际中。Okumura Hata 模型适用频率的范围为 150 MHz~1 500 MHz,基站天线有效高度为 30 m~200 m,移动定位终端的天线有效高度为 1 m~10 m,其基本部分为

$$L_{city} (dB) = 69.55 + 26.16\lg(f) - 13.82\lg(h_b) - a(h_m) + [44.9 - 6.55\lg(h_b)]\lg(d) \quad (1)$$

其中, f 为频率; h_b 为基站天线有效高度; h_m 为移动定位终端天线有效高度; d 为移动定位终端与基站之间的距离; $a(h_m)$ 是移动定位终端天线的有效高度的修正因子,其近似值如下:

$$a(h_m) = (1.1\lg(f) - 0.7)h_m - (1.56\lg(f) - 0.8) \quad (2)$$

$$a(h_m) = 3.2[\lg(11.75h_m)]^2 - 4.97 \quad (3)$$

$$a(h_m) = 8.29[\lg(1.54h_m)]^2 - 1.1 \quad (4)$$

其中,式(2)适应中小城市;式(3)适应大城市, $f \geq 400$ MHz;式(4)适应大城市, $f \leq 200$ MHz。

为了得到郊区和农村或宽阔地区的长区间路径损耗中值,把 Hata 模型改为

$$L_{sub} = L_{city} - 2[\lg(f/28)]^2 - 5.4 \quad (5)$$

$$L_{vil} = L_{city} - 4.78[\lg(f)]^2 - 18.33\lg(f) - 40.98 \quad (6)$$

其中,式(5)适应郊区;式(6)适应农村或开阔地区。

使用此模型可以根据移动台实际接收到的场强值来计算移动台与基站的距离,然后进一步进行计算。

2.2 HATA模型修正

在实际情况下,无线信道非常复杂,为了检测 Hata 模型与杭州市的无线传播环境是否吻合,本文对杭州市区的一些基站进行了实地测量和计算,发现它们与模型的理论值存在着一定的差距,选取的点分别是文二路与万塘路交叉口、浙江科技产业大厦、浙江大学玉泉校区正大门、浙江大学 6 号学生宿舍门口、古荡汽车站、三维通信事业公司、华星时代广场、孩儿巷长寿桥小学门口等。这些地点有的高楼林立,人流量很大,属于典型的市区环境;有的则地形开阔,属于郊区环境。具体情况如图 2 所示。由图 2 可见, Hata 理论值与实测值的拟合有一定差距,这是因为每个地区都有自己独特的地貌和无线传播环境。为了更准确地进行计算,本文对 Hata 模型进行了修改,使之更加符合杭州实际无线环境情况。笔者对 Hata 模型添加了修改因子:一般当 $\sigma=9.1$ 时,曲线拟合更加理想。其中, σ 是修改因子,这样就有:

$$r_{new} = r_{old} + \sigma \lg d \quad (7)$$

其中, r_{new} 是修改后的终端场强值; r_{old} 是修改前的终端场强测量值; d 是基站到测试点的距离,单位为千米,如图 3 所示。从图 3 可以看出,修改后的理论值更加符合实际情况。这样,本文就通过实地测试对 Hata 模型进行了修正,为定位的进一步计算服务。

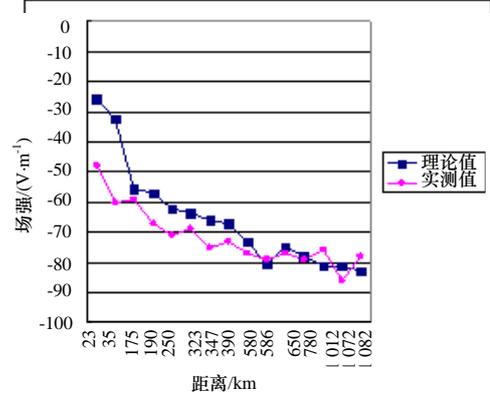


图 2 Hata 理论与实测值

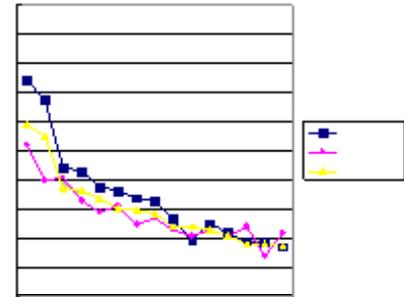


图 3 修改后的终端场强值与修改前对比

2.3 圆周算法

圆周定位方法是移动台定位方法的原始根据,若已知移动台到基站 i 的直线距离为 R_i ,则根据几何原理,移动台一定位于以基站 i 所在位置为圆心, R_i 位于半径的圆周上。

如图 1,基站 1~基站 3 到达手机的距离分别为 r_1, r_2, r_3 ,可以通过第 2 节讲的无线信道衰落的数学模型进行计算,这样就知道了移动台分别到 3 个基站的距离,若基站 1~基站 3 的坐标分别为 $(a_1, a_2), (b_1, b_2), (c_1, c_2)$,并设 MS 的坐标为 (x, y) ,然后通过下面的公式求解计算^[2-3]:

$$r_1 = \sqrt{(a_1 - x)^2 + (a_2 - y)^2} \quad (8)$$

$$r_2 = \sqrt{(b_1 - x)^2 + (b_2 - y)^2} \quad (9)$$

$$r_3 = \sqrt{(c_1 - x)^2 + (c_2 - y)^2} \quad (10)$$

则有以下公式^[4]:

$$x = \frac{(r_1^2 - r_2^2 + b_1^2 + b_2^2 - a_1^2 - a_2^2)(2c_1 - 2a_1) - (r_1^2 - r_3^2 + c_1^2 + c_2^2 - a_1^2 - a_2^2)(2b_2 - 2a_2)}{(2b_1 - 2a_1)(2c_2 - 2a_2) - (2c_1 - 2a_1)(2b_2 - 2a_2)} \quad (11)$$

$$y = \frac{(r_1^2 - r_2^2 + b_1^2 + b_2^2 - a_1^2 - a_2^2)(2c_1 - 2a_1) - (r_1^2 - r_3^2 + c_1^2 + c_2^2 - a_1^2 - a_2^2)(2b_1 - 2a_1)}{(2b_2 - 2a_2)(2c_1 - 2a_1) - (2c_2 - 2a_2)(2b_1 - 2a_1)} \quad (12)$$

即可解出手机的位置。事实上,只需测量基站 1 和基站 2 到 MS 的距离,就可以确定点 MS 的坐标。测量基站 3 到 MS 点的距离,仅起到修正作用^[5]。如果利用基站扇区方向等信

息,也可只用2个方程得到手机MS的位置。

2.4 孤立点检测算法

由于通过场强值进行定位受环境因素影响很大,因此本文提出可以多次对被定位移动台发送特殊短信,从而达到多次测算、多次定位的目的,这样往往会造成每次定位的结果都不一样,互相间有一定偏差。为了提高定位准确率,采用基于距离的孤立点检测算法对这些不同的位置估计点进行扫描,将其中的孤立点去除,将剩下的位置估计点的计算中心作为最后的位置估计结果输出。

2.4.1 基于距离的孤立点检测

在这种检测方法中,一般根据数据对象的最近邻居判断其是否为孤立点。基于距离的孤立点的优点在于不需知道数据的分布模型,因此,可以应用于任何用某种距离机制量度的特征空间。最常用的距离是绝对距离和欧氏距离。绝对距离又称曼哈顿距离,其公式如下:

$$d_{ij} = \sum_{k=1}^m |x_{ik} - x_{jk}| \quad (13)$$

而欧氏距离为

$$d_{ij} = \sqrt{\sum_{k=1}^m (x_{ik} - x_{jk})^2} \quad (14)$$

在数据集 S 中, o 是一个孤立点,仅当 S 中至少有 P 部分对象与 o 的距离大于 d 。换句话说,如果 o 在 d 范围内有不多于 M 个邻居,则 o 是一个带参数 p 和 d 的 $DB(p, d)$ 孤立点。这里 $M=n \times (1-p)$, n 为数据对象的个数。

2.4.2 基于距离和的孤立点检测

为降低孤立点检测对用户的要求,本文借鉴了基于距离和的孤立点定义。与 $DB(p, d)$ 孤立点一样,使用绝对距离或者欧氏距离,但不根据 p 和 d 判定孤点,而是计算数据集中对象两两之间的距离,然后累计每个对象与其他对象的距离,设 M 为用户期望的孤立点个数,则距离之和最大的 M 个对象即为孤立点。

基于距离和的孤立点检测可以描述如下:

对原始数据集进行标准化后,计算 n 个对象两两之间的距离 d ,形成距离矩阵:

$$R = \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{n1} & d_{n2} & \cdots & d_{nn} \end{bmatrix} \quad (15)$$

令 $P_i = \sum_{j=1}^n d_{i,j}$, 即 P_i 为矩阵 R 中第 i 行的和, P_i 值越大,

说明对象 i 与其他的对象距离越远, P_i 值最大的 M 个对象即为孤立点。

本文使用基于距离和的孤立点检测算法可以轻松地检测出那些因为偶然的误差而造成偏离其他点的孤立点。将它们从结果集中删除后,对剩下的结果集中的点求取算术平均值,就是移动台定位的估计位置。

设结果集中的点分别是 S_1, S_2, \dots, S_n , M 为要抛弃的点的个数,求应用孤立点检测算法后的移动台估计位置。

(1) 计算距离矩阵 R , 两两距离是 $d_{i,j}$, 表示第 i 点到第 j 点的距离。

(2) 如式 $P_i = \sum_{j=1}^n d_{i,j}$, 依次求取 P_i 。

(3) 删除 P_i 最大的 M 个点。

最后结果如下:

$$S0 = \frac{\sum S_{\text{余下}}}{n - M} \quad (16)$$

图4是本算法在定位中的应用示例。



图4 基于距离和的孤立点排除算法

通过对被定位的移动台多次发送特殊短信,获取测量报告中的场强参数后,先根据式(1)计算出MS(移动台)与基站的估计距离,然后可以根据2.3节介绍的圆周定位算法进行计算后得到一些位置估计点。

假设图中的10个点都是位置估计点,若设 $M=2$,则通过孤立点检测算法很容易地将孤立点检查出来(图4),然后计算余下8个点的平均值即是最后的位置估计值。

3 算法测试

为了对本算法的效果进行测试,笔者在杭州地区随机选取了11个点作为测试点,分别是城南凤山门、后市街、采荷路、四季青大酒店、东方假日宾馆、浙二医院龙翔服饰城、少年宫门口、西湖白堤、火炬大道、滨江创业路。在这些地点,使用专用的基站信号强度测试仪,对周围的基站信号强度一一进行记录。在误差的估计上,将GPS导航仪与算法估计位置的距离作为算法的误差。此外,在对该算法的误差进行分析时,采用了2种对比算法,分别是CELL ID算法和圆周算法。前者是现在移动运营商普遍采用的定位算法,实现简单;后者是本文算法的前半部分。测试结果如图5所示,图中,a为Cell Id定位误差; b为圆周算法定位误差; c为多次触发再孤立点排除定位误差。

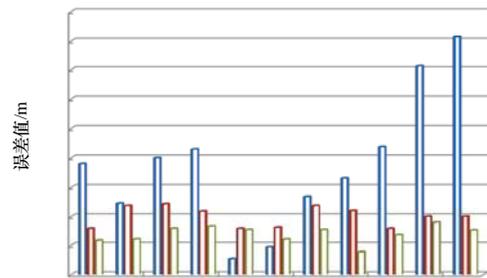


图5 几种定位算法误差比较

经过统计,采用本算法,平均误差达到70.27m;圆周算法平均误差为100.36m;传统的CELLID算法平均误差为190m。由此,可见本算法的精度高于其他2种定位算法。

4 原型实现

本文将本算法应用到作者设计的移动位置信息支撑系统中。该位置信息支撑系统属于B/S架构,这样可以方便地部署在Internet和Intranet上,极大地方便了用户的访问和使用。同时系统使用了Web Service技术,使得用户可以直观地获取到被定位移动台的位置信息,此外系统还对每个访问的用户进行鉴权,确保被定位移动台的安全隐私。系统总体架构如图6所示,采取三层结构模型。系统原型如图7所示,图中圆型圈即表示定位结果。(下转封三)