|  |
| --- |
| 北京优锐科技有限公司 |
| **基于概率计算和支持向量机计算的WiFi定位系统** |
| 丁贵金 朱韬 袁万尚 |
| 2013年7月6日 |

|  |
| --- |
| [在此处键入文档的摘要。摘要通常是对文档内容的简短总结。在此处键入文档的摘要。摘要通常是对文档内容的简短总结。] |

基于概率计算和支持向量机计算的WiFi定位系统

北京优锐科技有限公司 丁贵金 朱韬 袁万尚

2013年7月6日

# 1摘要

基于概率和支持向量机原理的 WiFi定位技术，不需要依赖专用设备，部署简单使用便捷，对环境无强制依赖，可以在复杂 WiFi 环境下实现移动设备精确定位。该 WiFi 定位技术的核心原理是支持向量机，辅助正态分布的概率计算来优化支持向量机计算过程。

本文针对的设备：是带有WiFi功能的移动设备。针对的环境：是分布着大量WiFi接入设备的室内环境。实现的主要目标：是通过WiFi移动设备，在分布着大量WiFi接入设备的室内环境中，实现精确定位。

**关键词：**

WiFi定位，室内定位，支持向量机。

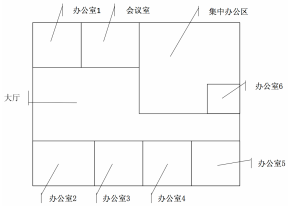
# 2引言

本文的“专用名词、概念和常识说明”部分，主要解释全文中出现的各种专用名词，涉及到的所有概念，以及阅读本文所需的技术常识。“实现方法”部分则具体阐述技术实现过程，“讨论”部分主要分析了实现过程中可能出现的各种问题，以及处理问题的方法。“技术创新”部分具体说明该技术的先进性，与同类技术相比下的优势，“应用前景”部分介绍了该技术实际应用的具体形式，以及对采用该技术的行业所产生的积极作用。

# 3专用名词、概念和常识说明

## 定位空间（CR）

定位空间（calibration room），指的是提供定位功能的空间。例如采用WiFi定位的公司内部空间，公司外部无限大的空间就是不可定位空间。如图：



## 移动终端设备

本文中的“移动终端设备”，是指拥有WiFi连接能力的移动设备，如有WiFi功能的智能手机、平板电脑和其他便携的移动终端设备。

## WiFi接入设备（AP）

提供WiFi接入服务的硬件设备，如无线路由器。本文中对WiFi接入设备的要求，是必须能够被移动终端设备识别到的WiFi接入设备。

## AP名（MAC）

AP名就是AP的MAC地址，是WiFi接入设备的唯一标示。

## AP信号强度（RSS）

AP信号强度，是移动终端设备对WiFi接入设备信号强度的识别结果。该结果是一个从0到-100的值，该值是一个指数值，代表信号的强度指数，并不是真正的物理信号量。不同的移动终端设备，对同一个WiFi接入设备，信号强度值的识别结果是不同的。

## AP信号强度平均值（mRSS）

同一个移动终端设备，对同一个WiFi接入设备，每次信号强度的识别结果也是不同的。一个移动终端设备，对同一个WiFi接入设备，进行多次采集，识别到一组信号强度，对该组信号强度的求平均值，就得到了信号强度平均值，计算公式如下：

，N为该组信号强度的个数。



## WiFi接入设备数据（APD）

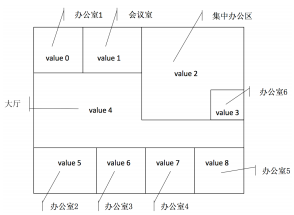
WiFi接入设备数据包括，WiFi接入设备名（MAC）和WiFi接入设备被识别到的信号强度。

## 待定位点（LP）

带定位点，是指一个移动终端设备，出现在定位空间内，该移动终端设备自身的位置。

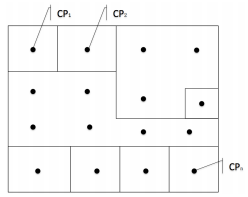
## 定位区域（value）

定位空间是一个完整的封闭空间，要在这个空间内实现WiFi定位，就必须将这个完整的空间划分为多个小区域，每个区域都有一个ID。如图：



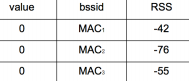
## 定位校准点（CP）

在定位区域中，指定一个位置点，这个点就是定位校准点。如图：



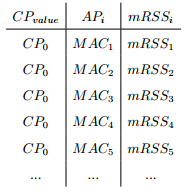
## 定位区域数据

在定位区域内的定位校准点上，采集到的WiFi接入设备数据，就是定位区域采集数据。该数据的标示是区域ID，内容是WiFi接入设备数据。如图：



## 校准数据

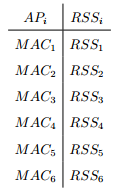
校准数据，是对所有定位区域数据的整理计算结果。如图：



Value是定位区域ID，AP是收集到的所有WiFi接入设备名，mRSS是该WiFi接入设备被识别到的，信号强度平均值。

## 待定位数据

移动设备在待定位点，某一时刻采集到的WiFi接入设备数据。如图：



AP是识别到的所有WiFi接入设备名，RSS是识别到的所有WiFi设备的信号强度。

## 概率

### 正态分布

正态分布是一种常用的概率分布函数，本文阐述的WiFi定位技术，在实现过程中采用了这种数学方法。公式如下：



### 概率值（Pr）

本文中所讲的概率值，是指一个待定位点，出现在某一定位区域中的概率。

## 支持向量机（SVM）

支持向量机，是一种机器学习原理，是采用数学方法，实现对某种向量值进行特定分类的工具。

### 样本

本文所说的样本，是指WiFi接入设备的数据。

### 样本空间

多个样本生成的集合，叫做样本空间。

### 特征向量

本文所说的特征向量，是指可以代表一个定位区域的样本。是校准数据中，一个定位区域中所有样本的子集。

### 核函数

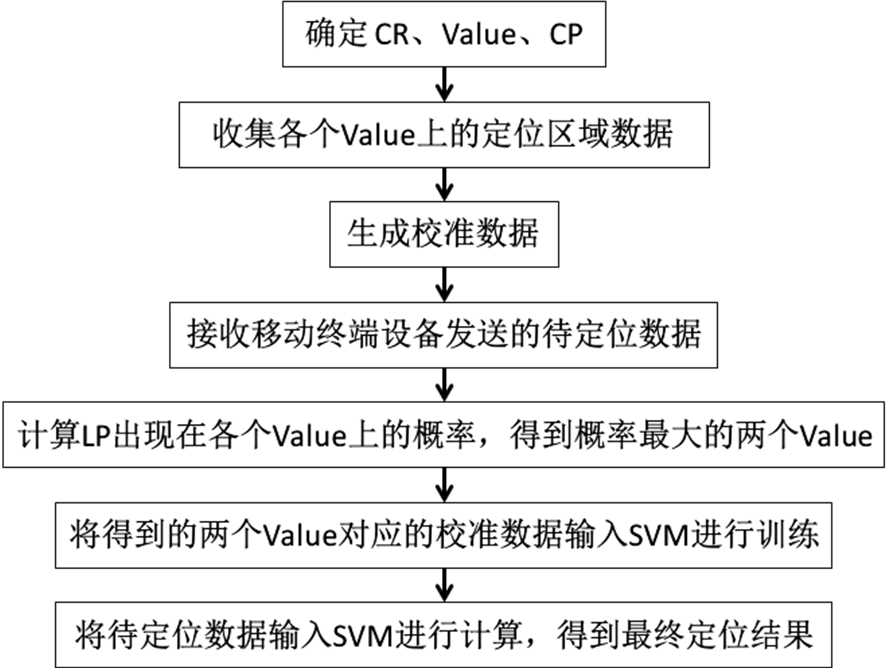
核函数，是指支持向量机所采用的某种特征向量训练模式。本文中的核函数采用RBF核函数。

### 惩罚因子

惩罚因子，是指在支持向量机计算中，分割两个样本空间时，对不可分样本的包容程度。

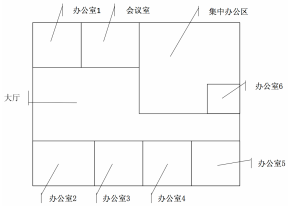
# 4实现方法

## 4.1总体流程

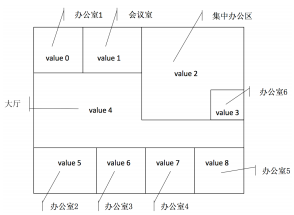


## 4.2 确定CR、Value、CP

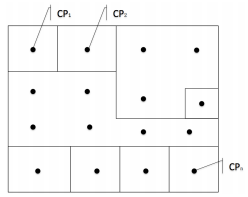
1. 根据实际需求，确定定位空间整体范围。如图：



1. 在定位空间内划分出各个定位区域。如图：



1. 在每个定位区域内确定一个或多个CP。如图：

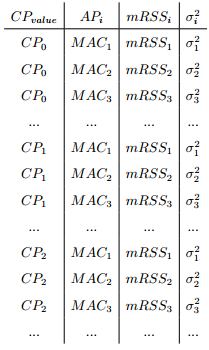


## 4.3 收集各个Value上的定位区域数据

用移动终端设备，在定位空间内，收集所有定位区域中的多组定位区域数据，将这些数据按定位区域ID，统一保存。

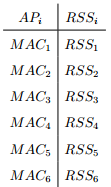
## 4.4 生成校准数据

将收集到的区域定位数据，进行处理，计算出其中每个AP的信号平均值（mRSS）和信号强度值的标准差。按照各AP所对应的区域ID（CP value）组织计算结果，统一保存生成校准数据。如图：



## 4.5 得到移动终端设备的待定位数据

一个移动终端设备，在定位空间内移动，在不同的位置点上会收集到不同的待定位数据，这些不同的待定位数据都对应不同的待定位点（LP）。将移动终端设备在LP上收集到的带定位数据保存。如图：



## 4.6 计算LP出现在各个Value上的概率，得到概率最大的两个Value

从校准数据中按照不同Value，依次取出每个Value对应的一组校准数据，将LP的带定位数据与每组校准数据中共有的AP，分别进行正态分布概率计算，计算公式如下：



再将计算得到的概率值相加，最终得到了该LP出现在每个Value中的概率。

其中概率值最大的两个Value，就是该LP可能出现的位置范围。

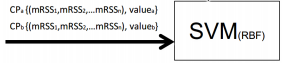
## 4.7将得到的两个Value对应的校准数据输入SVM进行训练

得到了概率最大的两个Value，说明LP出现在这两个Value中的可能性最大。将LP的待定位数据的AP名和这两个Value对应的校准数据的AP名做交集，得到一组大家都包含的AP。

从LP的待测数据中挑出这个几个共有AP的待定位数据，作为LP的特征数据。

两个Value的校准数据中分别挑出这几个共有的AP校准数据，与该Value的ID一起，作为SVM的特征向量。

将特征向量输入SVM进行训练。如图：



## 4.8将待定位数据输入SVM进行计算，得到最终定位结果

将LP的特征数据，输入训练完成的SVM，得到一个Value的ID值。该Value所对应的区域就是LP所在的位置。如图：



# 5讨论

# 6技术创新（权利保护）

## 6.1

## 6.2

## 6.3

# 7应用前景

WiFi 定位技术可以应用在多种领域，主要为行业提供室内定位基础服务。本文阐述的 WiFi定位技术，不依赖特定硬件和特殊环境要求，部署应用方便灵活，成本低，可以结合多种具体的业务实现垂直服务。

## 7.1 室内定位和导航

室内定位和导航，是 WiFi 最主要的应用方式，也是最直接的服务提供模式。

1. 公共室内空间的位置服务，如商场、超市和大规模的综合购物中心。为了方便顾客找到所需商品位置，或快速找到某个区域 (例如卫生间、出口、餐饮区，等等)，结合 WiFi 定位和电子地图技术，实现顾客自身位置确定和导航服务功能。
2. 室内区域流量统计服务，如在商场、超市和大规模的综合购物中心，有时业主需要统计一段时间内，某个区域或所有区域的客流信息，或者某个客流最大的热点区域。结合WiFi 定位和电子地图以及大数据技术，可以实现室内区域的流量统计，顾客使用移动设备上的 WiFi 定位客户端软件，经过的路径会推送至服务器，服务器端根据统计整理，最后分析出某个或所有区域的客流数据。

## 7.2 地理围栏

地理围栏是一个新兴的移动互联网服务概念，其主要目的是实现互联网应用和移动设备位置信息的结合。

1. 设备位置开启移动应用。当移动设备进入某个特定区域后，便自动开启一些应用软件。这种服务适合应用在商场导购，医院就医指导，机场火车站等公共服务场所。
2. 特定移动应用的使用范围。移动设备上的某些软件，只能在某些特定的环境中才可以使用。这种服务适合公司企业，用来实现现代移动化的企业管理和自动化办公。当员工进入企业办公区域后，相关工作的信息服务和业务软件才可以工作，保证了企业数据的安全，也将企业管理简单有效的实现。

### 7.2.1 移动设备防盗

某些非个人使用的移动设备，如餐厅的点餐设备，高级场馆的自助服务设备，都是为了本地服务存在的。为了防止个别人将其私自带出，又要实现该设备的自由使用，就必须采用一种防盗

的技术做应用保证。本文阐述的 WiFi 定位技术，可以划出移动设备的安全范围，当设备离开这个安全区域后，便可以通过报警、监控和追踪等方式及时发现并处理。

### 7.2.2 区域信息安全

信息隔离，是信息安全中的一项重要内容，指信息内容与信息的有效区域之间的对应关系。采用本文阐述的 WiFi 定位技术，可以通过软件安全策略，规定信息的安全范围，当移动设备需要读取某些带有安全级别的数据时，会根据移动设备所处的位置进行操作合法性判断，当发现移动设备所处的位置不合法时，便会组织数据的读取。对于已经读取到的数据，当设备离开该数据的安全区域后，通过软件安全策略谁自动删除该数据，保证数据存在的有效范围与数据的安全范围一致。

### 7.2.3 地理信息隔离

移动设备的最大特点是可以随身携带，可以随时随地的产生数据，如照相、笔记、下载或应用软件生成数据。利用本文所阐述的 WiFi 定位技术，可以在产生数据的同时，将数据与其产生时的位置做关联。无论在以后的搜索查找，或是统计，位置信息会提供跟多的有效途径，版主使用者对数据的管理。也可以将数据的存储或显示方式，与位置信息关联，是某些数据只在特定位置区域是才可以被操作。

## 7.3 游戏

电子游戏的场景通常是虚拟的空间，但是随着技术的发展，电子游戏的实现也可以利用现实空间。本文阐述的 WiFi 定位技术，可以实现真实空间在设备上的体现，电子游戏的方式也就从单一的虚拟空间扩展到了真实空间，游戏玩家除了操作设备，还需要亲身行动来完成游戏。这样的组合不仅扩展了电子游戏的功能，甚至可以改变电子游戏的传统运行模式。