

# 上海师范大学

## 研究生课程

院（系、所）信息与机电工程学院 专业电子信息

科目 无线网络与物联网应用 第 2 学期

研究生姓名 张天旭 学号 222503089

成绩\_\_\_\_\_

评语：

签字

年 月 日

## 一、问答题

### 1. 简述 wifi1~wifi7 无线网络通讯性能。(10 分)

(1) Wi-Fi 1(802.11b)是于 1999 年推出的首个被广泛使用的 Wi-Fi 标准。它基于 2.4GHz 频段,并提供最高 11Mbps 的传输速度。Wi-Fi 1 使用了 Direct Sequence Spread Spectrum (DSSS) 调制技术,通过在发送数据时引入扩频码来提高抗干扰性能。主要应用于家庭网络、小型办公室等场景。尽管 Wi-Fi 1 的传输速率低,但其开创了无线网络通信的先河,催生了无数基于无线网络的应用。

(2) Wi-Fi 2(802.11a)也于 1999 年发布, Wi-Fi 2 主要采用 5GHz 频段,采用了 Orthogonal Frequency Division Multiplexing (OFDM)调制技术,它将数据分成多个子载波进行传输,提高了传输效率和可靠性。提供最高 54Mbps 的传输速度,但信号在穿透墙壁和障碍物方面的能力较差,因此其覆盖范围相对较小。主要应用于一些高密度环境如大型办公楼、公共场所,可以在多媒体传输等应用中体现出其优越性。

(3) Wi-Fi 3(802.11g)于 2003 年发布,它结合了 Wi-Fi 1 与 Wi-Fi 2 的优点,使用 2.4 GHz 频段,提供最高 54Mbps 的传输速度,同时保持与 Wi-Fi 1 设备的向后兼容性。Wi-Fi 3 同样采用 OFDM 调制技术,提供更高的传输速率和可靠性。

(4) Wi-Fi 4(802.11n)于 2009 年发布。它可以在 2.4GHz 和 5GHz 两个频段上通信,因此它具有更高的灵活性,可以根据具体环境选择最佳频段进行通信。Wi-Fi 4 引入了 Multiple Input Multiple Output(MIMO)技术,通过同时使用多个天线来提高传输效率和可靠性。它可以提供最高 300Mbps 的传输速率。

(5) Wi-Fi 5(802.11ac)于 2013 年发布。它采用了 OFDM 和 Multi-User MIMO(MU-MIMO)技术,允许同时与多个设备进行通信。Wi-Fi 5 支持 20MHz、40MHz 和 80MHz 三种信道宽度,更宽的信道宽度可以提供更高的传输速率,最高传输速率可以达到 1.3Gbps。

(6) Wi-Fi 6(802.11ax)于 2019 年发布,引入了 Orthogonal Frequency Division Multiple Access(OFDMA)技术,将频谱划分为多个子载波,使得多个设备能够同时进行数据传输。Wi-Fi 6 支持 20MHz、40MHz、80MHz 和 160MHz 四种信道宽度进行数据传输,最高传输速率可达 10Gbps,因此能够支持更多的高带宽和低延迟应用,如 8K 视频流媒体、虚拟现实(VR)和增强现实(AR)应用。

(7)Wi-Fi 7(802.11be)是即将到来的下一代 Wi-Fi 标准。它在 Wi-Fi 6 的基础上引入了 320MHz 带宽、4096-QAM、Multi-RU、多链路操作、增强 MU-MIMO、多 AP 协作等技术,使得 Wi-Fi 7 相较于 Wi-Fi 6 将提供更高的数据传输速率和更低的时延。Wi-Fi 7 预计能够支持 30Gbps 的吞吐量。

### 2. 简述 Starlink 的原理与当前应用现状。(10 分)

Starlink 是 SpaceX 公司推出的一个卫星互联网项目,旨在通过在低地球轨道上部署数千颗卫星,为全球范围内的用户提供高速、低延迟的卫星互联网服务。

Starlink 网络由上万颗小型卫星组成,这些卫星以低地球轨道运行,相互之间形成一个星型网络。地面用户需要安装终端,用于接收和发送卫星信号。这些终端设备包括一个接收天线和调制解调器,负责与卫星通信和与用户设备连接。

StarLink 立足于解决当前地面互联网所存在的两大主要问题:

1、偏远地区的上网问题: Starlink 的应用重点是覆盖农村和偏远地区,这些地区通常缺乏传统的有线互联网基础设施。通过 Starlink,居住在这些地区的用户可以获得高速的互联网连接。

2、全球网络延迟问题: 光纤在近距离访问的网络时延可以忽略不计,但如果需要访问

较远距离的数据与设备，就会造成延时的产生。StarLink 在短距离网络数据传输方面没有明显优势，但全球范围内的长距离传输，它的延迟优势就非常明显，延迟仅仅 25 到 35 毫秒，能用很短的时间将数据传输地球的另一面，相当于点对点直连光缆的延迟。

## 二、课程论文

请撰写一篇无线网络相关小论文，题目自拟。（80 分）

### 基于机器学习和 Wi-Fi 指纹的用户位置预测

#### 摘 要

目前，Wi-Fi 设备在各类场景中被广泛应用，其中基于 Wi-Fi 信号的定位算法引起了广泛研究。在各类定位算法中，基于 Wi-Fi 指纹数据的定位能够在较为复杂的环境，尤其是室内环境中具有较好的适应度。本文通过对某复杂场景下 Wi-Fi 的指纹定位信息进行数据挖掘，使用有监督学习的学习方法，对比支持向量机、K 近邻、随机森林等经典机器学习算法，在用户位置预测（回归问题）和用户所在楼层预测（分类问题）上进行对比实验。在公开数据集 UJIIndoorLoc 上的实验结果表明，通过对收集的 Wi-Fi 指纹定位数据进行机器学习建模，能够准确、高效地预测用户位置。

**关键词：**Wi-Fi 定位；机器学习；用户定位预测

# 第 1 章 引言

随着无线网络技术的发展，目前 Wi-Fi 设备已经在各大领域广泛应用，基于 Wi-Fi 设备的延伸研究也逐渐开始进行。为便于人们的生活工作，目前 Wi-Fi 设备基本已经覆盖大大小小办公楼、商场、广场等地，而对于这些复杂场所，亟需一种定位算法来得知每个用户所处的位置。因此，基于 Wi-Fi 信号的定位算法引起了广泛关注。

区别于 GPS 定位，基于 Wi-Fi 信号的定位算法主要具有以下特点：

(1) 环境复杂。Wi-Fi 覆盖的场景大多为生活化的室内场景，这类场景中具有不同楼层、同一楼层中也存在各类障碍，导致定位困难。

(2) Wi-Fi 设备本质上用于信号传输，不利于使用传统位置计算方法进行定位。

(3) 实际环境中，Wi-Fi 设备品种多样，Wi-Fi 设备的发射功率区别也较大，需要将以上因素纳入考虑。

基于 Wi-Fi 信号定位的特殊性，一种基于 Wi-Fi 指纹数据的定位算法脱颖而出。基于指纹的定位（Fingerprinting-based Localization）方法<sup>[1]</sup>的基本原理是利用无线信号在不同位置上的差异性，将空间中特定位置上的无线信号特征作为该位置的指纹，建立位置-指纹关系数据库，从而通过指纹匹配的方式对用户位置进行估计。

实际场景中，主要用到的算法有 KNN (K-Nearest Neighbour) 算法、加权 KNN 算法、支持向量机算法等<sup>[2]</sup>，以上算法各有优缺点。本文的目的是基于 UJIIndoorLoc 数据集提供的 Wi-Fi 指纹定位数据，对用户位置、所在楼层进行定位，并综合比较各类机器学习算法在该数据集上的表现。

## 第 2 章 模型建立

### 2.1 数据预处理

UJIIndoorLoc 数据集<sup>[3]</sup>提供了来自 Jaume I 大学三座建筑物中的 Wi-Fi 指纹信息，由 2000 多个不同用户和 20 个 Android 设备创建。该数据集共包括 529 个特征和 19937 个数据样本。

前 520 个特征分别表示 520 个 Wi-Fi 信号的强弱，其属性值为一个从-104 到 0 的负整数或正整数 100（正整数 100 表示无信号）。第 521 列和 522 列分别表示用户的经纬度坐标，第 523 列表示用户所在的楼层，第 524 列表示用户所在的建筑物 ID，第 525 列表示用户所在的房间类型（如办公室、教室等），第 526 列表示用户所处的相对位置，第 527、528、529 列分别表示用户 ID、Android 设备 ID 以及时间戳。

首先对以上原始数据进行数据预处理，首先查看是否存在缺失值，如存在缺失值，则需要去除存在缺失值的样本。通过观察，该数据中不存在任何缺失值。

下面对数据进行降维。去除 521-529 列，原始数据共包含 520 个特征，需要剔除冗余的特征，以降低后续处理的复杂度。这里使用方差过滤算法，对方差为 0 的特征进行剔除。方差为 0 代表所有样本在该特征上的属性值均相同，即不具有区分性。经过方差过滤后，剩余 466 个特征。

## 2.2 模型参数调整

通过对数据集中经纬度属性进行进一步处理，可以得到距离数据，对该距离量进行预测可以认为是回归问题；该数据集中还给出了每个用户所在的楼层信息，对用户所在楼层进行预测可以认为是分类问题。因此，下面将分别将距离和楼层信息作为标签，对比各机器学习算法在回归预测与分类预测方面的准确性。

在回归问题方面，需要先对经纬度进行处理。这里使用了 Haversine 公式对经纬度进行转换，转换公式如下：

$$a = \sin^2\left(\frac{lat}{2}\right) + \cos(lat) \sin^2\left(\frac{lng}{2}\right) \quad (1)$$

$$distance = 2R \arcsin(\sqrt{a}) \quad (2)$$

其中， $lat$ ， $lng$  分别表示经纬度， $R$  表示地球半径，取 6371。经过处理后，经纬度属性转换为  $distance$  属性，作为这一回归问题的预测目标。

下面通过网格搜索方法对各模型的最佳参数进行搜索。首先对数据集进行训练集和测试集的划分，70%的样本作为训练集，剩余 30%样本作为测试集。由于网格化搜索调参过程耗时较长，数据集规模较大可能导致调参过慢，因此在训练集中取出 2800 个训练样本进行调参。

网格化搜索结果显示，对于 SVR 模型（支持向量回归），其最优 C 值为 20.0，

采用 rbf 内核, 惩罚系数为 0.3。对于 KNN 回归模型, 其最优 k\_neighbor 值为 3。对于随机森林回归模型, 其最优 n\_estimators=30, 最大深度为 24。

在分类问题方面, 以用户所在楼层作为标签, 选用 KNN、支持向量机和随机森林分类器进行对比实验, 并同样使用网格搜索方法进行参数调整。搜索结果显示, 对于 KNN 模型, 其最优 k\_neighbor 值为 3; 对于 SVC 模型, 其最优 C 值为 2; 对于随机森林分类模型, 其最优 n\_estimators=26, 最大深度为 16。

## 2.3 数据拟合与模型训练

对于回归任务, 将初始化完毕的 SVR、KNN 和随机森林回归模型对划分完毕的数据集进行拟合, 然后对测试集进行测试。测试指标选择平均绝对误差 MAE (Mean Absolute Error)。MAE 计算公式如下:

$$MAE = \frac{1}{m} \sum_{i=1}^m |y_i - \hat{y}_i| \quad (3)$$

其中,  $\hat{y}_i$  表示预测值,  $y_i$  表示真实值,  $m$  为样本总数。

对于分类任务, 将初始化完毕的 SVC、KNN 和随机森林分类模型对划分完毕的数据集进行拟合, 然后对测试集进行测试。测试指标选择准确率 ACC (Accuracy Score)、精确率 (Precision) 和召回率 (Recall)。计算公式如下:

$$ACC = \frac{TP + TN}{TP + FN + TN + FP} \quad (4)$$

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

其中,  $TP$  表示正类判断为正类的个数,  $FP$  表示负类判断为正类的个数,  $TN$  表示负类判断为负类的个数,  $FN$  表示正类判断为负类的个数。

在对模型数据拟合完毕, 评估指标制定完毕后, 准备开始实验。

## 第 3 章 实验结果分析

实验环境为 Python3.8, Windows 10 操作系统, 16G 内存。

对于回归任务, 实验结果如表 1 所示。

表 1 各回归模型实验指标

模型	MAE ↓
支持向量机	2361.80
随机森林	413.61
KNN	<b>337.42</b>

可以看出，KNN 模型在该数据集上的回归预测精度相对较高，其平均绝对误差在 330 左右，处在可接受的范围内。相比之下，随机森林模型的误差略高一些，而支持向量机的误差超过了 2000，表现相对较差。

对于分类任务，实验结果如表 2 所示。

表 2 各分类模型实验指标

模型	ACC↑	Precision↑	Recall↑
支持向量机	0.978	0.981	0.972
随机森林	<b>0.986</b>	<b>0.990</b>	<b>0.982</b>
KNN	0.976	0.735	0.731

在分类任务方面，随机森林分类器的表现最优，支持向量机表现次佳，而 KNN 模型在准确率指标方面差距不大，但在精确率和召回率方面表现较差。考虑到随机森林是集成学习的代表性算法，在该分类任务中能够体现出其优越性。

## 第 4 章 总结与展望

本文主要结合 Wi-Fi 设备目前的应用现状与 Wi-Fi 指纹定位的相关理论，基于 UJIIndoorLoc 公开数据集，对各用户进行距离和所在楼层的预测，分别对应了机器学习的回归预测与分类预测两大领域。

本文通过对比和分析 KNN 算法、支持向量机算法和随机森林算法在该数据集上的回归和分类任务的表现，证明了机器学习算法结合 Wi-Fi 指纹信息在用户定位预测问题上的可行性。

后续将会进一步对比和比较更多机器学习算法在该问题上的效果，对数据集预处理方面进行更全面的实验比较（如结合主成分分析、结合具体场景进入新的限制等），也将考虑结合深度学习对该问题进行更深入的分析 and 实验。

## 参考文献

- [1] 杨铮,吴陈沭,刘云浩.位置计算: 无线网络定位与可定位性[M].北京:清华大学出版社,2014:118-120.
- [2] 潘琢金,郑玮,杨华.一种改进的 WiFi 指纹定位算法[J].沈阳航空航天大学学报,2021,38(05):66-72.
- [3] Joaquín Torres-Sospedra, Raúl Montoliu, Adolfo Martínez-Usó, Tomar J. Arnau, Joan P. Avariento, Mauri Benedito-Bordonau, Joaquín Huerta. UJIIndoorLoc: A New Multi-building and Multi-floor Database for WLAN Fingerprint-based Indoor Localization Problems. In Proceedings of the Fifth International Conference on Indoor Positioning and Indoor Navigation, 2014.