



中国研究生创新实践系列大赛
“华为杯”第二十一届中国研究生
数学建模竞赛

学 校 北京邮电大学

参赛队号 24100130040

队员姓名

1. 沈飞
2. 汪金锋
3. 张睿涵

中国研究生创新实践系列大赛
“华为杯”第二十一届中国研究生
数学建模竞赛

题 目： WLAN 组网中网络吞吐量建模

摘 要：

2024 年格式变化说明

今年的格式变化如下：

1. 论文第一页为标识替换；
2. 合并 <https://github.com/andy123t/GMCMthesis> 更新. 修复首页页码；
3. 调整摘要、标题、关键词和浮动体标题等字体。

这是研究生报名官方网站，点击[这里](#)进入。

欢迎大家到 QQ 群里沟通交流：91940767/478023327/640633524。我们也开通了问答区交流 L^AT_EX 技术：<https://ask.latexstudio.net>，欢迎大家前来交流。

关键词： 模板 L^AT_EX 数学建模

1 问题重述

1.1 问题背景

无线局域网（Wireless Local Area Network, WLAN）以人们的业务诉求为驱动力，部署场景日益增多，在无线化办公、教育、医疗、工业制造、仓储等场景都有广泛应用。尽管下一代 Wi-Fi7 标准的峰值速率达到 30Gbps，但是在例如小区、商场、学校等高密部署场景下，相邻节点之间覆盖范围重叠，信号干扰碰撞问题突出，使得部署带宽和数据传输速率大幅下降。为了进一步提升系统吞吐量，对 WLAN 系统的进一步优化具有重要意义。

精准和快速的吞吐量预测是 WLAN 优化问题的基础问题，能够大幅度提升 WLAN 系统的鲁棒性和性能（吞吐量指节点单位时间内成功发送的比特数）。当前的一些研究通过机器学习方法，提取各节点之间的基本信息、架构信息以及节点动态位置、动态干扰等临时信息，进行训练和建模。但是由于研究中使用的仿真方法并不能精确反应实际部署中快速变化的信道环境、干扰以及复杂多样的业务流，其结果在真正商用中并不能达到预期实用价值。

因此，为了对 WLAN 进行优化，在工业、教育、医疗等新场景中，提升用户体验，设计一种能够基于 WLAN 实测数据，分析 WLAN 网络拓扑、节点间 RSSI、信道接入机制、干扰等因素对 WLAN 数据发送、速率的影响，进一步地，对 WLAN 系统吞吐量进行精确预测的模型具有重要研究意义。

1.2 问题描述

基于上述研究背景，题目提供了 WLAN 组网场景下的实测数据，数据集提供了包括测试基本信息，以及测试中收集的数据帧相关统计信息，围绕 WLAN 组网中网络吞吐量建模，提出以下问题：

问题 1：根据附件 WLAN 网络实测训练集中所提供的网络拓扑、业务流量、门限、节点间 RSSI 的测试基本信息，分析其中各参数对 AP 发送机会的影响，并给出影响性强弱的顺序。

问题 2：根据附件提供的实测训练集中的测试基本信息，特别是节点间 RSSI 信息和门限信息，结合问题 1 中对 AP 发送机会的分析，对测试中 AP 发送数据选用最多次数的（MCS, NSS）进行建模，并通过测试集（仅提供模型输入信息）预测（MCS, NSS）。

问题 3：结合问题 1 和问题 2 的分析，对系统吞吐量进行建模，并通过测试集预测网络吞吐量。

2 模型假设与符号说明

2.1 模型假设

1. 假设题目所给的数据真实可靠；
- 2.
- 3.

注意：假设对整篇文章具有指导性，有时决定问题的难易。一定要注意假设的某种角度上的合理性，不能乱编，完全偏离事实或与题目要求相抵触。注意罗列要工整。

2.2 符号说明

（对文章中所用到的主要数学符号进行解释）

.....

尽可能借鉴参考书上通常采用的符号，不宜自己乱定义符号，对于改进的一些模型，符号可以适当自己修正（下标、上标、参数等可以变，主符号最好与经典模型符号靠近）。对文章自己创新的名词需要特别解释。其他符号要进行说明，注意罗列要工整。注意格式统一，不要出现零乱或前后不一致现象，关键是容易看懂。

符号	意义
a	符号 1 的意义
b	符号 2 的意义
c	符号 3 的意义符号 3 的意义
d	符号 4 的意义
e	符号 5 的意义
f	符号 6 的意义符号 6 的意义
g	符号 7 的意义
h	符号 8 的意义
i	符号 9 的意义符号 9 的意义
k	符号 10 的意义
l	符号 11 的意义

3 问题分析

主要是表达对题目的理解，特别是对附件的数据进行必要分析、描述（一般都有数据附件），这是需要提到分析数据的方法、理由。如果有多个小问题，可以对每个小问题进行分别分析。（假设有 2 个问题）

3.1 问题 1 的分析

对问题 1 研究的意义的分析。

问题 1 属于。。。。数学问题，对于解决此类问题一般数学方法的分析。

对附件中所给数据特点的分析。

对问题 1 所要求的结果进行分析。

由于以上原因，我们可以将首先建立一个。。。。的数学模型 I, 然后将建立一个。。。。的模型 II,。。。。。。。。对结果分别进行预测，并将结果进行比较.

3.2 问题 2 的分析

对问题 2 研究的意义的分析。

4 基本原理

4.1 接入机制

WLAN 中工作在同一信道的各节点共享信道，节点通过载波侦听多址接入/退避的机制避免冲突。接入过程可分为以下 3 个步骤：

- (1) 信道可用评估 (Clear Channel Assessment, CCA)：节点有数据要发送时，首先对工作信道进行固定时长的载波侦听，这个固定时长被称为分布式协调帧间距 (Distributed Coordination Function Inter-frame Space, DIFS)。如果 DIFS 时段内接收到的信号强度 RSSI 低于 CCA 门限，判断信道为空闲，否则判断信道为繁忙。
- (2) 随机回退：判断信道为空闲时，为避免节点间碰撞，每个节点根据其竞争窗口 (Contention Window, CW)，从 $[0, CW-1]$ 的整数均匀分布选取一个随机整数作为回退数，将其乘以时隙 slotTime ($9\mu s$)，称为随机回退时段。如果信道在随机回退时段保持空闲，则节点开始一次数据传输。如果期间信道变繁忙，节点将暂停回退，直到信道重新在 DIFS 内空闲，再继续前面的回退。随机回退采用二进制指数退避算法确定。
- (3) 数据传输：回退到 0 的节点发送一个数据帧，接收节点在成功接收到数据之后等待短帧帧间距 (Short Inter-frame Space, SIFS) $16\mu s$ ，回复确认帧 (Acknowledgement, ACK) $32\mu s$ 。

4.2 聚合

为了提升发送小包的效率，协议允许通过聚合一次发送多个具有相同目的地址的数据包。来自网络上层（如以太网逻辑链路层 LLC 层），具有相同接收地址的同服务类别的 MAC 服务数据单元 (MAC Service Data Unit, MSDU)，可以在 MAC 层顶端被拼接起来，加上一个共同的 MAC 头，封装成一个 MAC 协议数据单元 (MAC Protocol Service Data Unit, MPDU)，这个过程叫作聚合 MSDU (Aggregated MPDU, AMSDU)。组装好的 MPDU 在 MAC 层底端被聚合起来，每个 MPDU 前加一个短分隔符，随后聚合成为一个 PHY 协议数据单元 (PHY Protocol Data Unit, PPDU) 送给 PHY 层。每次发送数据时，缓存里有包则聚，遵循 AMSDU 个数最多为 7，AMPDU 个数最多为 21，同时，一个 PPDU 时长不超过 4.5ms。

在进行聚合时，由于聚合的子 MSDU 之间，及子 MPDU 之间需要插入冗余字段，实际一个 PPDU 传输时长里，仅 MSDU 的总字节数是有效传输数据，进行吞吐量计算时才被计入，因此，AMSDU 和 AMPDU 聚合的准确评估对吞吐量预测也有影响。如附图 3 (a) 所示，应用层报文 payload 长度为 1420 Bytes，包含 40 Bytes 长的 IP 协议头和 1380 Bytes 长的有效载荷。经过以太网有线网到达 Wi-Fi 的 MAC 层后，不聚合，加上 8 Bytes 的 LLC 层（以太网逻辑链路层）头，封装上 30 Bytes 的 MAC 头 (802.11 header)，形成一个 MPDU。如 4.1 所示，缓存里包增多时，进行聚合，MSDU 之间插入 0-3 字节的填充字节进行 4 字节

对齐，一个 MPDU 中的 MSDU 共用一个 30 Bytes 的 MAC 头。MPDU 进行聚合时，前面加上分隔符 delimiter，再进行 4 字节对齐，一个 PPDU 封装一个 PHY 头。

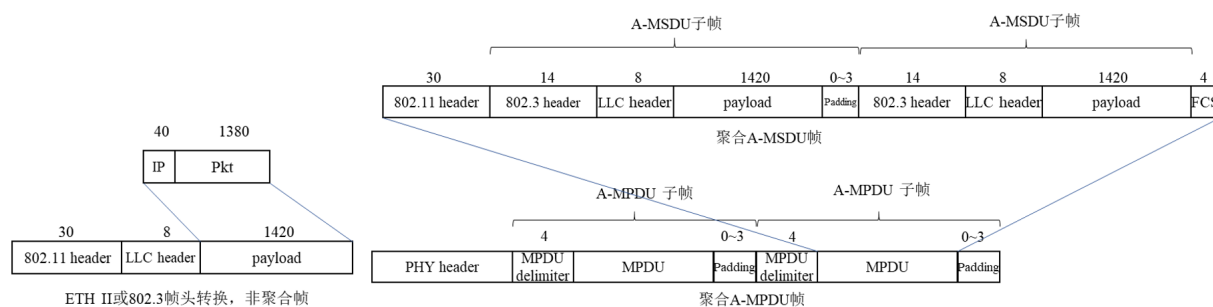


图 4.1 (a) 以太网帧不聚合 (b) AMSDU 聚合和 AMPDU 聚合

4.3 传输方式

当两个节点间的 $RSSI > ED$ 时，一个节点传输数据时，另一节点检测信道为繁忙，称两个节点能“听”到。那么多数情况下二者的数据传输将交替进行，只在偶然同时回退到 0 时同时开始数据传输。节点同时发送数据时产生碰撞，可能导致传输失败。这种在节点间可侦听情况下发生的传输称为同步传输，包括交替进行和同时发生的传输。

WLAN 实际部署中，受覆盖范围和可用信道数约束，工作在相同信道的同频 AP 间 RSSI 大多处于 $[PD, ED]$ 区间，一个区域内同频 AP 数量通常为 3 5 个。受业务类型影响，包长差异较大，如 4.2 所示，在某次同步传输过程中，先结束传输的 AP2 进行 CCA 时，由于已经错过侦听 AP1 的 Preamble，将采用 ED 作为 CCA 门限，从而判定信道为空闲，在回退到 0 时开始一次新的传输。称为异步传输。若 AP1 结束传输时，AP2 的第二次传输已开始，则 AP1 同样原因可能会开始第二次传输，那么两个 AP 可能进入长时间的异步传输状态。



图 4.2 一次异步传输和异步传输状态示意图

4.4 自适应调制编码算法

发送数据的 PHY 层速率 (PHY Rate) 由调制编码方案 (Modulation and Coding Scheme, MCS) 和空间流数 (Number of Spatial Stream, NSS) 表征，一组 (MCS, NSS) 对应一个 PHY Rate (见 4.2 节)。MCS 和 NSS 越高，发送时携带的有效比特数越多，即 PHY Rate

越高。节点的 PHY 层对信号进行解调时，要求一定的 SINR，其 SINR 越高，可支持成功解调的 MCS 和 NSS 越高。WLAN 采用经典的自适应调制解调（Adaptive Modulation and Coding, AMC）算法，根据信道条件动态调整发送数据使用的（MCS, NSS）。具体地，初始化时采用默认值（如 MCS6, NSS2）发送数据，并持续统计和更新丢包率（Packet Error Rate, PER）。当 PER 高于一定阈值时，代表当前 SINR 下，以该（MCS, NSS）发送数据的解调成功率低，需要降低 MCS 和 NSS；否则提高 MCS 和 NSS。AMC 算法自适应寻找最优的 PHY Rate 发送数据，使得吞吐量最大。

5 问题一的分析与求解

5.1 问题分析

5.1.1 问题分析

问题一主要研究 AP 发送的机会，旨在探索 WLAN 网络中的节点在测试场景下发送数据包的机会，即发送数据帧列的总时长 (seq_time)，并构建预测模型，该问题包括两个子问题：

(1) 分析训练集中各参数对 seq_time 的影响并排序：

首先，基于训练集中 `training_set_2ap_loc0_nav82` 等 11 个数据表，对 AP=2 和 AP=3 分类进行分析。对于每一种同频 AP 类型，研究表中的所提供的网络拓扑、业务流量、门限、节点间 RSSI 的测试基本信息，分析其中各参数对 AP 发送机会的影响，即对 seq_time 数值的相关性系数，并根据其影响性强弱给出重要性排序。

(2) 构建模型预测每个 AP 的 seq_time ：

根据训练集中 `training_set_2ap_loc0_nav82` 等 11 个数据表，分别对 AP=2 和 AP=3 构建模型预测测试集 `test_set_1_2ap` 和 `test_set_1_3ap` 中 AP 发送数据帧序列的总时长。这个模型将分别使用每一种同频 AP 的训练集进行训练，然后将模型应用于测试集中的 AP 参数，输出的 seq_time 表示每个 AP 在一定条件下发送数据机会的大小。

通过解决问题一的两个子问题，我们将更好地了解哪些因素能够影响 AP 发送数据的机会，并能够进一步研究各参数对 WLAN 网络吞吐量的影响。这对于无线局域网系统的进一步优化具有重要意义。

5.1.2 求解思路

针对上述问题，提出如下求解思路：

由题目可知，WLAN 部署后，节点基于信道竞争接入机制进行 CCA 和随机回退，并发送数据。在节点间 RSSI、信道竞争接入机制、CCA 门限、NAV 机制等共同影响下，节点以一定的概率发送数据。因此，节点在某一功率下的某一时刻听到的 **RSSI 处于什么范围（是否超过门限）**，会使得节点因此选择传输、等待（DIFS/回退）或静默，乃至导致可能的测试组内的传输方式（同步/异步），这一机制能够很大程度上影响对节点数据帧发送时长的预测，因此除了使用训练集给定的参数外，我们还需要从这些给定数据中挖掘对应新特征表示节点状态以更好地进行预测。

* 机制分析

通过简单观察训练集中各接收到 RSSI 参数的取值大小，发现其在测试中出现的可能范围于 -98 -29dBm 之间，结合文中提到的 CCA 门限以及测试中出现的两种 NAV 门限，某一 AP 节点在某一时刻根据接入机制对信道忙闲判断如下（对 RSSI 取值范围）：

1. $[-100, NAV)$: 信道空闲。
2. $[NAV, PD)$: 若未检测到 RTS, 信道空闲, 否则繁忙。
3. $[PD, ED)$: 若未检测到 RTS, 且未接收到 Preamble, 则信道空闲, 否则信道繁忙。
4. $[ED, -20)$: 信道繁忙。

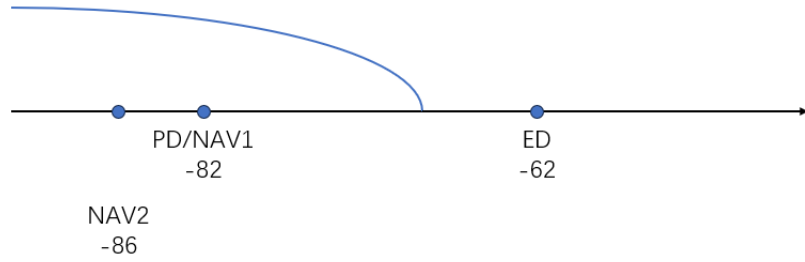


图 5.3 CCA 门限、NAV 门限示意图 单位 (dBm)

由此机制可总结:

(1) 当 RSSI 较小时, 即其他节点信号影响较小时, AP 根据对信道忙闲的判断更大可能性为空闲, 并发送数据。尽管此时仍有随即回退机制的限制, 但可以推断的是, 此时 AP 发送数据的机会更大, 并导致更长的 seq_time。因此, 我们可以根据对 AP 所听到的 RSSI 强弱判断, 将其作为新特征变量, 用于模型预测。

(2) 由题可知, 当两个节点间的 $RSSI > ED$ 时, 两节点互听, 此时进行同步传输, 每个 AP 交替抢到信道, 偶然同时发送, 发送时长基本上各占一半。当 RSSI 大多处于 $[PD, ED]$ 区间, 此时进行异步传输, 两个 AP 总是不互听时, 每个 AP 独立持续抢到信道, 发送时长基本占满, 此时 NAV 失效。因此按照 RSSI 处于门限区间, 将不同数据展示分类属性, 进一步推测其处的传输方法, 以根据此机制预测发送时长。

为了具象化这一区间关系, 我们将某一 AP 在一次测试中的 RSSI 信息数列中的每个值进行区间判断, 并分别得出在测试期间 RSSI 处于各门限区间的概率, 具体见 5.2.1 (3) 节。基于上述已知变量及新变量, 我们将 AP 发送机会的计算表示为:

$$asfda$$

5.2 子问题一: 各参数影响强弱分析

5.2.1 数据预处理

针对所提供的网络拓扑、业务流量、门限、节点间 RSSI 的测试基本信息包括共 35 个可选特征变量, 我们进行预处理, 进行以下步骤操作。

(1) 数据清洗, 我们将相同同频 AP 个数的所有训练表合并为一个主表, 对于表中观测到的异常空值, 和异常数值 (例如在 training_set_3ap_loc30_nav86 中部分列数据为空, 以

及 training_set_2ap_loc2_nav82 最下面两行的数据与其他数据相比明显异常，数据与表头内容对不上)，对此处理为删除整行数据避免对分析结果的影响。

(2) 变量删除，进行部分特征变量的移除，这些变量部分是常量（即在所有数据中保持不变），部分可以以行数奇偶性为隐藏属性表示，部分只是固有属性根据经验判断与因变量 seq_time 无关。删除变量包括：test_dur、loc_id、pkt_len、bss_id、ap_name、ap_mac、ap_id、pd、ed、sta_mac、sta_id 共 11 个变量。

(3) 创建新变量。题目中提到 AP 传输方式对 AP 发送时长具有直接影响，考虑到 pd, ed 表现为信道传输门限，它们决定了 AP 的 CCA 机制，进而导致不同的可能传输机制，因此我们采用差值变量，以概率形式表示在某一时间点 AP 的传输方式可能性。变换公式如下：

(4) 数据变换。对 protocol 变量，取 01 变换。对表示 RSSI 信息的 20 个变量，对数列中的各值取平均。例如：

(5) 主表拆分，观测到训练数据中，每个测试由 AP 个数行组成，考虑到这些行之间由于在同一个测试场景中，部分 RSSI 数据具有互补性，将其分割为 AP 个数个子表，对于 AP=2，按照行数对 2 的余数分为两个子表，对于 AP=3，按照行数对 3 的余数分为 3 个子表。每个子表对应一个 AP 在测试中的表现。此时观察到 test_id 在子表无意义，移除该变量。

(6) 相关性分析，接着我们对特征变量：协议，RSSI，nav，pro_pd, pro_ed 等进行相关性分析（热力图）再次剔除了一些具有强线性相关的特征变量。其中主要剔除的变量是各类 RSSI 中的 max 值和 mean 值，因为我们认为 sum 值相较于前两者有更多信息，更具有代表性。

(7) 子表合并，为了提升预测模型输入数据集的大小，我们将 AP=2 和 AP=3 的子表合并，形成一个新的数据集，以便于后续建模。合并过程中，对主表中互补的特征变量进行对应合并，例如，在 AP=2 的情境下，AP0 的特征 ap_from_ap_1_sum_ant_rssi 与 AP1 的特征 ap_from_ap_0_sum_ant_rssi 合并为 ap_from_ap_other_sum_ant_rssi。合并示意如图 5.4，具体合并规则请见附录。

5.2.2 模型构建

考虑到目标预测数据 seq_time 是连续变量，并且为了充分求解特征变量对目标变量的重要性，我们采用了三种不同的建模方法：随机森林、XGBoost 和岭回归。

1、随机森林 (Random Forest):

随机森林是一种集成学习方法，通过随机选择特征子集训练多个决策树，最终通过投票机制得到最终结果。

2、XGBoost:

XGBoost 是一种梯度提升算法，通过迭代训练多个决策树，每次训练都会根据上一次

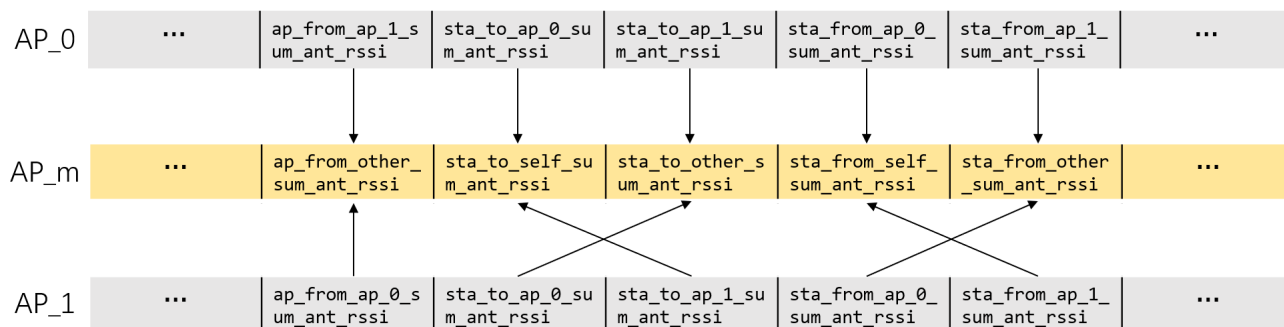


图 5.4 AP 子表合并示意图

的结果调整数据的权重，最终得到最终结果。

3、岭回归（Ridge Regression）：

岭回归是一种线性回归方法，通过加入正则项，可以有效防止过拟合。

三种方法的结果（分组条形图）均显示 xxxxx（可列表）。

5.3 子问题二：数据帧发送总时长预测

5.3.1 模型训练

然后我利用上面训练出的随机森林、XGBoost、Lasso 模型进行 seq_time 的预测。采用交叉验证法评古模型性能。

1、数据集划分

2、超参数选择

5.3.2 模型对比

在进行模型训练后，我们分别对其产生的结进行交叉验证并计算出准确率，结果如表 3.4 所示。

由表上表数据对比可知，随机森林模型的准确率最高，因此我们选择随机森林模型对 test 表进行预测。根据题目要求对测试集 test_set_1_2ap 和 test_set_1_3ap 中 AP 发送数据帧序列的总时长进行预测，并填写在对应数据位中。

数学建模求解算法示例：

算法 1 算法的名字

输入: input parameters A, B, C

输出: output result

```
1: some description 算法介绍
2: for condition do
3:     ...
4:     if condition then
5:         ...
6:     else
7:         ...
8: while condition do
9:     ...
10: return result
```

6 问题二的分析与求解

6.1 问题分析

6.1.1 问题简述

问题二主要研究 AP 在发送过程中选用的调制编码方案 (Modulation and Coding Scheme, MCS) 和空间流数 (Number of Spatial Stream, NSS) 表征。题目中要求, 通过分析实测训练集中的测试基本信息, 特别是节点间 RSSI 信息和门限信息进行分析, 针对测试中 AP 发送数据选用最多次数的 (MCS, NSS) 进行建模, 使用得到的预测模型预测 test_set_2_2ap 和 test_set_2_3ap 中的 (MCS, NSS)。

“选用最多次数的 (MCS, NSS)” 理解: 经观察, 训练集中每一行测试数据都给定了一组 (MCS, NSS), 我们认为这是 AMC 算法在测试初始阶段快速收敛后所得到的, 并在余下的大部分测试期间都使用这一组 (MCS, NSS) 用于解调信号, 即选用最多次数的 (MCS, NSS)。

通过得到这一数值对 (MCS, NSS), 可以一定程度上表示 AP 发送数据的 PHY 层速率 (PHY Rate), 其可以作为求解第三问系统吞吐量的前置影响条件。

6.1.2 求解思路

针对上述问题, 提出如下求解思路:

由题目可知, AP 发送数据时的 PHY 层速率 PHY Rate 取决于调制编码方案 MCS 和空间流数 NSS, 而这两者 (MCS, NSS) 又由信号解调时的 SINR 决定。WLAN 采用 AMC 自适应调制解调算法, 平衡当前 SINR 下所支持的 (MCS, NSS) 数值大小, 以保证丢包率不高于一定阈值。因此, 处理本题时, 需要通过测试信息, 初步计算信道 SINR 信息, 以进一步预测对应的 (MCS, NSS)。

AP 的 AMC 所选用的 (MCS, NSS) 不仅与 SINR 相关, 同时也与 AP 的传输方式相关。可以总结的, 两个 AP 之间的传输方式有三种可能: 同步传输、异步传输、同步异步混合传输。在各个传输方式下, 因为 STA 接收数据时受到的干扰来源不同, 导致 SINR 的计算方法也不同。因此, 为了计算信道 SINR 信息, 需要确定各测试数据所处的传输方式, 而这一过程需结合问题一中使用节点间 RSSI 信息和门限信息进行分类的方法。

综上所述, 为解决问题二, 我们首先根据节点间 RSSI 信息和门限信息对其所处**传输方式**进行分类, 随后根据其实际传输方式用不同的计算方式得到对应 **SINR 信息**, 进一步地, 利用 SINR 对 (MCS, NSS) 数值进行预测。

6.1.3 机制分析

分析各节点之间的 RSSI 信息和门限信息, 可以得出在 WLAN 信道争夺机制下, AP 可能处于如下三种传输方式:

(1) **同步传输**: 此时 AP 间 $RSSI > PD$ 且 $RSSI > NAV$, 两 AP 之间互听, 交替抢到信道, 偶然同时发送, 且 NAV 生效。STA 接受数据时干扰主要来自环境底噪, 此时 SINR 计算如下:

(2) **异步传输**: 此时 AP 间 $RSSI < PD$ 且 $RSSI < NAV$, 两 AP 之间不互听, 邻区干扰小, 同时进行数据发送, NAV 不生效。此时 STA 接收数据时可能受到较小的邻区干扰和环境底噪, 因此, SINR 计算如下:

(3) **同步与异步传输混合**: 此时两个 AP 的 $RSSI > PD$ 且 $RSSI < NAV$, 两 AP 之间互听, 邻区干扰大, NAV 不生效。当传输中出现由于错过 Preamble 时, 无法利用 NAV 更新通知其他 AP, 导致异步传输。此时 STA 接收数据时受到的干扰根据具体处于的传输方式决定, SINR 计算如下:

根据各节点间 RSSI 信息处于门限区间, 计算得出对应 SINR 值, 以预测对应的 (MCS, NSS)。

6.2 问题求解

6.2.1 数据处理

在第一问处理的基础上, 本问题进一步确定了因为 AP 间 RSSI 所处门限区间, 用以分类其在一次测试中的主要传输方式, 以此进一步计算 SINR 数值。

通过上题的数据处理, 我们创建了 3 个新的特征值 (在 3AP 情况是两组共 6 个): `pro_pd`、`pro_ed`、`pro_nav`, 分别表示了某 AP 在一次测试期间接收到来自其他组内 AP 的 RSSI 高于 PD、ED、NAV 门限的频率, 以此近似表示在以此次传输中各测量时刻该 AP 根据接受到的 RSSI 数值进行判断, 进而进行对应接入操作的概率。由于新特征呈现连续性, 取值处于 $[0, 1]$ 之间, 且具有不均匀性, 为了在进行门限区间分类时使用准确的阈值, 我们采用 K-means 聚类辅助分析。K-means 介绍:

聚类结果如上图。创建新的特征变量 'SINR', 根据分类类型, 使用 6.1.3 节提到的计算方法分别计算对应 SINR 值。

6.2.2 模型构建

6.2.3 结果分析

7 问题三的分析与求解

7.1 问题分析

7.1.1 问题简述

问题三主要研究 WLAN 网络系统吞吐量预测。题目要求结合问题一和问题二的分析，利用前两问的预测指标，包括 AP 发送数据帧序列的总时长、(MCS, NSS) 以及其他测试基本信息，共同构建预测模型，通过测试集 test_set_1_2ap 和 test_set_1_3ap 预测网络吞吐量。其中问题二所预测得到的 (MCS, NSS) 无法获得很高精度，允许采用实测中统计的数据帧真实 (MCS, NSS) 作为模型输入变量。

问题三是本课题的最终研究目标，前两问通过分析 WLAN 网络拓扑、节点间 RSSI、信道接入机制、干扰等因素对 WLAN 数据发送、速率的影响，本问题进一步地利用其预测结果作为重要影响因素与支撑，对 WLAN 系统吞吐量进行精确预测。基于该吞吐量预测模型，对 WLAN 进行优化，有望突破工业、教育、医疗等新场景，为用户提供极致的业务体验。

7.1.2 求解思路

根据题目背景，给出如下定义：

- 吞吐量：吞吐量指节点单位时间内成功发送的比特数。仅 MSDU 的总字节数是有效传输数据，进行吞吐量计算。
- 丢包率：发送数据帧的失败个数与总个数的百分比。
- 数据的帧序列的总时长：所有成功或失败的数据帧发送，帧序列时长从开始发送 RTS 到收到 ACK，或者超时。
- ppdu_dur(s)：一个数据帧的时长。一次测试里，统计每个数据帧的时长，取平均值。仅包括 PPDU 的传输时间，不包括 RTS，ACK 等。
- TCP 的 ACK 算是数据帧，算作吞吐量，个数是下行数据帧的 3:1 到 2:1 左右。根据丢包率改变传输速率。TCP ACK 的大小为 64Bytes。
- UDP 仅计算 AP 发往 STA 的数据帧，发送时间间隔服从泊松分布，发送数据的速率为 290 Mbps。

根据题目可知，WLAN 网络吞吐量主要取决于 AP 发送机会、发送时所选用的 PHY Rate 以及数据帧的比特数。

(1) 数据包聚合计算：

为了提升发送小包的效率，协议允许通过聚合一次发送多个具有相同目的地址的数据包。在一次数据包聚合 (AMSDU) 过程中，多个相同接收地址的同服务类别的 MAC 服务数据单元 (MSDU) 封装成一个 MAC 协议数据单元 (MPDU)。组装好的多个 MPDU 进一

步聚合（AMPDU），这一过程包括将多个 MPDU 聚合成一个 PHY 协议数据单元（PPDU）。此处 PPDU 即为在 WLAN 网络中进行传输的一个数据帧。

其中，组成 MPDU 的 MSDU 包长计算如式7.1（单位：Bytes）：

$$\begin{aligned} MSDU &= 802.3header + LLCheader + payload + padding \\ &= 14 + 8 + 1500 + 2 \\ &= 1524 \end{aligned} \quad (7.1)$$

其中，组成 PPDU 的 MPDU 包长计算如式7.2（单位：Bytes）：

$$\begin{aligned} MPDU &= 802.11header + MSDU \times AMSDU - padding + FCS \\ &= 30 + 1524 \times AMSDU - 2 + 4 \\ &= 32 + 1524 \times AMSDU \end{aligned} \quad (7.2)$$

最终，计算一个完整数据帧 PPDU 的包长（在 WLAN 协议（如 802.11 标准）中，**PHY 头（Physical Layer Header）** 的长度取决于所使用的物理层（PHY）类型，不是固定的。不同的 PHY 模式会有不同的 PHY 头大小。）（单位：Bytes）式7.3：

$$PPDU = (4 + MPDU + padding) \times AMPDU + PHY(\text{长度不固定}) \quad (7.3)$$

7.2 问题求解

7.2.1 数据预处理

7.2.2 模型构建

7.2.3 结果分析

8 表格和图形

8.1 表格

三线表

题目要求建立模型描述折叠桌的动态变化图，由于在折叠时用力大小的不同，我们不能描述在某一时刻折叠桌的具体形态，但我们可以用每根木条的角度变化来描述折叠桌的动态变化。首先，我们知道折叠桌前后左右对称，我们可以运用几何知识求出四分之一木条的角度变化。最后，根据初始时刻和最终形态两种状态求出桌腿木条开槽的长度。题目要求建立模型描述折叠桌的动态变化图，由于在折叠时用力大小的不同，我们不能描述在某一时刻折叠桌的具体形态，但我们可以用每根木条的角度变化来描述折叠桌的动态变化。首先，我们知道折叠桌前后左右对称，我们可以运用几何知识求出四分之一木条的角

表 8.1 某校学生升高体重样本

Numbers		身高	体重
Item	14	156	42
	16	158	45
3	14	162	48
4	15	163	50
平均	15	159.75	46.25

度变化。最后，根据初始时刻和最终形态两种状态求出桌腿木条开槽的长度。题目要求建立模型描述折叠桌的动态变化图，由于在折叠时用力大小的不同，我们不能描述在某一时刻折叠桌的具体形态，但我们可以用每根木条的角度变化来描述折叠桌的动态变化。首先，我们知道折叠桌前后左右对称，我们可以运用几何知识求出四分之一木条的角度变化。最后，根据初始时刻和最终形态两种状态求出桌腿木条开槽的长度。题目要求建立模型描述折叠桌的动态变化图，由于在折叠时用力大小的不同，我们不能描述在某一时刻折叠桌的具体形态，但我们可以用每根木条的角度变化来描述折叠桌的动态变化。首先，我们知道折叠桌前后左右对称，我们可以运用几何知识求出四分之一木条的角度变化。最后，根据初始时刻和最终形态两种状态求出桌腿木条开槽的长度。

某行业产量与生产费用的数据

表 8.2 某行业产量与生产费用的数据

企业编号	产量 (台)	生产费用 (万元)	企业编号	产量 (台)	生产费用 (万元)
1	40	130	7	84	165
2	42	150	8	100	170
3	50	155	9	116	167
4	55	140	10	125	180
5	65	150	11	130	175
6	78	154	12	140	185

研究生数学建模 2019 年 F 题结果示例

表 8.3 问题 1 结果 1 (左) 与问题 2 结果 (右)

数据集 1	数据集 1	数据集 2
AP_0	AP_1	AP_merge
503	503	163
294	200	114
91	80	8
607	237	309
540	170	305
250	278	123
340	369	45
277	214	160
B	397	92
	B	93
		61
		292
		B
104861	103518	109342

表 8.4 2AP 子表变量合并

	AP_0	AP_1	AP_merge
0	protocol	protocol	protocol
1	nav	nav	nav
2	ap_from_ap_1_sum_ant_rssi	ap_from_ap_0_sum_ant_rssi	ap_from_other_sum_ant_rssi
3	sta_to_ap_0_sum_ant_rssi	sta_to_ap_1_sum_ant_rssi	sta_to_self_sum_ant_rssi
4	sta_to_ap_1_sum_ant_rssi	sta_to_ap_0_sum_ant_rssi	sta_to_other_sum_ant_rssi
5	sta_from_ap_0_sum_ant_rssi	sta_from_ap_1_sum_ant_rssi	sta_from_self_sum_ant_rssi
6	sta_from_ap_1_sum_ant_rssi	sta_from_ap_0_sum_ant_rssi	sta_from_other_sum_ant_rssi
7	pro_pd	pro_pd	pro_pd
8	pro_ed	pro_ed	pro_ed
9	pro_nav	pro_nav	pro_nav

表 8.5 问题 1 结果 1 (左) 与问题 2 结果 (右)

数据集 1	数据集 1	数据集 2	数据集 1	数据集 1	数据集 2
A 问题 1	A 问题 1	A 问题 1	A 问题 2	A 问题 2	A 问题 2
503	503	163	503	503	163
294	200	114	294	200	114
91	80	8	91	80	8
607	237	309	607	237	309
540	170	305	540	170	305
250	278	123	250	278	123
340	369	45	340	369	45
277	214	160	277	214	160
B	397	92	B	397	92
	B	93		B	93
		61			61
		292			292
		B			B
104861	103518	109342	104917	103563	109427

表 8.6 问题 3 结果

数据集 1	数据集 1	数据集 2 (无问题点)		数据集 2 (有问题点)	
A 问题 3	A 问题 3	A 问题 3		A 问题 3	
503	503	169	73	169	73
69	69	322	249	322	249
506	506	270	274	270	274
371	371	89	12	89	12
183	183	236	216	236	216
194	194	132	16	132	16
450	450	53	282	53	282
286	113	112	84	112	141
485	485	268	287	268	291
B (9D)	248	250	99	250	161
	B (10D)	243	B (21D)	243	B (21D)
104861m	103518m		168924m		161650m

8.2 图形

图形并列

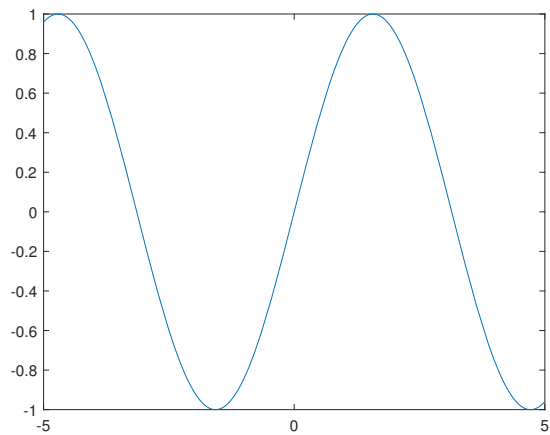


图 8.5 fig1

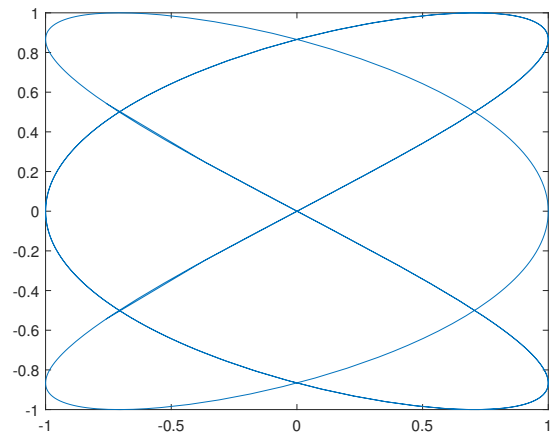


图 8.6 fig2

这是一个算法流程图

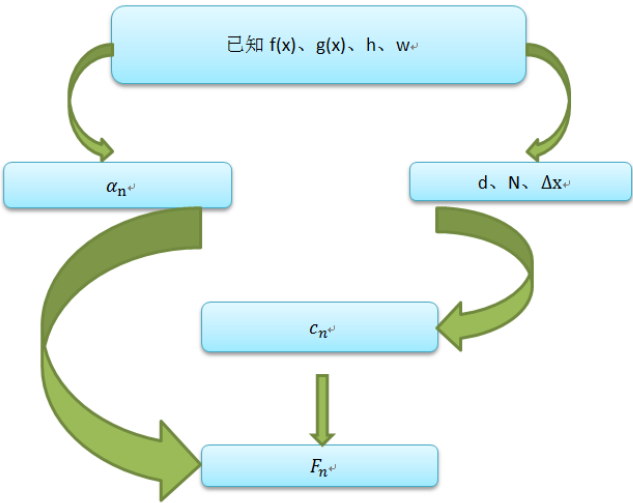


图 8.7 算法流程图

多图并排

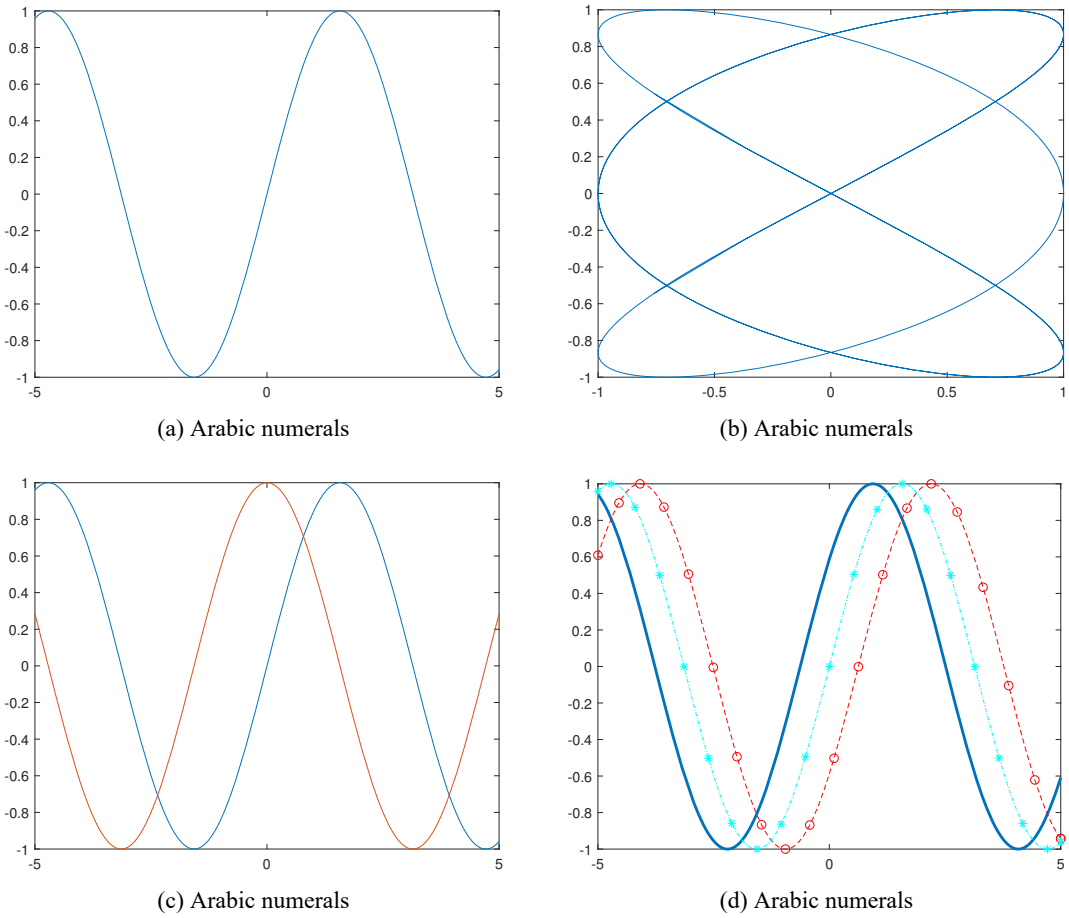


图 8.8 多图示例

8.3 问题三分析

题目要求制作软件的意思就是客户给定折叠桌高度、桌面边缘线的形状大小和桌脚边缘线的大致形状，将这些信息输入程序就得到客户想要的桌子。我们在求解最优设计加工参数时，自行给定桌面边缘线形状（椭圆、相交圆等），桌脚边缘线形状，折叠桌高度，应用第二问的非线性规划模型，用 MATLAB 软件绘制折叠桌截面图，得到自己设计的创意平板折叠桌。

9 模型评价

这里是模型评价

[1, 2, 3, 4, 5]

参考文献

- [1] Mittelbach F, Goossens M, Braams J, et al., The \LaTeX Companion, 2nd ed., Reading, MA, USA: Addison-Wesley, 107-109, 2004.
- [2] Wright J, \LaTeX 3 programming: External perspective, TUGboat, 30(1):107-109, 2009.
- [3] Beeton B, Freytag A, Sargent III M, Unicode support for mathematics, <http://www.unicode.org/reports/tr25/>, 2018-07-21.
- [4] Vieth U, Experiences typesetting mathematical physics, Proceedings of EuroTeX, 13, 2009.
- [5] TeXer W, \LaTeX 工作室, <https://www.latexstudio.net/>.

附录 A MATLAB 源程序

A.1 第 1 问程序

code.m

```
clear all
kk=2;
[mdd, ndd]=size(dd);
while ~isempty(V)
    [tmpd, j]=min(W(i, V));
    tmpj=V(j);
    for k=2:ndd
        [tmp1, jj]=min(dd(1, k)+W(dd(2, k), V));
        tmp2=V(jj);
        tt(k-1, :)= [tmp1, tmp2, jj];
    end
    tmp=[tmpd, tmpj, j; tt];
    [tmp3, tmp4]=min(tmp(:, 1));
    if tmp3==tmpd,
        ss(1:2, kk)=[i; tmp(tmp4, 2)];
    else
        tmp5=find(ss(:, tmp4)~=0);
        tmp6=length(tmp5);
        if dd(2, tmp4)==ss(tmp6, tmp4)
            ss(1:tmp6+1, kk)=[ss(tmp5, tmp4); tmp(tmp4, 2)];
        else, ss(1:3, kk)=[i; dd(2, tmp4); tmp(tmp4, 2)];
        end
    end
    end
    dd=[dd, [tmp3; tmp(tmp4, 2)]];
    V(tmp(tmp4, 3))=[];
    [mdd, ndd]=size(dd); kk=kk+1;
end;
S=ss; D=dd(1, :);
```

附录 B Python 源程序

B.1 第 2 问程序

mip1.py

```
# This example formulates and solves the following MIP model:
# maximize
#      x + y + 2 z
# subject to
#      x + 2 y + 3 z <= 4
#      x + y      >= 1
#      x, y, z binary

# import gurobipy as gp
from gurobipy import * #GRB
try:
    # Create a new model
    m = Model("mip1")
    # Create variables
    x = m.addVar(vtype=GRB.BINARY, name="x")
    y = m.addVar(vtype=GRB.BINARY, name="y")
    z = m.addVar(vtype=GRB.BINARY, name="z")
    # Set objective
    m.setObjective(x + y + 2 * z, GRB.MAXIMIZE)
    # Add constraint: x + 2 y + 3 z <= 4
    m.addConstr(x + 2 * y + 3 * z <= 4, "c0")
    # Add constraint: x + y >= 1
    m.addConstr(x + y >= 1, "c1")
    # Optimize model
    m.optimize()
    for v in m.getVars():
        print('%s %g' % (v.varName, v.x))
    print('Obj: %g' % m.objVal)

except GurobiError as e:
    print('Error code ' + str(e.errno) + ': ' + str(e))

except AttributeError:
    print('Encountered an attribute error')
```