

分类号: TP391

10710-2020124011



长安大学

硕士学位论文

风机故障诊断系统中的
监测点优选与特征选择算法研究

梁雅博

导师姓名职称	冯兴乐 教授		
申请学位级别	工学硕士	学科专业名称	信息与通信工程
论文提交日期	2023 年 3 月 31 日	论文答辩日期	2023 年 6 月 5 日
学位授予单位	长安大学		

Research on monitoring point optimization and feature Selection Algorithm in Fan fault diagnosis system

A Thesis Submitted for the Degree of Master

Candidate: Liang Yabo

Supervisor: Prof. Feng Xingle

Chang'an University, Xi'an, China

答辩委员会名单

答辩委员会	姓名	职称	工作单位
主席	宋焕生	教授	长安大学
委员	张绍阳	教授	长安大学
委员	齐达巍	高级工程师	西安银河电信技术有限责任公司
委员	刘博	高级工程师	西安微控网络技术有限公司
委员	冯兴乐	教授	长安大学
秘书	陈亮亮	讲师	长安大学

摘要

高速公路的轴流风机是长大隧道通风系统的重要组成部分,预防性维修对于提高风机寿命、保障隧道安全运营具有重要意义。对存量风机增加故障诊断系统,实时检测并反馈设备的功能和性能,可有助于实现预防性维修。但针对不宜加装基于有线传输的故障诊断系统的场景,无线方式存在的窄带宽传输条件是限制采样信息量的重要瓶颈。为此,本文研究风机状态监测点优选方法和故障特征选择算法,使采样信息量满足窄带宽传输的客观条件。本文的研究内容如下:

1. 状态监测点优选的目的是在带宽有限时优先采样关键监测点,缩减监测点数。当前基于有线传输的故障诊断系统仅通过有效性因素选择监测点,即只考虑安装的监测点是否有助于故障诊断,虽提高了诊断精度,但采样的大量信息以无线方式难以实时传输。对此,针对无线场景提出面向可靠性的监测点优选方法。该方法依据可靠性反向故障分析理论来确定风机故障的危害度、概率重要度、易检测度、易状态监测度及维修经济因素,然后对故障进行等级排序,根据带宽条件优先对排序高的故障设置监测点。最终,在风机的 33 类故障中确定电机轴承断裂、转子不平衡等故障需优先安装监测点。该方法确定的监测点为下文研究特征选择算法提供了研究对象。

2. 特征选择的目的是寻找量少、质优、通用的特征传输到后台进行故障诊断,在带宽有限时优先采样重要故障特征,减少单个监测点的采样信息量。当前基于有线传输的故障诊断系统常用的包裹式(Wrapper)方法所选特征量少、质优,但通用性差;过滤式(Filter)方法所选特征量少、通用性强但特征质量差。这两种方法单独应用到无线场景时会影响故障诊断模型的更新换代和诊断精度。为平衡特征质量与通用性的矛盾,本文设计 Filter 和 Wrapper 相结合的两步特征选择算法,使用 Filter 方法选择量少、通用性强的特征进行传输,接收到数据后使用 Wrapper 方法对特征进行二次优选,筛选出质优的特征。实验表明,本文算法在采样信息量缩减 65.385%时,故障诊断精度达到 98.33%,比单纯使用 Filter 方法的精度提高了 1.33%,证明所选特征具有量少、质优的优点。在通用性方面,本文算法在 k 近邻(K-Nearest Neighbor, KNN)和支持向量机(Support Vector Machine, SVM)故障诊断模型上的精度误差为 2.98%,低于 Wrapper 方法的 9.34%,通用性更强。本文方法在带宽有限时优先采样了关键监测点的量少、质优、通用性强的重要故障特征,减少了监测点的采样信息量。

关键词: 故障诊断, 窄带宽传输, 可靠性分析, 特征选择, 包裹式, 过滤式

Abstract

The axial fan of expressway is an important part of the ventilation system of long and big tunnel. Preventive maintenance is great significance to prolong the life of fan and ensure the safe operation of tunnel. Adding fault diagnosis system to the stock fan, real-time detection and feedback of the function and performance of the equipment, can be helpful to realize preventive maintenance. However, for the scenario that it is not suitable to install fault diagnosis system based on wired transmission, the narrow bandwidth transmission condition of wireless mode is an important bottleneck to limit the sampling information. Therefore, the optimization method of fan status monitoring points and fault feature selection algorithm are studied in this dissertation, so that the sampling information can meet the objective conditions of narrow bandwidth transmission. The research content of this dissertation is as follows:

1. The purpose of selecting condition monitoring points is to sample key monitoring points first and reduce the number of monitoring points when the bandwidth is limited. The current fault diagnosis system based on wired transmission only selects monitoring points based on validity factors, that is, only considers whether the installed monitoring points are helpful for fault diagnosis. Although the diagnosis accuracy is improved, it is difficult to transmit a large amount of sampled information in real time through wireless. Therefore, a monitoring point optimization method based on reliability analysis is proposed for wireless scenarios. Based on the reliability reverse fault analysis theory, this method determines the risk degree, probability importance degree, easy detection degree, easy condition monitoring degree and maintenance economic factors of fan faults. Then, the faults are ranked, and monitoring points are set for the faults with high ranking priority according to the bandwidth conditions. Finally, among the 33 types of fan faults, it is determined that the motor bearing fracture, rotor unbalance and other faults need to install monitoring points first. The monitoring points determined by this method provide a research object for the following research feature selection algorithm.

2. The purpose of feature selection is to find features with small quantity, high quality and general purpose and transmit them to the background for fault diagnosis. When the bandwidth is limited, important fault features are sampled preferentially to reduce the sampling information of a single monitoring point. Currently, Wrapper methods commonly used in fault diagnosis systems based on wired transmission. The features selected by Wrapper method are less in quantity and better in quality, but poor universality. The features selected by the Filter method are few and universal, but the feature quality is poor. When these two methods are applied to wireless scenario separately, the updating of fault diagnosis model and diagnosis

accuracy will be affected. In order to balance the contradiction between feature quality and universality, this dissertation designed a two-step feature selection method combining Filter and Wrapper. The Filter method was used for transmission of features with less selection and strong universality, and the Wrapper method was used for secondary optimization of features after receiving data. Experimental results show that the fault diagnosis accuracy of the proposed algorithm reaches 98.33% when the sampling data is reduced by 65.385%, which is 1.33% higher than that of the Filter method alone, proving that the selected features have the advantages of less quantity and better quality. In terms of universality, the accuracy error of this algorithm on K-Nearest Neighbor (KNN) and Support Vector Machine (SVM) fault diagnosis model is 2.98%, which is lower than 9.34% of Wrapper method. The proposed method preferentially samples important fault features such as small quantity, good quality and strong universality of key monitoring points when the bandwidth is limited, and reduces the sampling information of monitoring points.

Key words: Fault diagnosis, Narrow bandwidth transmission, Reliability analysis, Feature selection, Wrapper, Filter

目 录

第一章 绪论.....	1
1.1 研究背景及意义.....	1
1.2 国内外研究现状.....	2
1.2.1 基于无线传输的风机故障诊断系统研究现状	2
1.2.2 状态监测点优选方法研究现状.....	5
1.2.3 特征选择算法研究现状.....	6
1.3 研究内容与章节安排.....	8
1.3.1 本文研究内容.....	8
1.3.2 本文章节安排.....	9
第二章 基于无线传输的风机故障诊断系统方案设计及相关技术概述	11
2.1 基于无线传输的风机故障诊断系统方案设计	11
2.1.1 终南山隧道通风系统分析.....	11
2.1.2 故障诊断系统物理拓扑结构.....	12
2.1.3 故障诊断系统逻辑功能结构.....	14
2.2 监测点优选方法概述.....	16
2.2.1 可靠性理论概述.....	16
2.2.2 AHP 方法概述.....	18
2.3 特征选择算法概述.....	20
2.3.1 Filter 方法概述	20
2.3.2 Wrapper 方法概述.....	23
2.4 本章小结.....	25
第三章 面向可靠性的风机状态监测点优选方法	27
3.1 引言.....	27
3.2 风机结构功能分析.....	27
3.3 面向可靠性的多因素计算	28
3.3.1 基于 FMECA 的危害度计算.....	28
3.3.2 基于 FTA 的概率重要度计算	32
3.3.3 基于调研的因素计算.....	39
3.4 基于 AHP 的多因素权重分配模型	41
3.5 状态监测点优选结果.....	43
3.6 本章小结.....	45
第四章 基于 Filter 和 Wrapper 相结合的两步特征选择算法研究..	47
4.1 引言.....	47

4.2 特征选择算法研究对象分析	47
4.2.1 结构原理分析.....	47
4.2.2 常见故障形式及故障演变.....	48
4.2.3 本文特征选择算法的研究对象.....	48
4.3 基于 Filter 和 Wrapper 相结合的两步特征选择算法.....	48
4.3.1 两步特征选择算法设计.....	49
4.3.2 实验设计思路.....	51
4.4 实验结果分析.....	53
4.4.1 样本数据及特征集构建.....	53
4.4.2 最优特征子集生成.....	54
4.4.3 实验结果.....	57
4.5 本章小结.....	60
第五章 总结与展望	63
5.1 本文工作总结.....	63
5.2 未来工作展望.....	63
参考文献.....	65
附录 A.....	69

第一章 绪论

1.1 研究背景及意义

随着我国国民经济的不断发展,居民车辆保有量的不断增加,我国高速公路建设事业也获得了长足的发展。隧道的出现解决了制约高速公路发展的线路顺直、坡度平缓、路面宽敞的建设难题,在高速公路的快速发展过程中发挥着举足轻重的作用。在长大隧道中,竖井和轴流风机的组合提升了隧道内的通风功能,弥补了射流风机改善隧道空气能力不足的缺陷。但是由于隧道内外气压、烟囱效应以及风机所处环境中粉尘污染和热辐射的影响,轴流风机容易发生突发故障。故障引起的非计划停机使得隧道内氧气浓度下降、车辆排放尾气浓度上升,对驾驶员的人身安全产生了一定的隐患。因此,需要对轴流风机进行故障检测及诊断,在故障初期对设备进行及时的预防性维修,降低非计划停机的概率,提高风机寿命,保障隧道的安全运营。由于数据传输稳定、传输带宽大等优点,隧道中现役风机的故障诊断系统在数据传输时使用的是有线方式。但针对不宜加装有线传输线路的场景,研究基于无线传输的风机故障诊断系统可有效提高设备的智能化水平,保障风机的安全运行。因此,本文对暂未使用、不宜加装有线传输线路的存量风机在布设基于无线传输的故障诊断系统时存在的关键问题进行研究。

近年来随着物联网技术、云计算等的发展,无线传输方式基于应用场景拓展了许多通信技术。目前广泛使用的无线传输技术有第四代移动通信技术(4th Generation Mobile Communication Technology, 4G)、第五代移动通信技术(5th Generation Mobile Communication Technology, 5G)、ZigBee、远距离无线电(Long Range Radio, LoRa)技术、无线局域网(Wireless Fidelity, WiFi)等,其各有优劣且在隧道照明、车道控制、环境监控等各个业务子系统中均有广泛应用,但是在风机故障系统中应用较少。因为无线方式存在的窄带宽传输条件是限制采样信息量的重要瓶颈,而风机故障需要采样的信息量较多,具体来说:

从空间上来看,风机在发生故障时,往往会在设备的多个部位产生高频振动,有线传输故障诊断系统为了精准捕捉到多个部位存在的高频振动,往往会在风机机身不同角度、底座、轮毂、电机等多个部位安装大量的监测点采集其状态^[1],其采集对象产生的大量数据难以通过单个无线设备进行实时传输,间歇性传输将会损失故障信息,一定程度上降低了故障诊断的精度,增加无线设备往往带来了信道同频干扰以及建设成本问题。从时间上来看,奈奎斯特采样定理要求^[2],在不损失故障信息的前提下,采样频率要大

于信号最高频率的两倍以上，在面对高频振动时，单个监测点需要采样的信息量大，导致需要传输的信息量给窄带宽传输的无线故障诊断系统带来了数据传输挑战。

针对建立基于无线传输的风机故障诊断系统时存在的传输带宽窄与采样信息量多的矛盾，本文研究风机状态监测点优选方法和特征选择算法，在带宽有限的条件下优先采样关键监测点的重要故障特征，缩减监测点数并减少单个监测点的采样信息量，从而使得采样的信息量能够满足窄带宽传输的客观条件。风机状态监测点优选方法依据可靠性反向故障分析理论，通过对风机故障进行分析，筛选出重要度较高的故障优先安装状态监测点，以此来缩减监测点数；特征选择算法选择数量少、质量高、适用于不同故障诊断模型的故障特征进行采样，在有限的带宽条件下采样重要特征，减少单个监测点需要采样的信息量。通过本文方法，使得建立的基于无线传输的故障诊断系统实现了对风机重要故障的监测，对不宜加装有线传输线路的风机提高智能化水平具有重要意义。

1.2 国内外研究现状

1.2.1 基于无线传输的风机故障诊断系统研究现状

1. 无线传输在隧道中的应用现状

目前，工业领域广泛使用的无线传输方式有 4G、5G、ZigBee、LoRa、WiFi 技术等。4G、5G 依赖于移动通信基站，在隧道内的偏僻地方通信质量不佳，因此在隧道中应用较少。ZigBee、WiFi、LoRa 技术由于其可以自组网的特性使得数据通信不受基站的限制，被广泛应用到隧道中设备的无线数据传输。本节对 ZigBee、WiFi、LoRa 在隧道中的应用现状进行分析。

ZigBee 是一种低速短距离传输的无线网上协议，底层是采用 IEEE 802.15.4 标准规范的媒体访问层与物理层^[3]，其工作在 2.4GHz 频段，传输速率为 20~250kb/s，传输距离为 10~75m。大量学者及技术人员将 ZigBee 技术应用到隧道中设备的数据传输过程中。2018 年潘国荣等^[4]将 ZigBee 应用到隧道内机器人通信系统设计。同年曾熙鸿等^[5]在隧道断面监测系统中采用 ZigBee 与 GPRS 结合的数据传输方式，利用 GPRS 远距离传输的特性，减少了 ZigBee 的中继节点数量，提高了数据传输稳定性，拓展了 ZigBee 在隧道中的使用范围。文献[6-8]使用 ZigBee 进行隧道中的设备除冰、照明控制、环境温湿度校准等工作，均验证了 ZigBee 技术在隧道中应用的可行性。

WiFi 是一种短距离的无线局域网数据传输技术，其组网结构由无线网卡和众多基站接入点（Access Point，AP）组成，能够在其覆盖范围内形成蜂窝小区，并且基站在各蜂窝小区内可以实现全网覆盖^[9]。大量学者及技术人员将 WiFi 技术应用到隧道中设备

的数据传输中。田青等^[10]利用 WiFi 技术进行隧道中应急多媒体通信。覃中顺等^[11]使用 WiFi 技术进行隧道内的人员定位，其通过让工作人员手持 WiFi 终端设备，计算终端设备与 AP 节点之间的信号强度和位置来判断人员的位置。两人的研究均验证了 WiFi 技术在隧道中应用的可行性。

LoRa 是低功率广域网络（Low-Power Wide-Area Network, LPWAN）通信技术中的一种，传输距离为 2-5km，其特点主要有低功耗、远距离、抗干扰这三种。王大涛等^[12]使用 LoRa 建立隧道内的低功耗健康监测系统，验证了在隧道中使用 LoRa 技术的可行性。

总的来说，随着物联网技术、云计算等的发展，无线传输方式在隧道中被广泛使用，为建立基于无线传输的风机故障诊断系统提供了技术支持。

2. 风机故障诊断方法研究现状

随着通风系统的快速发展，国内外对于风机的故障诊断技术也在急速进步。根据研究进程可将风机故障诊断技术划分为三个时期：基于知识^[13]的时期、基于解析模型^[14]的时期以及基于数据驱动^[15]的时期。基于知识的方法需要较强的专业知识以及相关经验；基于解析模型的方法由于受到模型误差和未知因素干扰以及所建模型内部耦合的影响使得建模较难，因此，这两种方法在风机故障诊断应用中逐渐减少。

基于数据驱动的方法是近年来最常用的风机故障诊断方法，其以风机上的各种传感器采集到的数据为基础，经过数据挖掘^[16]以及数据分析等相关方法实现数据中深层的信息显现，从而实现故障检测以及故障诊断^[17]。对于风机来说，由于风机所处高噪声环境以及旋转机械的性质，最常用的是振动监测的方法，通过采集风机振动信号，来判断风机运行状态以及故障位置、故障程度^[17]，本文所说的故障诊断方法均是基于振动信号的故障诊断方法。

基于数据驱动的故障诊断方法从最开始计算机技术不成熟时的人工数理统计分析方法，发展到计算机辅助处理的信号处理方法，再到主要依靠计算机进行数据处理的人工智能的方法。由于计算效率高、故障诊断精准、节约人力成本等优点，基于人工智能的风机故障诊断方法是当前研究最热门的方法，其主要分为两大类：基于机器学习的方法和基于深度学习的方法。

机器学习算法赋予机器像人一样学习新知识的能力，使得机器可以对不同类的风机故障数据进行思考并分类，最终建立起故障诊断模型，对风机进行故障诊断。常用的方法有支持向量机（Support Vector Machine, SVM）、k 近邻（K-Nearest Neighbor, KNN）算

法等。文献[18]通过提取风机轴承的时域、频域等多尺度特征作为 SVM 算法的学习样本，建立起基于 SVM 模型的风机轴承故障诊断系统，最终实现了轴承的故障诊断。文献[19]将振动数据进行小波降噪处理后提取时频域特征和本征模态特征作为 KNN 算法的学习样本，建立起基于 KNN 模型的风机故障诊断系统，最终完成了风机关键零部件的故障诊断。

深度学习的方法作为机器学习方法的重要分支，其通过增加网络层数，使得建立的故障诊断模型能够捕捉到故障样本中的深层次特征，从而提高了模型的故障诊断精度，常用的方法有：卷积神经网络（Convolutional Neural Networks, CNN）^[17]、长短期记忆网络（Long Short-Term Memory, LSTM）^[20]等。但是深层次特征的捕捉需要大样本的情况，样本数据量较少时，模型容易出现过拟合的情况，即模型在训练数据上表现很好，但在测试数据上表现不佳。这是因为模型在训练数据上过度拟合，学习到了训练数据中的噪声和细节，而忽略了数据中的真正规律和特征。隧道风机在实际使用时，风机正常运行数据多、故障数据少的特点使得能够生成的故障样本较少，因此深度学习的方法面对小样本情况下难以发挥有效的作用。目前深度学习的方法一定程度上局限在实验室场景下，其通过人为方式生成故障去研究设备的故障诊断。

本文以机器学习算法中的 SVM、KNN 算法为故障诊断模型，在此模型基础上，验证本文的风机故障特征选择算法的可行性。

3. 无线传输在隧道风机故障诊断系统中的应用

目前，将无线传输技术应用到隧道风机故障诊断系统中的案例还相对较少。主要原因是：1）隧道环境影响无线数据传输稳定性。轴流风机所处的隧道环境使得数据在传输的过程中受到固定的多径效应影响以及变化的车载设备同频干扰、温度影响、电机噪声等隧道环境干扰，为了数据传输的稳定性，风机故障诊断系统在建设初期很少考虑使用无线方式进行数据传输。2）存在传输带宽窄与采样信息量多的矛盾。风机故障需要的监测点位多、单点需要采样的信息量大，这两个原因制约了窄带宽传输的无线传输技术在风机故障系统中的应用。

近年来伴随着物联网与云计算技术的发展，无线传输在技术上获得了极大的提高，在数据传输稳定性方面获得极大进步。ZigBee、WiFi、LoRa 等多种无线传输方式均实现了隧道内数据的稳定传输，隧道环境对无线数据传输稳定性的影响，已经不再成为制约无线传输技术应用到风机故障诊断系统中的关键因素。但是传输带宽窄与采样信息量多之间矛盾始终限制着无线传输技术应用到隧道风机故障诊断系统中。

本文借鉴基于有线传输的风机故障诊断系统的既有研究成果,针对建立基于无线传输的风机故障诊断系统时存在的传输带宽窄与采样信息量多的矛盾,通过研究风机状态监测点优选方法以及特征选择算法来降低采样信息量,使采样信息量满足窄带宽传输的客观条件。

1.2.2 状态监测点优选方法研究现状

当前状态监测点优选方法的研究主要基于不同的考虑因素对监测点的安装个数、安装位置进行最优配置。欧阳森等^[21]人针对电压暂降对供用电双方造成的危害,从经济指标和故障不确定性的角度对监测点安装个数进行了研究。刘一凡^[22]针对空调供暖状态下室内温度监测质量过分依赖于监测点个数的情况,基于环境因素对监测点个数及安装位置进行研究。上述研究指出状态监测点的确定往往针对不同的应用场景及需求选择不同的考虑因素。

在隧道风机故障诊断领域,基于有线传输的故障诊断系统在设置状态监测点时仅需增加成本较低的传感器,其对状态监测点的确定往往仅考虑有效性因素,即只考虑安装的状态监测点是否有助于故障诊断,在这种情况下,危害度小、故障频率低、故障后易检测的故障也安装了状态监测点,因此往往安装的监测点较多。

以风机叶片裂纹损伤故障为例,该故障发生时会导致风机风力减弱、寿命缩短、长期会导致风机停机,故障危害度大。但是由于出、入口防护网的存在阻挡了异物撞击事件的发生,使得叶片裂纹损伤故障发生概率较低。同时在风机定期维修时,维修人员会使用专业设备对叶片进行故障检测,能够很好的进行故障排查,易检测度较高,对风机叶片安装状态监测点的预期收益相对较低。而电机中的轴承等零部件发生故障时将导致风机出现非计划停机,故障危害度高。同时由于风机工作环境潮湿、工作时间长、工作负荷大,导致轴承容易发生故障,故障概率较高。最后轴承位于风机电机内部,位置隐蔽,维修人员难以排查故障,故障不易检测,因此对电机安装状态监测点来监测轴承是否发生故障的预期收益相对较高。

基于无线传输的风机故障诊断系统中需要尽可能少的状态监测点,还需要考虑其他因素缩减监测点数,避免在有限带宽资源条件下采样到危害度小、故障频率低、故障后易检测的故障,消耗了有限的带宽资源。

1.2.3 特征选择算法研究现状

在对设备进行故障诊断过程中,受到设备自身结构和实际工况的影响,提取的故障信号呈现非线性、非平稳的特点,进而导致难以使用信号的单一特征对设备进行精确的故障诊断。为提高故障诊断的精度,工程上往往从信号时域、频域、时频域等角度提取一系列故障特征构成特征集来进行故障诊断。但是特征集中往往存在冗余特征,冗余特征一方面相互影响^[23]导致诊断精度下降,另一方面也消耗了基于无线传输的风机故障诊断系统中的有限带宽资源。因此,通过合理的方法对特征集进行选择,降低冗余特征对诊断精度以及传输带宽的影响,是对设备进行故障诊断的必要步骤之一。

特征选择也称为特征子集选择,其目的在于“取其精华,去其糟粕”,“精华”指的是特征集中能更好的反映风机故障信息的敏感特征,“糟粕”指的是冗余特征。特征选择算法的目标是从特征集中选取一个最优特征子集,以剔除不相关或者冗余的特征,从而达到减少特征个数、提高故障诊断准确度的目的^[24]。特征选择在几十年来一直是方法论和实践的研究课题,被应用于图像识别^[25]、图像检索^[26]、文本挖掘^[27]、入侵检测^[28]、生物信息数据分析等多个领域。故障诊断领域常用的特征选择方式共有两种类型:包裹式(Wrapper)、过滤式(Filter)^[29]。

Wrapper 方法是将特征选择过程和故障诊断算法相结合^[30],特征选择过程通过全局搜索、随机搜索或者序列搜索的策略预选一个特征子集,故障诊断算法基于特征子集来计算诊断精度去评价所选特征子集的优度,诊断精度越高,认为所选的特征子集越好。它的优点在于诊断精度高、去除冗余特征的能力强;缺点是计算量大、计算复杂度高,因此难以适用于大规模、特征量庞大的数据集,同时 Wrapper 方法与故障诊断算法的耦合度较高,当故障诊断模型发生更换时需要重新选择特征子集以确保故障诊断的精度,通用性差。风机故障诊断系统提取的特征数据量有限,以分类精度为评价指标的 Wrapper 方法在这样的场景下能够很好的提取特征子集,剔除特征集中的冗余特征,而计算量大、计算复杂度高的缺点造成的影响并不显著。因此在风机故障诊断领域中,Wrapper 方法被广泛使用。文献[31]使用随机搜索策略中的 GWO 算法预选特征子集,以 SVM 故障诊断模型的分类精度来评价特征子集的优度,筛选出了低冗余度、高故障辨识度的低维故障特征子集,极大的提高了地铁塞门故障诊断的准确性。文献[32]使用全局搜索策略中的穷举法和随机搜索策略中的粒子群算法相结合的方式预选特征子集,最终完成了设备的故障诊断。除以之外,文献[33-37]使用随机搜索策略中的萤火虫算法、布谷鸟搜索算法、遗传算法、蚁群优化算法、麻雀搜索算法进行了特征子集的选择,均利用故障诊断

模型的分类精度来剔除冗余特征，实现了高精度的故障诊断。总的来说，在故障诊断领域中，Wrapper 方法由于去除冗余特征的能力强、故障诊断精度高的优势被广泛使用；其计算量大、计算复杂度高的缺点在特征数量较少的故障诊断领域中造成的危害不明显，因此缺点很少被考虑。

Filter 方法通过分析特征值自身的特点来对特征值进行排序，并筛选出排序靠前的特征值作为特征子集。它的特点是计算效率高、复杂度低且计算过程不依赖故障诊断算法、通用性强；缺点是对冗余特征的去除能力有限，故障诊断精度较低。常用的方法有 Fisher Score 方法、最大信息系数（The Maximal Information Coefficient, MIC）、最大相关最小冗余（Max-Relevance and Min-Redundancy, mRMR）等方法。风机故障诊断系统提取的特征数据量有限，采用 Filter 的方式难以体现计算效率高的优势，而诊断精度较低的缺点成为制约 Filter 方法在故障诊断领域中发展的关键因素，因此很少被使用。Filter 方法常应用于大规模数据集进行特征选择，以降低计算时间。文献[38]将 Fisher Score 方法应用到了大规模图像数据集的特征选择过程中，极大的提高了人脸识别模型的计算效率。文献[39]将 Fisher Score 和 mRMR 方法应用到遥感影像分类的特征选择过程中，利用 Fisher Score 评价特征重要度，利用 mRMR 算法评价特征与类之间的相关度、特征相互之间的冗余度，从一定程度上提高了图像分类的计算效率和诊断精度。总的来说，Filter 由于计算效率高的优势被广泛应用于大规模数据集的特征选择过程中，在故障诊断领域这样小规模、特征数量少的数据集中难以发挥其计算效率高、复杂度低的优势。

Wrapper 方法与 Filter 方法的优缺点对比如表 1.1 所示。

表 1.1 Wrapper 与 Filter 方法优缺点

特征选择算法	优点	缺点
Wrapper 方法	诊断精度高、去除冗余特征的能力强	计算量大、计算复杂度高、通用性差
Filter 方法	计算量小、计算复杂度低、通用性强	诊断精度低，去除冗余特征的能力有限

在建设基于无线传输的风机故障诊断系统时，当前广泛使用的 Wrapper 方法因为其通用性差的缺点存在一定的缺陷，具体来说：随着技术的发展，当前计算精度高、计算时间短、模型体积小的故障诊断模型将会被计算精度更高、计算时间更短、模型体积更小的模型取代。Wrapper 方法通用性差的缺点导致当故障分类模型发生更换时需要重新选择特征子集。但是，窄带宽传输条件下由于带宽限制只传输 Wrapper 特征选择后的特征子集到隧道管理所的后台服务器中，而新数据会逐渐覆盖旧数据的隧道数据存储特点将导致原有的用于训练模型的风机历史数据被 Wrapper 所选特征逐渐覆盖。在这样的场景下，新的故障诊断模型由于没有其他特征子集，难以实现模型的更新换代。因此，

Wrapper 方法通用性差的缺点极大的限制了其在窄带宽传输条件下的隧道风机故障诊断系统中的应用。Filter 方法虽然通用性较强,可以选择更通用的特征进行传输,但是其故障诊断的精度有限。

在这样的场景下,本文研究适用于基于无线传输的风机故障诊断系统的 Filter 和 Wrapper 相结合的两步特征选择算法,在数据传输前使用 Filter 方法进行特征初选以提高通用性,数据传输之后使用 Wrapper 方法提高诊断精度,充分利用两种算法的优点。

1.3 研究内容与章节安排

1.3.1 本文研究内容

本课题依托于河北省交通运输厅的《高速公路机电系统全生命周期监管技术研究》项目,研究基于无线传输的风机故障诊断系统中的相关算法。具体如下:

对隧道中不宜加装有线传输线路的存量风机增加基于无线传输的故障诊断系统,可实时检测并反馈设备的功能和性能,提高风机的智能化水平,保障隧道的安全运营。但是无线方式存在的窄带宽传输条件是限制采样信息量的重要瓶颈。对此,本文以基于无线传输的风机故障诊断系统为研究对象,分别研究监测点优选方法和故障特征选择算法,目的是在带宽有限的条件下优先采样关键监测点的重要故障特征,使得采样的信息量能够满足窄带宽传输的客观条件。文章具体工作如下:

首先,对终南山隧道轴流风机所在的隧道通风系统进行充分分析,提出本文的基于无线传输的风机故障诊断系统设计方案。该方案从故障诊断系统的物理拓扑结构、逻辑功能结构阐述了针对终南山隧道的轴流风机如何建立基于无线传输的故障诊断系统,为下文研究监测点优选方法以及特征选择算法提供应用场景。

其次,在监测点优选方法中,基于有线传输的故障诊断系统在设置状态监测点时仅需增加成本较低的传感器,其对状态监测点的确定往往考虑有效性因素,即只考虑安装的状态监测点是否能够帮助进行故障诊断,在这种情况下往往安装的监测点较多。本文依托于可靠性分析理论中的反向故障分析理论,通过可靠性分析模型计算风机故障的危害度、概率重要度、易检测度、易状态监测度及维修经济因素这五种因素,然后使用 AHP 方法对这五种因素进行权重分配,最终获得故障的等级排序,根据带宽条件优先对排序高的故障设置监测点。该方法实现了带宽有限时优先采样关键监测点的目的,缩减了监测点数。

最后,在特征选择算法相关研究中,当前有线故障诊断系统常用的 Wrapper 方法所选特征虽能实现量少、质优的目标,但其通用性差。有线场景将所有的特征都传输到后

台进行存储,但无线场景仅传输 Wrapper 所选特征,隧道以新代旧的数据存储特点导致存储器中仅存储 Wrapper 所选特征。这样的数据存储差异导致 Wrapper 所选特征通用性差的缺点不会影响有线故障诊断系统,但会限制无线故障诊断系统中故障诊断模型的更新换代。Filter 方法所选特征虽量少、通用性强但所选特征质量差,导致故障诊断精度低。为平衡特征质量与通用性的矛盾,本文提出基于 Filter 和 Wrapper 相结合的两步特征选择算法。该方法使用 Filter 方法选择量少、通用性强的特征进行传输,接收到特征后使用 Wrapper 方法对特征进行二次优选,提高故障诊断的精度,实现有限的带宽条件下尽可能的采样重要特征的目的。该方法实现了在带宽有限时优先采样重要故障特征的目的,减少了单个监测点的采样信息量。

1.3.2 本文章节安排

文章章节安排如图 1.1 所示,具体的章节安排如下:

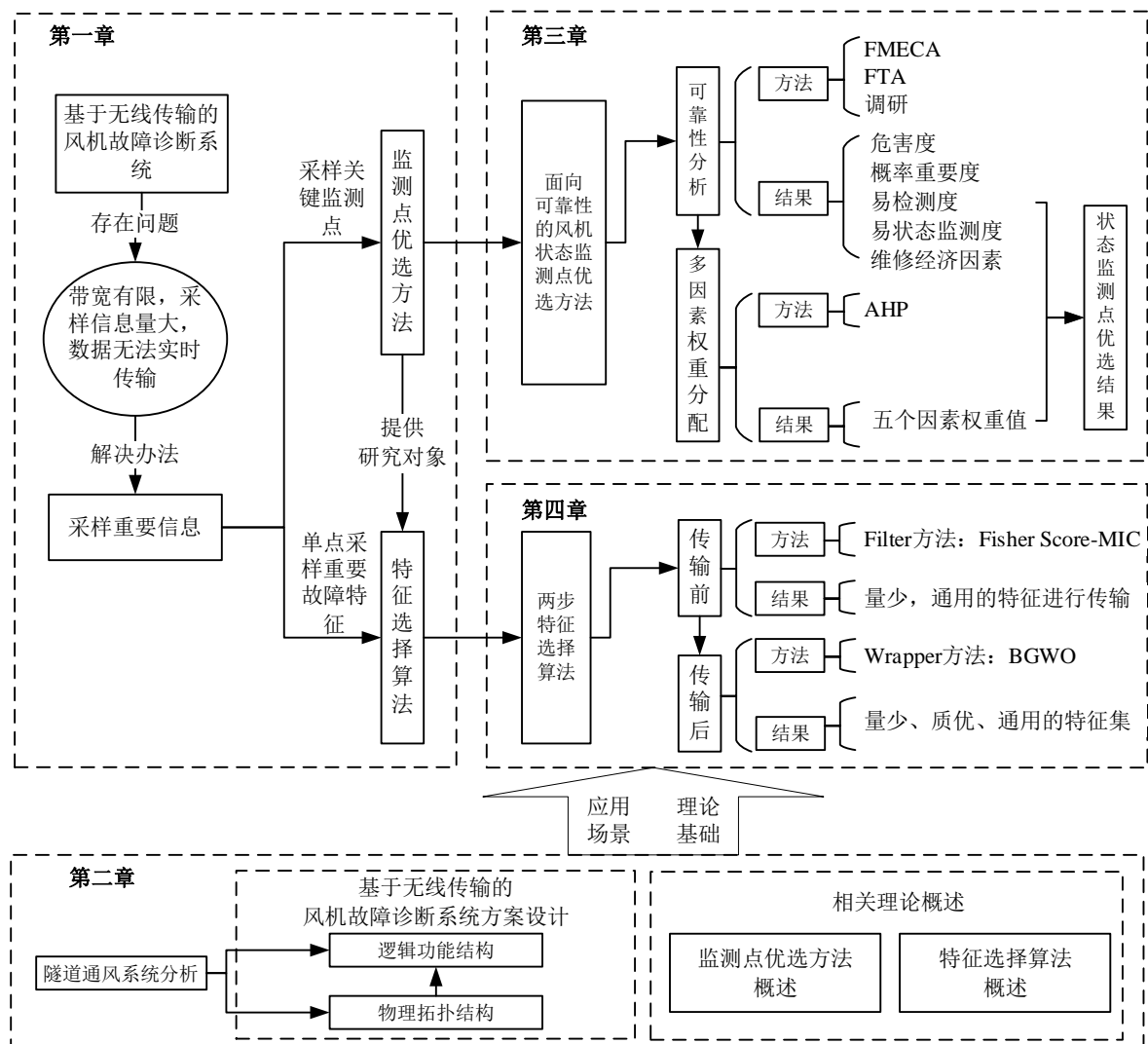


图 1.1 文章章节安排图

第一章：绪论。主要阐述本文的研究背景及意义，然后分别对基于无线传输的风机故障诊断系统、状态监测点的优选方法以及特征选择算法的国内外研究现状进行了分析，最后介绍了本文的研究内容和章节安排。

第二章：基于无线传输的风机故障诊断系统方案设计及相关技术概述。本章通过对终南山隧道的通风系统进行分析，然后从物理拓扑结构、逻辑功能结构两个角度介绍了本文基于无线传输的风机故障诊断系统的方案设计。接着介绍了监测点优选方法以及特征选择算法的技术概述。本章为第三章和第四章的研究提供了应用场景和理论基础。

第三章：面向可靠性的风机状态监测点优选方法。本章通过可靠性分析理论中的 FMECA、FTA 和调研这三种方法计算影响监测点选择的五个因素。然后通过权重分配模型的 AHP 方法对五个因素进行权重分配。最终基于因素值和权重值计算状态监测评价价值进而获得监测点优选结果。本章为第四章特征选择算法提供研究对象。

第四章：基于 Filter 和 Wrapper 相结合的两步特征选择算法研究。以第三章确定的监测点为研究对象，建立基于 Filter 和 Wrapper 相结合的两步特征选择算法。在数据传输前使用 Filter 方法选择量少、通用的特征进行传输，接收到数据后使用 Wrapper 方法对特征进行二次优选，提高故障诊断的精度。

总结和展望。总结了本文的工作内容和工作意义，并分析了本论文存在的不足，对隧道基于无线传输的风机故障诊断系统的未来研究做出了展望。

第二章 基于无线传输的风机故障诊断系统方案设计及相关技术概述

本章首先从终南山隧道的通风系统、故障诊断系统的物理拓扑结构和逻辑功能结构对本文的基于无线传输的风机故障诊断系统方案设计进行阐述,为研究风机状态监测点优选方法以及特征选择算法提供应用场景。之后分别介绍了监测点优选方法和特征选择算法的技术概述,为后续研究奠定理论基础。

2.1 基于无线传输的风机故障诊断系统方案设计

2.1.1 终南山隧道通风系统分析

秦岭终南山公路隧道是陕西境内一条连接西安市与商洛市的穿山隧道^[40],其双洞全长 36.04 公里,是世界第一长双洞公路隧道,也是世界上交通量较多的隧道之一。在终南山隧道中,进出两洞口形成的自然风流以及车辆通行时产生的活塞风不能起到有效的隧道通风效果,而大量车辆的行驶又在隧道中产生了大量的汽车尾气,这对驾驶员及隧道内管理人员的人身安全产生了严重的隐患。

终南山隧道利用射流风机、通风竖井和配套的轴流风机来进行隧道通风。射流风机以 150 米一个为间隔安装在隧道车道的顶部,共安装射流风机 72 组共 144 台,通过接力的方法,将新鲜空气从进/出洞口吸收进/排送出隧道。通风竖井共有三个,一、二、三号竖井高度依次是:230.4m、701.1m、430.6m;位置依次位于秦岭北坡石砭峪沟中游左侧约 40 米、秦岭北坡水洞子沟中上游右侧约 30 米、秦岭南坡大东沟中游左侧约 50m。通风竖井分为向上通风和向下通风两个风道,彼此相互独立,由于烟囱效应的影响,竖井可以将隧道中的废气通过向上风道排出,而轴流风机主要安装在竖井下的风机机房中,可将隧道外的新鲜空气通过向下风道吸收进隧道内。3 个竖井共安装轴流风机 32 台,其中一号竖井 9 台,二号竖井 12 台,三号竖井 11 台。如果把隧道比作一个长笛,竖井及轴流风机的组合就是笛子上的小孔。利用竖井和轴流风机的组合,提升了隧道内的通风效果,弥补了射流风机改善隧道空气能力不足的缺陷。

隧道管理中心负责对风机进行功能控制,其位于西安市长安区太乙宫立交桥附近的秦岭终南山公路隧道公司。三个竖井距离隧道管理中心 10~30 公里左右。

Flaktwoods 公司根据终南山公路隧道距离长,风量需求大的通风要求设计了 Flaktwoods 型轴流风机,将隧道中易燃、易爆、腐蚀性高的气体排出隧道,有效满足了交通运营中的通风需求。该型号风机实物图片如图 2.1 所示,风机结构如图 2.2 所示。



图 2.1 风机实物图

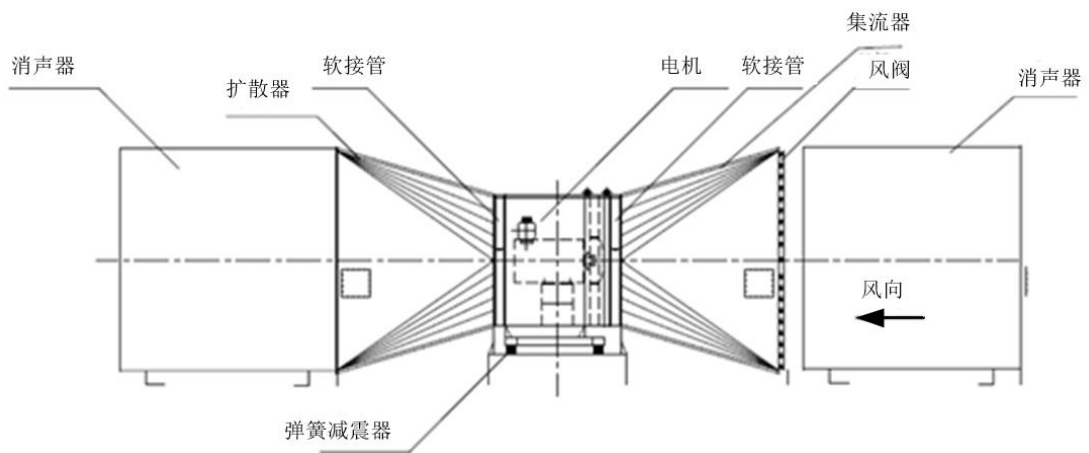


图 2.2 风机结构图

2.1.2 故障诊断系统物理拓扑结构

隧道中，轴流风机远离隧道出口，沿着隧道进行无线传输需要多个设备进行接力，传输的过程中信号干扰严重且资源利用率低。沿着竖井进行数据传输是信号向外界传输最短、最便携的路径，系统几何位置如图 2.3 所示。



图 2.3 系统几何位置图

本文基于无线传输的风机故障诊断系统由三部分组成：感知终端设备、中继设备以及上位机。故障诊断系统物理拓扑图如图 2.4 所示。

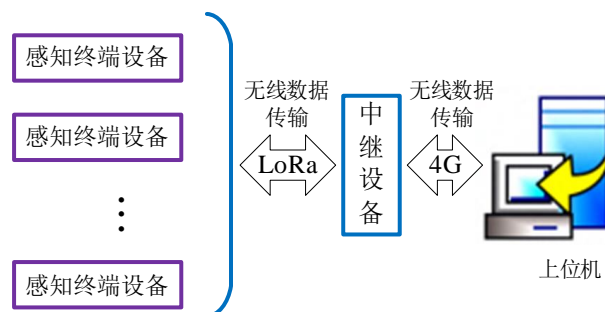


图 2.4 基于无线传输的风机故障诊断系统物理拓扑图

1. 感知终端设备处在故障诊断系统的底层，是整个系统的“感官”部分，负责风机振动数据的采集、预处理及传输工作。它安装在轴流风机外壳上，通过螺丝安装的方式将振动传感器安装在风机上，通过有线方式将数据传输到风机上的数据采集模块，然后通过数据采集模块上的 LoRa 无线传输模块将数据传输到中继模块。根据风机机房中轴流风机个数，每个风机机房共安装 9~12 个感知终端设备。

2. 中继设备是指能够实现一定计算功能的网关设备，其实现信号中继接力和边缘计算的功能，安装在竖井出口处的铁塔上。其中继接力的功能体现在：中继设备作为感知终端设备与上位机间的枢纽，需要支持感知终端设备使用的 LoRa 通信方式，也需要能够与上位机之间进行通信。由 2.1.1 节可知，三个竖井距离隧道管理中心 10~30 公里左右，远大于 LoRa 的 5km 传输距离。经过实地测试，竖井周围具有 4G 信号，而 4G 信号可以通过基站转发数据，大大提高了传输距离。因此中继设备除了需要具备 LoRa 模块来接收感知终端设备的数据之外，还需要 4G 模块来将数据发送到上位机。其边缘计算的功能体现在可对 9~12 个感知终端设备的数据进行处理，并进行实时的故障诊断。中继设备仅将诊断结果以及故障特征值传输到上位机，相比于传输所有特征值，其传输的信息量少，缓解了传输带宽的压力。

3. 上位机位于隧道管理所的监控中心的电脑上，其以隧道管理所的大型服务器为核心，为故障诊断系统提供故障诊断算法训练、数据可视化平台、故障数据存储以及风机控制的功能；

1.2.1 节中提到隧道中常用的窄带宽传输方式有 ZigBee、WiFi、LoRa 三种，本文感知终端与中继设备之间的数据传输方式使用 LoRa 的原因是：相比于 ZigBee 和 WiFi 技术，LoRa 抗干扰能力更强，能够充分满足隧道风机故障诊断系统对于数据传输稳定性的要求；其次，2.1.1 节中提到风机竖井最高 701.1m，加上风机到竖井下井口的位置，

距离在 1000m 左右，LoRa 传输距离远的特性能够满足该传输距离，而 ZigBee 和 WiFi 则需要进行中继传输，增加了建设成本。

2.1.3 故障诊断系统逻辑功能结构

基于无线传输的风机故障诊断系统逻辑功能结构如图 2.5 所示。

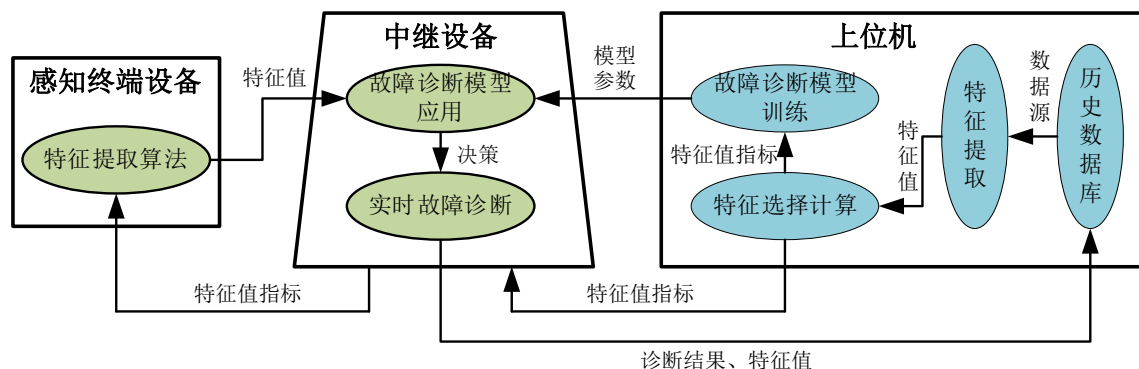


图 2.5 基于无线传输的风机故障诊断系统逻辑功能结构

从硬件系统实现的角度可将图 2.5 分为三部分：负责采集数据的感知终端设备、负责故障决策、数据中继传输的中继设备以及负责模型训练、特征选择算法计算的上位机设备，三者之间通过无线数据传输的方式连接。

从故障诊断功能实现的角度可将图 2.5 分为两部分：故障诊断模型训练以及模型应用，详细阐述如下：

1. 模型训练过程

模型训练过程在隧道管理所的上位机进行，包括历史数据库、特征提取、特征选择计算、故障诊断模型训练这四项内容，下面分别进行阐述：

1) 历史数据库

风机历史库中最初存储着现场采集的风机振动数据，包括风机正常运行时的振动数据以及发生故障时的异常振动数据。在基于无线传输的风机故障诊断系统开始应用之后，数据库中的原始振动数据会被中继设备传输的特征值数据覆盖，最终数据库中只包含风机故障特征值。

2) 特征提取

特征提取能够将有用信号从原始振动信号中提取出来，降低噪声对有用信号的干扰。具体来说：风机在运行时，由于风机制造装配误差、自身振动、监测点外其他部件传递的力和运动的影响，风机局部故障引起的故障信号易被淹没在背景噪声中，故障信号往往呈现非平稳、非线性特点，导致难以通过单一特征对设备进行精确的故障诊断，严重影响故障诊断算法诊断出故障的精度。为提高故障诊断的精度，工程上往往从信号时域、

频域等角度提取一系列故障特征构成多维度的原始特征集来代替振动数据进行故障诊断。

3) 特征选择

故障诊断的精度会受到特征数量的影响,特征数量较少时,会导致故障诊断模型处于“欠拟合”的状态,此时故障诊断模型不能学习到故障数据中反应故障特征的一般规律,模型诊断能力较差,随着特征数量的增加,诊断能力会越来越高。特征数量较多时,会导致模型处于“过拟合”的状态,此时模型会将非常规的、偶然的故障具有的独特特点认为是风机故障的一般规律,导致模型诊断能力变差,随着特征数据量增加,诊断能力会越来越低。特征选择的目的在于从特征提取步骤构造的原始特征集中选出一个最优特征子集,使得建立的故障诊断模型处于“欠拟合”与“过拟合”的中间状态,从而实现较高的故障诊断精度,除此之外,在基于无线传输的故障诊断系统中特征选择还需要选择冗余度低的特征子集,降低从感知终端设备到中继设备、中继设备到上位机的数据传输量。上位机特征选择过程计算的特征值指标经过中继设备的中继传输之后,传输到感知终端设备。

4) 故障诊断训练模型

故障诊断模型建立初期诊断能力较差,其以特征选择后的特征子集为输入、以提高故障诊断能力为目的,不断迭代运算来优化模型参数,最终输出最优模型参数,基于最优模型参数建立最佳故障诊断模型,从而实现较高的故障诊断精度。同时将模型参数传输到中继设备。根据 1.2.1 节的分析,本文使用 SVM、KNN 两种故障诊断模型验证特征选择算法的可行性。

2. 模型应用过程

模型应用过程的数据采集部分在风机旁的感知终端设备,其余部分在中继设备中进行。模型应用过程包括振动数据采集及特征提取、故障诊断模型应用这两部分,下面分别进行阐述:

1) 振动数据采集与特征提取

感知终端设备从风机上采集振动数据之后,基于模型训练过程中特征选择计算出的特征值指标,明确需要提取的特征值,使用附录 A 中表 A.1、表 A.2 的特征提取公式提取对应的特征值,然后通过无线的方式将特征值传输到中继设备。

2) 故障诊断应用模型

故障诊断应用模型使用训练模型输出的最优模型参数建立而成，通过感知终端设备传输的特征值对风机进行实时的故障诊断。并将故障诊断结果、相应的特征值传输到上位机的历史数据库中存储起来。

2.2 监测点优选方法概述

1.2.2 节中阐述了状态监测点的确定往往针对不同的应用场景及需求选择不同的考虑因素，在隧道风机的场景下，风机运行的可靠性影响着隧道的运营安全。因此本文基于可靠性反向故障分析理论中的故障模式、影响及危害性分析（Failure Mode, Effects and Criticality Analysis, FMECA）和故障树分析（Failure Tree Analysis, FTA）模型，计算风机故障的危害度、概率重要度、易检测度、易状态监测度及维修经济因素，确定故障是否需要安装状态监测点。

可靠性反向故障分析理论得到的五种因素对监测点选择的影响权重不同，需要进行权重分配，本文所用的权重分配方法是层次分析法（Analytic Hierarchy Process, AHP）。

2.2.1 可靠性理论概述

可靠性理论的反向故障分析模型常用的方法有 FMECA 和 FTA 的方法，本节对这两种方法的定义、使用流程、与本文的关系进行阐述，为第三章研究风机状态监测点优选方法提供理论基础。

1. FMECA 模型

FMECA 是研究风机所包含零部件的可能故障及其引发的影响并确定危害性的一种故障分析方法。通过 FMECA 可获得风机故障的危害度，通过对危害度较大的故障安装状态监测点，可有效降低其故障危害。

对隧道轴流风机进行 FMECA 的流程如图 2.6 所示：

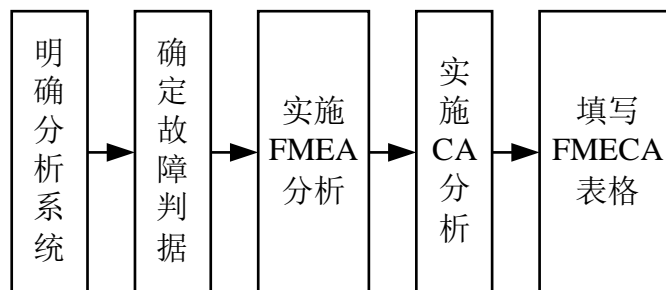


图 2.6 FMECA 模型流程图

1) 明确分析系统

FMECA 首先需要明确研究的系统，本文研究的系统是终南山隧道风机机房中的 Flaktwoods 型轴流风机，需要理清该型号轴流风机的结构组成、各个结构的功能以及轴流风机可能发生的故障模式。

2) 确定故障判据

故障判据是指确定轴流风机是否正常运行的判断标准。确定故障判据是确定轴流风机故障模式的前提，其在 FMEA 的故障模式确定前进行。

3) 实施 FMEA 分析

实施 FMEA 分析，即完成轴流风机的故障模式及故障影响的研究。故障模式是指分析轴流风机可能发生的故障类型；故障影响分析是指故障模式发生后对该风机零部件、子系统及系统产生的影响。

4) 实施危害度分析 (Criticality Analysis, CA)

实施 CA 分析，即进行高速公路轴流风机的危害度分析，是对风机故障导致的严重程度度的综合分析。

5) 填写 FMECA 表格

最后根据 FMECA 结果，填写 FMECA 表格，如表 2.1 所示。表中危害度严重等级由高到低依次为 I-III。I 级会造成风机停机；II 级长期对风机有害；III 级：对风机没有影响或者影响很小。

表 2.1 FMECA 表格

部件名称	故障模式	底层故障 (故障原因)	故障影响			危害度
			零部件级	子系统级	系统级	
叶轮子系统	叶片质量不平衡	叶片材料老化	载荷增加	叶片子系统与电机子系统之间的连接轴损坏	风机风力减弱、寿命缩短，长期将导致风机停机	II 级

2. FTA 模型

FTA 是一种采用逻辑方法将系统中各种因果关系绘制成树状结构的故障分析方法。其通过将某一个故障作为顶事件去演义推理出引起该故障发生的各种深层次故障，然后向下继续通过推理分析找到所有的基本故障（基本事件），并基于基本故障计算故障概率重要度的一种分析方法。通过 FTA 模型，即可获得风机各个故障的概率重要度，通过对概率重要度较大的故障安装状态监测点，可有效降低其故障发生概率。概率重要度是零部件故障概率与风机故障概率的比值，其表示相对概率，相比于使用绝对概率，其更

加能够反应风机单个故障对风机系统造成的影响大小。对终南山隧道轴流风机进行 FTA 分析的流程如图 2.7 所示。

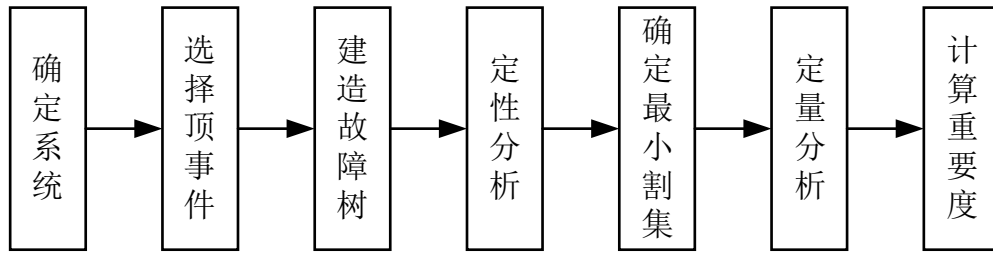


图 2.7 FTA 流程图

建立轴流风机故障树的步骤如下：

- 1) 明确系统是风机。
- 2) 将风机故障作为顶事件，将各个子系统故障作为中间事件，将导致子系统故障的底层故障作为基本事件。例如：风机故障是顶事件，机壳子系统故障、叶片质量不平衡故障是中间事件，叶片材料老化是导致叶片质量不平衡的底层故障，其是基本事件。
- 3) 使用规定的逻辑符号，表示出系统中各个故障事件的因果关系，形成故障树。
- 4) 对故障树定性分析，确定最小割集。
- 5) 基于最小割集对故障树进行定量计算，其通过各个基本事件发生的概率计算出顶事件的概率，最后计算出各个基本事件的概率重要度。

故障树主要符号如表 2.2 所示。

表 2.2 FTA 主要符号及基本含义

符号	名称	含义
	顶事件	位于故障树的顶端，表明分析的最终目标
	中间事件	用以连接上下级逻辑的中间事件
	基本事件	故障树底层的事件，通常为已熟知的单一故障
	非基本事件	发生概率极小，和其他基本事件不在一个数量级上
	转移事件	用于故障树与子树的连接
	或门	任意一个或门输出因素均是事件故障的原因
	与门	所有与门输出下的因素同时存在才引起事件故障

2.2.2 AHP 方法概述

AHP 方法建立多级层次，通过两两比较、综合计算的方式获得各个层次之间的指标权重，然后基于权重获得底层内容对于最上层的评价结果。通过 AHP 方法，计算故障的危害度、概率重要度、易检测度、易状态监测度、维修经济因素这五个因素对状态监测点安装的影响权重，基于权重计算以及五个因素值即可计算风机状态监测评价值，通过评价值来考虑对风机某个故障是否安装状态监测点。

AHP 方法流程如下:

1. 建立层次结构模型

通常为三级结构: 目标层、准则层、方案层。目标层是需要获得的结果, 本文是针对风机某个故障安装状态监测点的评价值。准则层是评价准则, 本文是风机故障的危害度、概率重要度、易检测度、易状态监测度、维修经济因素。方案层是可选择的方案, 本文是风机的一系列故障。

2. 建立判断矩阵并填写矩阵

判断矩阵目的是计算同一层次的各个元素关于上一层中某一准则的权重。通常分为两部分, 一部分是方案层中各个元素关于准则层中某一指标的判断矩阵; 另一部分是准则层中各个元素关于目标层的判断矩阵。判断矩阵填写准则如表 2.3 所示。

表 2.3 判断矩阵填写准则

编号	标度	含义
1	1	表示两个因素相比, 同样重要
2	3	表示两个因素相比, 前者稍微重要
3	5	表示两个因素相比, 前者明显重要
4	7	表示两个因素相比, 前者强烈重要
5	9	表示两个因素相比, 前者极端重要
6	2、4、6、8	上述两相邻判断的中值
7	倒数	前一个因素相比后一个因素的标度为 3, 那么后一个因素和前一个因素相比就是 1/3

3. 一致性检验

判断矩阵的填写通过两者相互比较进行填写的, 当判断矩阵的维数增加之后, 经常会出现 $A > B$, $A = C$, 而 $B > C$ 等自相矛盾的情况。一致性检验是检测判断矩阵是否出现自相矛盾的一种方法, 当出现自相矛盾情况时, 需要修改权重矩阵。一致性检验流程如下:

① 计算一致性指标 CI

$$CI = \frac{\lambda_{\max} - n}{n - 1} \quad (2.1)$$

式中 λ_{\max} 为判断矩阵最大的特征值, n 为判断矩阵的维数。

② 查找对应的平均随机一致性指标 RI

表 2.4 平均随机一致性指标 RI

矩阵维数 n	1	2	3	4	5	6	7
平均随机一致性指标 RI	0	0	0.52	0.89	1.12	1.26	1.36

③ 计算一致性比例 CR

$$CR = \frac{CI}{RI} \quad (2.2)$$

如果 $CR < 0.1$ ，则可认为判断矩阵没有出现自相矛盾的情况，否则需要对判断矩阵填写的值进行修改。

4. 求组合权重

使用算术平均法的方式计算组合权重。

$$\omega_i = \frac{1}{n} \sum_{j=1}^n \frac{a_{ij}}{\sum_{k=1}^n a_{kj}}, (i = 1, 2, \dots, n) \quad (2.3)$$

式中： ω_i 为第 i 个权重值， n 为判断矩阵的维数， a_{ij} 为判断矩阵中第 i 行 j 列的值， a_{kj} 为判断矩阵中 k 行 j 列的值。

5. 计算结果

2.3 特征选择算法概述

2.3.1 Filter 方法概述

由 Fisher 判别分析^[41]（Fisher Discriminate Analysis, FDA）演化的 Fisher Score 方法用来度量一个特征值辨别样本所属类的能力大小，其认为重要度高的特征能够方便故障诊断模型进行故障诊断，其要求选择的特征在相同的类之内表现出的差异尽可能的小，在不同的类之间表现出的差异尽可能的大。本文参照前人的方法，使用 Fisher Score 的方法度量特征的重要度。

但是，Fisher Score 方法只能确定风机不同故障特征的重要度，在基于无线传输的风机故障诊断系统不仅需要传输的特征尽可能重要，还需要特征之间相互独立、低冗余，实现有限的带宽资源传输尽可能多的精华特征。MIC 的方法依托于互信息思想，解决了互信息计算时联合概率密度函数难以计算的难题，同时继承了互信息的特点，可以评价特征与特征之间的冗余度。

因此，本文在 Fisher Score 的基础上，使用 MIC 的方法度量特征与特征之间的冗余度。Fisher Score 和 MIC 方法的理论如下：

1. Fisher Score 方法概述

Chen^[42]等给出了基于二分类问题的 Fisher Score 的表达式，作为特征选择准则来选择得分高的特征。在此基础上，Salih^[43]等提出了多分类问题下的 Fisher Score 特征选择方法，并用评分的均值作为阈值进行特征选择。但是 Salih 提出的多分类 Fisher Score 特

征选择方法应用到各类中心相同、分属于平衡和不平衡的两个数据集时 Fisher Score 取值相同,不能很好表达出不同的数据集在类上的差异,图 2.8 中(a)为类平衡的数据集,(b)为类不平衡的数据集。针对这个问题, Song^[44]等在 Salih 等人的基础上,在原有的 Fisher Score 计算方法上进行了改进,进而考虑到了类与类之间的差异。

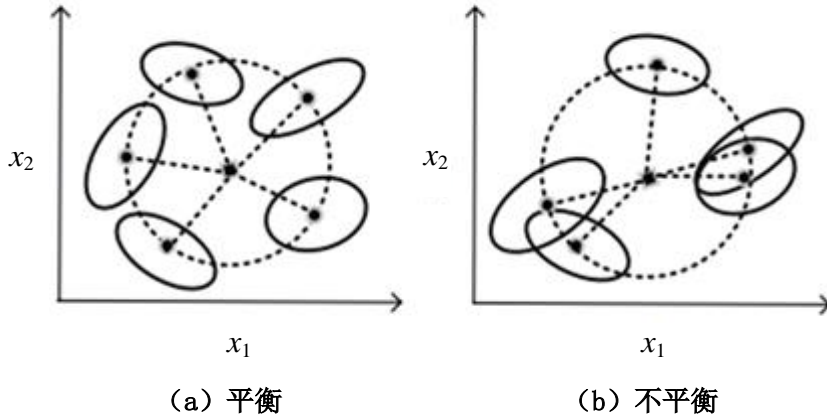


图 2.8 类平衡与类不平衡的数据集

风机故障通常表现出多类情况,且不能确定各类属于平衡或不平衡数据集中的哪一种,因此,本文基于 Song 等提出的 Fisher Score 计算方法,理论如下:

使用振动传感器在风机的某个安装位置进行一段时间的数据采样,将生成的一组数据进行特征提取构成一个样本。假设共有 N 个风机故障样本 $\{x_i, i = 1, 2, \dots, N\}$, N 个故障样本共包含 $c(c \geq 2)$ 种故障类别,每个样本由特征集 $\{f_1, f_2, \dots, f_j, \dots, f_M, M \geq 2\}$ 表示, x_{ij} 为第 i 个样本在第 j 个特征 f_j 下的取值。例如:有 3 个风机故障样本,包含 0 和 1 两种故障类别,每个样本由特征集 $\{f_1, f_2, \dots, f_7\}$ 表示。如表 2.5 所示。

表 2.5 风机故障样本示意表

样本编号	f_1	f_2	f_3	f_4	f_5	f_6	f_7	类别编号
1	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}	0
2	x_{21}	x_{22}	x_{23}	x_{24}	x_{25}	x_{26}	x_{27}	1
3	x_{31}	x_{32}	x_{33}	x_{34}	x_{35}	x_{36}	x_{37}	1

定义样本第 j 个特征 f_j 的类间散度 $S_b(f_j)$ 和第 k 类样本的第 j 个特征 f_j 的类内散度 $S_t^{(k)}(f_j)$ 为:

$$S_b(f_j) = \sum_{1 \leq p < q \leq c} \frac{n_p + n_q - n_{pq}^{(j)}}{N} (\mu_j^{(p)} - \mu_j^{(q)})^2 \quad (2.4)$$

$$S_t^{(k)}(f_j) = \sum_{i=1}^{n_k} (x_{ij}^{(k)} - \mu_j^{(k)})^2 \quad (2.5)$$

式中, n_p 和 n_q 分别为第 p 、 q 类样本的数目, $n_{pq}^{(j)}$ 为类别 p 、 q 在特征 f_j 下取值相同的数据个数, $\mu_j^{(p)}$ 和 $\mu_j^{(q)}$ 分别为第 j 个特征下第 p 、 q 类样本的均值, $x_{ij}^{(k)}$ 为第 k 类样本中第 i 个样本在第 j 个特征 f_j 下的取值, $\mu_j^{(k)}$ 为第 j 个特征下第 k 类样本的均值。

$$FS(f_j) = \frac{S_b(f_j)}{\sum_{k=1}^c n_k S_t^{(k)}(f_j)} = \frac{\sum_{1 \leq p < q < c} \frac{n_p + n_q - n_{pq}^{(j)}}{N} (\mu_j^{(p)} - \mu_j^{(q)})^2}{\sum_{k=1}^c n_k \sum_{i=1}^{n_k} (x_{ij}^{(k)} - \mu_j^{(k)})^2} \quad (2.6)$$

式中, n_k 为第 k 类样本的数目。

上式中第 j 个特征 f_j 的类间散度越大, 类内散度越小, 则 $FS(f_j)$ 的值越大, 说明 f_j 的重要度越高, 辨识力越强。

2. MIC 方法概述

MIC 利用互信息和网格划分的方法来进行计算。

首先互信息的相关理论如下: 假设 $X = \{x_i, i = 1, 2, \dots, N\}$ 、 $Y = \{y_i, i = 1, 2, \dots, N\}$ 是风机的两个特征变量, N 为样本数量, S 、 T 分别是是风机特征变量 X 、 Y 的取值范围, $p(x)$ 是特征 X 的边缘概率密度函数, $p(y)$ 是特征 Y 的边缘概率密度函数, $p(x, y)$ 表示特征 X 和特征 Y 的联合概率密度, 互信息的计算公式如 (2.7) 所示:

$$I(X, Y) = \sum_{x \in S} \sum_{y \in T} p(x, y) \log_2 \frac{p(x, y)}{p(x)p(y)} \quad (2.7)$$

其次网格划分的理论如下: 将风机的两个特征变量 X 、 Y 按顺序组成集合 $D[\{x_i, y_i\}, i = 1, 2, \dots, N]$ 。将集合 D 中的 x_i 和 y_i 绘制二维散点图, 然后对散点图进行 $x \times y$ 的网格划分, x 、 y 最小为 2, 最大通常为 $N^{0.6}$ 。

最后统计各个散点落入各个方格的情况, 代替计算互信息中的联合概率密度, 计算网格划分情况下的互信息 $I(X, Y)$ 。针对 $x \times y$ 的网格划分有很多中, 分别计算互信息, 取最大的互信息值进行归一化操作得到 MIC 。

$$MIC(X, Y) = \arg \max_{|x|, |y| < N^{0.6}} \frac{I(X, Y)}{\log_2(\min(|x|, |y|))} \quad (2.8)$$

例如: 假设有风机的两个特征变量取值为 $\{(15, 24), (22, 47), (37, 28), (43, 49), (38, 74), (74, 42), (71, 84), (97, 77), (8, 104), (101, 105)\}$, 对其绘制二维散点图如图 2.9 所示。以红、绿、蓝三种不同的网格划分方法为例阐述互信息计算方法, 实际划分时将网格划分为 2×2 至 $10^{0.6} \times 10^{0.6}$ 。

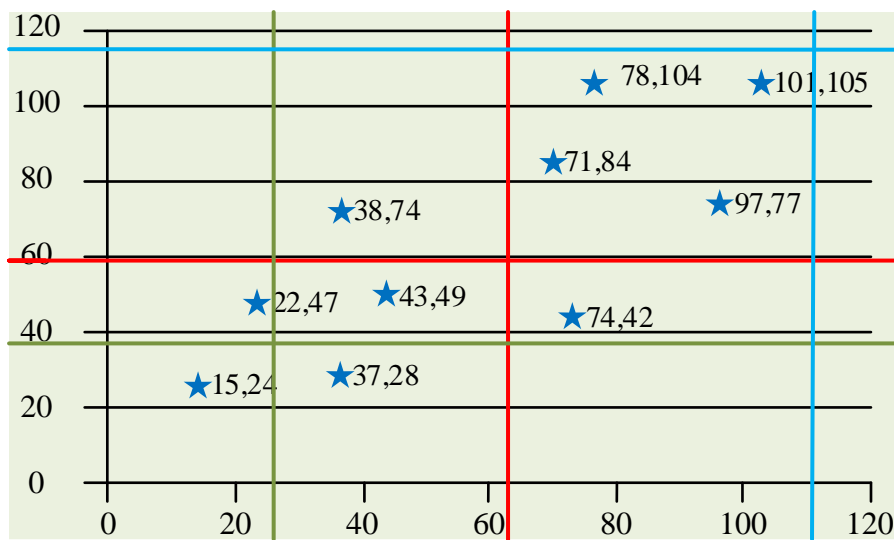


图 2.9 风机特征向量二维散点图

以红色的网格划分为例计算互信息，在图中落入左上、右上、左下、右下区域的概率为 0.1、0.4、0.4、0.1。

根据式 (2.7) 互信息的计算公式得到红色网格划分情况下互信息值：

$$\begin{aligned}
 I(X,Y) &= 0.1 \times \log\left(\frac{0.1}{0.5 \times 0.5}\right) + 0.4 \times \log\left(\frac{0.4}{0.5 \times 0.5}\right) \\
 &\quad + 0.4 \times \log\left(\frac{0.4}{0.5 \times 0.5}\right) + 0.1 \times \log\left(\frac{0.1}{0.5 \times 0.5}\right) \\
 &= 0.08370
 \end{aligned}$$

然后计算绿、黄以及其他未划分处的网格情况下的互信息，选择最大的互信息使用式 (2.8) 计算 MIC。

MIC 描述了风机两个特征之间含有共同信息的多少，MIC 的大小能够表征两个特征之间的相关程度，通过计算 MIC，剔除传输特征中 MIC 较高的特征，即冗余的特征，实现在窄带宽条件下尽可能传输有用信息的目的。

2.3.2 Wrapper 方法概述

本文所用的 Wrapper 方法是二进制灰狼优化算法 (Binary Grey Wolf Optimizer, BGWO) 算法。

BGWO 起源于灰狼优化算法 (Grey Wolf Optimizer, GWO)，GWO 算法是澳大利亚学者 Mirjalili^[45-47]于 2014 年提出以来，已经逐渐应用到了机器人路径规划^[48,49]、设备故障诊断^[50,51]、车间任务调度^[52,53]等领域，获得了极大的发展。其算法灵感来源于灰狼的捕食过程。灰狼是犬科、群居型动物，种群数量在 5-12 只。狼群中有严格的等级划分，

领导者 α 狼负责制定规则，决定种群狩猎计划；从属狼 β 负责协助领导者进行决策；下属狼 δ 服从于领导者和从属狼，是狼群中的中坚力量；等级最低的 ω 狼是狼群中最弱的狼。GWO 算法的优化过程由 α 、 β 、 δ 狼引导， ω 狼跟随其他三类狼。狼群等级划分如图 2.10 所示：

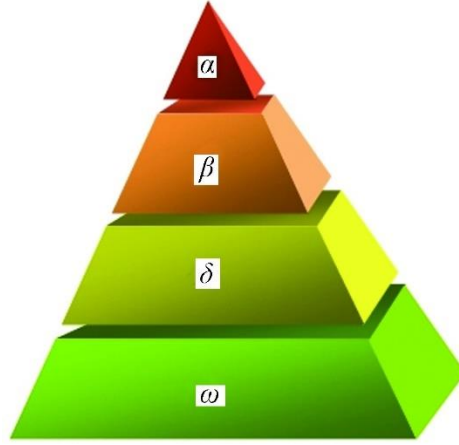


图 2.10 狼群等级划分

BGWO 算法设计如下：

1. 环绕猎物

灰狼在围猎猎物时，首先会包围猎物，对其进行建模，如式 (2.9)、(2.10) 所示。

$$D = |C \cdot X_p(t) - X(t)| \quad (2.9)$$

$$X(t+1) = X_p(t) - A \cdot D \quad (2.10)$$

其中： t 为当前迭代； A 和 C 为系数向量； X_p 为猎物的位置向量； X 为灰狼的位置向量。 A 和 C 向量的计算过程如下：

$$A = 2 \cdot a \cdot r_1 - a \quad (2.11)$$

$$C = 2 \cdot r_2 \quad (2.12)$$

其中， a 在迭代过程中从 2 线性下降到 0， r_1 、 r_2 是 [0,1] 间的随机向量。

2. 搜索（狩猎）

灰狼开始捕食时，并不知道最佳猎物所在的位置。搜索（狩猎）过程指的是灰狼寻找猎物的过程，在这个过程中， α 、 β 、 δ 狼由于狩猎经验丰富或者体力充足，对猎物潜在的位置有更好的预判，因此要求 α 、 β 、 δ 依次保存当前获得的最佳的三个解决方案， ω 狼根据 α 、 β 、 δ 的位置更新自己的位置。其数学模型抽象如下：

$$D_\alpha = |C_1 X_\alpha - X|, D_\beta = |C_2 X_\beta - X|, D_\delta = |C_3 X_\delta - X|, \quad (2.13)$$

$$X_1 = X_\alpha - A_1 \cdot D_\alpha, X_2 = X_\beta - A_2 \cdot D_\beta, X_3 = X_\delta - A_3 \cdot D_\delta, \quad (2.14)$$

$$X(t+1) = \frac{(X_1 + X_2 + X_3)}{3} \quad (2.15)$$

上述的灰狼优化算法主要应用于搜索空间连续的场景，例如在某个三维空间中寻找传感器的最优位置，其 x 、 y 、 z 的位置变化均是连续的。本文的特征选择是搜索空间离散的场景，主要考虑是否选择某一个特征值，0 为不选择，1 为选择，所以需要将搜索从连续空间转变为离散空间。由此 Emary 等人^[54]提出了 BGWO。BGWO 算法基于 GW O 算法，其在式 (2.15) 之后引入 (2.16)、(2.17)，实现搜索从连续空间到离散空间的转变。

$$\text{Sigmoid}(X(t+1)) = \frac{1}{1 + e^{-10 \cdot (X(t+1) - 0.5)}} \quad (2.16)$$

$$X_{out}(t+1) = \begin{cases} 1 & \text{if Sigmoid}(X(t+1)) \geq rand \\ 0 & \text{otherwise} \end{cases} \quad (2.17)$$

2.4 本章小结

本章首先对终南山隧道的通风系统进行分析。然后从系统位置分布、硬件构成等角度阐述了本文设计的基于无线传输的风机故障诊断系统的物理拓扑结构。接着从故障诊断模型训练及模型应用的角度阐述了本文设计的基于无线传输的风机故障诊断系统的逻辑功能结构，为第三章、第四章的研究提供应用场景。最后对第三章的监测点优选方法、第四章的特征选择算法的技术理论进行了阐述。

第三章 面向可靠性的风机状态监测点优选方法

3.1 引言

根据 1.2.2 节分析，基于有线传输的故障诊断系统由于安装状态监测点成本较低，其对状态监测点的确定往往仅考虑有效性因素，即只考虑安装的状态监测点是否有助于故障诊断，导致安装的监测点较多。在带宽受限的条件下，应安装尽可能少的监测点，安装的状态监测点应能够对故障频次高、危害度大的故障进行监测，同时也应能够对故障后不易检测、维修成本高的故障进行监测。因此，本章基于可靠性分析理论，提出面向可靠性的风机状态监测点优选方法，通过可靠性分析的反向故障分析理论计算故障的危害度、概率重要度、易检测度、易状态监测度、维修经济性这五种因素，然后使用 AHP 方法对这五种因素进行权重分配，最终获得故障的等级排序，根据带宽条件优先对排序高的故障设置监测点。根据本文方法，实现了在带宽有限的条件下，优先采样关键监测点的目标，缩减了监测点数。

3.2 风机结构功能分析

轴流风机工作时，叶轮转动引起风机内部压力变化，使气体在大气压的作用下沿轴方向从集流器进入风机内部。气流在叶轮的转动中被加速获得能量，再进入出风筒，被出风口的收缩状扩散器加压而增加风速和静压，再由出风筒内安装的导流叶片将偏转的气流流动转变为轴向流动，最后被加速的气体在出风筒的出口处进入隧道^[55]。

其主要结构由三部分组成：叶轮子系统、机壳子系统、电机子系统，具体包含的零部件及功能如表 3.1 所示。

表 3.1 风机结构及功能表

子系统	零部件	功能
叶轮子系统	轮毂	用于安装叶片并与机壳子系统主体相连
	叶片	对空气做功，产生风量和气压
	轮毂与叶片之间的连接螺栓	连接轮毂和叶片
机壳子系统	进口消声器（含消音过滤棉）	1）减弱风机运行时产生的噪声；2）过滤空气杂质，提高输送空气的清洁度
	出口消声器	减弱风机运行时产生的噪声
	扩散器	降低出口处风机动压以提高静压，减少气流动能损失
	进口软接管	减少风机主体与进口风道之间产生的共震，起到连接结构、静音降噪的作用
	出口软接管	减少风机主体与出口风道之间产生的共震，起到连接结构、静音降噪的作用
	进口软接管紧固件	连接风机主体和风筒
	出口软接管紧固件	连接风机主体和风筒

表 3.1 风机结构及功能表（续）

子系统	零部件		功能
机壳子系统	集流器		保证叶轮进口断面的气流分布均匀，从而使气流在损失最小的情况下进入叶轮
	风阀		开启或者关闭风机，发生火灾时可防止空气进入风道，避免火势增大
	弹簧减震器		降低风机振动对地基的影响
	减震器紧固件		连接减震器和风机主体，防止风机晃动
	进风口		将山体外的空气引入风机中，同时起到降低噪音的作用
	出风口		将风机产生的风排入隧道内，同时起到降低噪音的作用
	入口防护网		除尘过滤，防止异物进入风机
	出口防护网		除尘过滤，防止异物进入风机
电机子系统	接线盒	三相电缆	为风机中的电机提供电能
		接线盒盖	对供电接口进行密封，防止异物损伤供电接头导致供电异常
		接线盒体	对供电接口进行密封，防止异物损伤供电接头导致供电异常
	电机机壳	机壳连接支架	起到支撑电机主体的作用
		电机紧固件	将电机与机壳子系统主体连接
		非驱动端端盖	防护和支撑电机
		驱动端端盖	防护和支撑电机
	定子	端板	支撑定子结构
		支撑板	支撑定子结构
		定心铁芯+线圈	提供磁场
	转子	轴承	支撑和旋转
		轴承润滑脂	润滑轴承，防止摩擦生热使得轴承温度过高
		转子铁芯	旋转和传递扭矩
		风扇	为电机散热
	轮毂与电机之间的连接轴		用于连接电机和轮毂，电机通过连接轴带动轮毂上的叶片转动，进而产生风能
	其他辅助部件	O 型密封圈、平面密封圈等	密封

3.3 面向可靠性的多因素计算

3.3.1 基于 FMECA 的危害度计算

基于 3.2 节对轴流风机结构及功能的分析，结合对终南山隧道轴流风机的调研结果以及相关的资料查阅，采用 2.2.1 节 FMECA 流程，分析轴流风机各零部件包含的故障模式、导致故障模式发生的底层故障以及对风机零部件、子系统以及系统产生的影响，绘制 FMECA 表格如表 3.2 所示。

表 3.2 风机 FMECA 表格

部件	故障模式	底层故障	故障影响			危害度
			零部件级	子系统级	系统级	
叶轮子系统	叶片质量不平衡	叶片材料老化	载荷增加	叶片子系统与电机子系统之间的连接轴损坏	风机风力减弱、寿命缩短，长期将导致风机停机	Ⅱ级
		叶片结冰腐蚀污垢				
		叶片潮湿腐蚀污垢				
		叶片安装角度不正确				
	叶片裂纹损伤	叶片生产加工不良	叶片断裂	叶片子系统与电机子系统之间的连接轴损坏	风机风力减弱、寿命缩短，长期将导致风机停机	Ⅱ级
		叶片材料老化				
		叶片雷击损伤				
		叶片潮湿腐蚀污垢				
	轮毂裂纹损伤	轮毂材料老化	轮毂损坏，叶片掉落	长期导致叶轮子系统停止工作	长期导致风机停机	Ⅱ级
		叶片剧烈转动				
		轮毂生产加工不良				
	叶片与轮毂之间的连接螺栓塑性变形	叶片剧烈转动	叶片松动	叶片子系统与电机子系统之间的连接轴损坏	风机风力减弱、寿命缩短，长期将导致风机停机	Ⅱ级
		连接螺栓材料老化				
	叶片与轮毂之间的连接螺栓滑扣	叶片频繁装拆	叶片松动	叶片子系统与电机子系统之间的连接轴损坏	风机风力减弱、寿命缩短，长期将导致风机停机	Ⅱ级
	叶片与轮毂之间的连接螺栓断裂	叶片剧烈转动	叶片脱落	叶轮子系统停止工作	风机停机	Ⅰ级
		连接螺栓生产加工不良				
		连接螺栓材料老化				
机壳子系统	机壳内空气杂质较多	入口防护网损坏	叶片损伤	叶片子系统与电机子系统之间的连接轴损坏	风机风力减弱、寿命缩短，长期将导致风机停机	Ⅱ级
		出口防护网损坏	叶片损伤			

表 3.2 风机 FMECA 表格 (续)

部件	故障模式	底层故障	故障影响			危害度
			零部件级	子系统级	系统级	
机壳子系统	机壳内空气杂质较多	进口消声器丧失空气过滤功能	叶片损伤	叶片子系统与电机子系统之间的连接轴损坏	风机风力减弱、寿命缩短, 长期将导致风机停机	II 级
	机壳部位噪声异常增大	进口消音器损坏	消音器损坏	机壳子系统异常	风机开机后有异常噪音	III 级
		出口消音器损坏	消音器损坏	机壳子系统异常	风机开机后有异常噪音	III 级
		进口软接管损坏	进口软接管损坏	机壳子系统集流器工作异常	风机风力减弱, 风机开机后有异常噪音	II 级
		出口软接管损坏	出口软管管损坏	机壳子系统扩散器工作异常	风机风力减弱, 风机开机后有异常噪音	II 级
	机壳振动过大	弹簧减震器损坏	弹簧减震器损坏	机壳子系统晃动, 寿命减少, 加速老化	加速零件损耗, 风机寿命减少, 加速老化, 最终风机停机	II 级
		减震器紧固件损坏	减震器紧固件损坏	机壳子系统晃动, 寿命减少, 加速老化	加速零件损耗, 风机寿命减少, 加速老化, 最终风机停机	II 级
	机壳内风机风能减弱	扩散器损坏	扩散器损坏	机壳子系统运行异常	风机风力减弱	II 级
		集流器损坏	集流器损坏	机壳子系统运行异常	风机风力减弱	II 级
		风阀损坏	风阀损坏	机壳子系统运行异常	火灾发生时无法关闭风机	I 级
电机子系统	电机过热	转子轴承磨损	轴承损坏	电机寿命减少, 加速老化	风机风力减弱、寿命缩短, 长期将导致风机停机	II 级
		风机高速运行, 转速过高	轴承等相关零部件加速老化			
		转子风扇故障	风扇故障			
		转子轴承润滑脂过多或者不足	轴承加速老化			

表 3.2 风机 FMECA 表格 (续)

部件	故障模式	底层故障	故障影响			危害度
			零部件级	子系统级	系统级	
电机子系统	电机振动过大	定子支撑板松动	定子加速老化	电机寿命减少, 加速老化	风机风力减弱、寿命缩短, 长期将导致风机停机	II 级
		电机机壳连接支架松动或者损坏	电机和轮毂连接轴加速老化			
		电机紧固件松动或者损坏	电机和轮毂连接轴加速老化			
		转子平衡不好	转子轴承损坏			
	通电后电机未启动	电机接线盒三相电缆接头损坏	电机三相电缆损坏	电机停止运行	风机停机	I 级
		电机转子轴承断裂	轴承断裂			

为方便在 3.5 节中使用危害度因素计算状态监测评价值, 将危害度 I-III 的等级进行数值化, 通过对故障危害度进行分析, 定义 I 级数值为 1, II 级数值为 0.5, III 级为 0.1。同时为防止危害度因素的值过高或过低以至于覆盖了或淹没于其他因素, 对数值化后的值进行归一化, 后续对各个因素进行归一化均是由于这个原因, 之后不再阐述。根据式 (3.1) 对数值化后的危害度值进行归一化, 结果如表 3.3 所示。

$$P_{norm_i} = \frac{P_{harm_i}}{\sum_{i=1}^{33} P_{harm_i}}$$

(3.1)

式中: P_{norm_i} 为第 i 个故障的危害度归一化后的值; P_{harm_i} 为第 i 个故障的危害度。

表 3.3 各个故障危害度数值化及归一化值

编号	故障	危害度			编号	故障	危害度		
		等级	数值化	归一化值			等级	数值化	归一化值
1	叶片材料老化	II	0.5	0.02785	18	出口软接管损坏	II	0.5	0.02785
2	叶片结冰腐蚀污垢	II	0.5	0.02785	19	弹簧减震器损坏	II	0.5	0.02785
3	叶片潮湿腐蚀污垢	II	0.5	0.02785	20	减震器紧固件损坏	II	0.5	0.02785
4	叶片安装角度不正确	II	0.5	0.02785	21	扩散器损坏	II	0.5	0.02785
5	叶片生产加工不良	II	0.5	0.02785	22	集流器损坏	II	0.5	0.02785
6	叶片雷击损伤	II	0.5	0.02785	23	风阀损坏	I	1	0.05571

表 3.3 各个故障危害度数值化及归一化值（续）

编号	故障	危害度			编号	故障	危害度		
		等级	数值化	归一化值			等级	数值化	归一化值
7	轮毂材料老化	II	0.5	0.02785	24	转子轴承磨损	II	0.5	0.02785
8	轮毂生产加工不良	II	0.5	0.02785	25	风机高速运行，转速过高	II	0.5	0.02785
9	连接螺栓材料老化	I、II	0.75	0.04178	26	转子风扇故障	II	0.5	0.02785
10	叶片频繁装拆	II	0.5	0.02785	27	转子轴承润滑脂过多或者不足	II	0.5	0.02785
11	连接螺栓生产加工不良	I	1	0.05571	28	定子支撑板松动	II	0.5	0.02785
12	入口防护网损坏	II	0.5	0.02785	29	电机机壳连接支架松动或者损坏	II	0.5	0.02785
13	出口防护网损坏	II	0.5	0.02785	30	电机紧固件松动或者损坏	II	0.5	0.02785
14	进口消声器丧失空气过滤功能	II	0.5	0.02785	31	转子平衡不好	II	0.5	0.02785
15	进口消音器损坏	III	0.1	0.00557	32	电机接线盒三相电缆接头损坏	I	1	0.05571
16	出口消音器损坏	III	0.1	0.00557	33	电机转子轴承断裂	I	1	0.05571
17	进口软接管损坏	II	0.5	0.02785					

由下文故障树的分析可知叶片剧烈转动不是导致风机故障的底层故障，为和下文保持一致，本处未计算叶片剧烈转动的危害度。同时叶片与轮毂之间的连接螺栓材料老化会造成叶片脱落和叶片松动两种不同的故障影响，分别会造成 I 级和 II 级两种故障危害度，这里取二者的均值 0.75。

3.3.2 基于 FTA 的概率重要度计算

根据 2.2.1 节中介绍的 FTA 流程，进行基于 FTA 模型的风机故障概率重要度计算。

FTA 的第 1)、2)、3) 步骤是明确风机系统，确定顶事件、中间事件、基本事件，建造故障树，内容如下：

将风机故障作为故障树的顶事件，根据风机系统的结构和功能的分析，建造风机系统故障树如图 3.1 所示。

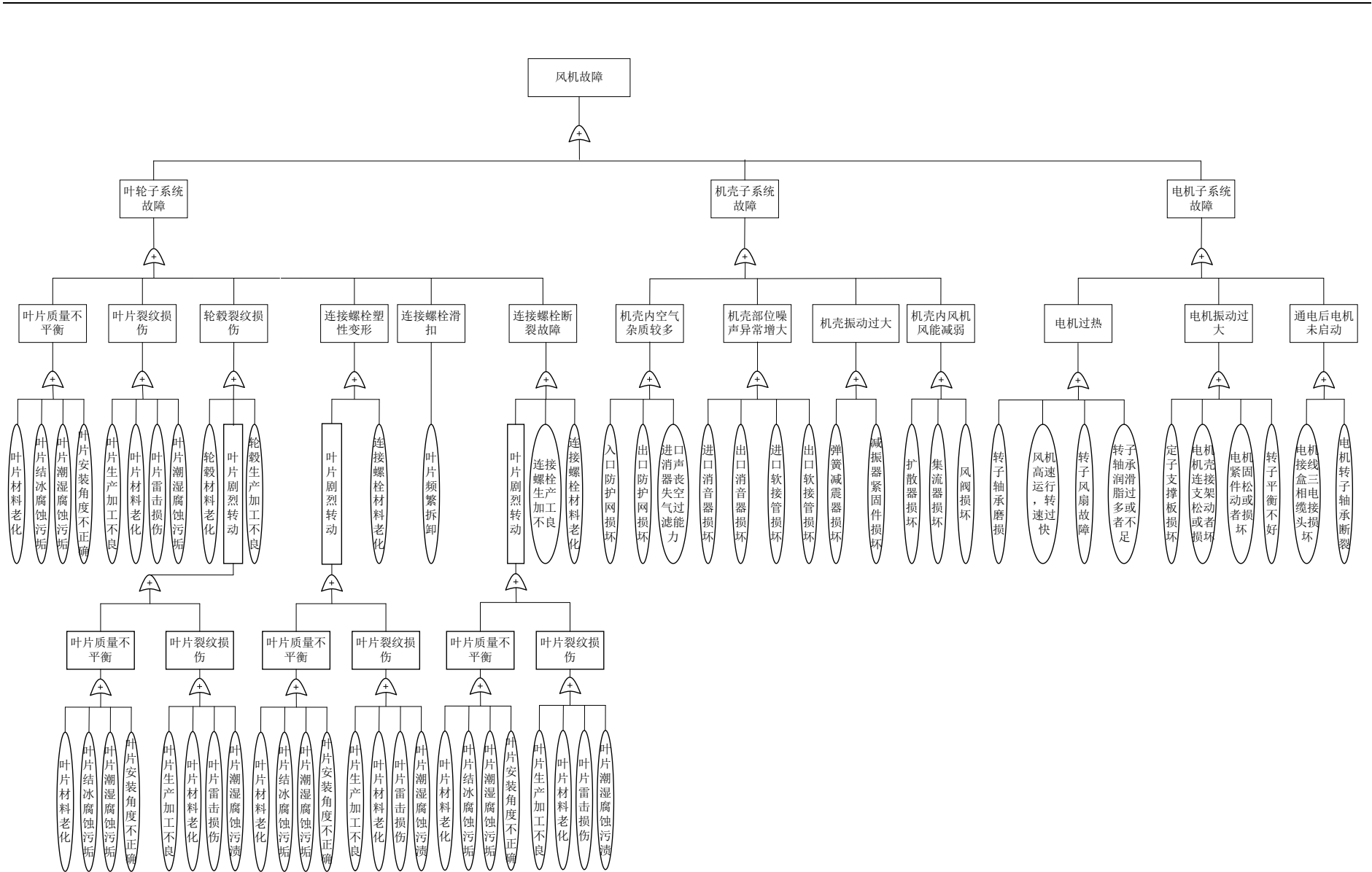


图 3.1 风机系统故障树

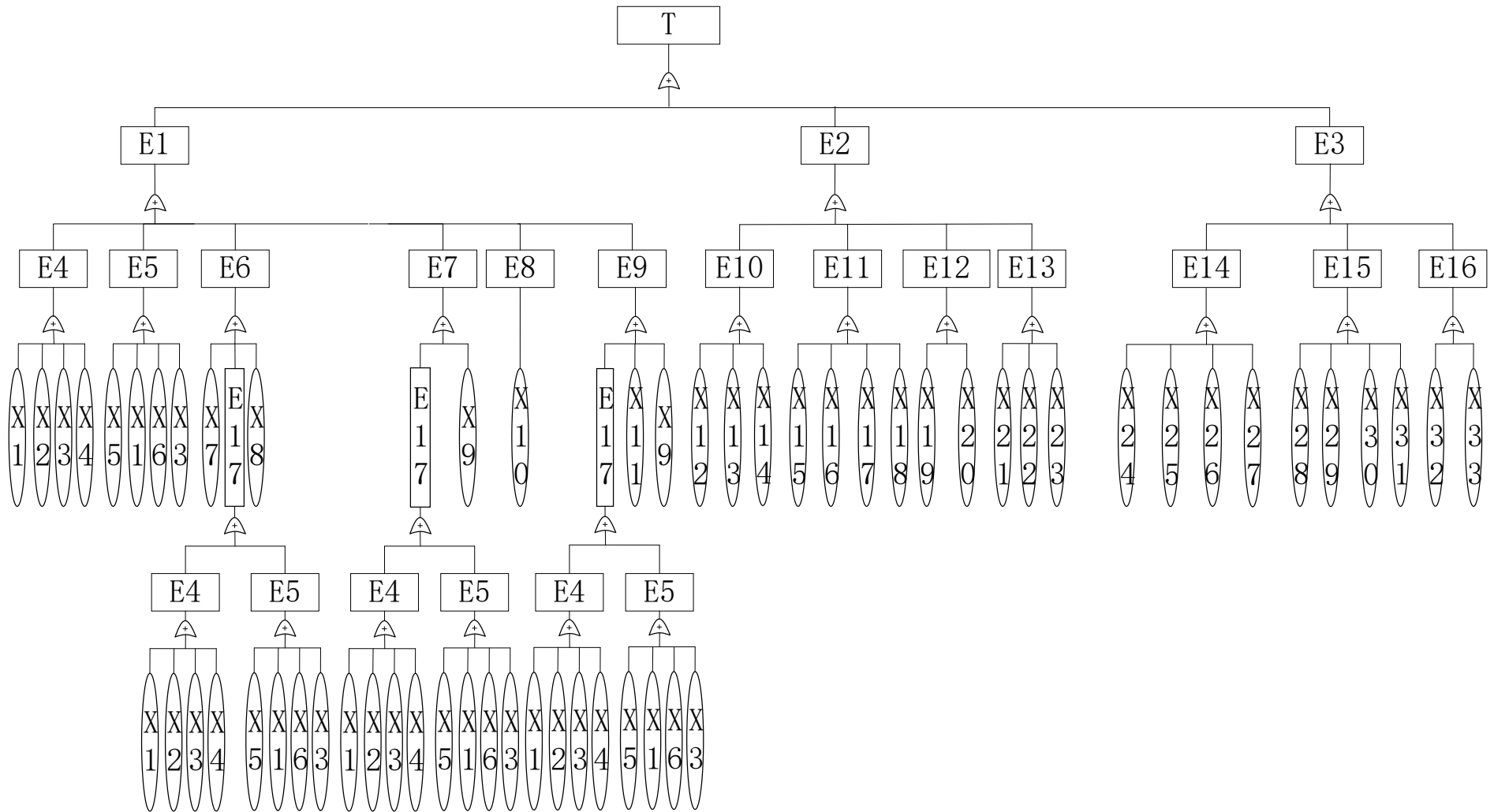


图 3.2 风机故障树符号图

图 3.1 中，风机故障具体可分为叶轮子系统故障、机壳子系统故障、电机子系统故障。

叶轮子系统故障可能由叶片质量不平衡、叶片裂纹损伤、轮毂裂纹损伤、连接螺栓塑性变形、连接螺栓滑扣、连接螺栓断裂故障导致，其中导致叶片质量不平衡的底层故障是叶片材料老化、叶片结冰腐蚀污垢、叶片潮湿腐蚀污垢、叶片安装角度不正确；导致叶片裂纹损伤的底层故障是叶片生产加工不良、叶片材料老化、叶片雷击损伤和叶片潮湿腐蚀污垢；导致轮毂裂纹损伤的底层故障是轮毂材料老化、叶片剧烈转动和轮毂生产加工不良；导致连接螺栓塑性变形的底层故障是叶片剧烈转动、连接螺栓材料老化；导致连接螺栓滑扣的底层故障是叶片频繁拆卸；导致连接螺栓断裂故障的底层故障是叶片剧烈转动、连接螺栓生产加工不良、连接螺栓材料老化。其中，叶片剧烈转动故障的发生是因为叶片质量不平衡和叶片裂纹损伤两个故障导致。

机壳子系统故障可能由机壳内空气杂质较多、机壳部位噪声异常增大、机壳振动过大和机壳内风能减弱导致。导致机壳内空气杂质较多故障发生的底层故障是入口防护网损坏、出口防护网损坏、进口消声器丧失空气过滤能力；导致机壳部位噪声异常增大的底层故障是进口消音器损坏、出口消音器损坏、进口软管管损坏、出口软管管损坏；导致机壳振动过大的底层故障是弹簧减震器损坏、减震器紧固件损坏；导致机壳内风机风能减弱故障的底层故障是扩散器损坏、集流器损坏、风阀损坏。

电机子系统故障可能由电机过热、电机振动过大、通电后电机未启动导致的。导致电机过热故障发生的底层故障是转子轴承磨损、风机高速运行导致转速过快、转子风扇故障、转子轴承润滑脂过多或者不足；导致电机振动过大故障发生的底层故障是定子支撑板松动、电机机壳连接支架松动或者损坏、电机紧固件松动或者损坏、转子平衡不好；导致通电后电机未启动故障发生的底层故障是电机接线盒三相电缆接头损坏、电机转子轴承断裂。

建立风机故障树后，再将故障树以表 2.2 的 FTA 符号表示，如图 3.2 所示。

图 3.2 中的符号含义见表 3.4。

表 3.4 故障树图中符号含义

符号	含义	符号	含义	符号	含义
T	风机故障	E17	叶片剧烈转动	X17	进口软接管损坏
E1	叶轮子系统故障	X1	叶片材料老化	X18	出口软接管损坏
E2	机壳子系统故障	X2	叶片结冰腐蚀污垢	X19	弹簧减震器损坏
E3	电机子系统故障	X3	叶片潮湿腐蚀污垢	X20	减振器紧固件损坏

表 3.4 故障树图中符号含义 (续)

符号	含义	符号	含义	符号	含义
E4	叶片质量不平衡	X4	叶片安装角度不正确	X21	扩散器损坏
E5	叶片裂纹损伤	X5	叶片生产加工不良	X22	集流器损坏
E6	轮毂裂纹损伤	X6	叶片雷击损伤	X23	风阀损坏
E7	连接螺栓塑性变形	X7	轮毂材料老化	X24	转子轴承磨损
E8	连接螺栓滑扣	X8	轮毂生产加工不良	X25	风机高速运行, 转速过快
E9	连接螺栓断裂故障	X9	连接螺栓材料老化	X26	转子风扇故障
E10	机壳内空气杂质较多	X10	叶片频繁拆卸	X27	转子轴承润滑脂过多或者不足
E11	机壳部位噪声异常增大故障	X11	连接螺栓生产加工不良	X28	定子支撑板损坏
E12	机壳振动过大	X12	入口防护网损坏	X29	电机机壳连接支架松动或者损坏
E13	机壳内风能减弱	X13	出口防护网损坏	X30	电机紧固件松动或者损坏
E14	电机过热	X14	进口消声器丧失空气过滤能力	X31	转子平衡不好
E15	电机振动过大	X15	进口消音器损坏	X32	电机接线盒三相电缆接头损坏
E16	通电后电机未启动	X16	出口消音器损坏	X33	电机转子轴承断裂

4) 定性分析确定最小割集

创建故障树后, 再完成 FTA 的定性分析, 主要进行最小割集分析, 目的是理清各个底层故障的串并联关系。串联时多个故障同时发生才会导致上一级故障发生; 并联时其中一个故障发生, 就会导致上一级故障发生。这两种不同情况影响风机故障概率的大小, 进一步将影响各个故障的概率重要度。

本文进行 FTA 的定性分析时从故障树的底事件出发, 自下而上逐层进行事件集合运算, 最后得到风机故障树的全部最小割集。

在图 3.2 中, 该故障数的底层的逻辑关系是:

$$\begin{aligned} E_4 &= X_1 \cup X_2 \cup X_3 \cup X_4 \\ E_5 &= X_5 \cup X_1 \cup X_6 \cup X_3 \end{aligned} \quad (3.2)$$

第二层逻辑是:

$$E_{17} = E_4 \cup E_5 = X_1 \cup X_2 \cup X_3 \cup X_4 \cup X_5 \cup X_6 \quad (3.3)$$

第三层逻辑是:

$$\begin{aligned}
E_4 &= X_1 \cup X_2 \cup X_3 \cup X_4 \\
E_5 &= X_5 \cup X_1 \cup X_6 \cup X_3 \\
E_6 &= X_7 \cup E_{17} \cup X_8 \\
E_7 &= E_{17} \cup X_9 \\
E_8 &= X_{10} \\
E_9 &= E_{17} \cup X_{11} \cup X_9 \\
E_{10} &= X_{12} \cup X_{13} \cup X_{14} \\
E_{11} &= X_{15} \cup X_{16} \cup X_{17} \cup X_{18} \\
E_{12} &= X_{19} \cup X_{20} \\
E_{13} &= X_{21} \cup X_{22} \cup X_{23} \\
E_{14} &= X_{24} \cup X_{25} \cup X_{26} \cup X_{27} \\
E_{15} &= X_{28} \cup X_{29} \cup X_{30} \cup X_{31} \\
E_{16} &= X_{32} \cup X_{33}
\end{aligned} \tag{3.4}$$

第四层逻辑是：

$$\begin{aligned}
E_1 &= E_4 \cup E_5 \cup E_6 \cup E_7 \cup E_8 \cup E_9 \\
E_2 &= E_{10} \cup E_{11} \cup E_{12} \cup E_{13} \\
E_3 &= E_{14} \cup E_{15} \cup E_{16}
\end{aligned} \tag{3.5}$$

最上层逻辑是：

$$\begin{aligned}
T &= E_1 \cup E_2 \cup E_3 \\
&= X_1 \cup X_2 \cup X_3 \cup X_4 \cup X_5 \cup X_6 \cup X_7 \cup X_8 \cup X_9 \cup X_{10} \cup X_{11} \\
&\quad \cup X_{12} \cup X_{13} \cup X_{14} \cup X_{15} \cup X_{16} \cup X_{17} \cup X_{18} \cup X_{19} \cup X_{20} \cup X_{21} \\
&\quad \cup X_{22} \cup X_{23} \cup X_{24} \cup X_{25} \cup X_{26} \cup X_{27} \cup X_{28} \cup X_{29} \cup X_{30} \cup X_{31} \\
&\quad \cup X_{32} \cup X_{33}
\end{aligned} \tag{3.6}$$

由式 (3.6) 的结果可以得到，一共有 33 个最小割集：

$$\begin{aligned}
&\{X_1\}, \{X_2\}, \{X_3\}, \{X_4\}, \{X_5\}, \{X_6\}, \{X_7\}, \{X_8\}, \{X_9\}, \{X_{10}\}, \{X_{11}\}, \{X_{12}\}, \{X_{13}\}, \{X_{14}\}, \\
&\{X_{15}\}, \{X_{16}\}, \{X_{17}\}, \{X_{18}\}, \{X_{19}\}, \{X_{20}\}, \{X_{21}\}, \{X_{22}\}, \{X_{23}\}, \{X_{24}\}, \{X_{25}\}, \{X_{26}\}, \{X_{27}\}, \\
&\{X_{28}\}, \{X_{29}\}, \{X_{30}\}, \{X_{31}\}, \{X_{32}\}, \{X_{33}\}
\end{aligned}$$

由以上结果可知，最小割集全部为一阶割集，这也意味着只要这些割集事件有一个发生，那么顶事件必然发生，即风机会发生故障。

基于最小割集即可计算风机零部件的概率重要度。

5) 基于最小割集进行定量计算基本事件的概率重要度

设 $X_1, X_2, X_3 \cdots X_{33}$ 发生的概率为 $p_1, p_2, p_3 \cdots p_{33}$ ，则风机发生故障的概率为：

$$P_T = 1 - \prod_{i=1}^{11} (1 - p_i) \quad (3.7)$$

其中, P_T 指风机发生故障的概率, p_i 表示 33 个最小割集发生故障的概率, $i=1,2,\dots,33$ 。

通过问卷收集 30 名高速公路机电设备管理专家、隧道风机维修人员、隧道设备管理人员、隧道仓库管理人员的建议,并通过对同类型风机的故障情况进行调研,同时研究风机零部件的产品说明书,确定了最小割集的发生概率,如表 3.5 所示。其中概率值与事件发生的可能性间的对应关系如下: 0.01: 一年内几乎不可能发生; 0.1: 一年内发生的可能性较小; 0.3: 一年内可能发生; 0.5: 一年内可能发生且概率较大; 0.7: 一年内有极大的可能发生; 0.9: 近乎一定会发生。

表 3.5 最小割集发生概率表

最小割集编号	发生概率	最小割集编号	发生概率	最小割集编号	发生概率
X1	0.3	X12	0.3	X23	0.01
X2	0.1	X13	0.1	X24	0.5
X3	0.3	X14	0.3	X25	0.1
X4	0.01	X15	0.1	X26	0.3
X5	0.1	X16	0.1	X27	0.5
X6	0.1	X17	0.1	X28	0.1
X7	0.3	X18	0.1	X29	0.1
X8	0.1	X19	0.3	X30	0.1
X9	0.3	X20	0.3	X31	0.5
X10	0.3	X21	0.1	X32	0.1
X11	0.1	X22	0.1	X33	0.7

基于式 (3.7) 计算风机发生故障的概率为:

$$P_T = 0.9982$$

其表示风机在一年内发生故障的可能性是 0.9982。

最小割概率集重要度指的是最小割集对顶事件发生的贡献。最小割集的概率重要度可通过下式进行计算:

$$P_{B/T} = \frac{P_B}{P_T} \quad (3.8)$$

其中, $P_{B/T}$ 表示最小割集概率重要度, P_B 表示最小割集发生的概率, P_T 表示风机故障发生的概率。通过式 (3.8) 计算出最小割集概率重要度,同时为方便下文使用概率重要度计算状态监测评价值,采用式 (3.9) 对数值进行归一化,结果如表 3.6 所示。

$$P_{\text{norm}_i} = \frac{P_{B/T_i}}{\sum_{i=1}^{33} P_{B/T_i}} \quad (3.9)$$

式中： P_{norm_i} 为第 i 个最小割集归一化后的值； P_{B/T_i} 为第 i 个最小割集的概率重要度。

表 3.6 最小割集概率重要度及归一化值

编号	最小割集	概率重要度	归一化	编号	最小割集	概率重要度	归一化	编号	最小割集	概率重要度	归一化
1	X1	0.3005	0.04336	12	X12	0.3005	0.04336	23	X23	0.0100	0.00144
2	X2	0.1001	0.01444	13	X13	0.1001	0.01444	24	X24	0.5009	0.07227
3	X3	0.3005	0.04336	14	X14	0.3005	0.04336	25	X25	0.1001	0.01444
4	X4	0.0100	0.00144	15	X15	0.1001	0.01444	26	X26	0.3005	0.04336
5	X5	0.1001	0.01444	16	X16	0.1001	0.01444	27	X27	0.5009	0.07227
6	X6	0.1001	0.01444	17	X17	0.1001	0.01444	28	X28	0.1001	0.01444
7	X7	0.3005	0.04336	18	X18	0.1001	0.01444	29	X29	0.1001	0.01444
8	X8	0.1001	0.01444	19	X19	0.3005	0.04336	30	X30	0.1001	0.01444
9	X9	0.3005	0.04336	20	X20	0.3005	0.04336	31	X31	0.5009	0.07227
10	X10	0.3005	0.04336	21	X21	0.1001	0.01444	32	X32	0.1001	0.01444
11	X11	0.1001	0.01444	22	X22	0.1001	0.01444	33	X33	0.7012	0.10117

3.3.3 基于调研的因素计算

1. 故障易检测度

故障易检测度是风机发生的故障所处风机位置、层次是否容易进行故障检测。易检测度根据等级由高到低依次为 I-III。I 级：故障部位位于风机内部、深层次处，维修时不易被首先发现故障；II 级：故障部位位于风机中部位置，维修时容易被发现故障；III 级：故障部位位于风机表层，维修时很容易被发现故障。以轴承断裂故障为例，当轴承断裂故障发生后，风机会停机，需要全面的进行故障排查才可以找到故障，同时轴承位于风机内部的电机里，根据故障排查由外到内、由表层到深入的原则，位于最内层的电机轴承往往最后被发现故障，维修耗时耗力。

为方便下文使用故障易检测度计算状态监测评价值，对故障易检测度 I-III 的等级进行数值化、归一化。通过对故障易检测度进行分析，定义 I 级数值为 1，II 级数值为 0.5，III 级为 0.1。

2. 故障易状态监测度

故障易状态监测度是指风机故障能否进行状态监测或者说当前是否存在对其进行状态监测的技术。易状态监测度根据等级由高到低依次为 I-III。I 级：存在状态监测技术且容易实施；II 级：存在状态监测技术但是实施存在一定困难；III 级：不存在状态监

测技术或者由于风机工作条件限制,难以安装状态监控技术。以风机叶片老化为例,受叶片自身工作时高速转动的工作条件限制,对其安装传感器的技术难度大,同时受到风机振动的影响,叶片老化的故障特征容易淹没在噪声中,故障信号难以提取、难以进行状态监测,其属于 III 级。

故障易检测度是指故障后对风机进行故障检测。故障易状态监测度是在风机故障前考虑是否可以状态监测,两者具有本质区别。

为方便下文使用故障易检测度计算状态监测评价值,对故障易状态监测度 I-III 的等级进行数值化、归一化。通过对故障易状态监测度进行分析,定义 I 级数值为 1, II 级数值为 0.5, III 级为 0。

3. 维修经济因素

维修经济因素指风机故障导致零部件损坏后的维修/更换成本。如果维修/更换成本较大的话,需要安装状态监测点在故障早期发现故障,降低风维修/更换成本。

为方便下文使用维修经济因素计算状态监测评价值,对故障经济因素 I-III 的等级进行数值化、归一化。通过对经济进行分析,不同零部件的成本差距较大,因此定义 I 级数值为 10, II 级数值为 1, III 级为 0.1。

4. 基于调研的风机故障的易检测度、易状态监测度、维修经济因素分析

通过问卷收集 30 名高速公路机电设备管理专家、隧道风机维修人员、隧道设备管理人员、隧道仓库管理人员的建议,并通过对同类型风机的故障情况进行调研,同时研究风机零部件的产品说明书,确定了风机各个故障的易检测度、易状态监测度、经济因素。整理出结果如表 3.7 所示:

表 3.7 风机故障易检测度、易状态监测度、经济因素表

编号	故障	易检测度			易状态监测度			经济因素		
		等级划分	数值化	归一化值	等级划分	数值化	归一化值	等级划分	数值化	归一化值
1	叶片材料老化	II	0.5	0.0357	III	0	0	I	10	0.1412
2	叶片结冰腐蚀污垢	III	0.1	0.0071	III	0	0	I	10	0.1412
3	叶片潮湿腐蚀污垢	III	0.1	0.0071	III	0	0	I	10	0.1412
4	叶片安装角度不正确	III	0.1	0.0071	III	0	0	I	10	0.1412
5	叶片生产加工不良	III	0.1	0.0071	III	0	0	I	10	0.1412

表 3.7 风机故障的易检测度、易状态监测度、经济因素表 (续)

编号	故障	易检测度			易状态监测度			经济因素		
		等级划分	数值化	归一化值	等级划分	数值化	归一化值	等级划分	数值化	归一化值
6	叶片雷击损伤	III	0.1	0.0071	III	0	0	I	10	0.1412
7	轮毂材料老化	II	0.5	0.0357	III	0	0	II	1	0.0141
8	轮毂生产加工不良	III	0.1	0.0071	II	0.5	0.025	II	1	0.0141
9	连接螺栓材料老化	II	0.5	0.0357	I	1	0.05	III	0.1	0.0014
10	叶片频繁装拆	III	0.1	0.0071	III	0	0	III	0.1	0.0014
11	连接螺栓生产加工不良	III	0.1	0.0071	I	1	0.05	III	0.1	0.0014
12	入口防护网损坏	III	0.1	0.0071	II	0.5	0.025	III	0.1	0.0014
13	出口防护网损坏	III	0.1	0.0071	II	0.5	0.025	III	0.1	0.0014
14	进口消声器丧失空气过滤功能	III	0.1	0.0071	II	0.5	0.025	III	0.1	0.0014
15	进口消音器损坏	III	0.1	0.0071	I	1	0.05	II	1	0.0141
16	出口消音器损坏	III	0.1	0.0071	I	1	0.05	II	1	0.0141
17	进口软接管损坏	II	0.5	0.0357	I	1	0.05	III	0.1	0.0014
18	出口软接管损坏	II	0.5	0.0357	I	1	0.05	III	0.1	0.0014
19	弹簧减震器损坏	III	0.1	0.0071	I	1	0.05	II	1	0.0141
20	减震器紧固件损坏	III	0.1	0.0071	I	1	0.05	II	1	0.0141
21	扩散器损坏	II	0.5	0.0357	I	1	0.05	II	1	0.0141
22	集流器损坏	II	0.5	0.0357	I	1	0.05	II	1	0.0141
23	风阀损坏	II	0.5	0.0357	I	1	0.05	II	1	0.0141
24	转子轴承磨损	I	1	0.0714	I	1	0.05	III	0.1	0.0014
25	风机高速运行, 转速过高	II	0.5	0.0357	I	1	0.05	III	0.1	0.0014
26	转子风扇故障	I	1	0.0714	II	0.5	0.025	III	0.1	0.0014
27	转子轴承润滑脂过多或者不足	I	1	0.0714	I	1	0.05	III	0.1	0.0014
28	定子支撑板松动	I	1	0.0714	II	0.5	0.025	III	0.1	0.0014
29	电机机壳连接支架松动或者损坏	II	0.5	0.0357	II	0.5	0.025	III	0.1	0.0014
30	电机紧固件松动或者损坏	II	0.5	0.0357	II	0.5	0.025	III	0.1	0.0014
31	转子平衡不好	I	1	0.0714	I	1	0.05	III	0.1	0.0014
32	电机接线盒三相电缆接头损坏	I	1	0.0714	III	0	0	III	0.1	0.0014
33	电机转子轴承断裂	I	1	0.0714	I	1	0.05	III	0.1	0.0014

3.4 基于 AHP 的多因素权重分配模型

3.3 节中通过可靠性分析的方法确定了风机故障的危害度、概率重要度、易检测度、易状态维修度、维修经济因素这五种因素的对应值，本节对五种因素对于设置状态监测点的影响权重进行分配。

基于 2.2.2 节中 AHP 方法的分析流程确定本文的五种因素对于设置状态监测点的影响权重。设计流程如下：

1. 建立层次结构模型

建立三级层次结构：目标层、准则层、方案层。目标层是针对风机某个故障安装状态监测点的评价值。准则层是风机故障的危害度、概率重要度、易检测度、易状态监测度、维修经济因素这五种评价准则。方案层是风机的 33 种故障。

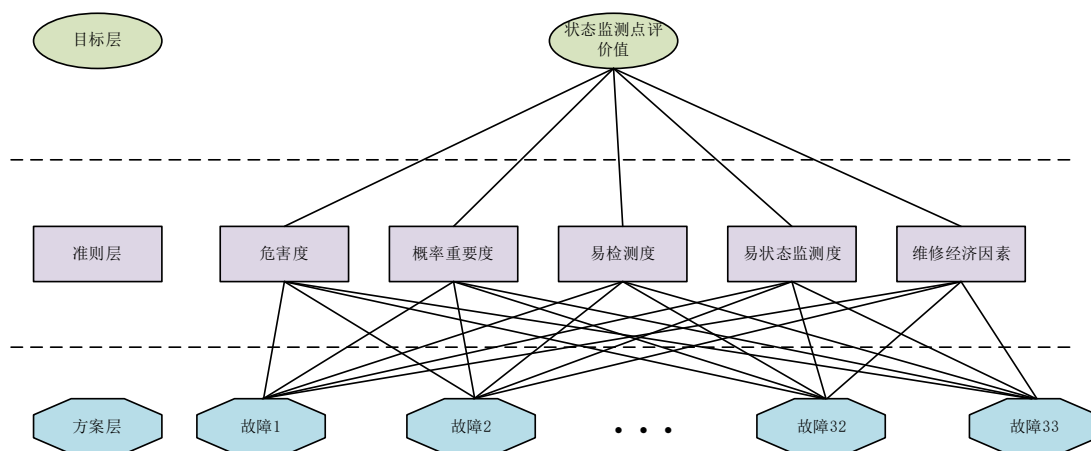


图 3.3 风机状态监测点优选的层级结构模型

2. 建立判断矩阵并计算权重

3.3 节中已经确定了方案层中各个故障关于准则层的权重。这里只需计算准则层的危害度、概率重要度、易检测度、易状态监测度、维修经济因素这五种因素关于目标层的判断矩阵。邀请相关领域的专家对风机故障的危害度、概率重要度、易检测度、易状态监测度、维修经济因素进行调研。首先，专家认为能否进行状态监测，是确定风机状态监测点的前提，因此易状态监测度要明显重要于其他几个指标；其次，专家认为易检测度指标反应了维修难度的变化，应尽量减轻维修难度，降低对工作人员的技能要求，因此易检测度指标要稍微重要于危害度、概率重要度、维修经济因素；最后，专家认为风机的危害度仅影响风机本身，并没有对人、社会造成影响，危害度直接反应了风机的维修成本，因此危害度与经济因素具有同样的重要性，但是应稍微高于概率重要度，因为概率重要度指标反应了维修频次的大小，影响着工作人员的工作量，而当前隧道维修人员充足，工作量不是很重要的考虑因素。根据上述分析以及表 2.3 判断矩阵的填写准则填写表 3.8。

表 3.8 准则层关于目标层的判断矩阵

状态监测点评价值	危害度	概率重要度	易检测度	易状态监测度	维修经济因素
危害度	1	2	1/4	1/6	1
概率重要度	1/2	1	1/4	1/6	1/2
易检测度	4	4	1	1/6	4
易状态监测度	6	6	6	1	6
维修经济因素	1	2	1/4	1/6	1

3. 一致性检验

判断矩阵 $A = \begin{bmatrix} 1 & 2 & \frac{1}{4} & \frac{1}{6} & 1 \\ \frac{1}{2} & 1 & \frac{1}{4} & \frac{1}{6} & \frac{1}{2} \\ 4 & 4 & 1 & \frac{1}{6} & 4 \\ 6 & 6 & 6 & 1 & 6 \\ 1 & 2 & \frac{1}{4} & \frac{1}{6} & 1 \end{bmatrix}$ 。

① 计算一致性指标 CI

假设判断矩阵 A 的特征值为 λ ，特征向量为 X ，由特征值与特征向量的关系可得： $(\lambda E - A)X = 0$ ，其中 E 为单位矩阵。由特征向量的定义可知其不为 0，因此 $\lambda E - A$ 组成的行列式 $|\lambda E - A| = 0$ ，求解可到 $\lambda_{max} = 5.32$ 。同时 A 的维度为 5，因此 $n=5$ 。根据式 (2.1) 计算得 $CI=0.08$ 。

② 查找对应的平均随机一致性指标 RI

当 $n=5$ 时，由表 2.4 可知 $RI=1.12$ 。

③ 计算一致性比例

由①、②可知， $CI=0.08$ ， $RI=1.12$ ，由式 (2.2) 可知 $CR=CI/RI=0.0714 < 0.1$ 。因此认为表 3.8 的判断矩阵没有出现自相矛盾的情况，不需要进行修改。

4. 求组合权重

根据式 (2.3) 以及表 3.8 准则层关于目标层的判断矩阵求得目标层与准则层之间的权重分别为 $\omega_1=0.0852$ ， $\omega_2=0.0554$ ， $\omega_3=0.2272$ ， $\omega_4=0.547$ ， $\omega_5=0.0852$ 。即图 3.3 中目标层的状态监测点评价值到准则层的危害度的权重为 0.0852、到概率重要度的权重为 0.0554、到易检测度的权重为 0.2272、到易状态监测度的权重是 0.547、到维修经济因素的权重是 0.0852。

5. 计算结果

3.5 状态监测点优选结果

基于 3.3 节对风机故障的危害度、概率重要度、易检测度、易状态监测度、维修经济因素这五种因素的计算值以及 3.4 节对五种因素的权重分配，计算针对风机故障设置状态监测点的评价值为表 3.9 所示。状态监测点评价值越高，说明该类故障越需要安装状态监测点。

表 3.9 状态监测点多因素评价值

编号	故障	危害度 归一化 值	概率重 要度归 一化值	易检测 度归一 化值	易状态 监测度 归一化 值	维修经 济因素 归一化 值	状态监 测点评 价值
1	叶片材料老化	0.02785	0.04336	0.0357	0	0.1412	0.024916
2	叶片结冰腐蚀污垢	0.02785	0.01444	0.0071	0	0.1412	0.016816
3	叶片潮湿腐蚀污垢	0.02785	0.04336	0.0071	0	0.1412	0.018418
4	叶片安装角度不正确	0.02785	0.00144	0.0071	0	0.1412	0.016096
5	叶片生产加工不良	0.02785	0.01444	0.0071	0	0.1412	0.016816
6	叶片雷击损伤	0.02785	0.01444	0.0071	0	0.1412	0.016816
7	轮毂材料老化	0.02785	0.04336	0.0357	0	0.0141	0.014087
8	轮毂生产加工不良	0.02785	0.01444	0.0071	0.025	0.0141	0.019662
9	连接螺栓材料老化	0.04178	0.04336	0.0357	0.05	0.0014	0.041542
10	叶片频繁装拆	0.02785	0.04336	0.0071	0	0.0014	0.006507
11	连接螺栓生产加工不良	0.05571	0.01444	0.0071	0.05	0.0014	0.034629
12	入口防护网损坏	0.02785	0.04336	0.0071	0.025	0.0014	0.020182
13	出口防护网损坏	0.02785	0.01444	0.0071	0.025	0.0014	0.01858
14	进口消声器丧失空气过滤功能	0.02785	0.04336	0.0071	0.025	0.0014	0.020182
15	进口消音器损坏	0.00557	0.01444	0.0071	0.05	0.0141	0.031439
16	出口消音器损坏	0.00557	0.01444	0.0071	0.05	0.0141	0.031439
17	进口软接管损坏	0.02785	0.01444	0.0357	0.05	0.0014	0.038753
18	出口软接管损坏	0.02785	0.01444	0.0357	0.05	0.0014	0.038753
19	弹簧减震器损坏	0.02785	0.04336	0.0071	0.05	0.0141	0.034939
20	减震器紧固件损坏	0.02785	0.04336	0.0071	0.05	0.0141	0.034939
21	扩散器损坏	0.02785	0.01444	0.0357	0.05	0.0141	0.039835
22	集流器损坏	0.02785	0.01444	0.0357	0.05	0.0141	0.039835
23	风阀损坏	0.05571	0.00144	0.0357	0.05	0.0141	0.041489
24	转子轴承磨损	0.02785	0.07227	0.0714	0.05	0.0014	0.050068
25	风机高速运行, 转速过高	0.02785	0.01444	0.0357	0.05	0.0014	0.038753
26	转子风扇故障	0.02785	0.04336	0.0714	0.025	0.0014	0.034791
27	转子轴承润滑脂过多或者不足	0.02785	0.07227	0.0714	0.05	0.0014	0.050068
28	定子支撑板松动	0.02785	0.01444	0.0714	0.025	0.0014	0.033189
29	电机机壳连接支架松动或者损坏	0.02785	0.01444	0.0357	0.025	0.0014	0.025078

表 3.9 状态监测点多因素评价值（续）

编号	故障	危害度 归一化 值	概率重 要度归 一化值	易检测 度归一 化值	易状态 监测度 归一化 值	经济因 素归一 化值	状态监 测点评 价值
30	电机紧固件松动 或者损坏	0.02785	0.01444	0.0357	0.025	0.0014	0.025078
31	转子平衡不好	0.02785	0.07227	0.0714	0.05	0.0014	0.050068
32	电机接线盒三相 电缆接头损坏	0.05571	0.01444	0.0714	0	0.0014	0.021888
33	电机转子轴承断 裂	0.05571	0.10117	0.0714	0.05	0.0014	0.054043

根据状态监测点评价值由高到低的顺序对故障的编号进行排序：33、24、27、31、9、23、21、22、17、18、25、19、20、26、11、28、15、16、29、30、1、32、12、14、8、13、3、2、5、6、4、7、10。在建立基于无线传输的风机故障诊断系统时，可根据传输带宽条件，由状态监测点评价值从高到低安装状态监测点，在有限的带宽条件下，尽可能的监测重要故障。由排序可知，33（电机转子轴承断裂）分值最高，应该首先安排状态监测点，轴承位于电机内部，通过在电机表面安装状态监测点，可有效对其进行状态监测。

3.6 本章小结

本章针对建立基于无线传输的风机故障诊断系统时存在的采样信息量受限于窄带宽传输条件的问题，提出面向可靠性的风机状态监测点优选方法。该方法首先对终南山隧道的轴流风机结构进行分析。然后基于分析结果使用可靠性分析模型的 FMECA、FTA 和调研的方法确定风机故障的危害度、概率重要度、易检测度、易状态监测度以及维修经济因素。接着使用 AHP 方法对这五种因素进行权重分配。最后基于因素值和权重值计算状态监测评价值，进而获得监测点重要度排序。可根据传输带宽条件，由状态监测点评价值从高到低的顺序安装状态监测点。通过本文的风机状态监测点优选方法可实现有限带宽条件下尽可能对重要故障安装监测点的目的，去除了不必要的状态监测点，缩减了采样的信息量，同时为第四章研究特征选择算法提供了研究对象。

第四章 基于 Filter 和 Wrapper 相结合的两步特征选择算法研究

4.1 引言

第三章通过可靠性分析理论实现了窄带宽条件优选重要监测点的目标,同时确定风机电机轴承断裂故障需要优先设置监测点。本章以电机轴承故障诊断为背景,重点研究如何利用有限的传输带宽,从众多的特征中遴选出关键特征,从而减少单点监测的采样信息量,将有限的传输带宽传输更关键的信息。

根据 1.2.3 节分析,当前常用的 Wrapper 特征选择算法所选特征虽能实现量少、质优的目标,但其通用性差。有线场景将所有的特征都传输到后台进行存储,但无线场景仅传输 Wrapper 所选特征,隧道以新代旧的数据存储特点导致存储器中仅存储 Wrapper 所选特征,这样的数据存储差异导致 Wrapper 所选特征通用性差的缺点不会影响基于有线传输的故障诊断系统,但会限制基于无线传输的故障诊断系统中的故障诊断模型的更新换代。Filter 方法虽通用性强但所选特征诊断精度低。为此设计基于 Filter 和 Wrapper 相结合的两步特征选择算法,使用 Filter 方法选择量少、通用性强的特征进行传输,避免 Wrapper 方法选择通用性低的特征对故障诊断模型更新换代的影响。接收到数据后使用 Wrapper 方法对特征进行二次优选,提高故障诊断的精度。

4.2 特征选择算法研究对象分析

3.5 节确定电机轴承断裂故障需要进行状态监测。本节先分析轴承结构原理、故障形式及故障演变规律,然后确定特征选择算法的研究对象。

4.2.1 结构原理分析

轴承是风机电机中的重要零部件之一,其在工作时将电机各部件之间的滑动摩擦力转化为滚动摩擦力,减小了部件之间的摩擦阻力,对电机的稳定运行发挥着重要作用。由于制造误差、高负载和高转速的影响,轴承容易发生故障,故障后严重影响风机的使用。

滚动轴承的主要组成零部件为外圈、滚动体、保持架和内圈这四部分^[56]。其中,外圈一般固定在滚动轴承的轴承座上或者整个装备的壳体上,起到支撑滚动体的作用^[56];滚动体承受整个负荷,其数量和大小等决定轴承的性能,它主要将内外圈的滑动运动转化为滚动运动^[56];保持架主要将滚动体均匀的分开,可以起到引导滚动体运动和减少滚动体相互之间摩擦的作用^[56];内圈一般与旋转设备的轴相套,与转轴一起转动。

4.2.2 常见故障形式及故障演变

1. 滚动轴承常见形式

常见的故障形式有断裂、胶合、腐蚀、磨损、疲劳点蚀等，其主要的故障原因如下：

1) 断裂：轴承运转过程中摩擦产生的过大热应力和负荷量过大都会使轴承内部保持架发生断裂。

2) 胶合：在转速过高、工作负荷过大或润滑不足的情况下，容易引起轴承灼伤；当摩擦产生的热量严重过高时，金属表面相互接触的部分逐渐粘接，最终胶合在一起^[57]。

3) 腐蚀：风机总是处于环境恶劣的隧道风机机房中，机组经常受到潮湿空气的影响，主轴轴承很容易受到水分的腐蚀，使得轴承表面生锈。

4) 磨损：润滑不良会使轴承与其接触件的摩擦增大或者滚动体不对中都会造成轴承表面磨损，尤其是风机工作在粉尘大的环境下，磨损更加严重^[57]。

5) 疲劳点蚀：风机长期使用时，轴承表层形成初始疲劳裂纹，随着工作时间的增长，裂纹不断扩张，最后达到一定深度时，轴承断裂。

2. 隧道风机场景下轴承故障演变

风机位于山体内部的风机机房中，山体内部常年潮湿，在这样的场景下，首先会导致轴承发生腐蚀故障。由于轴承表面出现腐蚀，进一步会导致轴承产生裂纹引发疲劳点蚀故障以及磨损故障，进一步导致轴承在旋转过程中接触不良，金属表面相接触的地方粘连，引发胶合故障，最终导致轴承断裂，彻底不能使用，导致风机停机。

4.2.3 本文特征选择算法的研究对象

3.5 节风机状态监测点优选方法实例分析中，确定电机轴承断裂故障需要进行状态监测。但是根据隧道风机场景下轴承的故障演变可知，轴承断裂时轴承已停止工作，将导致风机停机。对轴承断裂进行故障诊断已不具有实际意义，而轴承腐蚀故障是轴承故障的早期表现形式，对其进行故障诊断，有利于预防轴承断裂故障，提高风机的运行稳定性。因此，本章以轴承腐蚀故障为特征选择算法的研究对象，对轴承外圈、滚动体和内圈的腐蚀故障建立故障诊断模型并研究相应的特征选择算法，重点研究轴承外圈、滚动体和内圈在腐蚀面积较小、中等、大面积条件下的故障诊断，由于保持架位于轴承内部，受潮湿空气的影响较小，因此对保持架的腐蚀故障不进行重点阐述。

4.3 基于 Filter 和 Wrapper 相结合的两步特征选择算法

4.3.1 两步特征选择算法设计

基于 Filter 和 Wrapper 相结合的两步特征选择算法描述如表 4.1 所示，模型如图 4.1 所示，算法流程图如图 4.2 所示。

在图 4.1 中，第一步使用 Filter 方法选择量少、通用性强的预选特征子集，然后根据 2.1.3 节中阐述的基于无线传输的风机故障诊断系统逻辑功能框架，感知终端设备提取预选特征子集并传输到中继设备。具体来说，使用 Fisher Score 方法进行特征重要度分析，Fisher Score 方法度量特征值辨别类的能力，通过阈值筛选的方式，剔除辨别类的能力差的特征，然后利用 MIC 方法进行特征冗余度分析，MIC 方法评价特征子集中两两特征之间的冗余性，同样通过阈值筛选的方式，剔除冗余度较高的特征，得到量少、通用性强的预选特征子集 F_2 进行传输。通过先进行特征重要度分析的方式，降低了特征冗余度分析时需要分析的特征数量，提高了计算效率。

第二步中继设备接收到预选特征子集后使用 Wrapper 方法对特征子集进行二次优选，提高故障诊断的精度，得到最优特征子集。具体来说，根据 BGWO 的计算规则不断从预选特征子集 F_2 中选择一个新的特征子集，输入到故障诊断模型计算诊断精度，根据诊断精度评价所选特征子集的优度，将优度最高（诊断精度最高）的特征子集输出为最优特征子集。

表 4.1 基于 Filter 和 Wrapper 相结合的两步特征选择算法描述

算法 4-1: 基于 Filter 和 Wrapper 相结合的两步特征选择算法

输入：特征集 $F(x_1, x_2, \dots, x_M)$ ， M 为特征的个数；Fisher Score 算法的重要度阈值 FS_{th} ；MIC 算法的冗余度阈值 MIC_{th} ；BGWO 算法的狼群数量 N ，最大迭代次数 $iter_{max}$ 。

输出：输出最优特征子集 F_{out} ，故障诊断精度 ACC。

第一步：基于 Filter 方法的预选特征子集生成

Step 1: 通过式 (2.6) 依次计算特征集 F 中各特征对应的 Fisher Score，记为 FS_k ；

Step 2: 将 FS_k 值低于重要度阈值 FS_{th} 的特征从特征集中剔除，形成预选特征子集

$F_1(x_1, x_2, \dots, x_s), s \leq M$ ；

Step 3: 将预选特征子集 F_1 中的特征两两组合，每个特征与除自己外的特征使用式 (2.8) 计算最大信息系数 $MIC(x_i, x_j)$ ，其中 $i \neq j$ ；

Step 4: 将预选特征子集 F_1 中两两特征的 MIC 高于冗余度阈值 MIC_{th} 的两个特征中 FS_k 值较低的特征从 F_1 中剔除，形成预选特征子集 $F_2(x_1, x_2, \dots, x_m), m \leq s$ ；

第二步：基于 Wrapper 方法的最优特征子集生成

Step 5: 基于输入的狼群数量 N 、最大迭代次数 $iter_{max}$ 设置 BGWO 的参数；

Step 6: 初始化狼群。设置式 (2.9) 中灰狼个体 X_i 的位置向量为预选特征集 F_2 ，即 $X_i = [x_1, x_2, \dots, x_m]$ ，并进行随机初始化， x_i 随机为 0 或 1，1 代表选择该特征，0 代表不选择该特征；

Step 7: 计算初代的灰狼个体的适应度，确定 α 、 β 、 δ 狼；即将灰狼个体位置向量初始化后选择的特征值输入到故障诊断算法计算诊断精度，精度越高，适应度越高（精度=适应度）。将适应度最高的位置向量定义为 α 狼所在位置，次之为 β 狼，最后为 δ 狼。

Step 8: 更新狼群位置，根据式 (2.17) 更新狼群位置；

Step 9: 计算位置更新后灰狼个体的适应度，将适应度最高的位置向量设置为 α 狼；

表 4.1 基于 Filter 和 Wrapper 相结合的两步特征选择算法描述（续）

算法 4-1: 基于 Filter 和 Wrapper 相结合的两步特征选择算法

Step 10: 进行循环迭代, 不断执行 step 8 更新狼群位置, step 9 计算狼群个体的适应度, 将适应度最高的位置向量设置为 α 狼, 当达到迭代次数之后, α 狼的位置即为最优特征子集, α 狼的适应度即为故障诊断的最高精度;

Step 11: 输出最优特征子集 F_{out} , 故障诊断精度 ACC。

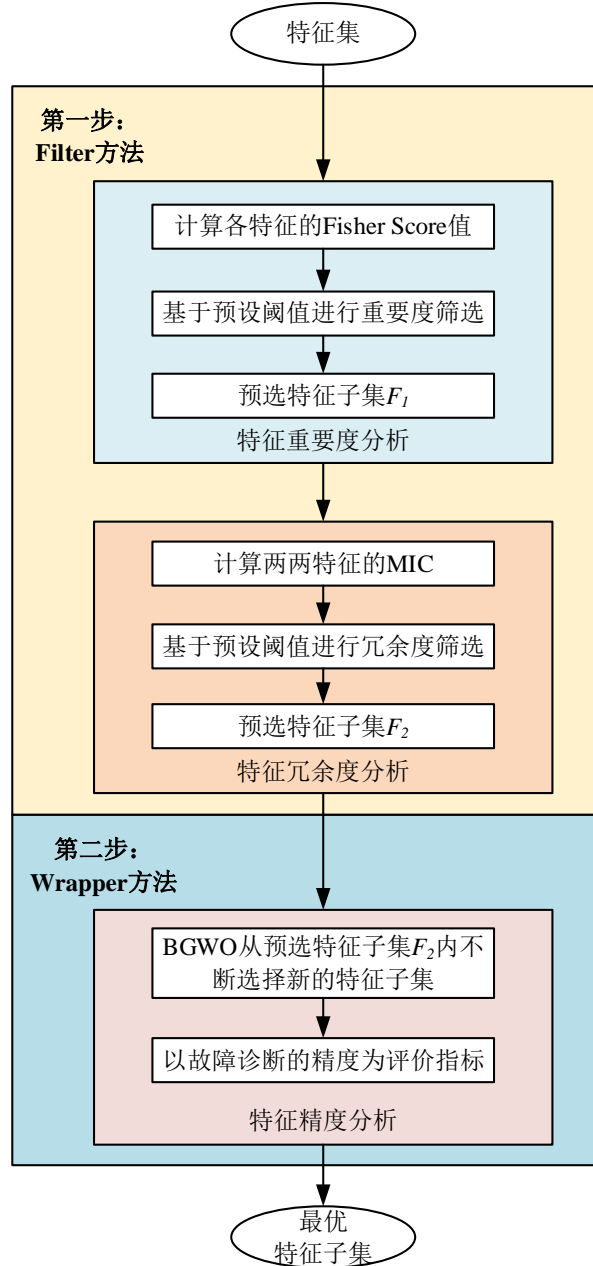


图 4.1 基于 Filter 和 Wrapper 相结合的两步特征选择模型

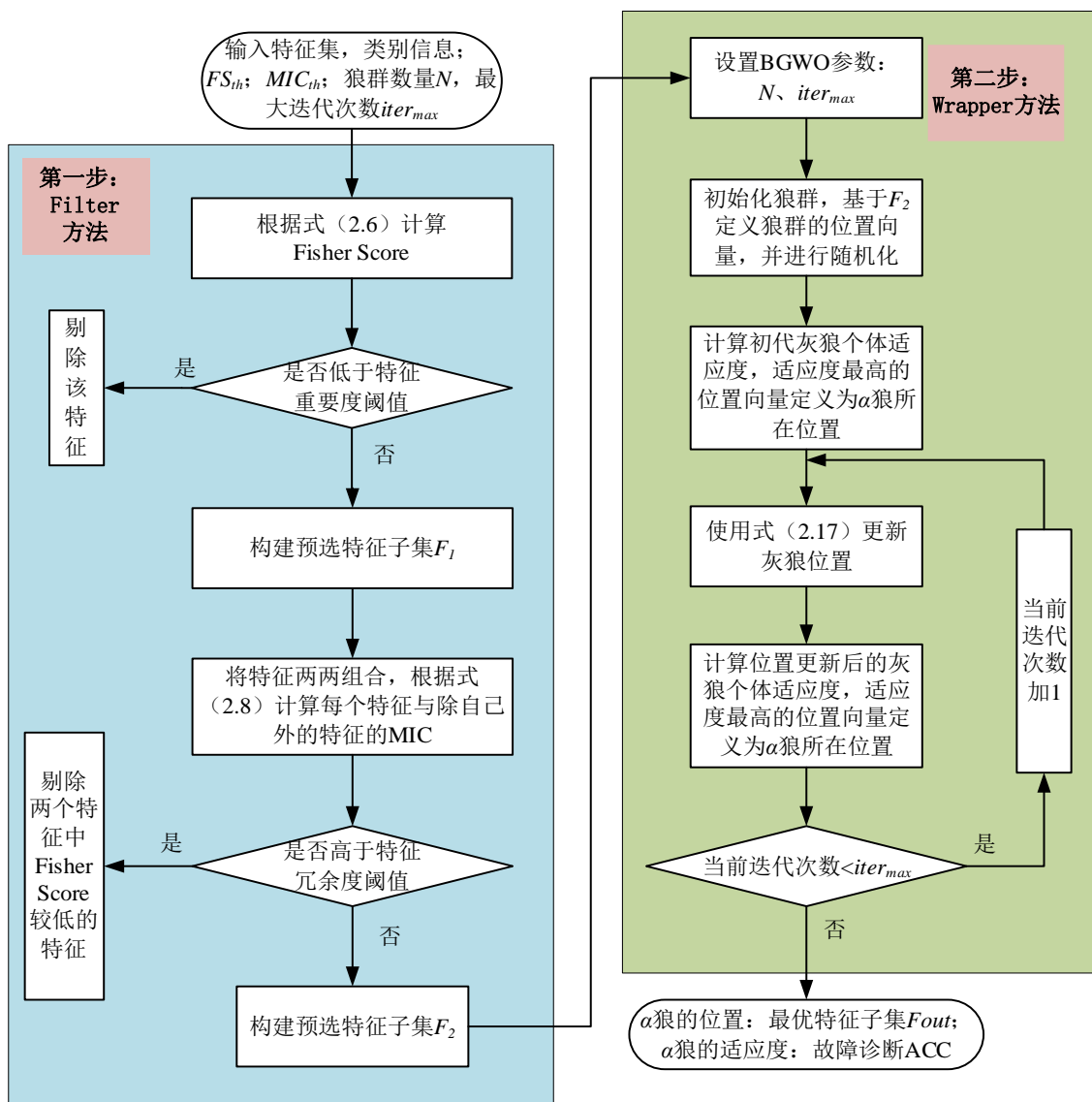


图 4.2 基于 Filter 和 Wrapper 相结合的两步特征选择算法设计流程

4.3.2 实验设计思路

1. 实验结果评价指标

文中所提出的基于 Filter 和 Wrapper 相结合的两步特征选择算法的目的在于选出量少、质优、通用性强的特征进行传输，缓解传输带宽的压力。因此，本文以所选特征数量与总特征数量的比值来衡量所选特征的数量多少；以故障诊断的精度（Accuracy, ACC）为指标评价所选特征子集的质量；通过在不同故障诊断模型上的精度（ACC）误差衡量所选特征的通用性。

ACC 表示能够正确诊断故障的样本数与样本总数的比值。如式（4.1）所示：

$$ACC = \frac{n_{correct}}{n} \quad (4.1)$$

式中： $n_{correct}$ 为能够正确诊断故障的样本数， n 为样本总数。

2. 实验验证内容

本文实验所用的故障诊断模型为 1.2.1 节阐述的 SVM、KNN 模型。实验设计思路如图 4.3 所示。具体来说，主要完成 5 项实验内容：

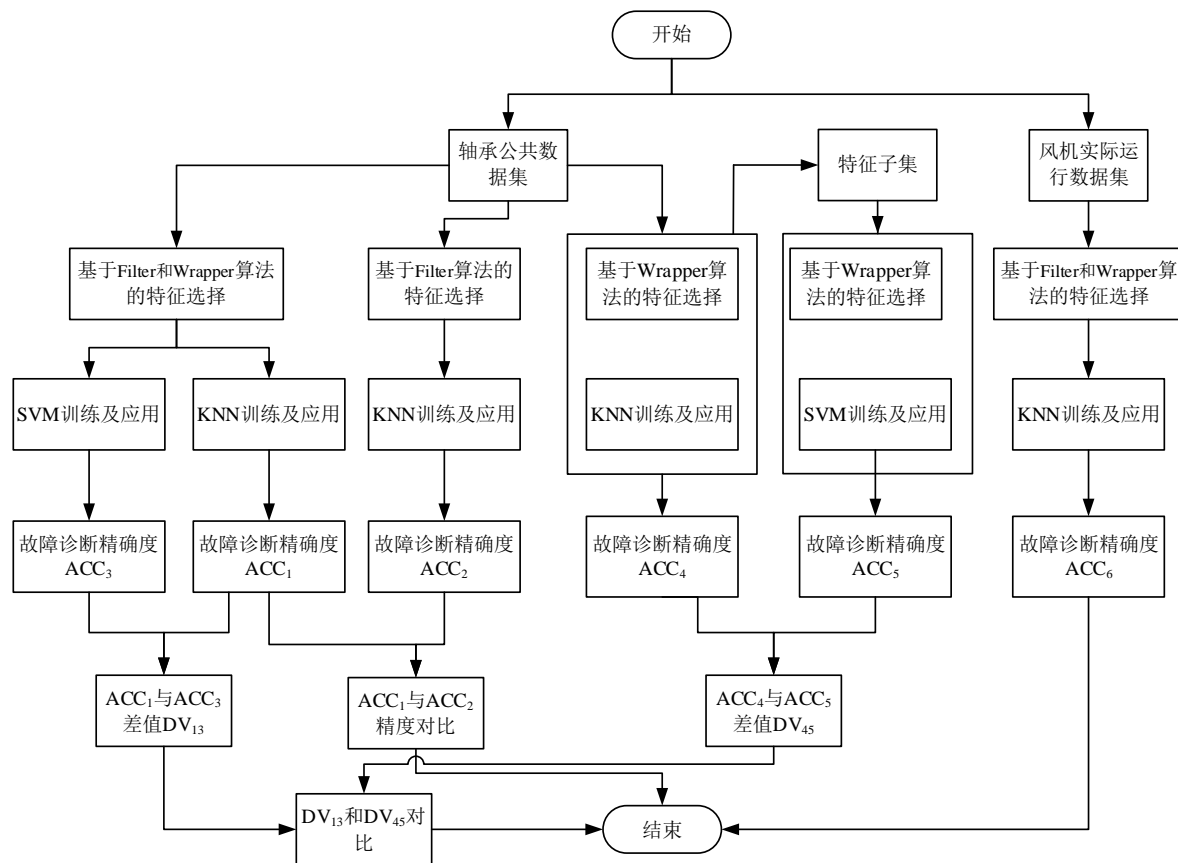


图 4.3 实验设计思路图

1) 计算本文方法在 KNN 故障诊断模型上的精度 ACC_1 ，该部分实验内容在 4.4 节中的 4.4.1、4.4.2 节中阐述；

2) 采样信息量分析。计算本文方法采样的信息量，由于其未涉及故障诊断精度计算，图 4.3 中未显示本部分，该部分实验内容在 4.4.3 节中阐述。

3) 故障诊断精度分析。计算仅使用 Filter 方法在 KNN 故障诊断模型上的精度 ACC_2 ，然后与本文方法在 KNN 故障诊断模型上的精度 ACC_1 进行对比，本文方法精度越高说明本文方法选择特征质量越好，该部分实验内容在 4.4.3 节中阐述。

4) 方法通用性分析。计算本文方法在 SVM 故障诊断模型上的精度 ACC_3 ，将 ACC_1 与 ACC_3 计算差值 DV_{13} ；计算使用 Wrapper 方法在 KNN 故障诊断模型上的精度 ACC_4 和特征子集，然后基于特征子集计算使用 Wrapper 方法在 SVM 故障诊断模型上的精度 ACC_5 ，将 ACC_4 与 ACC_5 计算差值 DV_{45} 。比较 DV_{13} 和 DV_{45} 的大小，越大说明该方法在不同故障诊断模型上的误差越小，该方法通用性越好。该部分实验内容在 4.4.3 节。

5) 本文方法在实际风机场景下的诊断效果 ACC₆，该部分实验内容在 4.4.3 节。

4.4 实验结果分析

4.4.1 样本数据及特征集构建

本文通过美国凯斯西储大学滚动轴承数据中心产生的、学术界普遍使用的轴承故障数据集，来验证本文所提的特征选择算法的可行性。该数据集的研究对象是 SKF 6205-2RS 型轴承，工作时轴承转速为 1797r/min。通过电火花的方式生成电腐蚀故障来模拟潮湿产生的腐蚀故障，通过采样频率为 12kHz 的传感器对轴承正常、故障状态采集 0.17s，获得近 2048 个数据组成一个样本，共采集 1200 个样本，其中轴承内圈、滚动体、外圈位置以及对应的轻度、中度、重度故障共九种状态各 120 个样本，共 1080 个样本；轴承正常 120 个样本。将 10 类样本按照 75%:25%的比例设置成训练集和测试集。

该试验数据集的样本划分及标签如表 4.2 所示。

表 4.2 实验数据集的样本划分及标签

编号	故障直径/mm	训练样本/个	测试样本/个	故障类型	样本标签
1	0.0000	90	30	正常	0
2	0.1778	90	30	内圈轻度故障	1
3	0.1778	90	30	滚动体轻度故障	2
4	0.1778	90	30	外圈轻度故障	3
5	0.3556	90	30	内圈中度故障	4
6	0.3556	90	30	滚动体中度故障	5
7	0.3556	90	30	外圈中度故障	6
8	0.5334	90	30	内圈重度故障	7
9	0.5334	90	30	滚动体重度故障	8
10	0.5334	90	30	外圈重度故障	9

将每个样本中的 2048 个样本点进行以附录 A 中表 A-1、表 A-2 的公式进行特征提取并按照表 4.3 所示构建为一个 26 维特征集。

$$X = \{F1, F2, F3, F4, F5, \dots, F26\}$$

表 4.3 原始特征集

特征值编号	特征值	特征值编号	特征值	特征值编号	特征值
F1	均值	F10	峭度	F19	中心频率
F2	标准差	F11	峰值指标	F20	频率标准差
F3	最大值	F12	波形指标	F21	频率均方根
F4	最小值	F13	脉冲指标	F22	P1
F5	绝对平均值	F14	裕度指标	F23	P2
F6	峰峰值	F15	偏态系数	F24	P3
F7	均方根	F16	峭度系数	F25	P4
F8	方根幅值	F17	变异系数	F26	P5
F9	偏斜度	F18	平均频率		

其中 P1-P5 无实际物理意义，其用来反应频域信号的离散程度。

4.4.2 最优特征子集生成

根据表 4.1 的算法流程进行基于 Filter 和 Wrapper 相结合算法的最优特征子集生成:

1. Fisher Score 计算

通过式 (2.6) 计算表 4.3 中 26 个特征的 Fisher Score, 计算结果如表 4.4 所示。

表 4.4 26 个时域、频域特征的 Fisher Score

特征编号	Fisher Score	归一化	特征编号	Fisher Score	归一化	特征编号	Fisher Score	归一化
F1	2.828	0.002013	F10	6.854	0.004878	F19	81.802	0.058218
F2	69.871	0.049726	F11	5.176	0.003684	F20	45.667	0.032501
F3	20.960	0.014917	F12	20.744	0.014763	F21	102.349	0.072841
F4	20.915	0.014885	F13	7.657	0.005449	F22	12.256	0.008722
F5	75.787	0.053937	F14	9.400	0.00669	F23	29.290	0.020845
F6	22.425	0.01596	F15	0.851	0.000606	F24	32.889	0.023407
F7	69.705	0.049608	F16	6.83	0.004861	F25	256.829	0.182782
F8	70.281	0.050018	F17	3.354	0.002387	F26	369.195	0.262752
F9	0.932	0.000663	F18	60.262	0.042888			

表 4.4 中 Fisher Score 归一化值越大说明该特征辨别类的能力越强, 重要度越高。

2. 设置特征重要度阈值, 生成预选特征子集 F_l

通过实验比较数据设置重要度阈值为 0.02, 对表 4.4 中的特征进行筛选, 归一化值低于 0.02 的特征剔除。并按特征值从高到低排序, 生成预选特征子集: $F_1 = \{F_{26}, F_{25}, F_{21}, F_{19}, F_5, F_8, F_2, F_7, F_{18}, F_{20}, F_{24}, F_{23}\}$ 。

3. MIC 计算

表 4.5 两个特征之间的 MIC 计算结果

特征编号	MIC	特征编号	MIC	特征编号	MIC	特征编号	MIC
F26、F25	0.4689	F25、F18	0.0001	F19、F18	0.0001	F2、F7	0.4735
F26、F21	0.4689	F25、F20	0.4689	F19、F20	0.4689	F2、F18	0.0001
F26、F19	0.4689	F25、F24	0.0001	F19、F24	0.0001	F2、F20	0.5356
F26、F5	0.0548	F25、F23	0.0001	F19、F23	0.0001	F2、F24	0.0001
F26、F8	0.0360	F21、F19	0.4981	F5、F8	0.3109	F2、F23	0.0001
F26、F2	0.0795	F21、F5	0.1152	F5、F2	0.4406	F7、F18	0.001
F26、F7	0.0795	F21、F8	0.1027	F5、F7	0.4406	F7、F20	0.7356
F26、F18	0.0001	F21、F2	0.1262	F5、F18	0.0001	F7、F24	0.1419
F26、F20	0.4689	F21、F7	0.1262	F5、F20	0.5523	F7、F23	0.0001
F26、F24	0.0001	F21、F18	0.0001	F5、F24	0.0001	F18、F20	0.0001
F26、F23	0.0001	F21、F20	0.4689	F5、F23	0.0001	F18、F24	0.0001
F25、F21	0.4689	F21、F24	0.0001	F8、F2	0.2385	F18、F23	0.0001
F25、F19	0.4689	F21、F23	0.0001	F8、F7	0.2385	F20、F24	0.0001
F25、F5	0.0548	F19、F5	0.1757	F8、F18	0.0001	F20、F23	0.0001
F25、F8	0.0360	F19、F8	0.1573	F8、F20	0.3398	F24、F23	0.0001
F25、F2	0.1419	F19、F2	0.2393	F8、F24	0.0001		
F25、F7	0.1419	F19、F7	0.2393	F8、F23	0.0001		

将预选特征子集 F_1 中的特征两两组合,每个特征与除自己外的特征使用式(2.8)计算 MIC,计算结果如表 4.5 所示。

表 4.5 中 MIC 取值越大说明该维特征冗余越高,反之冗余度越低。

4. 设置特征冗余度阈值,输出预选特征子集 F_2

通过实验比较数据设置冗余度阈值为 0.47,对预选特征子集 F_1 进行筛选,将 MIC 值高于 0.47 的特征中 Fisher Score 值较低的特征剔除,即剔除特征 F19、F20、F7。生成预选特征子集: $F_2=\{F26, F25, F21, F5, F8, F2, F18, F24, F23\}$ 。

5. 基于输入的狼群数量 N 、最大迭代次数 $iter_{max}$ 设置 BGWO 的参数

设置狼群数量 N 为 12,最大迭代次数 $iter_{max}$ 为 5。

6. 初始化狼群

设置狼群个体的初始化位置向量 $X_i=[F26_i, F25_i, F21_i, F5_i, F8_i, F2_i, F18_i, F24_i, F23_i]$, $i = \alpha, \beta, \delta, \omega 1, \omega 2, \dots, \omega 9$, 如 α 狼的位置为 $X_\alpha=[F26_\alpha, F25_\alpha, F21_\alpha, F5_\alpha, F8_\alpha, F2_\alpha, F18_\alpha, F24_\alpha, F23_\alpha]$ 。之后进行 0 或 1 的随机初始化,例如 $X_1=[1,1,1,1,1,0,0,0,]$,即选择 F26、F25、F21、F5、F8 这五个特征值。12 只狼随机化后的位置向量如表 4.6 所示。

表 4.6 灰狼个体初始化后的位置向量

狼群编号	位置向量								
	F26	F25	F21	F5	F8	F2	F18	F24	F23
X_α	0	0	0	0	1	0	1	1	0
X_β	0	1	0	0	1	1	1	0	1
X_δ	1	1	1	0	1	0	1	0	1
$X_{\omega 1}$	0	1	1	0	0	0	0	0	1
$X_{\omega 2}$	1	1	0	1	0	0	0	1	1
$X_{\omega 3}$	0	0	0	0	0	1	1	0	1
$X_{\omega 4}$	1	0	1	1	1	1	1	0	1
$X_{\omega 5}$	1	1	1	0	0	1	0	1	0
$X_{\omega 6}$	1	1	1	1	0	1	1	0	0
$X_{\omega 7}$	1	1	1	0	0	0	0	0	1
$X_{\omega 8}$	1	0	1	1	0	1	1	0	0
$X_{\omega 9}$	1	0	1	1	1	0	1	0	1

7. 使用 KNN 算法计算初代灰狼个体的适应度;

表 4.7 灰狼初始位置个体适应度

狼群编号	适应度	狼群编号	适应度	狼群编号	适应度
X_α	0.9278	$X_{\omega 2}$	0.7778	$X_{\omega 6}$	0.8639
X_β	0.8056	$X_{\omega 3}$	0.8444	$X_{\omega 7}$	0.8645
X_δ	0.8500	$X_{\omega 4}$	0.9611	$X_{\omega 8}$	0.8608
$X_{\omega 1}$	0.8699	$X_{\omega 5}$	0.8628	$X_{\omega 9}$	0.8621

将 12 只灰狼个体选择的特征值对应的训练集和测试集样本输入到 KNN 模型中进行模型训练和模型验证,计算故障诊断的精度,精度越高,适应度越高(精度=适应度),

12 只灰狼的适应度结果如表 4.7 所示。将适应度最高的位置向量更新 α 狼所在位置，次之为 β 狼和 δ 狼，即 X_{ω_4} 、 X_{α} 、 X_{ω_1} 依次为 α 狼、 β 狼、 δ 狼所在的位置向量。

8. 更新狼群位置；

根据 α 狼、 β 狼、 δ 狼所在的 X_7 、 X_1 、 X_4 位置向量，根据式 (2.17) 更新狼群的位置向量，灰狼个体第一次位置更新后的位置向量如表 4.8 所示。

表 4.8 灰狼个体第一次位置更新后的位置向量

狼群编号	位置向量								
	F26	F25	F21	F5	F8	F2	F18	F24	F23
X_{α}	1	1	0	0	1	1	1	0	1
X_{β}	0	1	1	1	1	0	1	1	1
X_{δ}	1	1	1	0	1	0	1	1	1
X_{ω_1}	0	0	0	0	1	0	1	1	0
X_{ω_2}	1	1	0	0	1	1	1	1	1
X_{ω_3}	0	1	0	0	1	1	1	0	1
X_{ω_4}	1	1	1	0	1	1	0	1	1
X_{ω_5}	1	1	1	0	0	0	1	1	1
X_{ω_6}	1	1	1	0	0	0	1	0	0
X_{ω_7}	1	1	1	1	0	1	1	0	1
X_{ω_8}	0	0	1	0	1	0	1	0	0
X_{ω_9}	1	1	1	1	1	0	1	0	1

9. 计算位置更新后灰狼个体的适应度；

与步骤 7 中的方法相同，计算灰狼个体位置更新后的适应度，12 只灰狼位置更新后的适应度结果如表 4.9 所示。将 X_{α} 、 X_{ω_6} 、 X_{ω_5} 依次设置为 α 狼、 β 狼、 δ 狼所在的位置向量。

表 4.9 灰狼个体第一次位置更新后的个体适应度

狼群编号	适应度	狼群编号	适应度	狼群编号	适应度
X_{α}	0.9810	X_{ω_2}	0.7778	X_{ω_6}	0.9634
X_{β}	0.9278	X_{ω_3}	0.8056	X_{ω_7}	0.9480
X_{δ}	0.9611	X_{ω_4}	0.9601	X_{ω_8}	0.9340
X_{ω_1}	0.9278	X_{ω_5}	0.9613	X_{ω_9}	0.9290

10. 循环迭代，不断更新狼群的位置，计算狼群适应度，更换 α 狼、 β 狼、 δ 狼的位置；

灰狼个体第五次位置更新后的位置向量如表 4.10 所示，位置更新后的适应度结果如表 4.11 所示。

表 4.10 灰狼个体第五次位置更新后的位置向量

狼群编号	位置向量								
	F26	F25	F21	F5	F8	F2	F18	F24	F23
X_{α}	1	1	1	0	1	0	1	0	1
X_{β}	1	0	1	0	1	1	1	0	1
X_{δ}	1	0	1	0	1	1	1	1	1
X_{ω_1}	1	0	1	0	1	0	1	1	1
X_{ω_2}	1	1	1	1	1	1	1	0	1

表 4.10 灰狼个体第五次位置更新后的位置向量 (续)

狼群编号	位置向量								
	F26	F25	F21	F5	F8	F2	F18	F24	F23
$X_{\omega 3}$	1	1	1	1	1	1	1	0	1
$X_{\omega 4}$	1	1	1	0	1	0	1	1	1
$X_{\omega 5}$	1	1	1	0	1	1	1	0	1
$X_{\omega 6}$	1	1	1	0	1	1	1	1	1
$X_{\omega 7}$	0	0	1	1	0	0	1	0	0
$X_{\omega 8}$	1	1	1	1	1	1	1	1	1
$X_{\omega 9}$	1	1	1	0	1	0	0	1	1

表 4.11 灰狼个体第五次位置更新后的个体适应度

狼群编号	适应度	狼群编号	适应度	狼群编号	适应度
X_{α}	0.9833	$X_{\omega 2}$	0.9818	$X_{\omega 6}$	0.9831
X_{β}	0.9790	$X_{\omega 3}$	0.9821	$X_{\omega 7}$	0.9831
X_{δ}	0.9832	$X_{\omega 4}$	0.9823	$X_{\omega 8}$	0.9790
$X_{\omega 1}$	0.9832	$X_{\omega 5}$	0.9801	$X_{\omega 9}$	0.9801

11. 输出最优特征子集 F_{out} , 故障诊断精度 ACC;

最优特征子集即为 α 狼的位置向量, 因此 $F_{out}=\{F26,F25,F21,F8,F18,F23\}$ 。在 KNN 故障诊断模型上的诊断精度 ACC 即为 α 狼的适应度, 即 $ACC=98.33\%$ 。

4.4.3 实验结果

根据 4.3.2 节的实验设计思路, 本文实验进行如下:

1. 采样信息量分析

2.1.3 节中阐述的基于无线传输的风机故障诊断系统, 其从感知终端设备到中继设备进行数据传输时的特征子集为 4.4.2 节步骤 4 生成的特征子集 F_2 , 共 9 个特征值, 与传输完整的 26 个特征值相比, 传输的特征数量缩减了 65.385%。通过附录 A 中表 A-1、表 A-2 的特征提取公式可知, 每个特征值需要传输的信息量相同。因此, 本文方法在采样信息量方面缩减了 65.385%。

2. 故障诊断精度分析

计算仅使用 Filter 方法在 KNN 故障诊断模型上的精度, 然后与本文方法在 KNN 故障诊断模型上的精度进行对比, 验证本文方法相比于 Filter 方法所选的特征质量更高。

Filter 方法生成的特征子集为 4.4.2 节步骤 4 生成的特征子集 F_2 : $\{F26,F25,F21,F5,F8,F2,F18,F24,F23\}$, 该特征子集在 KNN 故障诊断模型上的故障诊断结果如图 4.4 所示。

采用本文 Filter 和 Wrapper 相结合的特征选择算法在 KNN 上选择出的最优特征子集是: $\{F26,F25,F21,F8,F18,F23\}$, 该最优特征子集在 KNN 故障诊断模型上的故障诊断结果如图 4.5 所示。

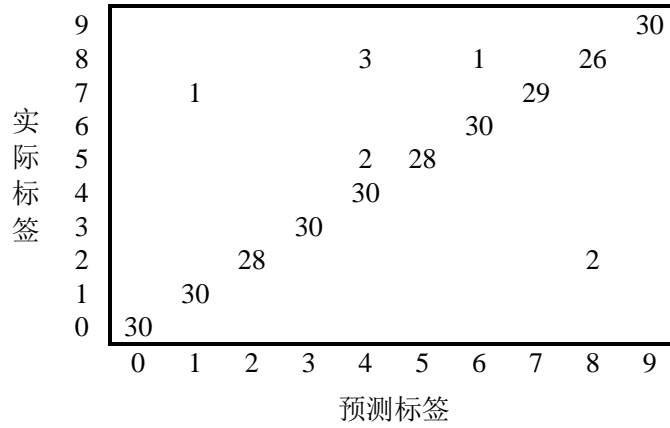


图 4.4 Filter 方法选择的特征子集对轴承不同故障的诊断情况

从图 4.4 中可看出, Filter 方法有 6 类轴承状态的识别率为 100%, 对于第 2 类故障 (滚动体轻度故障) 的诊断精度是 93.33% (28/30)、对于第 5 类故障 (内圈中度故障) 的诊断精度是 93.33% (28/30)、对于第 7 类故障 (外圈中度故障) 的诊断精度是 96.67% (29/30)、对于第 8 类故障 (内圈重度故障) 的诊断精度是 86.67% (26/30), Filter 方法共识别出 291 个样本, 模型诊断的准确率为 97.00% (291/300)。

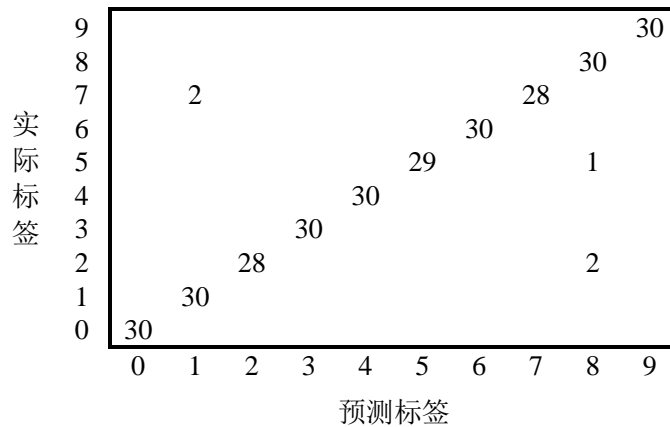


图 4.5 本文方法选择的特征子集对轴承不同故障的诊断结果

从图 4.5 中可看出, 本文方法在有 7 类轴承状态的识别率为 100%, 对于第 2 类故障 (滚动体轻度故障) 的诊断精度是 93.33% (28/30)、对于第 5 类故障 (内圈中度故障) 的诊断精度是 96.67% (29/30)、对于第 7 类故障 (外圈中度故障) 的诊断精度是 93.33% (28/30), 本文方法共识别出 295 个样本, 模型诊断的准确率为 98.33% (295/300)。

本文方法相比于 Filter 方法在精度方面提高了 1.33%。

3. 方法通用性分析

在方法通用性分析中, 本文方法先使用 Filter 方法中的 Fisher Score-MIC 方法选择特征子集, 然后在特征子集的基础上使用 Wrapper 方法中的 BGWO+KNN 方法 (表 4.12 中编号 1) 以及 Wrapper 方法中的 BGWO+SVM 方法 (表 4.12 中编号 2) 计算最优特征

子集和故障诊断精度。仅使用 Wrapper 方法是先使用 Wrapper 方法中的 BGWO 方法+KNN 选择特征子集并计算故障诊断精度（表 4.12 中编号 3），然后在最优特征子集的基础上使用 Wrapper 方法中的 BGWO 方法+SVM 方法计算最优特征子集和故障诊断精度（表 4.12 中编号 4）。不同特征选择算法的故障诊断结果如表 4.12 所示。

表 4.12 不同特征选择算法故障诊断结果表

编号	特征选择算法	故障诊断模型	最优特征子集	诊断精度	说明
1	Fisher Score-MIC 和 BGWO	KNN	{F26,F25,F21,F8,F18,F23}	98.33%	本文方法模
2	Fisher Score-MIC 和 BGWO	SVM	{F25,F21,F5,F24,F23}	95.35%	型更换
3	BGWO	KNN	{F14,F19,F20,F21,F26}	99.67%	Wrapper 方
4	BGWO+KNN 和 BGWO	SVM	{F19,F20,F26}	90.33%	法模型更换

表 4.12 中编号 1 和 2 对比表明本文方法将故障诊断模型从 KNN 更换到 SVM 时，精度下降了 2.98%，编号 3 和 4 对比表明仅用 Wrapper 方法将故障诊断模型从 KNN 更换到 SVM 时，精度下降了 9.34%，远高于本文的 2.98%，表明本文方法所选特征子集的通用性更强，在不同故障诊断模型上的精度误差小。

4. 风机运行数据实例验证

本文所设计的感知终端设备如图 4.6（a）所示，结构图如图 4.7 所示，其由一个数据采集板组成，数据采集板上包含 LoRa 无线传输模块。中继设备如图 4.6（b）所示，结构图如图 4.8 所示，包含 LoRa 无线传输模块以及 4G 通信模块。通过风机历史数据训练特征选择算法以及 KNN 故障诊断模型，通过本文所设计的感知终端设备和中继设备实时采集风机电机轴承的运行状态并进行故障诊断，算法选出的最优特征子集为 {F26,F25,F21,F5,F24,F23}，诊断的准确率为 97.13%，由于实际工作中噪声的影响，诊断精度略低于轴承公共数据集上的 98.33%。



(a) 感知终端设备



(b) 中继设备

图 4.6 感知终端设备及中继设备硬件实物图

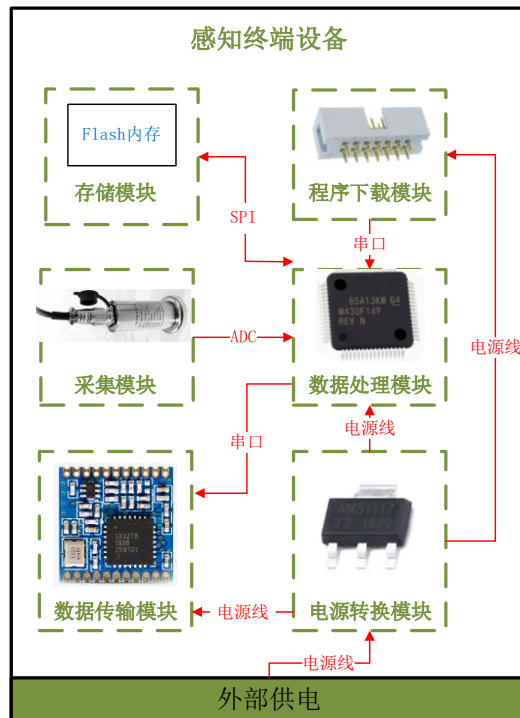


图 4.7 感知终端设备结构图

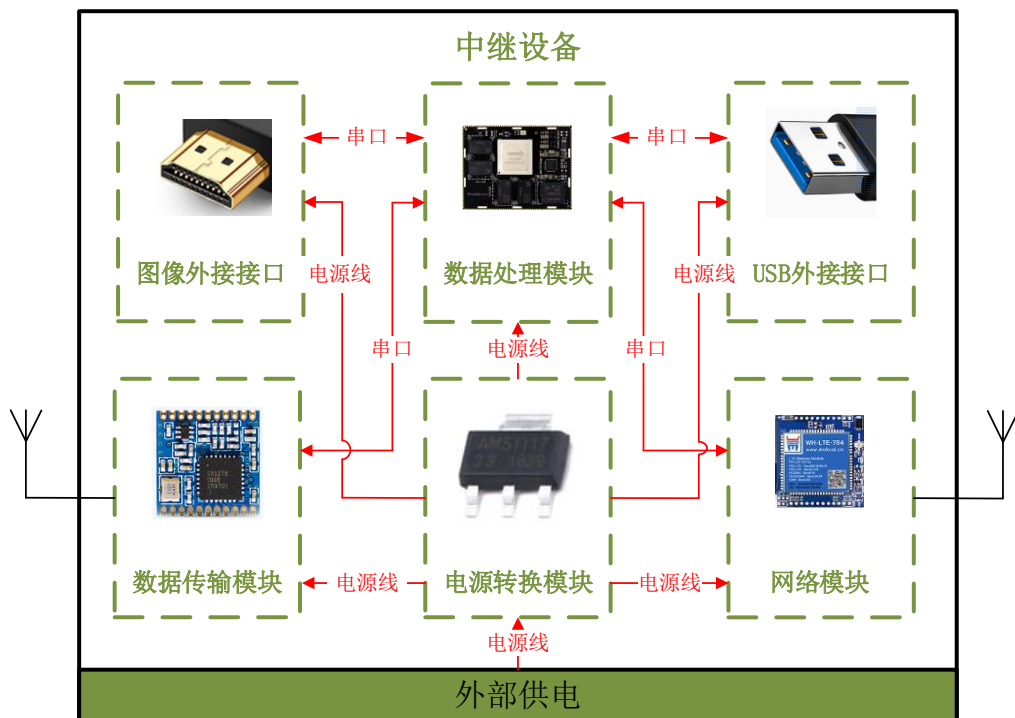


图 4.8 中继设备结构图

4.5 本章小结

本章首先对特征选择算法的研究对象轴承进行特点分析，重点分析其结构原理以及常见的故障形式和故障演变，最终确定研究轴承外圈、滚动体和内圈在腐蚀面积轻度、中度和重度条件下的故障诊断，并基于此研究特征选择算法。接着提出本文基于 Filter

和 Wrapper 相结合的两步特征选择算法，该方法使用 Filter 方法选择量少、通用性强的特征进行传输，接收到特征后使用 Wrapper 方法对特征进行二次优选，提高故障诊断的精度，实现有限的带宽条件下尽可能的传输重要特征。最后提出本文的实验思路及具体实验结果分析。实验结果表明：本文算法在采样信息量缩减 65.385% 时，故障诊断精度达到 98.33%，比单纯使用 Filter 方法的精度提高了 1.33%，证明所选特征具有量少、质优的优点。在通用性方面，本文算法在 KNN 和 SVM 故障诊断模型上的精度误差为 2.98%，低于 Wrapper 方法的 9.34%，通用性更强。通过本文方法的特征选择算法实现了在有限带宽条件下优先采样重要故障特征的目的，缩减了采样的信息量。

第五章 总结与展望

5.1 本文工作总结

本文针对建立基于无线传输的风机故障诊断系统时存在的采样信息量受限于窄带宽传输条件的问题,以终南山隧道的轴流风机为研究对象,提出研究风机状态监测点优选方法与适用于风机故障诊断的特征选择算法,在带宽有限的条件下优先采样关键监测点的重要故障特征,缩减监测点数并减少单个监测点的采样信息量,从而使得采样的信息量能够满足窄带宽传输的客观条件。虽然在基于有线传输的故障诊断系统中“监测点优选方法”与“特征选择”方法已经得到了广泛的应用,但是由于无线方式的窄带宽条件的限制,其难以应用到基于无线传输的风机故障诊断系统,本文的具体研究内容如下:

1. 基于终南山隧道轴流风机的位置及特点,提出适用于该场景的基于无线传输的风机故障诊断系统方案设计。该方案从故障诊断系统位置分布、硬件构成等角度阐述整个系统的物理拓扑结构,然后从故障诊断模型训练及模型应用的角度阐述了故障诊断系统的逻辑功能结构,为下文研究监测点优选方法以及特征选择算法提供应用场景。

2. 在监测点优选方法中,本文的研究目标是在有限带宽条件下,优先对关键的故障设置状态监测点。基于这个目标,本文深入分析了基于有线传输的故障诊断系统采用的状态监测点优选方法应用到基于无线传输的故障诊断系统存在的问题后,提出了一种面向可靠性的风机状态监测点优选方法。该方法通过可靠性分析模型的 FMECA、FTA 和调研的方法确定风机故障的危害度、概率重要度、易检测度、易状态监测度以及维修经济因素。接着使用 AHP 方法对这五种因素进行权重分配。最后基于因素值和权重值计算状态监测评价价值进而获得监测点重要度排序。在有限的带宽条件下,可根据状态监测点评价价值从高到低的顺序安装状态监测点,尽可能的监测重要故障。

3. 在特征选择算法研究中,本文的研究目标是选择量少、质优、通用的特征传输到后台进行故障诊断,在有限带宽条件下尽可能的采样重要特征。基于这个目标,本文提出基于 Filter 和 Wrapper 相结合的两步特征选择算法,该方法使用 Filter 方法选择量少、通用性强的特征进行传输,接收到特征后使用 Wrapper 方法对特征进行二次优选,提高故障诊断的精度,实现了有限的带宽条件下尽可能的采样重要故障特征的目的。

5.2 未来工作展望

本文虽然通过研究风机状态监测点优选方法以及适用于隧道风机的特征选择算法缩减了监测点数并减少了单个监测点的采样信息量,从而使得采样的信息量能够满足窄

带宽传输的客观条件。但是由于工作条件以及研究时间的限制，本文的研究内容还存在很多不足，后期可从以下角度进行研究：

1. 由于受工作条件的限制，作者未能对隧道风机实际工作细节进行完全充分的了解，影响状态监测点安装的危害度、概率重要度等各因素值仅能通过分级的方式进行评价，后续可根据实际情况进行详细分析，对评价因素进行更加详细的定性分析或者定量计算。

2. 目前特征选择算法仅研究了单一故障的特征选择，但是在实际应用时，风机更可能同时发生多类故障，因此，下一阶段可将基于 **Filter** 和 **Wrapper** 相结合的两步特征选择算法应用到混合故障情况下的特征选择中。

参考文献

- [1] 刘瑶. 基于时频分析的风机故障智能诊断方法研究[D]. 杭州: 浙江大学, 2021
- [2] 杨峰, 胡剑浩, 李少谦. 基于欠奈奎斯特采样的超宽带信号总体最小二乘重建算法[J]. 电子与信息学报, 2010, 32(06): 1418-1422
- [3] 张雨婷, 李策, 孔飞飞, 等. 基于 ZigBee 的电缆接头压接缺陷监测系统[J]. 传感器与微系统, 2023, 42(05): 87-90
- [4] 潘国荣, 范伟. ZigBee 技术在地下线型隧道工程自动导向中的应用[J]. 测绘通报, 2018(S1): 172-175
- [5] 曾熙鸿, 李富年, 颜永逸. 基于物联网的地铁隧道断面监测系统的设计[J]. 现代电子技术, 2018, 41(18): 182-186
- [6] 于燕平. 基于 ZigBee 的隧道电气化设备除冰系统研究[J]. 科技与创新, 2020(05): 1-2
- [7] 周华妹, 周华安, 李端峰, 等. 基于 ZigBee 的隧道照明智能控制系统研究与实现[J]. 测控技术, 2019, 38(02): 69-74
- [8] 王丽丹, 李舒, 朱光. 基于 ZigBee 技术的无线环境温湿度校准系统[J]. 计量与测试技术, 2018, 45(11): 17-19
- [9] 鲍胜文, 方拥军, 赵飞, 等. 基于 LoRa 技术的无线通信管理系统研究与实现[J]. 电子世界, 2019, 582(24): 120-122+126
- [10] 田青, 袁方周, 郑宣传, 等. WiFi 多级桥接应急多媒体通信技术研究[J]. 信号处理, 2019, 35(04): 656-666
- [11] 覃中顺, 赵四海, 胡云兰, 等. 煤矿井下应急导航系统设计[J]. 煤炭工程, 2020, 52(07): 49-52
- [12] 王大涛, 滕德贵, 李超. 基于低功耗无线传感网络的隧道健康监测系统[J]. 测绘通报, 2018(S1): 273-277
- [13] 盛林, 马波, 张杨. 基于知识图谱的旋转机械故障诊断方法[J]. 机电工程, 2022, 39(09): 1194-1202
- [14] 江雪晨, 王大志, 张翠玲, 等. 利用模型诊断降维的电网故障诊断完全解析方法[J]. 中国电机工程学报, 2016, 36(23): 6371-6378+6602
- [15] 牟宗磊, 王晨, 张媛, 等. 数据驱动的凸轮式绝对重力仪微小故障诊断[J]. 振动测试与诊断, 2022, 42(06): 1068-1075+1240
- [16] 王东, 张永伟, 南亚峰, 等. 数据挖掘的风力发电设备健康状态自动识别[J]. 自动化

技术与应用, 2022, 41(11): 32-35

- [17]张雪嘉. 基于经验小波与改进卷积神经网络的风机故障诊断方法研究[D]. 北京: 北京化工大学, 2021
- [18]常新宇, 李琦. 特征筛选与 SVM 结合的风机轴承故障诊断研究[J]. 机械设计与制造, 2022, 380(10): 71-74+80
- [19]吴斌, 奚立峰, 范思遐, 等. 基于改进的 K 最近邻分类器的风机故障诊断[J]. 机械设计与研究, 2016, 32(05): 163-167+174
- [20]徐敏, 王平. 基于深度 LSTM 残差网络的旋转机械故障诊断研究[J]. 机床与液压, 2023, 1(04): 84-190
- [21]欧阳森, 刘天马, 杨墨缘. 考虑综合经济损失的电压暂降监测优化配置方法[J]. 广东电力, 2021, 34(04): 34-42
- [22]刘一凡. 空调供暖状态下室内热环境监测点特性研究[D]. 重庆: 重庆大学, 2021
- [23]周家豪. 基于声电信号的铝合金零件焊接缺陷识别方法研究[D]. 成都: 电子科技大学, 2022
- [24]高建民, 高智勇. 一种复杂机电系统 LE-SVDD 异常监测方法[J]. 振动.测试与诊断, 2017, 37(03): 469-475+625-626
- [25]Khotanzad A and Hong Y H. Rotation invariant image recognition using features selected via a systematic method[J]. Pattern Recognition, 2019, 23(10): 1089-1101
- [26]Swets D L and Weng J J. Efficient content-based image retrieval using automatic feature selection[C]. International Symposium on Computer Vision, 2018, 14(03): 13-19
- [27]张人上, 曲开社. 一种基于新的特征选择的海量网络文本挖掘算法研究[J]. 计算机应用研究, 2014, 31(09): 2632-2634+2684
- [28]刘翔宇, 芦天亮, 杜彦辉, 等. 基于特征选择的物联网轻量级入侵检测方法[J]. 信息网络安全, 2023, 23(01): 66-72
- [29]胡逾航. 基于集成算法 LightGBM 的高血脂风险预测研究[D]. 武汉: 湖北工业大学, 2021
- [30]Hancer E, Xue B and Zhang M. Differential evolution for filter feature selection based on information theory and feature ranking[J]. Knowledge-Based Systems, 2018, 140(15): 103-119

- [31] Wang T, Liu D, Liu Y, et al. Fault feature selection of subway plug door based on reliefF and BGWO[C]. 2021 IEEE 16th Conference on Industrial Electronics and Applications (ICIEA), 2021: 1346-1351
- [32] 舒凡娣, 谢嘉晟, 廖晓娇, 等. 结合粒子群算法和穷举法的配电网故障诊断方法[J]. 智慧电力, 2019, 47(01): 94-99
- [33] Sun L J, Xin Y, Chen T F, et al. Rolling bearing fault feature selection method based on a clustering hybrid binary cuckoo search[J]. Electronics, 2023, 12(02): 459-481
- [34] Chen S L, Meng Y, Xie Z M, et al. Feature selection and detection method of weak arc Faults in photovoltaic systems with strong noises based on stochastic resonance[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 1-13
- [35] Wutzl B, Leibnitz K, Rattay F, et al. Genetic algorithms for feature selection when classifying severe chronic disorders of consciousness[J]. PLOS ONE, 2019, 14(07): 161-177
- [36] Ren B, Bai D, Xue Z, et al. Method for fault feature selection for a baler gearbox based on an improved adaptive genetic algorithm[J]. 中国机械工程学报(英文版), 2022, 35(03): 45-45
- [37] Shi H Y and Chen M X. A two-stage transformer fault diagnosis method based multi-filter interactive feature selection integrated adaptive sparrow algorithm optimised support vector machine[J]. IET Electric Power Applications, 2022, 17: 341-357
- [38] Benlamoudi A, Samai D, Ouafi A, et al. Face spoofing detection using local binary patterns and Fisher Score[C]. 2015 3rd International Conference on Control, Engineering & Information Technology (CEIT), 2015: 1-5
- [39] Lv C Z, Lu Y F, Lu M, et al. A classification feature optimization method for remote sensing imagery based on Fisher Score and mRMR[J]. Applied Sciences, 2022, 12(17): 1-19
- [40] 陈双庆. 山区公路隧道事故特点及对策——以柞水秦岭终南山为例[J]. 今日消防, 2022, 7(02): 33-35
- [41] 吴迪, 郭嗣琮. 改进的 Fisher Score 特征选择方法及其应用[J]. 辽宁工程技术大学学报(自然科学版), 2019, 38(05): 472-479
- [42] Chen Y W and Lin C J. Combining SVMs with various feature selection strategies[J]. Studies in Fuzziness & Soft Computing, 2006, 207: 315-324
- [43] Guenes S, Polat K and Yosunkaya S. Multi-class f-score feature selection approach to classification of obstructive sleep apnea syndrome[J]. Expert Systems with Applications, 2010, 37(02): 998-1004

- [44] Qingjun S, Haiyan J and Jing L. Feature selection based on FDA and f-score for multi-class classification[J]. Expert Systems with Applications, 2017, 81: 22-27
- [45] Mirjalili S, Mirjalili S M and Lewis A. Grey wolf optimizer[J]. Advances in Engineering Software, 2014, 69(03): 46–61
- [46] Mirjalili S. How effective is the grey wolf optimizer in training multi-layer perceptrons[J]. Applied Intelligence, 2015, 43(01): 150-161
- [47] 韩麟, 陈宏伟. 基于 Spark 的灰狼优化算法研究[J]. 湖北工业大学学报, 2019, 34(05): 60-63
- [48] 王霄. 基于改进灰狼优化算法的仓储物流机器人路径规划方法[J]. 自动化技术与应用, 2022, 41(12): 23-27
- [49] 刘志强, 何丽, 袁亮, 等. 采用改进灰狼算法的移动机器人路径规划[J]. 西安交通大学学报, 2022, 56(10): 49-60
- [50] 曾耀传, 林云树, 吴晓梅. 基于 EEMD 与 GWO-MCKD 的门座起重机回转支承故障诊断[J]. 机床与液压, 2022, 50(07): 170-175
- [51] 张振海, 王维庆, 王海云, 等. 基于 HCS-GWO-MSVM 的风电机组齿轮箱复合故障诊断研究[J]. 太阳能学报, 2021, 42(10): 176-182
- [52] 赵婷, 王爽, 段晓萌. 基于灰狼优化算法的智能电表嵌入式操作系统任务调度算法[J]. 单片机与嵌入式系统应用, 2022, 22(10): 55-57+78
- [53] 李开言, 王亚刚, 张涛. 基于改进灰狼优化算法的边缘计算任务调度方法[J]. 信息与控制, 2022, 51(04): 489-497+512
- [54] Emary E, Zawba H M and Hassanien A E. Binary grey wolf optimization approaches for feature selection[J]. Neurocomputing, 2016, 172(08): 371-381
- [55] 肖遥. 基于均匀试验设计法的射雾器轴流风机 CFD 数值模拟分析[D]. 秦皇岛: 燕山大学, 2016
- [56] 陈鹏. 滚动轴承故障诊断及性能退化评估方法研究[D]. 兰州: 兰州理工大学, 2021
- [57] 姜双阳. 基于 CPSO-BBO 优化 SVM 的风机主轴轴承的故障诊断分析[D]. 吉林: 东北电力大学, 2019

附录 A

在故障诊断领域，常用的特征值有时域、频域等。

1. 时域特征值

时域特征值分为有量纲和无量纲两种，有量纲是指特征具有实际的物理意义，其通常对风机的早期故障很敏感，当故障发生时，幅值会随着故障的出现而上升，但是容易受到工况的影响，在不同应用场景，有量纲特征差异很大；而无量纲特征优点在于其不易受复杂工况因素的影响，能够较为直接的描述和表征故障状态的变化。

本文选取的时域特征指标共有 17 个，其中前 8 个为有量纲特征值；后 9 个为无量纲特征值。下文公式中 $x(n)$ 为振动信号的时域序列， $n = 1, 2, \dots, N$ ， N 为样本点数。

表 A.1 时域特征值

编号	名称	公式	备注说明
F1	均值	$F_1 = \frac{\sum_{n=1}^N x(n)}{N}$	描述风机振动数据的集中趋势
F2	标准差	$F_2 = \sqrt{\frac{\sum_{n=1}^N (x(n) - F_1)^2}{N - 1}}$	描述风机振动数据的离散程度
F3	最大值	$F_3 = \max(x(n))$	
F4	最小值	$F_4 = \min(x(n))$	
F5	绝对平均值	$F_5 = \frac{\sum_{n=1}^N x(n) }{N}$	
F6	峰峰值	$F_6 = \max(x(n)) - \min(x(n))$	峰峰值是风机振动信号最大值与最小值之差，它反应了风机振动的剧烈程度
F7	均方根	$F_7 = \sqrt{\frac{\sum_{n=1}^N x(n)^2}{N}}$	又叫有效值，反应了风机振动的能量水平。
F8	方根幅值	$F_8 = \left[\frac{\sum_{n=1}^N \sqrt{ x(n) }}{N} \right]^2$	
F9	偏斜度	$F_9 = \frac{\sum_{n=1}^N [x(n) - F_1]^3}{(N - 1)F_2^3}$	风机的振动数据的频数有些是对称的，有些是不对称的，既呈现偏态，偏斜度是对风机振动数据偏斜程度的度量

表 A.1 时域特征值（续）

编号	名称	公式	备注说明
F10	峭度	$F_{10} = \frac{\sum_{n=1}^N [x(n) - F_1]^4}{(N-1)F_2^4}$	描述数据尖峰度的特征指标，对风机故障中的冲击信号特别敏感
F11	峰值指标	$F_{11} = \frac{\max x(n) }{F_7}$	衡量风机振动信号是否具有冲击，该指标既考虑了峰值，又考虑了均方根
F12	波形指标	$F_{12} = \frac{F_7}{F_5}$	波形指标的分子表征风机振动能量的大小，分母表征平均幅值，该指标可以描述渐变故障
F13	脉冲指标	$F_{13} = \frac{\max x(n) }{F_5}$	与峰值指标类似，对冲击类故障较敏感
F14	裕度指标	$F_{14} = \frac{\max x(n) }{F_8}$	对风机振动信号的冲击较为敏感
F15	偏态系数	$F_{15} = \frac{1}{N} \sum_{n=1}^N \left[\frac{x(n)}{F_2} \right]^3$	
F16	峭度系数	$F_{16} = \frac{1}{N} \sum_{n=1}^N \left[\frac{x(n)}{F_2} \right]^4$	
F17	变异系数	$F_{17} = \frac{F_1}{F_2}$	

2. 频域特征值

对信号进行频率分析可知不同频率成分的分布情况及不同故障对应特征频率的大小和位置，利用频率特征可以很好地描述频域中风机振动信号的变化。

传感器采集的振动信号是时域信号，通过快速傅里叶变换可以将其转化成频域信号。风机振动信号的时域序列为 $x(n)$ ， $n = 1, 2, \dots, N$ ， N 为样本点数，经过 FFT 变换后变成 $s = [s(1), s(2), \dots, s(N_{fft})]$ ， N_{fft} 表示频域信号的长度， f_k 为频率值， $s(k)$ 为对应的频率幅值。频域特征如表 A.2 所示。

表 A.2 频域特征

编号	名称	公式	备注说明
F18	平均频率	$F_{18} = \frac{\sum_{k=1}^{N_{fft}} s(k)}{N_{fft}}$	描述了风机振动信号在频域的能量大小
F19	中心频率	$F_{19} = \frac{\sum_{k=1}^{N_{fft}} f_k s(k)}{\sum_{k=1}^{N_{fft}} s(k)}$	描述主频带位置的变化

表 A.2 频域特征 (续)

编号	名称	公式	备注说明
F20	频率标准差	$F_{20} = \sqrt{\frac{\sum_{k=1}^{N_{fft}} (f_k - F_{19})^2 s(k)}{N_{fft}}}$	描述频谱的离散程度
F21	频率均方根	$F_{21} = \sqrt{\frac{\sum_{k=1}^{N_{fft}} f_k^2 s(k)}{\sum_{k=1}^{N_{fft}} s(k)}}$	描述主频带位置变化
F22	P1	$F_{22} = \sqrt{\frac{\sum_{k=1}^{N_{fft}} f_k^4 s(k)}{\sum_{k=1}^{N_{fft}} f_k^2 s(k)}}$	描述主频带位置变化
F23	P2	$F_{23} = \sqrt{\frac{\sum_{k=1}^{N_{fft}} f_k^2 s(k)}{\sum_{k=1}^{N_{fft}} s(k) \sum_{k=1}^{N_{fft}} f_k^4 s(k)}}$	描述主频带位置变化
F24	P3	$F_{24} = \frac{F_{20}}{F_{19}}$	描述频谱集中或者分散程度
F25	P4	$F_{25} = \frac{\sum_{k=1}^{N_{fft}} (f_k - F_{19})^3 s(k)}{N_{fft} F_{20}^3}$	描述频谱集中或者分散程度
F26	P5	$F_{26} = \frac{\sum_{k=1}^{N_{fft}} (f_k - F_{19})^4 s(k)}{N_{fft} F_{20}^4}$	描述频谱集中或者分散程度