## 摘要

## Abstract

## 目录

## 第一章 引言

### 1.1 项目背景

### 1.2 研究意义

飞机交流发电机是飞机电源系统中最重要的部件，它的工作状态直接影响到整个飞机机载设备的运行状况和飞机自身的飞行安全。因此，研究飞机交流发电机的异常检测技术具有十分重要的意义。对飞机交流发电机电源质量的评价指标主要有电压偏差、频率偏差、电压波动和闪变（浪涌电流）、谐波含量等几个方面。其中电压、电流等参数的波动利用传统方法较为容易检测，而频域中的谐波含量参数却缺少有效且直观的检测方法。本次研究主要面向的就是飞机交流发电机电源的谐波含量部分。本篇论文的以下部分也主要关注发电机电源谐波含量参数的异常检测。

飞机交流发电机电源的谐波属于噪声的一种，谐波含量应当维持在一个很小范围内，单次谐波含量应当小于基波的4%。尽管当前飞机电源普遍采用三相交流发电机供电，而且对于平衡的三相发电机而言，偶次谐波可以被消除，但是仍然存在若干含量较高的奇次谐波。如果谐波含量超出标准范围，会影响到飞机交流发电机的电源质量，严重时会导致电源品质严重下降，进而导致其他用电设备的工作异常，后果十分严重。因此，研究一种有效的电源谐波含量异常检测方法是十分必要的。

对于电源谐波的传统异常检测方法主要是采样电源的时域数据，再进行傅里叶变换转换到频域生成频谱，直接观察频谱是否出现异常。这种对频域的检测只是通过观察频谱图像的方法，比较耗费时间和人力，不够直观可靠且效率较低。

本次研究采用的异常检测方法是利用训练后的长短期记忆网络模型（LSTM），能够对电源品质中的基波与谐波含量参数进行实时可靠的预测及分析。与传统方法相比，本次研究采用的异常检测方法有着实时性、高可靠、高效率等优点。

/\*\*

**创新点及贡献阐述**

\*\*/

## 第二章 模型理论基础

### **2.1 时域-频域转换**

#### **2.1.1 傅里叶变换**

#### 2.1.2 采样定理

### 2.2 循环神经网络RNN

### 2.3 LSTM神经网络

### 2.4 其他变种RNN网络

### 2.5 时间序列的实时预测

## 第三章 电源频率数据异常检测方法的研究设计

### 3.1 现状与不足

正如本文引言中所叙述的，对于电源信号频域方面的异常检测，传统方法是先将时域数据进行快速傅里叶变换，得到电源信号的频谱，然后直接分析频谱来判断电源的基波与谐波含量是否异常。由于进行傅里叶变换需要单位时间的信号数据，即使忽略后续的频谱分析时间，也存在着至少一个单位时间的延迟。如果使用传统方法，当前时刻观察分析的频谱一定是迟滞的，是之前某个时刻的频谱，这就可能造成一定的隐患。因为如果延时过久，分析的结果是没有任何意义的，而且飞机交流发电机电源至关重要，应当做到对电源信号频域部分的实时检测与分析。简而言之就是这种传统方法不具备实时性，效率不高且不够可靠。

由本文第二章的理论基础部分可以知道，LSTM神经网络是十分适合用来处理时间序列的，而且可以通过预测的方式实现对时间序列数据的实时分析，但是它并不适用于处理频域数据。本文正是基于以上两个现状，思考并研究是否能够将LSTM神经网络模型利用在分析频域信号数据上，研究出一种新的具有实时性、高效率、高可靠的频域异常检测方法，来替代传统频域异常检测方法。

### 3.2 模型设计思想

如上所述，LSTM神经网络模型仅适用于处理分析时域上的信号数据，对频域信号数据却无能为力，这是由于时域与频域的特性所决定的，频域信号数据映射到时域中对应的是完整的时域信号数据，相当于LSTM神经网络模型只有一个单独的输入，而不是连贯的时间序列信号数据，这显然是不适用来获取理想结果的。

本次研究提出的基于时域-频域-时域转换与LSTM神经网络模型的电源频率数据异常检测模型如下图所示。（**模型总体结构图**）。如图所示，模型主要分为时域-频域-时域转换、数据预处理和LSTM神经网络三个部分，其中数据预处理部分又由数据降维、数据清洗、数据缩放、数据标记四个部分组成。

#### 3.2.1 时域-频域-时域转换

本文提出一种新的思路来综合利用时域-频域-时域转换和LSTM神经网络模型。其核心思想是对电源信号的电压-时间数据（V-t）进行分段，分为一个个单位时间（如1秒）的时域时间子序列片段。然后对这些时间序列片段分别进行快速傅里叶变换，获取这些子序列片段对应的频域信号数据（F）。再按原先的时序依次将这些频域信号数据串联，来生成一个完整的频率-时间数据（F-t），这个频率-时间数据在表现上是离散而非连续的，但每个离散的单位数据之间却具有时序上的连贯性。（**如图所示，时频时转换图**）。通过这样一个拆分后再组合的过程，完成一个从时域到频域，然后再到时域的转换，从而生成可以被LSTM神经网络模型使用的时间序列数据，这个时间序列数据实际上是由一系列时间上连贯的频域数据拼接而成的，而且可用于LSTM神经网络模型。

对原始的飞机交流发电机电源信号数据，先进行分段的时频转换，然后将频域按时序串联得到完整的频谱数据。这样得到的频谱数据还不能直接用于LSTM神经网络模型进行训练，需要进行一系列预处理操作。

#### 3.2.2 数据预处理

对于机器学习任务来说，数据预处理是十分重要的，它在很大程度上影响了模型最后的结果。对于本次处理电源信号数据的任务来说，数据预处理的流程可分为：数据降维、数据清洗、数据缩放和数据标记四个部分。

首先是数据降维。数据降维通常又被称为特征提取，在很多数据处理的任务中，需要被处理的数据都属于高维数据，具有非常多的特征数目。 如果不进行降维处理，直接进行分析通常会带来极大的计算量，一方面这会使后续的模型训练时间大大增加，另一方面，过多的特征数目会带来很大的噪声，影响模型的收敛和最终结果。 因此，在直接对数据进行分析之前需要对先提取出高维数据的特征，在保证结果不受影响的情况下，尽可能地减小原始数据的维数。本次实验中通过时域-频域-时域转换得到的完整频谱数据量是十分庞大的，这是因为转换后的频谱数据维度太高，且其中的绝大多数特征分量是我们不需要关心的。因此需要对完整频谱数据进行降维处理。

一般来说，降维算法有主成分分析法（PCA）、因子分析法（FA）、独立成分分析法（ICA）、奇异值分解法等等。以上降维算法适用于特征分量之间的关联性不易被显性发现的情形，其中最常用的是主成分分析法。但是本次研究中使用的频谱数据较为特殊，由于研究目的是进行基波与谐波的异常检测，因此不需要使用上述这些具有普适性的降维算法，而是采用人工剔除筛选特征的方法来降维。这里我们只需要关注电源信号的基波与主要高次谐波即可，其中次数过高的谐波由于其含量过低可以被直接剔除，而且由于本次研究采用的是平衡三相发电机，所以偶次谐波也忽略并剔除，只保留基波和若干低次的奇次谐波即可。

（**可附降维前后对比图**）

（**可尝试增加维度实验**）

然后是数据清洗。对降维后的频谱数据进行检查和清洗是十分重要的，数据清洗分为去除重复数据和填补缺失值两部分。对于本次采用的电源信号频谱的时序串联数据，去除重复数据可以检查时间序列，以时间序列为索引，去除采样时刻相同的重复数据。然后检查原始数据中是否有缺失值，如果有，就采用k最近邻算法进行缺失值的填补。这里对于k最近邻算法不再重点介绍。

接下来是数据缩放。本次研究使用的电源信号频谱数据的数值变化区间较大，既存在较大数量级的频率幅度数值，也存在较小数量级的频率幅度数值。因此需要将频谱数据缩放到指定的区间中，以增强数据的稳定性，同时也便于后续模型的计算处理。本次研究将电源信号的频谱数据缩放到0和1之间。

最后是数据标记。这一步的作用是规范LSTM神经网络模型的输入输出维度，以及生成符合LSTM神经网络模型要求的输入数据。机器学习根据标记信息的有无，可以大致分为监督学习和无监督学习两类。其中监督学习可以分为分类和回归两类问题。本次研究的内容是如何对飞机交流发电机电源数据的频率参数进行有效预测，来完成实时的异常数据检测，这实际上属于回归问题。然而，本次研究用到的电源信号频谱数据是没有标记信息的，无法直接用于模型训练，所以频谱数据应当先被标记。对于本次研究采用的电源信号频谱数据来说，由于频谱数据整体上是以时间序列为索引的，所以只需要把待预测的频率参数值序列前移一个时间标度，然后得到的新序列就是数据集的标记信息。这部分的详细原理可以参考第二章中对时间序列的实时预测部分的阐述。简单来说，每个时间标度的完整频谱数据就是LSTM神经网络模型的输入数据，对应的标签是前移后下一个时间标度的待预测频率参数数值，同时这部分也是LSTM神经网络模型的输出数据。

#### 3.2.3 LSTM神经网络模型

本文使用的LSTM神经网络模型结构如图所示（**LSTM结构图**），

## 第四章 **模型实现与实验结果分析**

### 4.1 原始数据介绍

### 4.2 时频转换结果

### 4.3 数据预处理

### 4.4 网络训练与调参

### 4.5 实验结果与分析

## 第五章 对比实验结果与分析

## 第六章 总结与展望