

# 基于大数据的动车组故障诊断方法研究

赵 伟

中国铁路济南局集团有限公司济南车辆段 山东济南 250000

**摘 要:** 随着我国的经济快速发展,铁路行业的发展也取得了令人骄傲的成绩。尤其以动车组的成就最为令人瞩目。动车因其优秀的安全性能、舒适性、准点和环保等优点得到了我国人民及世界各国的认可。随着动车的运营规模不断地扩大,相关单位收集了大量的故障诊断和状态监测的数据,通过这些数据进行分析和判断其背后的价值以及规律,并利用这些数据进一步指导后续的动车组的运用维修工作。这对于提升动车组的运行安全有着非常重要的作用。再分析过程中我们可以利用数据分析和数据挖掘等方法实现故障诊断分析。本文通过分析Hadoop平台及一些相对应的算法来解决动车组故障诊断中的维修问题,希望带来一定的参考价值。

**关键词:** 大数据; 动车组; 故障诊断

## 引言:

由于动车的运行环境复杂,其整体结构不仅复杂,包含的零件还非常的多。不同零件的使用寿命和负载情况也有所不同,零件出现损伤和一些故障都是不可不免的现象。这就需要动车机组的检修工作人员及时地发现故障,并且第一时间进行维修<sup>[1]</sup>。还需要相关工作人员定期巡查,及时的预防车辆发生事故,这对行车安全和运行效果有着非常重要的作用。随着我国信息化技术的成熟,传感器技术不断的成熟,并结合计算机技术,可以实现动车组监测数据的收集。现阶段已经实现了数据的快速收集,那么如何快速高效的处理这些技术就成了现阶段动车组故障诊断面临的一大挑战<sup>[2]</sup>。其实,只需要充分地利用大数据技术就可以从根本上解决数据增长过快和数据处理效率底下的问题。利用大数据技术不仅可以覆盖全列车,还可以全方位多个维度地对动车组出现的故障进行在线监测,通过数据挖掘和分析故障产生的原因和故障变化的规律。所以大数据和信息化技术应用在动车组运维服务中,可以使运维服务质量上升到一个新的高度,也是我国现阶段铁路高校运行的重要保障。

## 1. 数据挖掘理论

### 1.1 特征提取

特征提取简单地讲,就是在大量的动车运维数据中提取有价值的数据进行深度的分析并挖掘这些数据背后的价值<sup>[3]</sup>。从中这些原始数据中提取出一些可以使用并

且加以描述的数据特征并提取出来。在对这些数据进行数据特征提取的过程中,要尽可能地保留数据所在的动车组的原有的相关信息。除此之外还要尽可能地分析干扰因素,并减少干扰因素对数据的影响,将这些信息整合完成后输入数据模型分析系统中,利用计算机的学习算法得出最接近事实的数值。同时在进行大数据资源整合的过程中,要合理地利用数据关联规则,并找出与分析对象有关的数据<sup>[4]</sup>。除此之外,我们还要从庞大的历史数据中对不同的数据进行分析 and 分类,给算法建立起一个明确的系统和目标。

### 1.2 Hadoop平台介绍

Hadoop平台的构成主要包括开源分布式系统、主要的数据仓库、分布式文件系统、引擎等几个部分构成。在这部分中,最重要的核心就是分布式文件系统,这个系统是平台的运行基础,位于最底层,主要功能是用来存储其他节点位置上的文件。它不仅具有较高的容错性能,对运行他的设备硬件要求较低,可广泛地运用到各种廉价的硬件设备中。除了以上这几个优点,它还可以一流的形式访问平台中其他的应有程序,并形成数据集,从而提升整个系统数据的吞吐量。所以,许多大数据处理系统都会用到分布式文件系统。MapReduce和分布式文件系统不同,他是一个软件的框架系统,主要用在数据的并行计算中<sup>[5]</sup>。先把所有的计算任务进行分配,将任务分配到对应的节点上,同时进行多个节点的计算,然后将多个节点计算的结果进行合并得到最终的计算结果。并行计算的过程中,不论是改变任务,还是负载均衡的容错处理,只需要一个简单地变成都可以实现,因此该软件框架的使用范围非常的广泛。

**作者简介:** 赵伟,男,汉,1986.10.24,山东济南,学历:本科,职称:助理工程师,毕业于华东交通大学,主要研究动车组检修,邮箱:345059244@qq.com。

## 2. ADASYN 算法

ADASYN 算法是一种利用样本合成的方法进行计算的算法。这种方法可以利用少数的样本分布情况，合成该样本的少数类样本，并且在合理的容易分类的位置形成较少的样本，在难分类的地方合成较多的样本。

### 3. 高速动车组故障诊断

高速动车组的故障诊断主要指的是对列车实际运行状态和理想状态进行对比，通过两者出现的偏差判定高速动车的状态。高速动车组的故障诊断过程是一个非常庞大的过程<sup>[6]</sup>。建立起相关数据模型，并使用这些数据模型对动车组的故障做出提前预判，实现预防事故出现的目的。高速动车组在实际运行过程中呈现速度较快，变化较多的特点，通过数据组的有效挖掘和分析可以有效地提升动车组故障的诊断能力，这对动车组来说是非常重要的。

#### 3.1 基于关联规则进行故障诊断

首先要在庞大的数据库中找到出现故障的那一组数据，之后对故障数据进行预处理，并利用关联法则及相关算法对处理过后的故障数据生成对应的故障规则，最后根据故障规则进行决策<sup>[7]</sup>。系统会根据提交的故障规则进行决策并输出决策结果。接下来在对输出后的决策结果进行性能及各方面的评价，如果评审可以通过的话，那么该决策就可以实施，如果评审不通过的话那么使用该决策。

#### 3.2 高速动车组的故障数据处理

高速动车组的数据处理及特征提取是在故障诊断过程中非常重要的一个环节，如果数据特征提取出现错误，可能会导致整个故障诊断都发生问题。由于高速动车组在实际运行过程中受到的环境因素影响较大，所以我们获得的原始数据可能会存在一定的偏差，例如：原始数据中包含一定的噪声<sup>[8]</sup>。除此之外，原始数据还有冗余性的特点，这一特点可能导致数据在某些维度或者某些值的呈现连续性非离散性分布。为了提高所收集的信息的可靠性以及可用性，就需要我们对原始数据进行预处理。高速动车组所采集到的数据状态包括了动车组行驶过程中的故障信息、故障编号、动车组编号等。其中，故障编号作为动车组出现故障信息的主要表现形式，形成高速动车组的故障数据。利用这些故障数据及相关专业人士的经验，形成一个故障知识库。

## 4. 国外动车故障信息检修系统

### 4.1 德国动车组检修技术

德国的动车故障检修系统采用的是 COBRA 检测诊断

系统。该系统最大的优点就是目的性较为明确。该系统的形成主要是以车辆为单位，将不同车辆系统相互链接形成一个完整的车厢系统。该系统具有诊断功能，相关的工作人员可以根据该系统判断是否有故障，并根据不同的故障级别将故障信息输送到对应的维修人员处。使维修人员可以第一时间完成检修工作。

### 4.2 其他国家

从全世界各国来看，大数据的动车组。故障检修技术以及发展得较为成熟，国外许多故障信息管理系统不仅功能全面，而且操作非常的便捷，具有信息传输、接受、存储管理等各方面的功能。同时这类系统还具备了自动处理信息的功能，可以对数据进行深度挖掘并应用，不同类目的子系统模块的结构也非常的清晰，不同模块对应的系统的目标用户定位也非常的准确，真正做到了对高速动车组的全方位监控。

### 4.3 国内动车组检修系统与国外动车组检修系统的差距

通过对比国内外的故障检修系统的差距我们可以得出我国现阶段故障诊断中存在的问题。首先就是目标用户不准确。目标用户即系统所要对接的用户群，主要指的是某一部分或某一个设备所针对的用户不够明确。举一个最简单的例子，例如动车机组的空调出现了故障，那么需要知道这条消息的人应该是动车组的后勤维修人员，其他人员不需要掌握这条信息。但我国现阶段的系统在运行过程中对应的目标群关系不够明确，导致信息对接出现偏差，降低了系统中的数据对于动车检修工作的指导价值。第二点，历史故障信息的利用质量不高。故障信息本身是具有重要价值的，故障数据不仅可以侧面的体现动车组过去行驶的数据，同时还可以对未来的运行状态进行预测，但现阶段我国对于故障信息数据的研究不够深入，故障预测的质量较低。第三点，无法良好的协同检修与列车的运行计划。列车的运行计划虽然给动车组的检修留下了充足的时间，但在实际运行过程会遇到一些不确定因素，导致动车组的检修时间过长，这些都没有考虑在列车的运行计划过程中。除此之外列车的历史故障信息也没有得到较好地利用，难以实现合理地安排并制定有针对性的列车检修计划，这些都是我国现阶段大数据背景下动车组故障诊断技术存在的一些问题。我们要积极地寻求改进的方法，使我国的铁路事业有更好的发展。

## 5. 结语

综上所述，在大数据时代下的高速动车组故障诊断，

主要是依托动车组的历史故障信息库，并结合多年来专家的经验，运用关联规则下的挖掘算法对动车组故障信息进行计算并得出计算结果，并以此为关联依据得出动车组故障诊断的结果，为高速动车的检修提供具有参考价值的诊断依据。

#### 参考文献：

[1]赵鹏振，刘继.基于随机共振与蝙蝠算法的高速动车组滚动轴承故障诊断[J].中国测试，2021，47（3）：16-23.

[2]毕涛.基于CBR和RBR动车组车辆故障诊断专家系统探讨[J].建筑工程技术与设计，2021（8）：236.

[3]黄金根.基于大数据的高速动车组关键部件故障诊断技术研究[J].汽车博览，2021（3）：96.

[4]褚云博.基于神经网络动车组牵引电机定子匝间故障诊断系统设计[J].石家庄铁路职业技术学院学报，2021，20（2）：89-91.

[5]吴越，黄浩，陈科，等.温州市域动车组牵引系统接地故障诊断与保护应用研究[J].机车电传动，2020（4）：146-151.

[6]韩俊臣，崔利通，李晓峰，等.动车组轴承异音故障诊断与研究[J].机车电传动，2020（1）：149-153.

[7]胡文涛，孟建军.基于大数据的动车组故障诊断方法研究[J].工业控制计算机，2020，33（6）：31-32，35.

[8]宋宏智，李力，杨兴宽，等.高速机车轴承故障诊断与剩余寿命预测的发展及展望[J].轴承，2020（3）：61-67.