

面向工业4.0的直升机制造大数据平台的应用研究

张贝贝

(昌河飞机工业(集团)有限责任公司, 江西 景德镇 333000)

摘要: 随着信息化技术的不断发展, 大数据技术正以惊人的速度融入到人们的日常工作和生活中, 推动着工业4.0时代的到来。本文利用大数据技术在数据管理优化方面的优势, 根据直升机制造生产活动的特点构建了一个面向工业4.0的直升机制造大数据平台, 平台主要由生产过程监测与管理、直升机制造大数据平台和直升机制造生产环境三层结构组成。此平台不仅能够应用于直升机零部件加工生产的环节中, 而且在产品创新、产品供应链的分析与优化等环节也会发挥更大作用。

关键词: 直升机制造; 大数据; 工业4.0

中图分类号: TP311.13 **文献标识码:** A

The Application Study on the Helicopter Manufacture Big Data Platform for Industry 4.0

ZHANG Beibei

(ChangHe Aircraft Industry Group, Jingdezhen 333000, China)

Abstract: With the continuous development of information technology, big data technology is integrated into people's daily work and life at an amazing speed, driving the realization of Industry 4.0. With the advantages of big data in data management optimization, the paper builds a helicopter manufacture big data platform for Industry 4.0 based on the characteristics of the helicopter manufacture. It is mainly composed of the following three layers: production process monitoring and management, helicopter manufacturing big data platform and helicopter production environment. The platform can not only be used in helicopter parts processing, but also play an important role in product innovation, production supply chain analysis and optimization.

Keywords: helicopter manufacture; big data; industry 4.0

1 引言(Introduction)

早在2013年, 德国就提出了“工业4.0”生产模式的概念, 思科、通用电气和英特尔等企业纷纷响应组建了“工业互联网联盟”, 为制造业赋予了智能化、信息化和个性化生产的新特性, 促进了数据资源的利用。“工业4.0”生产模式以大数据技术为基础, 实现降低对劳动力的依赖、满足用户的个性化需求和减少流通成本的目的, 其包括智慧工厂、智能生产和智能物流三大部分, 其中智慧生产是关键核心。

直升机制造大体分为零件制造、部件装配、机身铆装和整机总装四大部分, 但包括多个生产环节, 每个环节都有各自的生产规律, 属于典型的离散型生产活动, 其特点包括: (1)产品规格多、结构复杂、技术难度大; (2)品种多、采用中小批量、混线生产模式; (3)产品物料需求量大而杂, 物流管理复杂; (4)过程中大量数据及信息分散在各部门技术人员、管理人员、甚至工人手中, 部门之间的信息交流不畅, 从而导致生产及质量问题不能及时得到反馈、解决, 造成大量的生产计划延期, 生产效率低下。

自我国工业4.0战略推行以来, 物联网技术和互联网等信

息技术与直升机制造业的不断融合, 直升机制造企业在生产过程中积累了大量的数据, 不仅包括指令、生产编号、图号等文本格式的数据, 还包括图片、声音、视频等复杂格式的信息, 这些数据的产生有规模性、多样性等特点。如何高效的获取、存储、分析处理海量数据, 并最大化地从这些数据资源中提炼价值成为直升机制造行业普遍关注的热点。

大数据技术是目前公认的处理海量数据的有效方法, 被誉为助推实体经济迈向“工业4.0”的“核动力”。利用大数据管理技术和分布式计算分别实现对直升机加工数据的存储和计算, 然后利用大数据分析技术对这些数据进行挖掘, 从中获取对生产决策有价值的信息, 帮助企业合理安排生产计划, 从而有效提高企业的综合竞争力^[1]。

2 国内外发展现状(The development situation at home and abroad)

近年来, 直升机在国防建设中扮演角色的重要性日益凸显, 直升机产业的发展已经被各国提升到战略高度。发达国家西科斯基直升机公司在产品制造与设计领域目前处于世界领先地位, 其利用数字化制造技术不仅完成了从传统的大批量生产

向精益生产模式的转变，而且创新了管理模式，实现了对成本的透明性和作业流程的标准化。此外，西科斯基公司S92项目中，利用CATIA软件将来自六个不同国家和地区供应商的部件实现了高效化无纸整合，建立了一个设计组件整合的“虚拟直升机”模拟系统，缩短了生产进度。

欧直公司在数字化技术方面的应用水平实际上并不亚于西科斯基公司，该公司在企业资源策划、数字化电子样机模拟等方面均达到了很高的水平。欧直公司采用了商务与决策公司同SAP公司联合为其推出的商务智能平台，主要用于绩效管理 with 知识管理。

随着“中国制造2025”战略的提出，中国制造业迎来了一次新的改革浪潮，我国直升机制造企业纷纷建立了一系列的企业信息化系统来为公司的经营决策、资源配置提供支撑，其中包括资源的优化配置和目视化管理等，基本实现了生产过程的数字化管理。但与发达国家相比，中国直升机制造现在还是以低端制造为主，主要表现在以下两个方面：

(1)没有充分利用数据资源，将数据资源的价值资源挖掘出来，例如：若能从海量的数据中提取设备故障相关征兆数据，采用预测与相关管理算法生成该设备将要或已经发生故障的相关信息，使得设备元件具备“自我感知”能力，分析预测自身状态，就能采取措施进行预防。

(2)缺乏面向产品制造的“知识库”管理，若能根据大数据技术分析不同生产环节的当前需求，通过推理将不同产品状态、设备信息相关联并呈现给管理层，就能对企业科学、有效的决策提供有力的技术支持。

总之，直升机制造生产活动是建立在精密、有序的生产工序的基础上，通过各工序之间的密切配合，在确保直升机制造质量的同时提升直升机生产效率。大数据在数据管理优化方面具有极强的优势，除了能够应用在直升机零部件的加工生产中，还能够对产品创新、产品供应链的分析与优化，以及工业物联网的应用中发挥出更大作用。将大数据技术与数字化和检测等技术相结合，能够提前对直升机生产环节中的各关键要素进行分析与预测，从而更好地提高产品的质量与生产效率^[2]。

3 直升机制造大数据平台的总体框架(The overall framework of helicopter manufacturing big date platform)

直升机制造过程中产生的数据具有规模性、多样性等特点，为了从这些数据资源中挖掘有用价值，首先要确保这些数据的存储的完整性和处理的高效性，本文构建了一个面向工业4.0的直升机制造大数据平台，平台的框架图如图1所示。

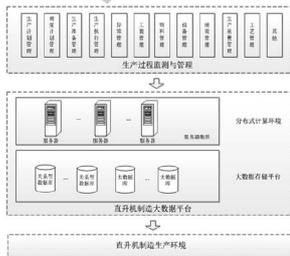


图1 直升机制造大数据平台总体框架图

Fig.1 The overall framework of helicopter manufacturing big date platform

从图1可以看出，直升机制造大数据平台主要由生产过程监测与管理、直升机制造大数据平台和直升机制造生产环境三

层结构组成^[3]，本文主要研究前两层。

直升机制造大数据平台的主要功能是大数据存储和分布式计算，大数据存储平台通过传统关系型数据库与大数据技术相结合的方式实现对直升机制造生产环境中所产生数据的存储与管理。在直升机制造过程中，有关系型数据库对实时性要求较高的数据进行处理，同时定期利用数据迁移技术将关系型数据库中的数据同步到大数据库中，实现直升机制造海量数据的存储。

随着时间的推移，直升机制造企业会积累大量的生产相关数据，传统的集中式处理方法已经无法满足数据快速处理的要求，本平台通过搭建分布式计算环境将处理任务分解到不同的服务器上，用“分治”的思想将大规模任务分解为若干个规模较小的任务，分解任务完成后汇总到同一个服务器上，分布式计算的实现不仅提高了数据处理的效率，数据备份的机制也避免了单个服务器崩溃导致整个系统瘫痪的危险^[4]。

生产过程的监测与管理主要是建立在对直升机制造海量历史数据分析的基础上，此层不仅能够帮助生产执行单位根据生产设备的运行状况、原材料的数量和产品的交付期等信息选择合适的生产方案，对生产设备的任务和产品的生产顺序分别进行合理分配和排序，充分利用生产环境中的资源进行优化排产，而且还能对产品的质量进行预测分析，为企业领导层的决策提供数据支持。

4 关键技术(The key technology)

4.1 大数据存储技术

大数据下，直升机制造业对于海量数据的存储将面临如下挑战：(1)存储数据量大，直升机制造工序复杂，积累的数据将达到PB级别甚至更高；(2)直升机制造数据来源广泛，数据形式和数据结构均比较复杂；(3)要满足数据的完整性、可扩展性。直升机制造大数据存储采用分布式文件系统HDFS和按列存储数据库HBase。

传统的数据存储方式已经不能满足业务的需求，分布式文件系统是被国内外一致认可的存储大数据最好的工具，比较有代表性的分布式文件系统当属Hadoopz的HDFS(Hadoop Distributed File System)。HDFS采用master/slave主从架构，即一个控制节点和若干个数据节点组成^[5]。一个文件信息通常被切分为若干部分分别存放在不同的数据节点，控制节点相当于整个集群的目录，用来存储数据节点的文件元数据。

HBase是基于HDFS是一种分布式的非关系型数据库，其本质是一张面向列的分布式多维映射表。传统的MySQL等关系型数据库通常被部署在一台服务器上，而HBase数据库是在多台服务器上分布存储的，如表1所示数据在Hbase的存储方式^[6]。

表1 数据在Hbase的存储方式

Tab.1 The storage way of data in Hbase

Row Key	Time Stamp	Column Family	
		Column1	Column2

如表1所示，HBase主要由Row Key(行键)、Time Stamp(时间戳)和Column Family(列族)组成，Row Key由任意部大于64kB的字符串组成^[7]；Time Stamp是主要是用于记录数据的版本，当有新的数据插入，系统会自动添加时间戳；Column Family必须在使用表之前预先定义，但列族成员可以根据需要随时添加。

4.2 大数据分析技术

Map/Reduce则是Hadoop的唯一分布式程序设计模型，主要用于海量数据集的并行处理。Map、Reduce来源于函数式编程模型中的“Map”“Reduce”两个核心操作。

当进行大数据分析时，此模型先对输入的文件进行切割，然后分发给Map函数，Map函数将其映射成另外一组中间文件，Reduce函数主要作用是整合Map函数映射成的中间文件，Map函数的映射规则和Reduce函数的整合处理是开发人员通过编写代码实现的。

4.3 大数据可视化技术

大数据可视化包括两层含义，不仅指利用数据挖掘技术从数据中挖掘有用的信息，而且还要将得到的信息向用户直观地展示。大数据可视化的主要步骤包括获取数据、分析数据、过滤数据、挖掘数据、展示数据和数据总结等。

将数据挖掘应用于直升机制造企业^[8]，一方面可以更加精准地把握产品的加工、质量和安全管理等方面的情况，另一方面可以预测未来可能遇到的生产和管理上的问题。数据挖掘常用的算法是关联分析法，通过分析数据和系统之间的关联，然后进行数据的分类和聚类，最后有目标型的挖掘数据潜在的价值。

可视化展示是建立在数据挖掘的基础上，将数据中的信息形象化展示出来，常用的方法是多维智能分析法，该方法可以对分析对象进行时间、空间和逻辑关系等多个维度进行分析。

5 大数据平台的实践应用(The practical application of big date platform)

5.1 系统功能简介

直升机制造大数据平台的有些功能是需要后台做处理的，而这些功能对于实时性的要求并不高，比如企业每月都要进行的加工数据宏观分析、预测企业加工效率增长趋势、直升机制造关键设备的健康管理等功能都属于后台数据处理，本文以直升机制造关键设备的健康管理为例进行说明^[9]。

A公司已经构建了一套具有状态感知、实时分析、自主决策特征的旋翼系统部件生产及装配的智能制造生产线执行系统，通过实时监控发现生产过程中的问题，及时采取应对措施。若能从海量的历史数据中提取故障征兆数据，采用预测与相关管理算法进行计算，从而生成该设备将要或已经发生故障的相关信息，分析预测自身状态，就能更有效提高产品加工效率。

系统由设备数据采集、设备专项体检、设备健康评估和建议、设备健康风险预警四大模块组成，如图2所示。

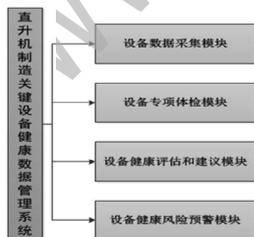


图2 直升机制造关键设备健康管理信息系统

Fig.2 The health management system for key equipment of helicopter manufacture

设备数据采集模块主要是针对设备的基本信息和历史状态数据进行管理，设备的基本信息包括设备编号、购买时间、加工零件类型、使用年限等；设备专项体检模块中的“专项”包括生产数量、产品种类、工艺难度、原料供应、废品率、产品要求精度、员工效率、设备生产率、设备故障率等，通过此

模块可以根据设备和产品的特点，有针对性分析影响设备利用率低下的原因；设备健康评估和建议模块能够对专项中的所有指标进行评估，并根据评估结果给予一定的建议，针对生产管控系统某个关键设备利用率低的问题，可以通过此模块进行分析，并采取措施；风险预警模块主要是对设备的近期若干次的健康体检报告的数据进行管理，若发现严重影响设备运行率的指标超低，会进行红色预警。

5.2 开发工具及环境

硬件工具及环境：本项目的Hadoop集群由五台普通的PC机组成，其中两台用作控制节点和第二控制节点，其余三台用作数据节点。

软件工具及环境：本项目基于Java EE开发模式和B/S架构，后台采用Struts、Spring和Hibernate混合框架，前台采用ExtJS框架；所有节点采用Linux操作系统、Java Web服务器采用Tomcat。

5.3 大数据存储

Hadoop集群搭建好后，通过MyEclipse平台开发大数据平台存储功能，主要包括的类如表2所示。

表2 大数据存储相关类

Tab.2 The related classes of big data storage	
类名	类的功能
FSImage	保存文件系统目录树
FileStatus	查询文件属性信息，包括文件大小、备份数、最后修改时间和文件权限等
DateStorage	管理整个存储系统的状态
FSEditLog	修改日志文件
JobConf	对文件数据进行压缩
FSNamesystem	维护数据节点信息
DataBlockScanner	定时对数据块文件进行数据校验
FSDataOutputStream	从本地文件系统中复制文件到文件系统中
FSDataInputStream	读取中的指定文件
Configuration	该类的对象封装了客户端或者服务器的配置

5.4 大数据挖掘和设备健康状态预测

本文在数据挖掘阶段采用神经网络、非结构数据关联分析、模糊识别等算法，形成了算法工具库。以非结构数据关联分析为例，直升机制造关键设备健康管理系统可以通过两种方式定义关联规则的规则，一种是对非结构数据库中的数据进行分析发现潜在的关联，另一种是根据直升机制造过程中积累的经验自定义关联规则^[10,11]。

本文采用回归分析法进行设备健康状态的预测，目的是根据重要数据的变化趋势，对未来一定时期内关键设备的健康状况进行推测，提前采取应对措施。

5.5 大数据的可视化分析

直升机制造关键设备健康的大数据可视化分析主要分为三个步骤^[12]，如图3所示。



图3 可视化分析处理流程

Fig.3 Visualization analysis process

(下转第18页)