

文章编号：1005-8451 (2013) 01-0058-04

动车组全生命周期数据可视化研究

李 燕，张惟皎，贾志凯

(中国铁道科学研究院 电子计算技术研究所，北京 100081)

摘要：本文将数据可视化技术引入到动车组数据管理范畴，介绍了动车组的数据分类和可视化现状，可视化技术的发展历程及其应用到动车组领域的主要问题和研究主题。对数据可视化技术在动车组管理的应用前景进行了展望。

关键词：动车组；全生命周期；数据可视化

中图分类号：U266.2 : TP39 **文献标识码：**A

Research on data visualization of EMUs whole life-cycle

LI Yan, ZHANG Weijiao, JIA Zhikai

(Institute of Computing Technologies, China Academy of Railway Sciences, Beijing 100081, China)

Abstract: In this paper, it was introduced the data visualization techniques into the management scope of EMUs data. Firstly, the paper described the data variety and the visualization actuality. Secondly, this paper introduced the development of visualization techniques and their applications in EMUs area, including the main problems and the potential research topics. Finally, the future application of visualization techniques in EMUs was discussed.

Key words: Electric Multiple Units (EMUs); whole life-cycle; data visualization

动车组的全生命周期是指动车组从设计、建造开始所经历的新造、备用、运用、检修、报废等全部生命历程。依附动车组全生命周期会产生并存在海量的运用、检修等业务数据，这些数据以时间为维度，隐含着全路动车组的设备状态、运用规律、检修问题、故障发生规律等信息。这些信息伴随动车组的全生命周期存在，甚至在所依附的动车组报废后，依然作为历史数据保留，作为各种潜在规律统计、分析的基础。尽管现有动车组管理信息系统已经通过图形、图像、图表等多种可视化方式为用户提供直观的信息展示，但数据之间的复杂关系和隐藏的信息、特征依然很难被发现和挖掘出来。数据可视化可以帮助用户在海量的数据中迅速发现隐含的特征、模式和趋势，为业务决策提供支持和依据，对于分析动车组运用、检修中的核心问题具有广阔的应用前景。

本文分析了目前动车组管理信息系统的数据特点和多现有的统计分析现状，指出系统在统计、分析方法的不足之处，对现有的可视化方法和发展历程进行了回顾，并结合动车组管理信息系统

对动车数据可视化模型的构建、面临的挑战等关键问题及未来的趋势展开了讨论。

1 动车组数据分类及统计分析现状

1.1 动车组数据分类

(1) 公用基础数据：包括需要人工整理录入或通过接口从其它信息系统中获取的各类基础代码、数据字典、基本技术参数、对照表、设施布局图等空间基础数据。这类数据由各个应用子系统共用，内容相对稳定，数据更新频度较低。

(2) 履历数据：包括动车组履历、关键配件履历和设备台帐等。这类数据的初始化需要人工批量整理和录入，需要结合相关业务处理实现动态更新维护，在动车组、配件和设备寿命期内长期保存。

(3) 生产计划类数据：包括上级调度指挥系统下达的不同时段的动车组交路、运用计划、检修计划（调度命令），动车基地、运用所向上级调度指挥系统提交的计划变更申请，铁路局和铁道部系统制定的各种轮廓计划及实绩，动车基地、运用所系统制定的各种生产计划及实绩、作业单

收稿日期：2012-11-12

作者简介：李 燕，助理研究员；张惟皎，副研究员。

等。这类数据主要作为生产组织、任务量统计、生产能力评估分析的依据，原始数据保存期一般为2年以上，其中动车组开行实绩等数据需长期保存，用于可靠性分析。

(4) 作业管理数据：包括生产作业过程中采集的各种作业派工单、检修记录、故障记录、质检和验收记录、人员考勤记录、现场视频监控数据、交接班日志等。主要用于生产作业过程的精细管理，并作为作业量统计、绩效考核、安全质量分析的原始数据；数据保存期一般为2年以上，其中的动车组和关键部件的检修记录和故障记录需要长期保存，用于可靠性分析。

(5) 技术资料数据：包括维修指导书、工艺卡片、图纸、资料目录等相关技术资料，主要供查询检索使用，数据更新周期较长。

(6) 安全监控管理数据：包括由各种检测、测试和实验设备自动采集的数据，由车载信息地面监控系统提供的动车组运行情况、安全状态监测、质量检测过程数据等。这类数据主要作为运行安全监控、安全质量评估、技术状态分析和指导维修的依据；试验数据还可作为维修质量检查的依据，支持趋势分析和故障规律分析等，数据保存期为10年以上。

(7) 物流信息：包括配件库存信息、出入库登记、配送记录、退料记录、物资采购订单、收货和付款凭证、货源信息、供应商资料、合同协议等；数据保存期一般5年以上，其中的收货和付款凭证记录等作为历史数据长期保存待查。

(8) 人力资源、办公类信息：包括人员基本信息、人事档案、人员培训、薪酬、绩效管理、公文、规章制度、总结报告、通知公告信息等；除通知公告信息外，大部分数据保存期一般较长，部分数据要求保密，其中历史数据长期保存待查。

1.2 动车组管理信息系统统计分析现状

现有动车组管理信息系统依托全路各运用所、基地数据的汇总机制和不同系统之间数据的共享机制，针对实际业务中的管理要求，采用报表、图形、图像、趋势曲线等多种可视化方式，展示数据沿时间、运用/检修单位、各种业务关键字等众多不同维度的统计、变化趋势，为决策管理提供了可靠的数据支持。

但上述统计、分析手段还相对简单，仅能够

在明确业务统计维度，已知维度几何属性的前提下，对数据进行多维分析。数据之间的隐含制约或协同关系，隐藏的信息和特征依然很难被挖掘出来，更没有通过可被迅速感知的图形化方式表现出来。本文研究的重点是如何通过多维数据的可视化发现动车组检修、运用等业务数据间的关系，以此提高运输能力、检修效率并保障运行安全。

2 可视化原理与实现

2.1 可视化方法

可视化是数据描述的图形表示，其充分利用人类对可视模式的快速认知能力，将数据、知识转化为图像、图形等视觉形式^[1]，达到一目了然地揭示数据中的复杂信息，产生对数据的有效洞察和新的理解，以达到清晰传达与沟通信息的目的^[2]。可视化参考模型(Reference Model for Visualization)，描述了原始数据、数据表格、可视化结构和视图之间的转换关系，以及用户进行的数据变化、可视化映射、视图变换等操作，如图1所示。

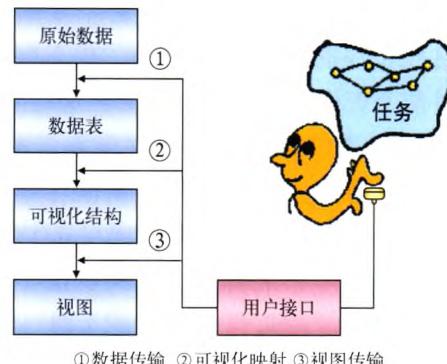


图1 可视化参考模型

可视化发展初期，大量的信息可视化都是基于标准格式的，基本的可视化展示方式包括条形图、折线图、散点图、饼图、组织流程图等。这些格式能够容易地通过软件生成，便捷且易于理解，是常见的可视化问题解决方案。随着计算机处理能力、信息技术手段、交互方式的不断更新和进步，更多的研究专注于冲破传统可视化的框架限制，以没有几何属性的抽象信息为研究对象，使用非传统的编排、组织方式，有效的用户界面浏览、过滤、抽象、交互等方法发现隐藏在信息内部的特点和规律。

(1) 基于几何投影的方法

几何投影法的基本思路是寻找多维数据的有意义投影，以几何化法或几何投影方式表示数据。此类方法以最早提出的平行坐标方法为代表，该方法将n维数据属性空间通过n条等距离的平行轴映射到二维平面上，每条轴线表示一个属性维，每个数据项都可以用一条折线段在n条平行轴上表示出来，相似的对象具有相似的折现走向。平行坐标技术可以将多维信息有效的表示在二维平面中，并能够对初始数据的相似性进行分析。针对平行坐标技术中折线容易重叠交错，根据轴线的取值范围与数据属性间不同的映射关系，平行分层坐标^[3]、坐标轴重排技术^[4]、小波刷^[5]等很多研究对该方法进行了改进。

(2) 面向像素技术

面向像素技术^[6]的基本思想是将每一个数据项的数据值对应于一个彩色的屏幕像素，并将属于某一个属性的数据映射在一个独立的窗口中，其关键问题是怎样排列这些像素点，可分为独立于查询的方法和基于查询的方法。独立于查询的方法根据某些属性对数据排序，使用某种屏幕填充模式在显示器上排列数据值。对于根据某个属性进行自然排序的数据特别有效。基于查询的方法依据对某个查询结果的数据相关性进行可视化，根据数据值与查询要求的符合程度来匹配颜色，通过自定义的距离函数反映每个数据与查询之间的匹配程度，距离越小越是用户希望的结果。这种方式下，用户不但能够得到查询数据，还能直观的观察到数据不同属性之间的相关性、功能依赖等。

(3) 基于图标的技术

基于图标数据的基本思想是将每个多维数据映射成为一个图标，即用一个简单图标的各个部分来表示多维数据的属性。Chernoff-faces是其中最著名的图标显示技术，它将数据项中的两维映射为用于显示的坐标维度，剩下的维则被映射为脸的各个器官以及脸本身的形状，但这种可视化方法能够应用的数据集十分有限。枝形图也是一种有效的可视化方法^[7]，其首先选取多维属性中的两种作为基本显示面坐标轴，然后利用小树枝长度或角度的不同表示出其它属性值的变化。

(4) 基于层次的可视化技术

基于层次的可视化技术^[8~10]主要针对具有层次结构的数据信息，其基本思想是将多维数据空间划分为若干子空间，对这些子空间以层次结构的方式组织并以图形表示出来。

除上述几大类外，还有很多研究致力于使用虚拟现实、数据挖掘等技术，强调动态的展示数据，增加数据与用户的交互程度等。

2.2 实现方式

数据可视化技术包含以下几个基本步骤：

(1) 分析目标主题。可视化要解决的具体问题是数据组织的基础，通过对目标主题的分析，确定哪些数据在可视化的过程起支撑作用。(2) 把数据转换成有意义的结构，即建模—获取数据与定义维度指标。根据维度和指标的概念，从现有的数据表中选择可对应维度或指标的字段，将这些字段的信息分别保存在维度表和指标表中，维度表的信息主要包括：维度名称、维度类别、表ID、字段名、规则等，指标表的信息主要包括：指标名称、指标类别、表的ID、字段名、规则、聚合类型等。(3) 对数据维度进行分析。根据目标主题分析的结果，确定最优的展示数据形式，并确定展示规则和展示样式。(4) 对数据进行可视化展示。

3 动车组可视化关键问题

3.1 可视维度的选择

数据可视化要解决的是多维数据的展示问题，目的是用图像更加清晰的描述数据间的复杂关系和隐含信息。在这个过程中，维度的选择十分重要，关系到数据的哪些属性可以被用来表示数据。通过对动车组运用、检修工作的充分了解和分析，针对不同的业务和可视化目标，可供选择的维度主要有以下几大类。

(1) 静态履历数据

动车组和关键配件静态履历包括车型、车辆和关键配件的各种技术参数，其中对运营、检修业务具有重要意义的参数都可作为该类数据的维度信息。如车型参数可具有的维度有列车总长、车体最大宽度、车体最大高度、编组总重、全列载重、最高试验速度、标记速度、最大制动距离、转向架中心距离、LKJ、动力分配方式、定员等；

车辆参数可具有的维度有全长、车体高度、定员、轴重、自重、轴式、载重、额定输出功率、车种、代号、端门形式、转向架型号、座椅形式等。

(2) 配属数据

动车组配属数据是动车组履历中最基本的信息，可供展示的维度包括历次配属操作的各级配属单位、配属时间、配属类型、配属电报发布时间、送车时间、送车单位、接车时间、接车单位等。

(3) 运用数据

动车组运用管理是动车组管理的核心业务，包括动车组的上线开行计划与实绩，随车机械师乘务担当计划与实绩，全路动车组每日车组状态信息和图定、临时交路的管理。可作为展示维度的数据属性包括动车组每日的走行公里数据、详细状态、停放位置、担当路局和运用所，担当车次目的地、运行时间、运行里程、运行线别等。

(4) 检修数据

检修数据目的是对全路动车组各修程检修工作进行综合、统一管理，对检修情况进行统计、对检修历史进行查询、并对检修进度进行实时监控。检修数据中可供选择的维度有检修单位、检修日期、检修时长、检修时车组的走行公里、检修项目种类、车组检修项目状态、各个不同工序的作业开始和完成时间、完成质量量化标识、各类检修/检测数据类别等。

(5) 故障数据

故障数据为多角度的统计分析，了解故障分布情况及变化趋势，为故障趋势分析及改进检修工作提供依据，可以分为检修故障和运行故障。运行故障的主要可视维度包括故障的发生时间、发生线别、故障发生时走行公里信息、故障发生部位、故障发生地理位置区间、故障对运输的量化影响、故障处理时间、故障处理方式、故障责任类型、责任单位等。检修故障的主要可视维度包括故障发生部位、故障发生频度属性、故障发现时间、故障处理方法、故障处理结果等。

根据不同的可视主题需要，可以对上述可视维度进行增加、减少及联合分析，以满足从不同角度剖析和展示数据内在关系的目的。

3.2 可视化研究主题

(1) 动车组和关键配件的全生命周期可视化

在动车组数据整个生命周期中，不同类型的

业务数据需要有不同水平的性能、可用性、保护、迁移、保留和相应处理。通常情况下，在其生命周期的初期，数据的重要性和使用频率会维持在一个很高的水平上，以达到高可用性和提供相当等级的服务。随着时间的推移，部分静态数据将不会再参与到统计分析中，数据的重要性会逐渐降低，使用频率也会随之下降。但以时间为坐标轴的动态数据本身所携带的历史、经验信息，随着时间的推移，对于统计、分析依然有着不可小觑的意义。利用动车组及关键配件的全生命周期数据，以时间为基本维度，对整个车组生命周期中的车组状态、开行历史、故障历史、检修履历等进行可视化展示，对于动车组运用状态变化规律的掌握，动车组可能发生的问题进行及时的预判和预警都将具有十分重要的意义。

(2) 运用—检修计划优化

动车组运输组织工作包括制定动车组交路计划，在保证动车组运用的基础上完成各修程的检修任务。动车运用所负责动车组的日常运用和检修工作，需要编制动车组运用计划，保证运用交路得以实施；编制动车日常检修计划，保证动车组的良好状态；根据运用计划和检修计划制定调车作业计划，组织运用所内部的运用检修过程^[11]。动车基地负责动车组的高级检修，编制动车组高级检修车间调度计划，合理组织动车组的高级检修过程。以列车运行图为已知条件，通过对动车组运用、检修数据间的关系进行分析，构建合理的可视化模型，能够反应运用—检修间的制约、合作关系，为提高动车组的使用效率，降低动车组的检修成本，提高动车运用所和动车基地的检修能力具有重要意义。

(3) 故障闭环管理及后续影响研究

通过对动车组故障规律的分析，可以进一步优化动车组检修规程，减少检修成本。建立完整的CRH系列动车组故障信息闭环管理，研究实现有效的故障分析方法，建立故障知识库指导动车组的检修过程，对于动车组的安全运行具有重要的意义。故障闭环管理及后续影响的可视化研究应以动车组累计的故障信息为基础，研究故障从发生到检修完毕全过程与动车组运行规律之间的可视化展示方式，对于掌握故障的发生规律，分析动车组的可靠性，预测配件的寿命等十分关键。

4 结束语

随着动车组各项业务信息化水平的不断提高，海量数据的不断累积，数据的分析、利用将是指导运用、检修、制造的重要参考和依据。随着数据可视化技术成为更便捷的交互方式，适应更大规模和多样化的数据以及对可视化数据所反映出的结构进行形式化等方面的发展，可视化技术在动车组数据中的应用必将对综合管理、分析全路动车组各类业务及关键问题发挥重要的作用。

参考文献：

- [1] Nahum D, Stephen G. Information visualization [J]. IEEE Computer Graphics and Applications, 1997, 17(4):29-31.
- [2] Steele J, Llinsky N. Beautiful Visualization. O'Reilly Media, Inc. 2010.
- [3] Fua Y.H., Ward M.O., Rundensteiner E.A. Hierarchical parallel coordinates for exploration of large datasets[C]. Proceedings of IEEE Conference on Visualization. Los Alamitos: IEEE Computer Society Press. 1999: 43-50.
- [4] Peng W., Ward M.O., Rundensteiner E.A. Clutter reduction in multi-dimensional data visualization using dimension reordering[C]. Proceedings of the IEEE Symposium on Information Visualization. Los Alamitos: IEEE Computer Society Press. 2004: 89-96.
- [5] Wong P.C., Bergeron R.D. Multi-resolution multidimensional wavelet brushing[C]. Proceedings of IEEE Conference on Visualization. Los Alamitos: IEEE Computer Society Press, 1996: 141-148.
- [6] Keim D.A., Kriegel H.P. VisDB: Database Exploration using Multidimensional Visualization[J]. Computer Graphics and Applications. Sept, 1994, pp:40-49.
- [7] Pickett R.M., Grinstein G.G. Iconographic Displays for Visualizing Multidimensional Data[C]. In: Proc. IEEE Conf. on Systems, Man and Cybernetics. IEEE Press, Piscataway, NJ, 1988. 514-519.
- [8] LeBlanc J, Ward M O, Wittels N. Exploring N-Dimensional Databases[C]. In: Visualization' 90, San Francisco, CA, 1990. 230-239.
- [9] Johnson B. Visualizing Hierarchical and Categorical Data[R]. Department of Computer Science, University of Maryland, 1993.
- [10] Robertson G G, Mackinlay J D, Card S K. Cone Trees: Animated 3D Visualizations of Hierarchical Information[C]. In: Proc. Human Factors in Computing Systems CHI' 91 Conf., New Orleans, LA, 1991:189-194.
- [11] 王忠凯. 动车组运用检修计划优化方法的研究 [D]. 北京: 中国铁道科学研究院, 2012.

责任编辑 陈 蓉

(上接 P57)

级修作业环节中的必换件点收作为检修的一个作业环节，如果配件没有点收则不可进入下一个作业环节。

质量检验时提供给质量检验车间查询检修更换配件情况，检验结果通过系统共享给检修部门，作为车辆出库联检结果的依据。

技术资料提供查询各种配件的技术指标、全方位图片等，减少配件的误申请和误发放。

4 结束语

本文从多个专业分工协同管理等方面阐述了动车组检修过程中各专业的分工及协同管理模式，该管理模式对于指导动车组检修物资管理软件的开发和应用，把握资源协同过程等具有积极的现

实意义。

参考文献：

- [1] 刘翔. 企业协同管理系统功能设计思想研究 [J]. 中共浙江省委党校学报, 2007 (5).
- [2] 杜栋. 协同管理系统 [M]. 北京: 清华大学出版社, 2007.
- [3] 邬斌波, 张玉鑫. 基于双向协同的物资需求计划管理在电网工程中的应用研究 [J]. 华东电力, 2012 (5).
- [4] 陈飚. 协同管理系统驱动企业低碳高效运转 [J]. 软件工程师, 2010 (5).
- [5] 王爱民. 供电企业物资管理的核心环节及控制方法 [J]. 企业技术开发, 2012 (8): 104-105

责任编辑 陈 蓉