

Design and Construction of High-Speed Train Pedigree Data Management Platform

Qinglong Wei¹, Ailin Yang¹, Na Wang¹, Yongzhi Jiang²

¹Zhongche Tangshan Locomotive & Rolling Stock Co., Ltd., Beijing

²State Key Laboratory of Traction Power, Southwest Jiaotong University, Chengdu Sichuan

Email: weiqinglong@tangche.com, yangailing@tangche.com, wangna@tangche.com

Received: Oct. 18th, 2018; accepted: Oct. 31st, 2018; published: Nov. 7th, 2018

Abstract

The data warehouse of high speed EMU pedigree is subject oriented, integrated, time-varying and relatively stable, and is a collection of high-speed train space data and non-spatial data. At present, the platform of high-speed train genealogized development has been built based on genealogical theory, and the integrated platform of eight subsystems is constructed, including the virtual visualization and rapid process file generation. The data management platform of high speed EMU pedigree is the basis for the smooth and efficient operation of the integrated platform. Based on the theoretical research of high speed EMU pedigree, this paper designs and constructs a data management platform for high speed EMU pedigree. Overall architecture of this platform is introduced, and the module composition and the functions, user interfaces and uses of each module are combined, which provide a foundation for the operation and management of high speed EMU pedigree platforms.

Keywords

High-Speed EMU, Pedigree, Data Warehouse, Data Management

高速列车谱系化数据管理平台设计与构建

魏庆龙¹, 杨爱林¹, 王娜¹, 蒋咏志²

¹中车唐山机车车辆有限公司科技管理部, 北京

²西南交通大学牵引动力国家重点实验室, 四川 成都

Email: weiqinglong@tangche.com, yangailing@tangche.com, wangna@tangche.com

收稿日期: 2018年10月18日; 录用日期: 2018年10月31日; 发布日期: 2018年11月7日

摘要

高速列车谱系化数据仓库是面向主题的、集成的、时变的、相对稳定的，并集合了高速列车空间数据和非空间数据。如今的高速列车谱系化研制平台已经在谱系化理论的基础上，构建了虚拟可视化、快速工艺文件生成等八项子系统为核心的一体化平台。高速列车谱系化数据管理平台是该一体化平台流畅、高效运行的基础。本文基于高速列车谱系化相关理论研究，设计并构建了高速列车谱系化数据管理平台。系统阐述了平台的整体架构、模块组成以及各个模块的功能，界面和用途，为高速列车谱系化平台的运行和管理提供了基础支撑。

关键词

高速列车，谱系化，数据仓库，数据管理

Copyright © 2018 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 前言

数据仓库是面向主题的、集成的、稳定而随时间不断变化、用来支持管理人员决策的数据集合。它是近年来为联机分析处理(OLAP, Online Analytical Processing)、数据挖掘等处理提供海量数据存储、数据组织的容器和解决数据集成问题的关键技术。

经历引进消化吸收再创新之后，我国高速铁路已经在“四纵四横”的基础上，建成了具有中国特色的高速铁道网络[1]。在一带一路的倡导下，我国在现有高速动车组研制平台的基础上，构建了需求多样化，设计一体化，技术体系可定制化，标准体型国际化的高速动车组设计制造平台[2] [3] [4]。

高速列车谱系化数据仓库是面向主题的、集成的、时变的、相对稳定的，并集合了高速列车空间数据和非空间数据[5] [6]，它可以将多个异种的、自治的、分布的信息源有机的组织起来，并提供对空间数据和非空间数据简便、有效的访问，为决策分析提供数据[7] [8] [9] [10]。谱系化的研究最早开始于 60 至 70 年代，Sutherland 在系统中提出利用约束作为辅助手段生成零件，但缺少约束定义和修改几何模型，因此，模型的修改为一个单行过程，当模型生成之后约束无法反过来限制模型随谱系化理论不断完善，因此谱系化方法广泛运用于实践中。

目前，谱系化理论研究主要代表有两个：1) Jae Yeol Lee 提出的利用图表示的基于知识的几何推理法；2) Xiao Shan Gao 提出的约束传播法。国内这几年来对谱系化建模的研究也逐渐成熟，此类方法多利用软件的二次开发功能，实现对三维几何模型的简单化生成。中南大学的周咏翔、谭灿枚等以基于编程环境下的二次开发，结合数据库技术，对货车转向架进行了谱系化设计系统的开发。中北大学的史丽娱等提到在软件中建模时，要对零件模型的几何约束和尺寸约束进行定义。随着高速列车快速发展，谱系化理论成为解决高速列车设计中需求多样化重要的工具。

文献[11]已经对高速列车谱系化数据仓库系统研究进行了系统的阐述，本文基于高速列车谱系化相关理论研究和数据仓库系统的研究，设计并构建了高速列车谱系化数据管理平台。文章系统阐述了平台的整体架构、模块组成以及各个模块的功能，界面和用途，为高速列车谱系化平台的运行和管理提供了基础支撑。

2. 高速列车谱系化数据管理平台架构

高速列车数据内容繁杂、涉及领域广泛、数据量庞大，目前仍是以分散、多源、异构、语义不一致等存在，没有达到为决策支持提供有效数据源的水平，因此结合谱系化高速列车平台采用数据仓库技术、面向对象思想、空间数据库等关键技术将现有的多域高速列车数据有机地集成起来，并发挥其应有的共享作用，进行相应的 OLAP 分析和数据挖掘等处理，最终以友好的管理界面显示给数据管理者。为此需要开发相应的数据管理平台用以支持高速列车谱系化一体化设计平台。

数据管理平台的基本架构如图 1 所示，主要包括数据管理可视化界面和数据管理功能模块，功能模块包括需求数据管理模块、需求元模型数据管理模块、产品结构树数据管理模块、产品元模型管理模块、映射规则数据管理模块、指标体系数据管理模块、产品生命周期数据管理模块等。数据管理者通过可视化界面操作各功能模块，实现对各类数据的管理操作，包括数据的新增项目、数据的录入、错误数据清理、数据备份等管理和维护工作。

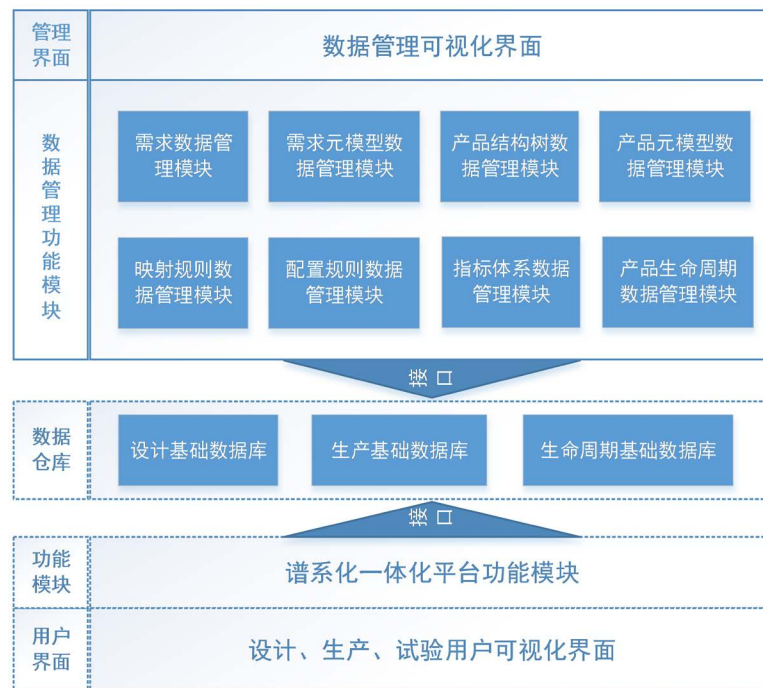


Figure 1. Basic structure of the data management platform

图 1. 数据管理平台的基本架构

这些功能模块的所有操作都通过接口直接影响数据仓库中的基础数据，包括设计、生产以及产品整个生命周期相关的基础数据，这些数据是谱系化一体化平台功能模块的基础，支撑整个一体化平台的运作。服务与谱系化平台的最终用户，如设计师、生成人员、试验人员等。

数据安全对于企业尤其重要性，因此数据管理者需要有足够的权限才可进行管理操作。

3. 高速列车谱系化数据仓库接口需求

数据仓库是面向主题的、集成的、稳定而随时间不断变化、用来支持管理人员决策的数据集合。同时数据仓库也是联机分析处理(OLAP, Online Analytical Processing)、数据挖掘等处理提供海量数据存储、数据组织的容器和解决数据集成问题的关键技术。由于数据仓库是数据管理平台的操作对象，因此需要

具有友好的、完整的接口。

数据仓库系统的一般体系结构如图 2 所示。

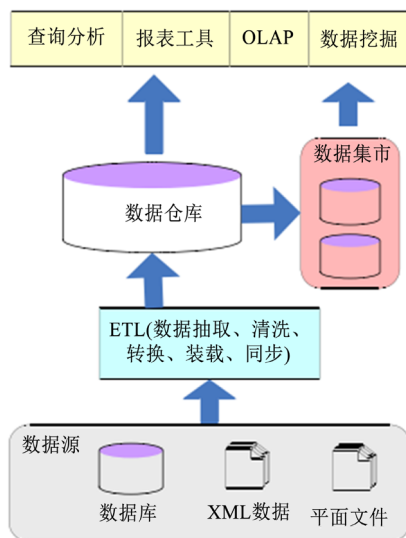


Figure 2. General architecture of the data warehouse system

图 2. 数据仓库系统的一般体系结构

数据仓库系统的体系结构一般包括四个部分：

1) 底层多种数据源，一般情况下，这些数据源可以是关系数据库或其他如平面文件、XML 文档等在内的数据类型。

2) 各种数据源通过 ETL (Extract-Transform-Load)，即数据抽取(Extract)、转换(Transform)、清洗(Cleansing)、装载(Load)。谱系化高速列车数据源众多，ETL 是保证数据一致性，实现跨多平台、多系统整合数据库，最大可能提高数据质量的关键技术。ETL 按照统一的规则从数据源抽取所需的数据，经过数据清洗和转换，其间可能还要进行过滤(Filtering)、汇总(Summary)，或与其它数据源的信息合并(Merging)，最终按照预先定义好的数据仓库模型，将数据加载到数据仓库中，是数据仓库实施成败的关键。

3) 经过 ETL 处理的数据进入企业级数据仓库，企业相关部门，如设计、制造、试验等可以通过数据集市(Data marts)，建立满足特定需求的数据存储，形成一个小型的部门级数据仓库。

4) 对数据仓库或数据集市应用多种数据处理技术，如查询分析、报表分析、OLAP 和数据挖掘，以多维度方式查询和分析数据，通过钻取、旋转、切片等操作将结果表现为综合性的决策信息，有效地支持大规模数据分析及统计计算，并以一种直观且易懂的形式提供给决策人员，对决策提供参考和支持。

谱系化高速列车数据仓库包含的基础数据库有需求数据库、需求元模型数据库、产品结构树数据库、产品元模型数据库、高速列车映射规则数据库、高速列车指标体系数据库、配置规则数据库和生命周期数据库等。这些数据库与数据管理平台的功能模块相对应，如图 3 所示。

数据管理平台需要对这些数据库进行管理和维护，因此数据仓库系统需求具备以下接口和功能，便于数据管理平台对数据库进行操作：

1) 数据导入接口。

谱系化高速列车数据仓库的数据信息来源于谱系化高速列车在设计、制造、试验、运行、检修、维护等涉及高速列车全生命周期各环节中且在承载、走行、动力及其它组份模块等各组份模块中产生的海

量数据，这些数据需按照相应的接口，以 XML 中性文件的形式导入到数据仓库中，并以标准化的方式进行存储。

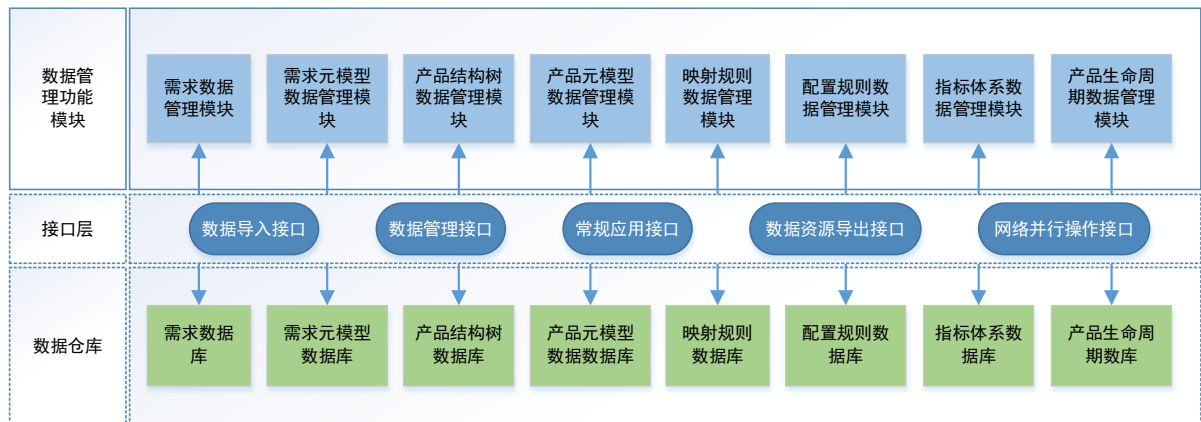


Figure 3. Data warehouse and data management platform interface correspondence and interface requirements

图 3. 数据仓库与数据管理平台接口对应关系及接口需求

2) 数据管理接口。

数据信息导入之后，用户可以从多方面对数据进行管理，主要包括设置数据信息、决策环节和分析模型，以及设置高速列车及其组份模块的实例对象与元模型的映射规则。

3) 常规利用接口。

系统需要为数据资源的利用实现部分常规利用功能，主要包括数据信息的浏览、查询和图形显示等；并在此基础上集成部分数据挖掘方法如聚类算法、多元线性回归算法等方法。

4) 数据资源导出接口。

若用户需对数据信息进行深入利用，可将其导出到其他分析工具。系统需要依据用户使用的主要分析软件的数据标准编制相应的信息接口，便于将信息资源按特定格式导出，如 Excel 文件或 CSV 文件。

5) 支持多用户网络操作并具备权限设置。

为了让系统使用用户有效地获取和利用信息资源，系统用户应可同时通过网络进行数据信息利用；同时系统用户只能访问具有权限的数据资源和分析模型。

4. 数据管理平台模块功能及界面设计

4.1. 需求数据管理模块

需求数据管理模块主要是对高速列车相关的需求数据进行构建和处理，具体的界面如图 4 所示。需求数据管理界面是用列表显示的所录入的需求数据名称字典列表。

在需求名称构建完成的基础上，对需求数据进行属性化的分类，包括全面质量控制(TQC)需求、主体需求、关键需求和旅客需求等。同时对需求数据进行初始信息的录入，包括了需求的初始值、单位、值类型等。将构建好的属性化的需求数据通过相应的可视化方式进行表达，包括了树形式以及列表形式的显示等。并且可以对需求属性、需求数据等进行添加、修改、删除以及更新等操作。如图 5 所示。

4.2. 需求元模型数据管理模块

需求元模型管理主要的功能是将构建好的需求数据与需求结构树进行关联，将需求数据进行结构化分解得到相应的高速列车需求元模型。然后通过一定的计算机显示方法对需求元模型数据进行可视化的

展示，包括了基于结构树的树状形式显示以及数据列表的显示等。主要功能包括需求名称字典管理、需求数据全集管理、需求元模型构建等。通过结构树中的结构查看出与之相关的所有需求数据信息。对于动车组而言需求元模型主要包括走行组份、动力组份、承载组份和其他组份，每个组份都会按照 TQC 需求、主体需求、关键需求以及旅客需求等分类。通过选择需求全集的节点，如选择组成新设计动车组的需求全集。



Figure 4. Demand data name dictionary list interface

图 4. 需求数据名称字典列表界面



Figure 5. Requirement data corpus visualization list interface

图 5. 需求数据全集可视化列表界面

在构建好需求元模型的需求结构节点，对其进行树状以及列表形式的显示。点击编码、名称、匹配类型等列表名，可以对需求数据列表进行修改、删除等动态管理操作。提供包括对需求数据进行添加、修改、删除以及更新等操作，具体的软件界面如图 6 所示。

序号	需求项名称	编码	匹配类型	取值	单位	数据类型	取值条件	操作
1	最高运营速度	001	等于	250	km/h	整数	0	编辑 删除
9	联挂运行时通过最小曲线半径	009	等于	180	m	整数	0	编辑 删除
10	单车调车运行时通过最小曲线半径	010	等于	130	m	布尔	0	编辑 删除
11	轴重	011	小于等于	14	t	小数	0	编辑 删除
15	环境温度最低值	015	等于	0	°C	整数	0	编辑 删除
17	海拔高度	016	小于等于	1500	m	整数	0	编辑 删除
21	脱轨系数		小于等于	0.8		小数		编辑 删除
22	准静态时轮重减载率		小于等于	0.65		小数		编辑 删除
23	动态时轮重减载率		小于等于	0.8		小数		编辑 删除
71	舒适性		小于	2		小数		编辑 删除
72	运行平稳性		小于等于	2.5		布尔		编辑 删除
73	相对湿度	004	小于等于	0		字符		编辑 删除
74	限界标准		等于	GB146.1		字符		编辑 删除
78	轴距	sss	等于	1435	mm	小数		编辑 删除
79	坡度		等于	30‰		字符		编辑 删除
80	欠超高		等于	90	mm	小数		编辑 删除

Figure 6. Demand meta model visualization interface
图 6. 需求元模型可视化界面

需求元模型应保证需求库的完备性，因此元模型管理模块可以方便实现需求元模型的扩展，同时在增加新需求元模型时可以进行需求冲突或包含关系验证，避免需求项之间的矛盾或冗余。

4.3. 产品结构树数据管理模块

产品结构树数据管理模块包含三个主要功能，产品结构层级管理、产品结构名称管理和产品结构全集。针对现有的高速列车实例进行分析，收集和归纳出相关的产品结构层级和名称，保证此类数据的唯一性，将这些结构数据作为字典进行管理并约束后续的与之相关的各类数据。基于产品结构层级和名称构建出产品结构树的全集，并提供结构树数据的添加、修改、删除、查询等动态管理操作。

产品结构树按照不同层级划分，包括列车级、车辆级、组份级、部件级、子部件级和零件级。产品结构树不同层级之间是包含和隶属关系。点击绿色新增按钮，即进入到结构层级的增加界面，提供给用户适时增加产品结构层级名称信息的功能，点击某一个产品层级名称编辑按钮，在此可以对此产品数据名称进行重新填写。

产品结构名称是为了统一产品或部件在整个设计、生产和运用过程中的名称。在主界面中点击产品结构名称管理按钮，即进入到产品结构名称管理界面。每个名称都有特定的编号、所属层级和在所属层级中的序号。

高速动车组产品包含的所有组份、部件、子部件和零件构成了产品结构全集。在进入产品结构全集界面后，可以看到一个高速列车产品的产品结构树，选定某一项后可以调整其包含和隶属关系。针对一个具体的产品，如果要增加一个子部件，如二系垂向减振器。可以首先在产品结构名称管理中增加一项二系垂向减振器，设置其层级为子部件级。在产品结构全集中将其挂靠到二系悬挂部件级项目下。

4.4. 产品元模型数据管理模块

产品元模型数据管理模块包含三部分，产品数据项名称管理、产品数据项全集和产品元模型构建。

进入产品数据名称管理界面后，首先所看到的就是用列表显示的所录入的产品数据名称字典列表。可以进行增加新数据项或编辑和修改已有数据项。

在产品数据名称构建完成的基础上，对产品数据进行属性化的分类，同时对产品数据进行初始信息的录入，包括了产品的初始值、单位、值类型等。将构建好的属性化的产品数据通过相应的可视化方式进行表达，包括了树形式以及列表形式的显示等。并且可以对产品属性、产品数据等进行添加、修改、删除以及更新等操作。

将构建好的产品数据与产品结构树进行关联，将产品数据进行结构化分解得到相应的高速列车产品元模型。然后通过一定的计算机显示方法对产品元模型数据进行可视化的展示，包括了基于结构树的树状形式显示以及数据列表的显示等。通过结构树中的结构查看出与之相关的所有产品数据信息。提供包括对产品数据进行添加、修改、删除以及更新等操作。

4.5. 映射规则数据管理模块

映射规则是从需求到技术指标自动转化的核心，通过映射规则将描述性的、模糊的需求转化成可实施的、具体的和可实现的技术指标，作为产品设计的约束边界。映射关系的类型分为一对一映射、一对多映射、多对一映射和多对多映射等。

1) 映射规则输入输出

映射规则是以需求作为输入，以技术指标作为输出的。在添加映射规则的界面中，提供给用户勾选需求数据作为输入，并填写信息。要添加技术指标输出，进入到映射规则技术指标添加界面，用户勾选技术指标数据作为输出项，并填写信息。

2) 映射规则表达

映射规则的类型分为直接映射、函数映射和知识映射。直接映射和函数映射相对简单，知识映射往往包含设计人员的经验知识。因此映射规则本质上是将设计人员的经验固化到数据管理平台的过程。映射规则会采用不同的标签按照前述的三种类型进行分类显示。在映射规则构建完成的基础上，将其通过列表形式进行表达，能够清晰的表示出映射规则信息包括了输入输出项，以及规则的本身信息。结合产品结构树，可以方便的通过结构树中的结构查看出与之相关的所有映射规则信息。对映射规则数据、以及数据中输入输出项、规则信息等进行添加、修改、删除以及更新等操作。

4.6. 配置规则管理模块

配置规则管理包括模板级选配参数管理、实例级选配参数管理、权重配置、选配参数赋值规则构建和选配参数赋值规则管理等。

模板级选配参数管理是对产品结构树全集进行配置，选中其中的某个产品结构节点后，显示出其相关的产品元模型数据信息，在此基础上可以进行选配参数的勾选，同时可以对选配参数的信息进行编辑，包括了重要性，选配的原则等。

实例级选配参数管理是对项目实例的全集进行配置，选中某个项目中的某个产品结构节点后，可以对其相关的实例级别的产品元模型的参数进行选配。

根据选配参数管理确定好选配参数后，对其进行赋值规则的构建，将选配参数作为目标参数，通过与其有关联的其他产品数据对其进行赋值。并对赋值规则的信息进行填写，包括了编码、名称，输入项等信息。并且对选配参数赋值规则进行增加、删除、修改功能。

4.7. 生命周期数据管理模块

为了是动车组设计形成闭环，实现可持续的优化迭代，需要掌握产品从生产完成和实际运用、维护

等所有过程监测到的数据。包括台架试验数据、滚振试验数据、模态试验数据、线路跟踪试验数据、车轮踏面演化数据、故障数据、维修数据等等。

5. 结论

本文对高速列车谱系化数据仓库系统研究进行了阐述，对其中的谱系化数据管理平台的系统架构进行了阐述和说明，对高速列车谱系化数据仓库接口需求进行论述，并提出了满足需求的数据管理平台功能模块，分别说明了其中的各个模块的功能，界面和用途。高速列车谱系化数据管理平台的研究、设计与开发，为高速列车谱系化平台的运行和管理提供了基础支撑。

参考文献

- [1] 铁路局完成铁路“十三五”发展规划[J]. 城市规划通讯, 2015(2): 13.
- [2] 都青华, 邢海英, 李瑞淳, 张中云. 高速动车组谱系化技术研究平台初探[J]. 兰州交通大学学报, 2014, 33(6): 141-144.
- [3] 张卫华. 高速列车顶层设计指标研究[J]. 铁道学报, 2012, 34(9): 15-19.
- [4] 谢宇迪. 谱系化动车组转向架 CAD/CAE 集成分析的研究[D]: [硕士学位论文]. 北京: 北京交通大学, 2016.
- [5] 吴文龙. 谱系化动车组转向架构架参数化建模及分析[D]: [硕士学位论文]. 北京: 北京交通大学, 2015.
- [6] 李陆星. 谱系化动车组车体参数化建模及分析[D]: [硕士学位论文]. 北京: 北京交通大学, 2014.
- [7] 郑建智, 段占祺, 应桂英. 数据仓库和 OLAP 技术在卫生统计决策支持系统中的应用[J]. 中国卫生信息管理杂志, 2012, 9(3): 47-51.
- [8] 宋旭东, 刘晓冰. 数据仓库 ETL 任务调度模型研究[J]. 控制与决策, 2011, 26(2): 271-275.
- [9] 吕海燕, 车晓伟. 数据仓库中数据粒度的划分[J]. 计算机工程与设计, 2009, 30(9): 2323-2325, 2328.
- [10] 赵虹. 动车组维修决策系统数据仓库的研究与应用[D]: [硕士学位论文]. 北京: 北京交通大学, 2009.
- [11] 王光新, 高浩. 高速列车谱系化数据仓库系统研究[J]. 交通科技, 2018, 7(2): 79-90.

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2326-3431, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>
期刊邮箱: ojtt@hanspub.org