

# 高校通用数据超市平台的设计与实现

王永超 来天平 罗 盘 高志同

北京大学计算中心 北京 100871

**摘 要** 针对高校对数据访问的多样性需求,在已有数据中心的基础上,研究开发了具备良好通用性、灵活性和安全性的数据超市平台,设计了基于 RBAC 权限模型的细粒度数据权限划分模式,实现了可灵活配置的数据组装技术。该平台具备快速响应复杂数据访问需求的能力,是对高校数据中心在应用层面的扩展,也是对提高数据利用率、发挥数据中心应用价值的重要实践。

**关键词:** 数据超市;权限模型;数据组装;角色管理;数据中心

**中图法分类号** TP315

## Design and Implementation of Universal Data Supermarket Platform in Universities

WANG Yong-chao, LAI Tian-ping, LUO Pan and GAO Zhi-tong

Computer Center, Peking University, Beijing 100871, China

**Abstract** According to the diverse requirement of data access within universities, based on the existing data center, this paper researchs and develops a data supermarket platform with good versatility, flexibility and security, designs a finegrained data permission division pattern based on the RBAC permission model, and realizes a flexibly configurable data assembly technology. This platform has the ability to quickly respond to the complex data access requirement. It is an extension of the data center of universities at the application level and it is also an important practice for improving data utilization and exerting data center application value.

**Keywords** Data supermarket, Permission model, Data assembly, Role management, Data center

## 1 引言

随着高校信息化的持续发展,高校内部建成的管理信息系统越来越多,由此积累了海量的基础数据和业务数据。为充分利用这些数据,各高校积极投入建设数据中心,将分散在各业务系统的数据进行集中存储管理<sup>[1]</sup>。解决数据集中存储管理的问题后,各高校又面临如何将数据中心的数据安全、便捷地共享给用户和业务系统的新挑战。

北京大学数据中心已完成建设并投入运行,可完成各类数据的转换、聚合、交换和集中存储;同时,基于数据中心,学校也开发了数据统计分析和数据分发的应用<sup>[2]</sup>。但随着高校信息化建设的不断深入,学校各级管理人员和各类业务系统对数据访问的需求日益强烈,对数据共享多样性的要求也越来越高,已有应用已无法满足当前需要,因此亟需建立一个可支持复杂需求的通用数据超市平台。

鉴于此,本文提出了通用数据超市平台的建设方案,同时对如何解决数据共享过程中的通用性和灵活性等具体难点进行了深入阐述。本文的研究内容和主要贡献包括:分析数据访问需求并分类;提出对数据再次结构化的数据组装技术;设计数据权限的多层次、细粒度的划分模式;建立数据申请—数据组装—数据发布—数据访问的工作流程。

通信作者:王永超(wangyc629@pku.edu.cn)

## 2 需求分析

从数据访问需求方面对校内各部门进行调研发现,校内各部门之间和各业务系统之间存在数据信息壁垒,各级管理单位遇到跨部门的数据统计和查询需求时,不但无法确认数据是否存在,还缺乏统一的数据申请和访问平台。因此,通用数据超市平台须满足统一数据申请、统一数据发布、统一数据检索和统一数据访问方式的要求,为数据访问提供全流程的支持。另一方面,通过对调研的总结,本文将数据访问需求分为管理需求和接口需求两种类别。

1)管理需求:管理需求是学校各部门管理人员对数据进行查询和统计的需求,该类需求的特点是需求提出频繁、查询口径多变、访问控制复杂。以往,当管理过程提出新的数据访问需求时,业务系统需要定制开发,这种模式费时费力,无法及时响应需求。

2)接口需求:接口需求来自校内业务系统接入外部数据的困难。校内业务系统在过去的开发过程中存在无法快速接入外部数据的问题,数据接入时需要对接不同业务系统不同类型的接口,而且在对接过程中还需要不同业务系统开发人员的支持。

在解决数据访问需求的同时,考虑当前数据标准文档在共享过程中存在的困难,数据超市平台须同时提供对数据标准文档的统一管理和发布功能。

为保证数据超市平台数据的安全性和通用性,系统需要通过人员角色对权限进行严格控制,在数据超市平台内部,权限需要分为功能权限和资源权限两种类型。功能权限用于限制用户对平台功能模块的访问权限;资源权限用于配置用户对数据的访问权限。

除此之外,数据超市平台还须关注非功能性需求。非功能性需求包括系统安全性、功能易用性和界面美观性。系统安全性要求平台采用安全的身份认证策略,具备完善的日志审计能力。功能易用性和界面美观性要求平台在建设时更多地考虑用户体验、优化页面交互设计。

基于上述描述,建设通用数据超市平台有如下要求:

1)具备完善、灵活的数据管理模块,支持数据申请和数据发布;

2)具备通用的数据访问功能,能够解决管理过程中复杂的数据查询需求,能够为业务系统提供统一的数据访问接口;

3)具备数据标准文档的管理和发布功能;

4)具备严格的权限控制机制和复杂的数据权限模型,支持数据权限的细粒度划分;

5)支持完善的角色管理能力,可对角色权限进行灵活的配置;

6)具备良好的系统安全性、功能易用性和界面美观性。

## 3 系统设计

### 3.1 平台架构

通过需求分析,可得到如图 1 所示的系统架构图。数据超市平台的主要功能模块为:系统管理模块、日志审计模块、角色管理模块、权限管理模块、数据管理和共享模块和数据标准文档管理和发布模块。其中,角色管理模块、权限管理模块、系统管理模块和日志审计模块是数据超市平台的基础模块,为其他功能模块提供基础功能的支持。除内部功能模块之外,数据超市平台还借助 IAAA(北京大学统一身份认证系统)完成用户登录和身份认证的功能。

系统管理模块主要对数据超市平台进行基础设置,支持发布系统公告,动态配置前台提示内容。

日志审计模块负责记录系统日志,并支持对系统日志进行复杂查询。

角色管理模块与权限管理模块之间具有相互依存的关系,共同构建了用户-角色-权限的权限控制机制。角色管理模块提供创建角色、用户授权和角色汇总查询等功能;权限管理模块则提供对功能权限和资源权限的维护、角色授权和权限多维查询等功能。

数据标准文档管理和发布模块负责在系统内配置和发布数据标准文档,数据标准文档发布后,已授权的用户可在系统内检索和下载文档。

数据管理和共享模块是数据超市平台的核心模块,该模块提供数据管理和数据共享两类功能。数据管理提供申请数据、组装数据、发布数据和多维度配置数据权限等功能;数据共享提供平台内数据检索、数据查询统计、嵌入式数据查询页面和数据接口等功能。

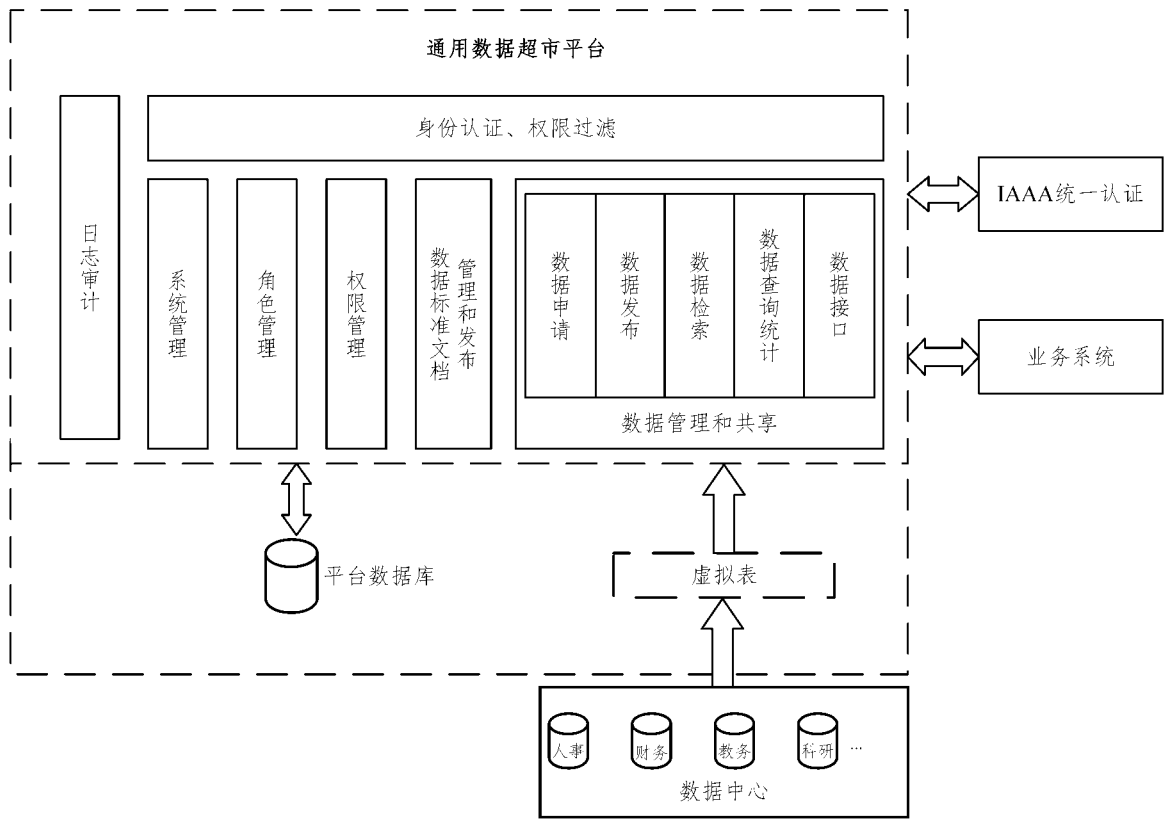


图1 数据超市平台架构

Fig.1 Architecture of data supermarket platform

3.2 角色管理与权限模型

数据超市的通用性基于灵活的角色管理和权限模型。图2展示了系统内基于RBAC权限模型<sup>[3]</sup>的用户-角色-权限的层次关系。在系统内部,用户不会被直接授予某一权限,所有权限都需要通过角色与用户进行关联,该模式可避免系统长时间运行后权限混乱的问题。系统内部的权限模型较为复杂,总体上权限被划分为功能权限和资源权限两大类,每类权限又有更细粒度的划分。

3.2.1 角色管理

系统设计了灵活的角色管理功能,且默认只创建Admin和User两个角色,User角色作为默认角色分配给所有登录用户,User角色登录系统后只能查看系统使用说明和系统公告;Admin角色作为系统管理员具有数据超市平台的全部权限,可以创建角色,为角色配置权限,为用户授予角色。为方便用户使用,对于具有多个角色的用户,系统会自动汇总所有角色的权限,用户登录后无需切换角色可获取不同权限。

3.2.2 功能权限模型

功能权限包含页面访问权限和接口调用权限,

一组页面的权限和一组接口的权限可以构成一个功能的权限。系统的每个页面都会作为一个权限,例如登录系统后的首页即表示一个页面权限;同时,每个页面对应的后台接口也会被作为权限,为减小授权的复杂度,系统支持对接口权限进行分组,例如数据标准文档相关的接口可被划分为文档维护接口权限(如新建、删除等接口)和文档查询接口权限(如检索、文档下载等接口)。

3.2.3 资源权限模型

资源权限包含数据的权限和数据标准文档的权限,其中数据的权限又包含数据访问权限、显示字段权限、查询字段权限和行权限,数据权限的细粒度划分增加了数据共享过程中的灵活性。对于具体角色而言,数据访问权限表示数据对该角色是否可见;显示字段权限表示对数据进行查询时数据提供哪些数据列;查询字段权限表示数据的哪些字段可以被作为检索条件;行权限则表示如何对数据进行过滤。

对于数据的行权限,系统更详细地划分为基于角色的行权限<sup>[4]</sup>、基于用户属性的行权限和基于用户行为的行权限。为解释上述3种行权限,假设有学生基本信息数据,该数据包含学生类别、

院系和专业等字段。在共享上述数据时,对于研究生院管理人员,需要为其增加学生类别为研究生的过滤条件,该过滤条件即为基于角色的行权限;对于院系教务员,则需要增加学生院系为教务员所在院系的过滤条件,由于教务员所在院系为

教务员的固有属性,因此该过滤条件被称为基于用户属性的行权限;若某一教务员需要被限制只能查询某一专业的学生数据,则需要为该教务员增加该专业的过滤条件,该过滤条件则称为基于用户行为的行权限。

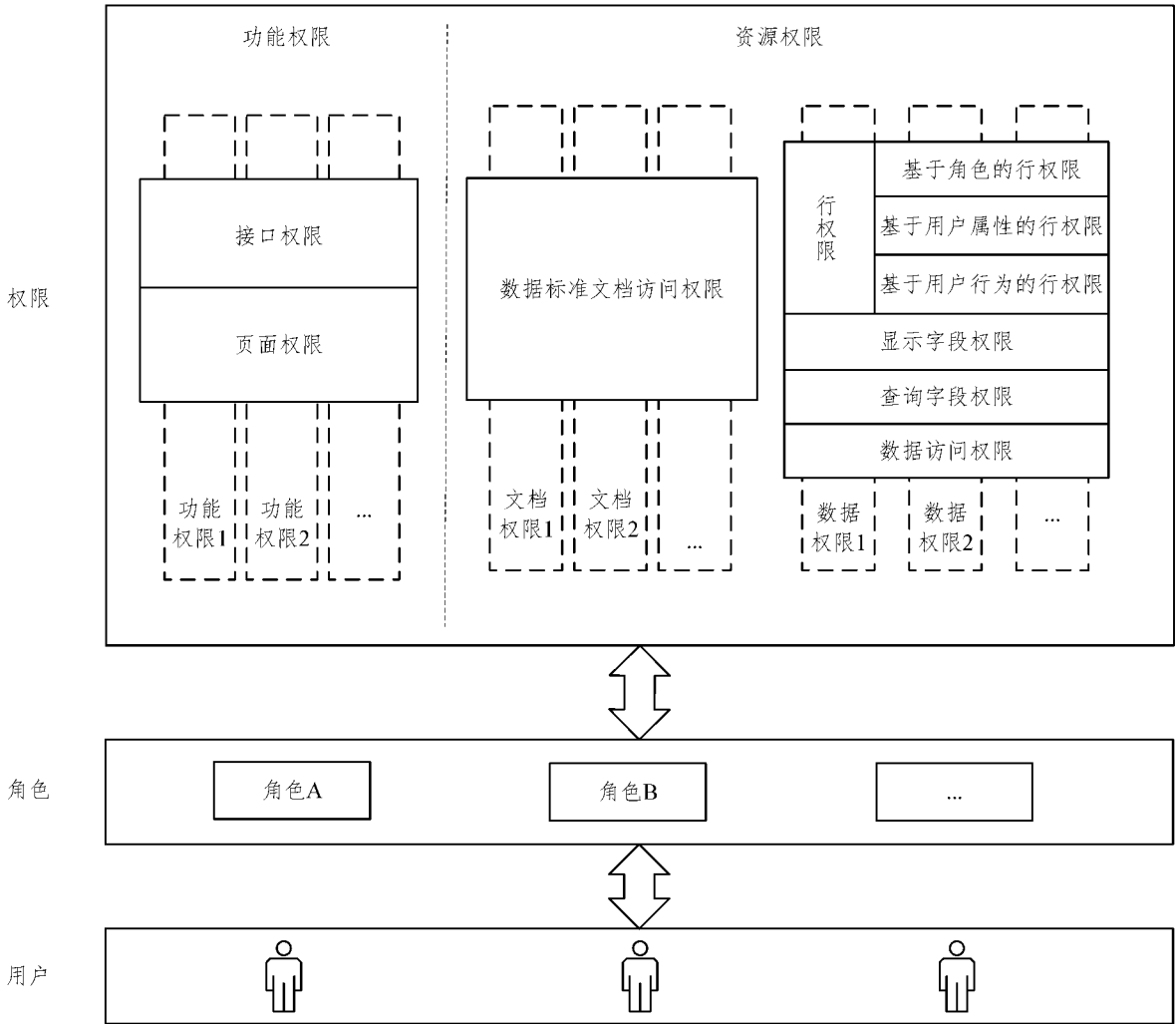


图 2 角色与权限模型

Fig. 2 Roles and the module of permissions

3.3 数据发布与数据组装

数据发布是数据超市平台的核心功能,一个完整的数据发布流程如下:用户提出数据访问需求;用户在数据超市平台检索数据,确认数据是否已提供,若尚未提供则发起数据申请;系统管理员根据申请信息进行需求分析和无权限过滤的数据检索,若申请数据不存在,则结合数据中心构建符合需求的虚拟表;具有数据管理权限的管理员在系统内组装数据,组装数据时可以在虚拟表的基础上定义数据的基本信息,显示字段的属性和查询字段的属性;完成数据组装后,具有数据管理权限的管理员对数据进

行各个维度的赋权;最后,已授权用户可以通过页面查询或数据接口进行数据访问。

基于数据中心的虚拟表是快速响应数据访问需求的基础,虚拟表实际上是在数据中心基础数据上创建的视图,在创建视图时管理员可以对基础数据进行自由关联、聚合和统计。

数据组装是保证系统灵活性和通用性的关键技术,数据组装实质上是对虚拟表在系统内的再次结构化。通过数据组装可以重新定义数据的结构,使数据的结构对系统用户可见,以此来支持字段级别的权限分配。在数据组装时还支持对字段进行属性

设置,以此来优化数据查询页面的显示,字段属性分为显示字段属性和查询字段属性,显示字段属性包括字段中文名、字段显示顺序、字段显示形式、字段显示宽度等用于数据显示的属性;查询字段属性包含查询框标题、查询框显示顺序、查询框显示宽度以及查询框下拉数据类型等属性。

对于管理人员的数据查询需求,系统提供了两种实现模式。第一种是提供数据超市平台内部的查询页面,用户登录系统后可以在系统内访问数据;第

二种是为业务系统提供嵌入式的数据查询页面,业务系统可以快速将数据查询页面集成到系统内部,用户在业务系统内即可访问需要的数据。

4 系统实现

数据超市平台采用了目前较为流行的前后端分离的开发模式,系统前台使用了 Vue + ElementUI 的开发框架,系统后台则使用了 Spring Boot + MyBatis 的开发框架。系统实现框架如图 3 所示。

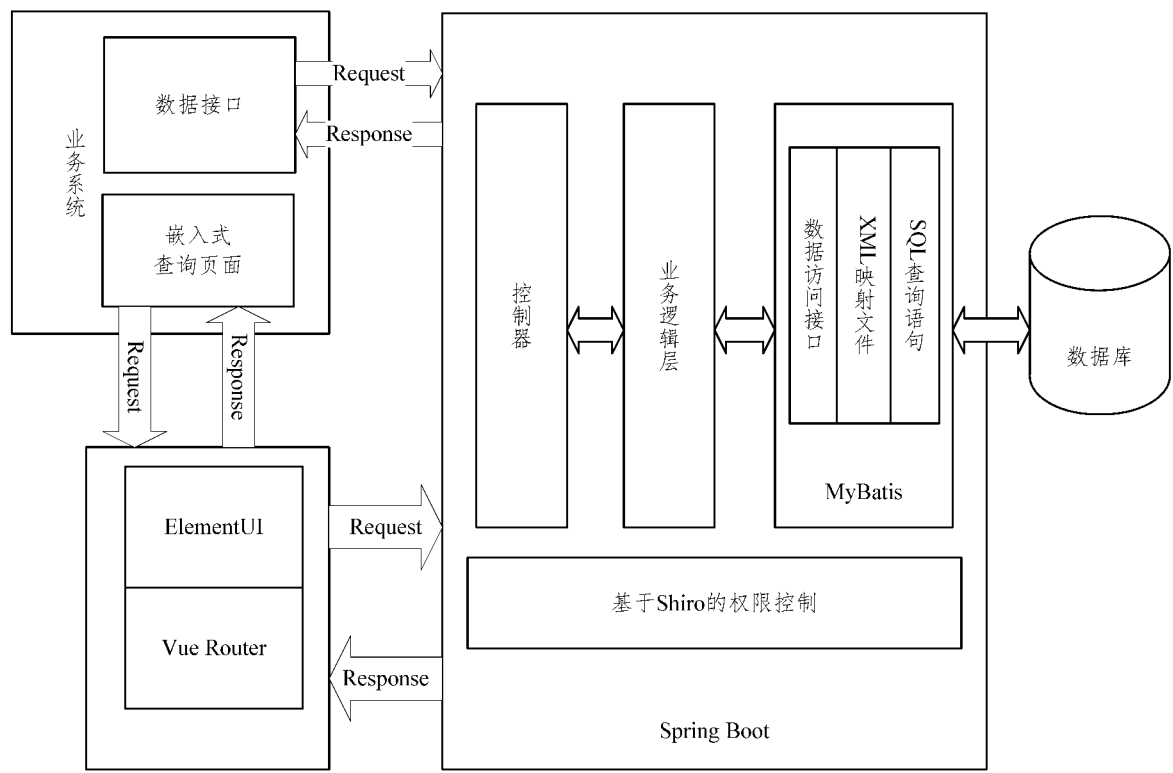


图 3 系统实现框架图

Fig. 3 Diagram of system implementation framework

Spring Boot 是对 Spring 开发框架的进一步封装,该框架基于“约定优于配置”的原则,将开发者从繁琐的配置工作中解脱出来,提高了开发者的开发效率<sup>[5]</sup>。

MyBatis 是一个灵活的持久层框架,集成了自定义 SQL 查询、存储过程以及高级映射等功能。MyBatis 提供注解和 XML 映射文件两种方式编写 SQL 查询语句<sup>[6]</sup>。

系统前台采用了基于 Vue 的 ElementUI 来构建 Web 页面,采用响应式布局方式,同时支持 PC 端和移动端。Vue 是一套用于构建用户界面的渐进式框架<sup>[7]</sup>,采用了 MVVM (Model-View-ViewModel) 的模式,在 MVVM 的架构下,Vue 通过双向数据绑

定将数据和页面关联起来,使页面与数据的变动可以响应式地自动同步。ElementUI 是基于 Vue 开发的组件库,具有设计美观、组件丰富、主题可配置等特点,通过应用 ElementUI,可以减少开发者在组件设计和页面美化方面花费的精力,将开发者的关注点放在业务逻辑上。

区别于传统的 MVC 模式,基于前后端分离开发模式的系统后台不再关注视图层,仅提供 RESTful API。后台请求由 Spring Boot 控制处理过程,处理请求时,首先需要通过 Shiro 过滤器来完成认证和鉴权,之后根据 URL 映射将请求交给对应的控制器,再由控制器分发给业务逻辑层来完成业务处理。请求处理过程中的数据访问则交给 MyBatis 来控

制,MyBatis 通过数据访问接口-XML 映射文件-SQL 查询语句的三级关系来关联 Java 代码与 SQL 语句。

**结束语** 针对当前复杂的数据访问需求,本文通过严谨的需求分析、科学的架构设计和合理的技术选型,提出了通用数据超市平台的建设方案。系统提供了多种形式的数据库访问方式,统一整合了数据申请、数据发布、数据检索和数据查询等功能,具有配置灵活、需求响应快速等特点。在实现需求的同时,系统也对数据权限模型进行了深入研究,支持数据权限的细粒度划分,同时兼顾灵活性和安全性。在交互设计和页面美化方面,本系统采用了当前流行的开发框架,提高了系统的易用性和美观度。通用数据超市平台是对高校数据中心在应用层面的实践和探索,基于其良好的通用性和灵活性,对高校范围内促进数据共享具有一定参考意义。

### 参 考 文 献

- [1] ZHAO W,LIU D,WANG H,Design and Realization of Shared Data Center in the University[J]. Journal of Changchun University of Science and Technology(Natural Science Edition),2015(1):132-135. (in Chinese)  
赵巍,刘丹,王欢. 高等学校共享式数据中心设计与实现[J]. 长春理工大学学报(自然科学版),2015(1):132-135.
- [2] LONG X Z,LI L,PENG Y M,et al, Research of university data statistical platform based on data warehouse [J]. Journal on Communications, 2013 (S2): 163-169. (in Chinese)  
龙新征,李丽,彭一明,等. 基于数据仓库的高校数据统计服务平台研究[J]. 通信学报,2013(S2):163-169.
- [3] FERRALOLO D F. Role-based access control (RBAC): Features and motivations[C]//Proc. 11th Annual Computer Security Applications Conference (ACSAC), New Orleans,1995.
- [4] OUYANG R B,LIU Y F, LONG X Z. Implementation of Parameter Model of PRBAC Based on Attribute Rules [J]. Journal of Zhengzhou University(Engineering Science),2017(2). (in Chinese)  
欧阳荣彬,刘云峰,龙新征. 基于属性规则的 PRBAC 参数模型研究与实现[J]. 郑州大学学报(工学版),2017(2).
- [5] Spring Boot Reference Documentation [EB/OL]. <https://docs.spring.io/spring-boot/docs/2.2.10.RELEASE/reference/html>.
- [6] MyBatis Introduction [EB/OL]. <https://mybatis.org/mybatis-3/index.html>.
- [7] Vue.js [EB/OL]. <https://cn.vuejs.org/v2/guide/index.html>.



**WANG Yong-chao**, born in 1991, postgraduate, engineer. His main research interests include educational informationization, network and database technology.