

**软件工程综合课程设计**

--开发手册

**题 目** 政府采购网中标数据采集工具

**院 系**  计算机科学与技术学院

**专 业**  软件工程

**学生姓名**高正东,温志豪,柳家胜,屈有昊,曾君宇

**指导教师** 张德平

**二**〇**一五 年 十二 月 三十 日**

目录

[第一章 引言 4](#_Toc439577887)

[1.1 编写目的 4](#_Toc439577888)

[1.2 项目背景 4](#_Toc439577889)

[1.3 开发工具 4](#_Toc439577890)

[1.4 参考资料 5](#_Toc439577891)

[第二章 总体设计 6](#_Toc439577892)

[2.1 需求概述 6](#_Toc439577893)

[2.2 软件结构 6](#_Toc439577894)

[第三章 总体设计 8](#_Toc439577895)

[3.1 模块 8](#_Toc439577896)

[3.1.1 登录模块 8](#_Toc439577897)

[3.1.2 系统管理模块 9](#_Toc439577898)

[3.1.3 网址配置模块 10](#_Toc439577899)

[3.1.4 数据采集模块 11](#_Toc439577900)

[3.1.5 数据查询模块 11](#_Toc439577901)

[3.6 统计分析模块 12](#_Toc439577902)

[3.1.7 接口 12](#_Toc439577903)

[3.2 功能 13](#_Toc439577904)

# 引言

## 1.1 编写目的

重庆市政府采购网中标数据采集工具详细设计说明书是设计的第二阶段，这个阶段的主要任务是在重庆市政府采购网中标数据采集工具概要设计说明书基础上，对概要设计产生的功能模块进行过程描述，设计功能模块的内部细节，包括算法和详细数据结构，为编写源代码提供必要的说明。

概要设计解决了软件系统总体结构设计的问题，包括整个软件系统的结构、模块划分、模块功能和模块间的联系等。详细设计则要解决如何实现各个模块的内部功能，即模块设计。具体的说，模块设计就是要为已经产生的重庆市政府采购网中标数据采集工具各子系统设计详细的算法。但这并不等同于系统实现阶段用具体的语言编码，它只是对实现细节作精确的描述，这样编码阶段就可以将详细设计中对功能实现的描述，直接翻译、转化为用某种程序设计语言书写的程序。

## 1.2 项目背景

随着大数据分析技术的兴起，数据成为大数据分析的重要基础，大数据的采集已成为直接影响大数据分析技术应用的瓶颈，政府采购网中标数据采集工具可用于对各省、市级政府采购网中标数据的采集，为企业景气指数分析、企业生产状况分析、企业融资风险分析以及股民股市投资提供分析数据，并在此基础上提取相应的辅助决策信息。

## 1.3 开发工具

下面将给出实际开发中所使用的平台和工具。

开发IDE：Eclipse for J2EE

WEB服务器：Apache Tomcat 7.2

JAVA：JDK 1.7

PYTHON: python 2.7

WEB框架：Struts2

操作系统：Windows 8/10

数据库：MySQL

开发语言：Java

版本管理工具：Git

## 1.4 参考资料

《软件工程（第三版）》 齐治昌等著 高等教育出版社.

# 第二章 总体设计

## 2.1 需求概述

**总体**：包括用户的登录、注销，以及相关权限的操作，如找回密码、注册等功能。

**对于普通用户**：包括在线查询数据采集结果、查询各公司中标情况等。

（需求1）对保存在数据库中的内容进行查询，包括：

1）按项目名称查询

2）按中标单位查询

3）按中标金额或中标金额范围查询

4）按中标时间或是中标时间范围查询

5）多项组合查询

（需求2）针对中标单位进行统计分析，包括：  
 1）可统计分析出指定中标单位历史中标金额数据趋势（直方图、曲线图表示）

2）可统计分析出各中标单位指定时间范围的中标金额汇总数据（表格给出）

**对于管理员**：爬取网址的管理，进行数据采集等功能。

（需求1）通过传入网址进行指定网站的抓取，其中包括多种网站配置方式，包括：手工输入、Excel表格批量导入、或Txt文档批量导入，导入后应在界面中可见，可修改、删除、增加等。

（需求2）数据采集部分，可以对已配置的网址进行指定，然后开启爬虫抓取。并且在爬取过程中，可显示爬取的当前网页和进度。

（需求3）将爬取的网站内容进行筛选，抽取有效地信息，主要包括：项目名称、中标单位、中标金额以及中标时间，并将抽取出的数据保存到数据库中。

## 2.2 软件结构绘图1

# 总体设计

## 3.1 模块

## 3.1.1 登录模块

|  |  |
| --- | --- |
| **用例名：** | login.jsp |
| **位置：** | 根目录/WebRoot |
| **描 述：** | 用户登录 |
| **一般事件过程：** | 将表单信息提交到LoginServlet进行数据库的查询验证操作 |
| **其它事件过程：** | 如果登陆信息不存在，则重新登陆 |
| **前导状态/条件：** |  |
| **完成后的状态：** | 进入index.jsp选择操作 |

## 3.1.2 系统管理模块

|  |  |
| --- | --- |
| **用例名：** | modify\_pw.jsp |
| **位置：** | 根目录/WebRoot |
| **描 述：** | 密码修改 |
| **一般事件过程：** | 将表单信息提交到ModifyPassWordServlet.java进行数据库的查询验证操作 |
| **其它事件过程：** | 如果密码格式错误，重新输入 |
| **前导状态/条件：** | 用户处于登录状态 |
| **完成后的状态：** | 进入index.jsp选择操作 |

## 3.1.3 网址配置模块

|  |  |
| --- | --- |
| **用例名：** | webAddressConfig.jsp |
| **位置：** | 根目录/WebRoot |
| **描 述：** | 网址配置 |
| **一般事件过程：** | 将表单信息提交到ReadExcelServlet.java、ReadTxtServlet.java，显示文件内容 |
| **其它事件过程：** | 文件不存在，重新输入 |
| **前导状态/条件：** | 管理员登录状态 |
| **完成后的状态：** | 进入webAddressConfig.jsp,显示网址信息 |

## 3.1.4 数据采集模块

|  |  |
| --- | --- |
| **用例名：** | dataAcquisition.jsp |
| **位置：** | 根目录/WebRoot |
| **描 述：** | 数据采集 |
| **一般事件过程：** | 显示爬取的网址和中标信息 |
| **其它事件过程：** | 若网址配置错误，显示NULL |
| **前导状态/条件：** | 管理员登录状态 |
| **完成后的状态：** | dataAcquisition.jsp显示信息 |

## 3.1.5 数据查询模块

|  |  |
| --- | --- |
| **用例名：** | projectQuery.jsp |
| **位置：** | 根目录/WebRoot |
| **描 述：** | 数据查询 |
| **一般事件过程：** | 输入查询信息，将信息提交到QueryServlet.java |
| **其它事件过程：** | 查询条件错误则重新查输入查询条件，未查询到显示为空 |
| **前导状态/条件：** | 用户登录状态 |
| **完成后的状态：** | 进入pro\_query.jsp并显示相应的查询结果 |

## 3.6 统计分析模块

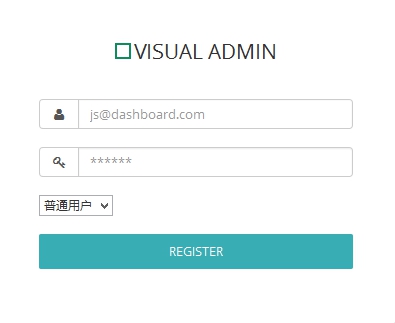
|  |  |
| --- | --- |
| **用例名：** | statisticalAnalysis.jsp |
| **位置：** | 根目录/WebRoot |
| **描 述：** | 用户登录 |
| **一般事件过程：** | 将表单信息提交到Stservlet.java进行数据库的查询验证操作 |
| **其它事件过程：** | 如果登陆信息不存在，则重新登陆 |
| **前导状态/条件：** |  |
| **完成后的状态：** | 进入login\_success.jsp选择操作查询成绩queryscore.jsp或选择科目考试slsubject.jsp |

## 3.1.7 接口

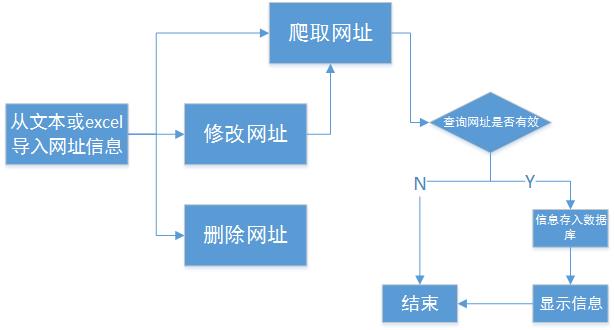
## 3.2 功能

* 登录
* 注销



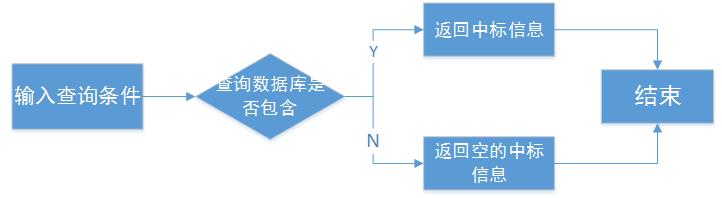


* 网址配置





* 查询中标信息





* 查看统计分析

