

1 引言.....	2
1.1 编写目的.....	2
1.2 背景.....	2
1.3 定义.....	2
1.4 参考资料.....	2
2 界面设计.....	2
2.1 用户类.....	错误!未定义书签。
2.2 账户信息类.....	错误!未定义书签。
2.3.....	错误!未定义书签。
2.4.....	错误!未定义书签。
3 事务逻辑设计（用 PDL）.....	5
3.1 用户类.....	错误!未定义书签。
3.2 账户信息类.....	错误!未定义书签。
3.3.....	错误!未定义书签。
4 数据库设计（用 PDL）.....	6

软件设计说明书

1 引言

1.1 编写目的

经过面向对象分析阶段，得到了系统的类图，细化的属性表和操作表，及三层精化类图；设计阶段的任务是给出各个类每层分解的实现方案，即界面显示方式、事务处理逻辑、数据库设计，不同任务的程序员可据此进行方案的实现。

预期读者为：软件开发的人员，项目评审人员，及软件测试人员。

1.2 背景

说明：

- a. 待开发软件系统的名称：网络爬虫；
- b. 本项目的任务提出者：李达；
- c. 本项目的任务开发者：李达，骆克云，吴钱胜，冯志远，陈睿进，邵月；
- d. 用户及实现该软件的计算中心或计算机网络：待定；

1.3 定义

爬取内容：存放中标公告中的项目名称、成交供应商、采购日期或评审日期、成交金额，用户可以按时间范围或成交金额范围筛选。

1.4 参考资料

[1] Scrapy0.25 文档：http://scrapy-chs.readthedocs.org/zh_CN/latest/index.html.

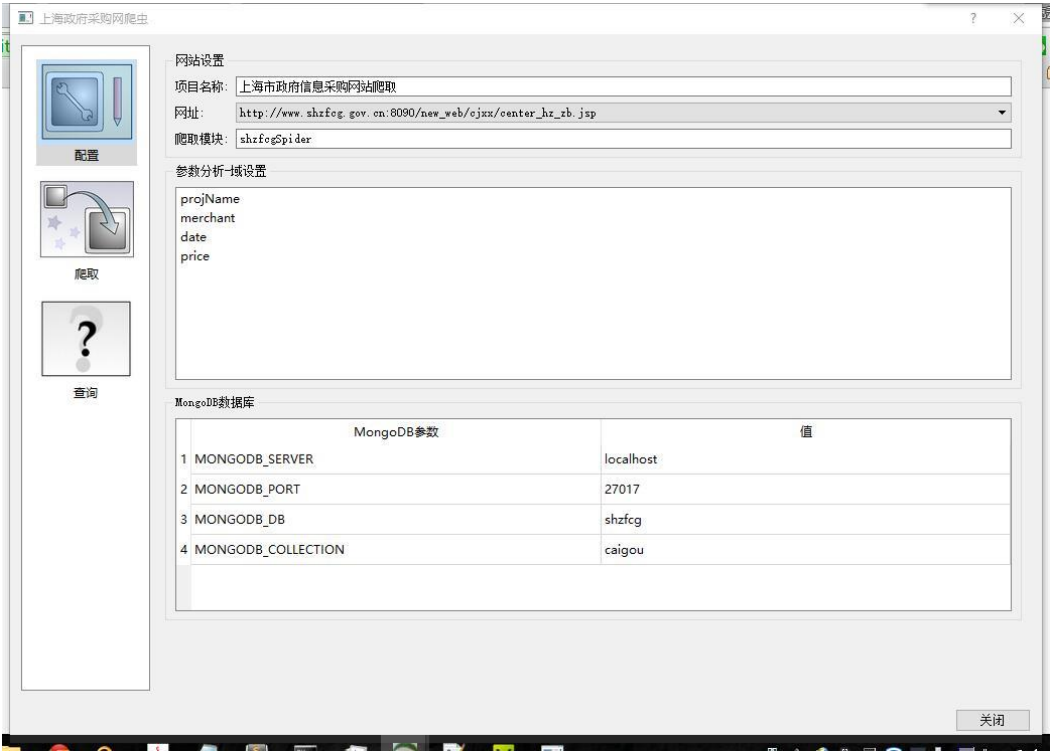
[2] Python 下用 Scrapy 和 MongoDB 构建爬虫系统：<http://python.jobbole.com/81280/>.

2 界面设计

（根据精化类图逐个给出每个类相关界面的设计结果）

2.1 配置类

class ConfigurationPage 爬虫配置界面



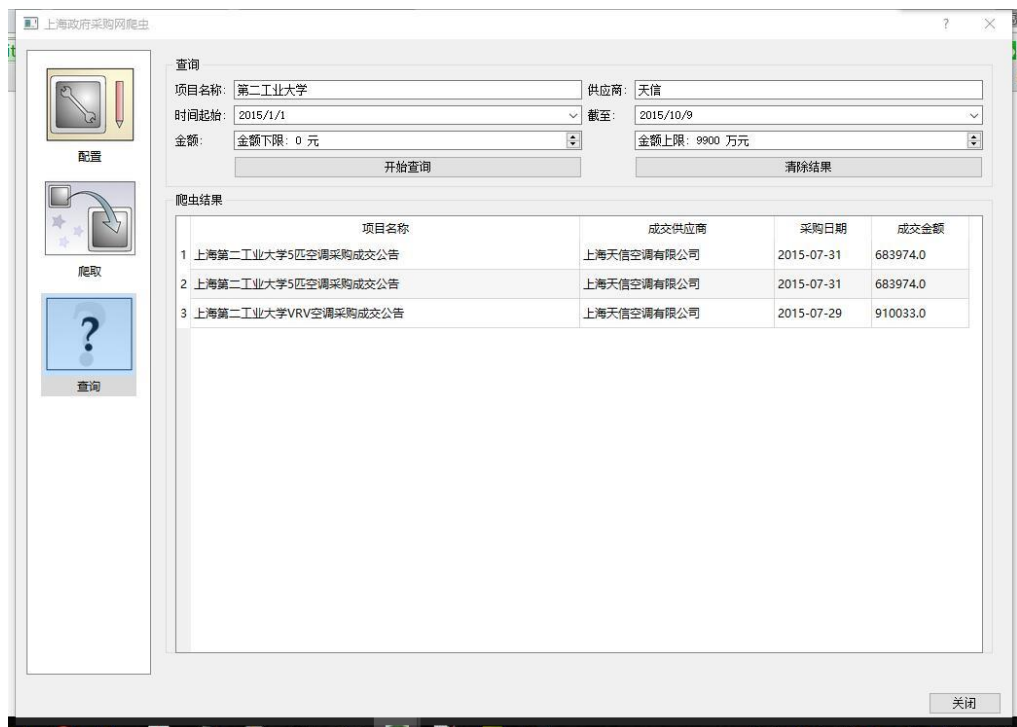
2.2 爬取类

class CrawlPage 爬取数据界面



2.3 查询类

class QueryPage 查询界面



3 事务逻辑设计

用详细设计方法之 PDL 完成，如已完成代码设计也可贴源码

配置：

```
configGroup = QtGui.QGroupBox(u"网站设置")
#项目名称
projectLabel = QtGui.QLabel(u"项目名称:")
projectEdit = QtGui.QLineEdit()
projectEdit.setText(u"上海市政府信息采购网站爬取")
projectEdit.setReadOnly(True)
#网址
serverLabel = QtGui.QLabel(u"网址:")
serverCombo = QtGui.QComboBox()
serverCombo.addItem("http://www.shzfcg.gov.cn:8090/new_web/cjxx/center_hz_zb.jsp")
#爬取模块
spiderLabel=QtGui.QLabel(u"爬取模块:")
spiderEdit=QtGui.QLineEdit()
spiderEdit.setText("shzfcgSpider")
spiderEdit.setReadOnly(True)
```

抓取：

```
crawlGroup = QtGui.QGroupBox(u"爬虫")
startButton=QtGui.QPushButton(u"开始")
self.textEdit = QtGui.QTextEdit()
self.textEdit.setMaximumHeight(100)
self.textEdit.setReadOnly(True)
startButton.clicked.connect(self.crawl)
crawlLayout = QtGui.QVBoxLayout()
crawlLayout.addWidget(startButton)
crawlLayout.addWidget(self.textEdit)
crawlGroup.setLayout(crawlLayout)
#结果布局
resultGroup=QtGui.QGroupBox(u"爬虫结果")
self.table=QtGui.QTableWidget(0,4)
self.table.setMinimumHeight(450)
#设置标题头
headerLabels = (u"项目名称", u"成交供应商",u"采购日期",u"成交金额")
self.table.setHorizontalHeaderLabels(headerLabels)
```

查询功能：

```
client=MongoClient()
db=client.shzfcg
```

```

collection=db.caigou
if projName!="" and merchant=="":
    queryCollection=collection.find({ "date":{"$gte":str(fromDate),"$lte":str(toDate)}, "proj
    name":{"$regex":projName}, "price":{"$gte":str(lowPrice), "$lte":str(highPrice)}})
elif projName !="" and merchant!="":
    print merchant
    queryCollection=collection.find({ "date":{"$gte":str(fromDate),"$lte":str(toDate)}, "proj
    name":{"$regex":projName}, "merchant":{"$regex":merchant}, "price":{"$gte":str(lowPr
    ice), "$lte":str(highPrice)}})
elif projName=="" and merchant!="":
    queryCollection=collection.find({ "date":{"$gte":str(fromDate),"$lte":str(toDate)}, "merc
    hant":{"$regex":merchant}, "price":{"$gte":str(lowPrice), "$lte":str(highPrice)}})
else:
    queryCollection=collection.find({ "date":{"$gte":str(fromDate),"$lte":str(toDate)}, "price
    ": {"$gte":str(lowPrice), "$lte":str(highPrice)}})
self.showResult(True, queryCollection)
client.close()

```

4 数据库设计

名称	类型
projName	String
merchant	String
date	Date
price	Float