# SuperSpider爬虫软件

# 需求分析与概要设计

## 项目说明

## 项目目标：

制作一个爬虫软件，可以根据关键词爬取知乎、百度贴吧、微博的相关信息（高赞答案、神回复）或热点信息，然后呈现在界面上。

## 软硬件环境需求

环境：.Net Framework

操作系统：Windows

需要网络

## 使用的关键技术：

分析URL、下载WebServer返回的HTML、分析HTML内容、构建HTTP请求的模拟、在爬虫过程中存储有用的信息

## 需求分析

## 系统用例

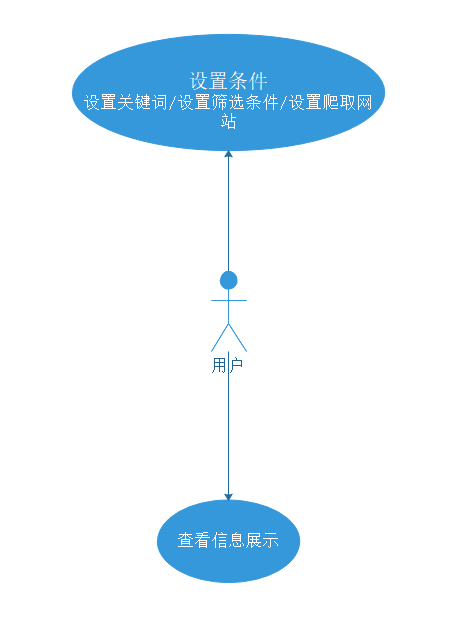


图 1 系统用例图

1. **设置关键词**

**参与者：**用户

**基本事件流：** 在输入框中输入想要查询的关键词，在复选框中勾选想要筛选的条件（热度、点赞数），在复选框中勾选想要爬取的网站（知乎、虎扑、微博）。

1. **查看信息展示**

**参与者：用户**

**基本事件流：** 单击开始，爬取成功后会在新的弹出框中可以查看爬取到的数据

## 业务流程

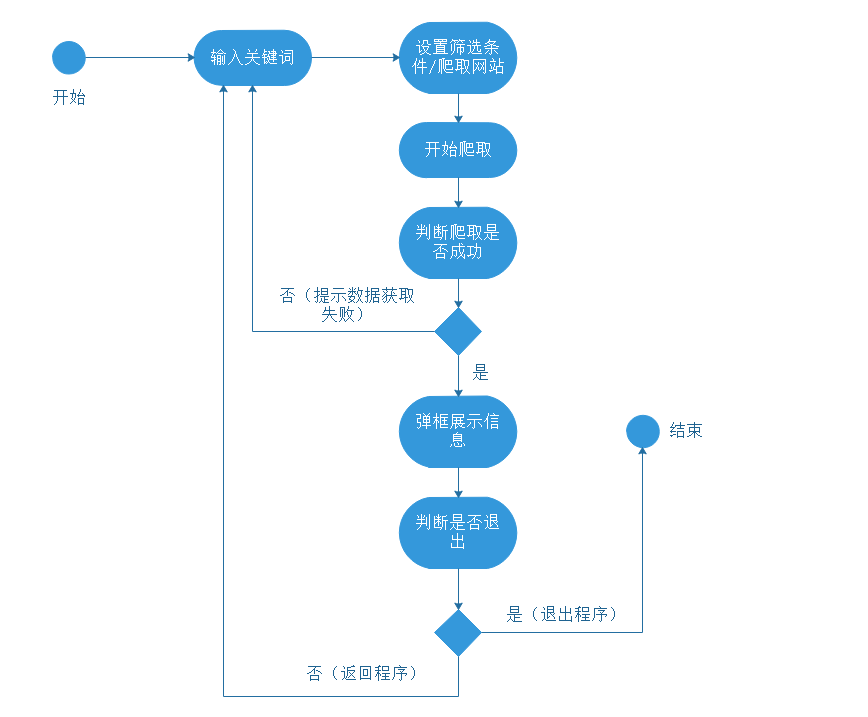


图 2软件使用流程图

## 概要设计

## 功能模块设计

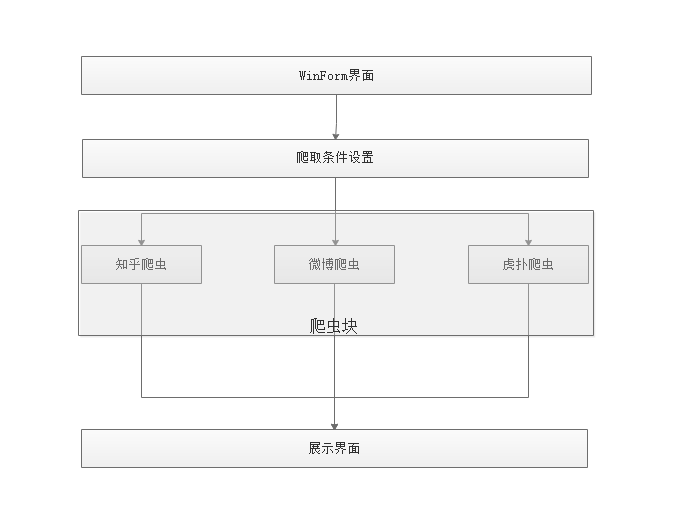


图 3功能模块图

1. **爬取条件设置**

**输入：**关键词、搜索优先条件（热度、赞同数、时间） 、想要查看的网站、热点查看

**输出：无**

**功能概述：** 设置爬取的条件，为爬虫做准备

1. **爬虫块**

**输入：**爬取条件

**输出：**爬取到的数据

**功能概述：**通过三个爬虫程序同时按照输入的条件进行数据爬取，并将数据输出到界面

**（3） 展示界面**

**输入：**爬取到的数据

**输出：**爬取到的数据

**功能概述：**将爬取到的数据展示在界面上

## 核心类图

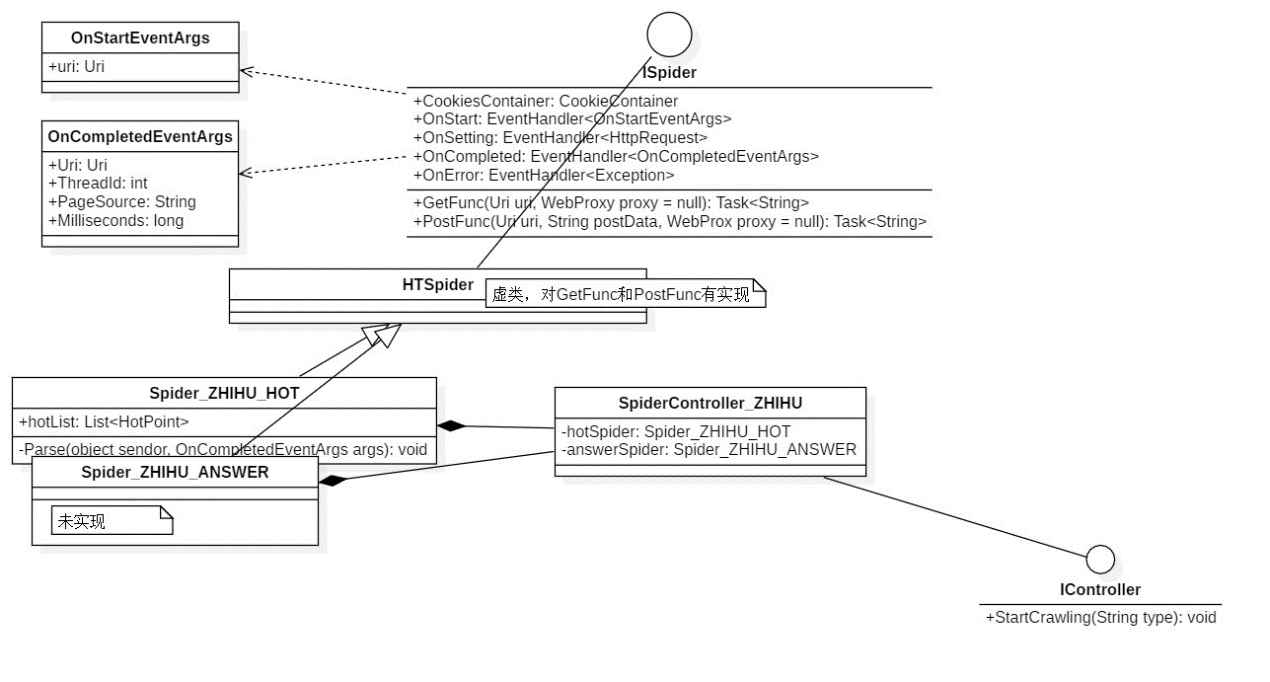
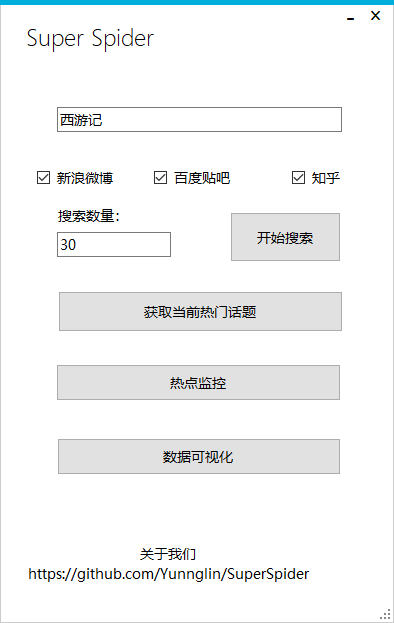


图 4核心类图

## 界面设计



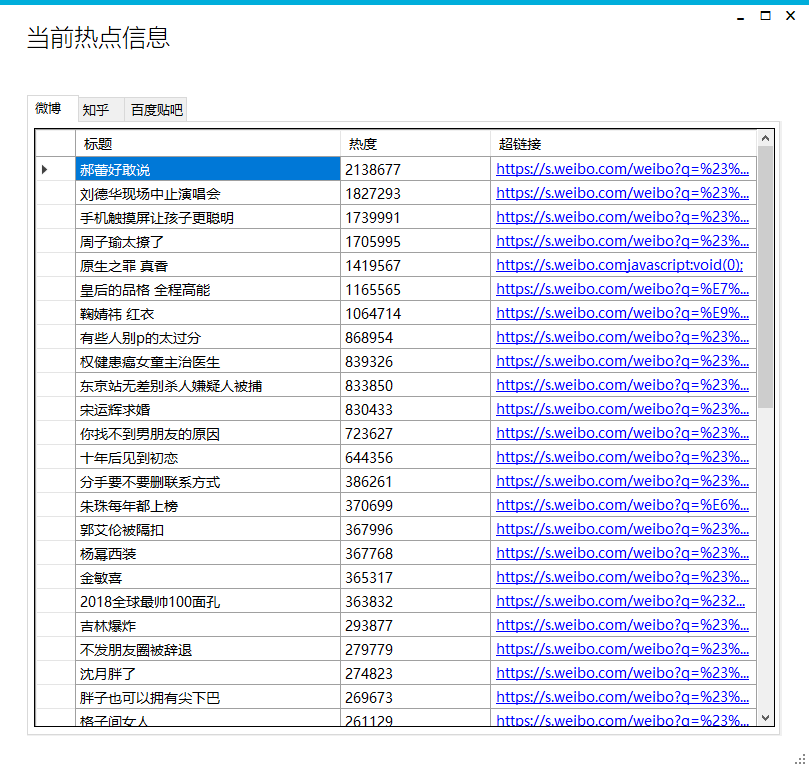
爬虫主界面



热点监控界面



信息获取界面



热榜界面