股票爬虫系统设计与实现

摘　　要

　　大数据时代来临，数据成为重要资产之一，各行各业纷纷将目光投向互联网，股票市场也不例外，越来越多的人开始通过互联网获取股票信息。面对网络上纷繁杂乱的数据，如何整合股票数据，从而使股票行业相关人士快速获取自己所需的信息成为需要探讨的问题。本文设计股票爬虫系统，旨在为股票行业相关人士提供一个上述问题解决方案。

　　首先，研究国内外同类型系统的成功案例和参考文献，结合股票网站，确定研究内容和方向；其次，学习 Python 和 Scrapy 框架，并根据股票网站的特点为网络爬虫制定详细的爬取方案；再次，根据用户需求为系统确定合理的架构，对系统的功能模块进行合理的划分；然后，对系统进行实验验证系统的功能和可行性；最后，对论文工作进行了总结并对未来的工作进行了展望。

　　关键词：网络爬虫；Scrapy；股票

Abstract

With the advent of the big data era, data has become one of the important assets, and all kinds of corporations have turned their eyes to the Internet, and the stock market is no exception. More and more people have embarked on obtaining stock information through the Internet. In the face of the numerous and messy data on the network, how to integrate the stock data so that the relevant people in the stock industry can quickly obtain the information they need becomes a problem to be discussed. The article designs a stock crawler system designed to provide a solution to the above problems for relevant people in the stock industry.

Firstly, we study the successful cases and references of the same type of system at home and abroad, combine the stock website to determine the research content and direction. Secondly, we study the Python and Scrapy framework, and make a detailed crawling scheme for web crawler according to the characteristics of stock website. Thirdly, we choose the logical structure for the system according to the user's requirements, and divide the functional modules of the system logically. Fourthly, we test the system to testify the function and practicability of the system. Finally, we sum up the paper and yearn for the promising future.

Key Words: Web crawler；Scrapy；Stocks

目　　录

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：https://d.book118.com/108061051131006120