



# 文本复制检测报告单(简洁)

№:ADBD2019R\_20190527123208446519102431

检测时间：2019-05-27 12:32:08

检测文献：基于python的分布式网络爬虫设计与实现

作者：邓凯

检测范围：中国学术期刊网络出版总库

中国博士学位论文全文数据库/中国优秀硕士学位论文全文数据库

中国重要会议论文全文数据库

中国重要报纸全文数据库

中国专利全文数据库

图书资源

优先出版文献库

大学生论文联合比对库

互联网资源(包含贴吧等论坛资源)

英文数据库(涵盖期刊、博硕、会议的英文数据以及德国Springer、英国Taylor&Francis 期刊数据库等)

港澳台学术文献库

互联网文档资源

CNKI大成编客-原创作品库

个人比对库

时间范围：1900-01-01至2019-05-27

## 检测结果

去除本人已发表文献复制比：4.4%

跨语言检测结果：0%

引 去除引用文献复制比：4.4%

总 总文字复制比：4.4%

单 单篇最大文字复制比：2.6%

去除本人已发表文献复制比：

重复字数：[ 990 ]

总字数：[ 22334 ]

单篇最大重复字数：[ 577 ]

总段落数：[ 2 ]

前部重合字数：[ 43 ]

疑似段落最大重合字数：[ 922 ]

疑似段落数：[ 2 ]

后部重合字数：[ 947 ]

疑似段落最小重合字数：[ 68 ]

指 标：☐ 疑似剽窃观点 ☒ 疑似剽窃文字表述 ☐ 疑自我剽窃

☐ 一稿多投

☐ 过度引用

表格：0

公式：没有公式

疑似文字的图片：0

脚注与尾注：0

9.1% (922) 基于python的分布式网络爬虫设计与实现.doc\_第1部分 (总10078字)

0.6% (68) 基于python的分布式网络爬虫设计与实现.doc\_第2部分 (总12256字)

(注释：☐ 无问题部分 ☒ 文字复制比部分 ☐ 引用部分)

## 指导教师审查结果

指导教师：何粒波

审阅结果：通过

审阅意见：通过，各功能模块实现成功，论文得结构安排合理，层次比较清晰，内容描述比较清楚，同意通过。

1. 基于python的分布式网络爬虫设计与实现.doc_第1部分		总字数：10078
相似文献列表	基于Scrapy框架的聚焦爬虫的设计与实现	5.7% ( 577 )
去除本人已发表文献复制比：9.1%(922)	韩颜隆 - 《大学生论文联合比对库》 - 2018-04-27	是否引证：否
文字复制比：9.1%(922)	疑似剽窃观点 ( 0 )	
2	基于Scrapy框架的聚焦爬虫的设计与实现	5.7% ( 577 )
	韩颜隆 - 《大学生论文联合比对库》 - 2018-05-02	是否引证：否
3	基于Scrapy框架的聚焦爬虫的设计与实现	5.7% ( 577 )
	韩颜隆 - 《大学生论文联合比对库》 - 2018-05-15	是否引证：否
4	基于Scrapy框架的聚焦爬虫的设计与实现	5.7% ( 577 )
	韩颜隆 - 《大学生论文联合比对库》 - 2018-05-29	是否引证：否
5	电信1403_141608060321_高明生_Python网络爬虫的设计与实现	5.7% ( 577 )
	高明生 - 《大学生论文联合比对库》 - 2018-04-03	是否引证：否
6	电信1403_141608060321_高明生_网络爬虫的设计与实现	5.7% ( 577 )
	高明生 - 《大学生论文联合比对库》 - 2018-04-12	是否引证：否
7	基于Android和J2EE的图书资讯系统设计 with 实现	5.7% ( 572 )
	李振全 - 《大学生论文联合比对库》 - 2018-04-25	是否引证：否
8	Scrapy分布式爬虫搜索引擎	5.6% ( 566 )
	刘思林;- 《电脑知识与技术》 - 2018-12-05	是否引证：否
9	基于半监督学习的微博水军识别系统的研究与实现	5.2% ( 528 )
	陶涛(导师：陶军;彭艳兵) - 《东南大学博士论文》 - 2018-06-01	是否引证：否
10	电商平台商品数据抓取系统的设计与实现	5.0% ( 508 )
	杨盼盼 - 《大学生论文联合比对库》 - 2018-05-11	是否引证：否
11	不同社交平台上公众情感与股市关系的对比研究	4.8% ( 480 )
	刘学斌(导师：刘肖凡;王会方) - 《东南大学博士论文》 - 2018-06-25	是否引证：否
12	基于scrapy框架的新浪网资讯爬虫	4.7% ( 476 )
	徐泽汐 - 《大学生论文联合比对库》 - 2018-05-14	是否引证：否
13	基于scrapy框架的新浪网资讯爬虫	4.3% ( 437 )
	徐泽汐 - 《大学生论文联合比对库》 - 2018-05-23	是否引证：否
14	基于scrapy框架的新浪网资讯爬虫	4.3% ( 437 )
	徐泽汐 - 《大学生论文联合比对库》 - 2018-05-30	是否引证：否
15	基于大数据和Mahout架构的毕业生就业智能推荐平台研究	2.2% ( 221 )
	张海华;杨秀波;张聪;- 《电脑编程技巧与维护》 - 2019-01-18	是否引证：否
16	传统佳节网站设计	1.3% ( 129 )
	包莉莉 - 《大学生论文联合比对库》 - 2014-05-15	是否引证：否
17	一些对HTTP协议的理解 - 博客频道 - CSDN.NET	0.6% ( 63 )
	- 《网络 ( <a href="http://blog.csdn.net">http://blog.csdn.net</a> ) 》 - 2017	是否引证：否
18	恶意程序 能将你电脑变成代理服务器	0.6% ( 61 )
	吴俊;- 《计算机与网络》 - 2016-01-12	是否引证：否
19	基于web的购书网站的设计和实现	0.5% ( 54 )
	陆佩 - 《大学生论文联合比对库》 - 2017-05-09	是否引证：否
20	定向网络爬虫	0.4% ( 42 )
	任伟超 - 《大学生论文联合比对库》 - 2017-06-02	是否引证：否
21	5472224_陈卓_非结构化数据库搜索软件设计	0.4% ( 40 )
	陈卓 - 《大学生论文联合比对库》 - 2018-06-01	是否引证：否
22	IIS服务器安全防护系统的设计与实现	0.4% ( 38 )
	贺永佳(导师：王伟平) - 《中南大学博士论文》 - 2012-05-01	是否引证：否

2. 基于python的分布式网络爬虫设计与实现.doc_第2部分		总字数：12256
相似文献列表		
去除本人已发表文献复制比：0.6%(68)      文字复制比：0.6%(68)      疑似剽窃观点 ( 0 )		
1	基于关联规则与孤立点的信息舞弊特征识别研究 黄慧连(导师：幸莉仙) - 《华北电力大学博士论文》 - 2013-03-01	0.6% ( 68 ) 是否引证：否

说明： 1.指标是由系统根据《学术论文不端行为的界定标准》自动生成的

2.本报告单仅对您所选择比对资源范围内检测结果负责

3.Email：[amlc@cnki.net](mailto:amlc@cnki.net)

 <http://e.weibo.com/u/3194559873>

 [http://t.qq.com/CNKI\\_kycx](http://t.qq.com/CNKI_kycx)

<http://check.cnki.net/>