


原 R语言爬虫：RSelenium包常用函数

2018年01月09日 09:05:50 [weixin_40628687](#) 阅读数：2172

 版权声明：本文为博主原创文章，未经博主允许不得转载。 https://blog.csdn.net/weixin_40628687/article/details/79008995

文字太多，一眼找不到要找的内容？ [Ctrl+F](#) 即可查找~

使用RSelenium包控制浏览器主要依靠remoteDriver系列函数：

```
remoteDriver( remoteServerAddr,  
  
              port,  
  
              browserName,  
  
              path,  
  
              version,  
  
              platform,  
  
              javascript,  
  
              nativeEvents,  
  
              serverURL,  
  
              sessionInfo)
```

其中，参数remoteServerAddr，表示远程服务器的IP地址，默认值是本机地址，是character型；

参数port，表示远程服务器连接端口，是numeric型；

参数browserName，表示浏览器名称，可以取chrome、firefox、[htmlunit](#)、internetexplorer或[iphone](#)，是character型；

参数path，表示远程服务器上命令的基本URL路径前缀，默认值是“/ wd / hub”；

参数version，表示浏览器版本，是character型；

参数platform，表示浏览器运行的系统，可以是WINDOWS、XP、VISTA、MAC、LINUX或UNIX，是character型；

参数javascript，表示会话是否支持在当前页面的上下文中执行用户提供的JavaScript，是logical型；

参数nativeEvents，表示会话是否支持本地事件，是logical型；

参数serverURL，表示JSON请求发送到的远程服务器的URL，是character型；

参数sessionInfo，表示会话相关信息，是list型。

通常情况下，只要设置browserName就可以对浏览器进行操作。

```
1 | library(RSelenium)  
2 | remDr <- remoteDriver(browserName ="chrome")
```

remoteDriver函数的返回值是一系列函数，即remDr是一系列函数，先将主要的操作总结如下：

1、打开浏览器、关闭浏览器

```
1 | # 打开浏览器  
2 | remDr$open()  
3 | # 关闭浏览器
```

```
4 | remDr$quit() # 直接退出
5 | remDr$close() # close用于关闭当前会话，也可以用作关闭浏览器
```

2、打开网页、关闭网页

```
1 | # 打开网页
2 | url <- 'http://www.baidu.com'
3 | remDr$navigate(url)
4 | # 关闭网页
5 | remDr$closeWindow()
```

3、定位元素

定位到网页上某个元素需要使用函数findElement(using, value)，其中using可以取"xpath", "css selector", "id", "name", "tag name", "class name", "text"或"partial link text"，value是其相应的值。

4、移动鼠标

移动鼠标使用函数mouseMoveToLocation(x = NA_integer_, y = NA_integer_, webElement = NULL)，其中x和y是坐标值，webElement是移向的网页元素。

5、单击

单击使用函数click(buttonId = 0)，其中buttonId为0表示单击左键，为1表示单击滚动条，为2表示单击右键。

但是通常我们单击需要单击在网页的某个元素上，要定位到这个元素需要使用函数findElement(using, value)，其中using可以取"xpath", "css selector", "name", "tag name", "class name", "link text"或"partial link text"，value是其相应的值。当然，单击某个元素也可以使用clickElement()函数。

```
1 | # 以点击百度的“百度一下”按钮为例
2 | # 打开网页
3 | url <- 'http://www.baidu.com'
4 | remDr$navigate(url)
5 | # 定位
6 | xpath <- '//*[@id="su"]'
7 | btn <- remDr$findElement(using = 'xpath', value = xpath)
8 | # 移动鼠标至该元素
9 | remDr$mouseMoveToLocation(webElement = btn)
10 | # 单击
11 | remDr$click()
```

另一种方法（使用clickElement函数）：

```
1 | # 打开网页
2 | url <- 'http://www.baidu.com'
3 | remDr$navigate(url)
4 | # 定位
5 | xpath <- '//*[@id="su"]'
6 | btn <- remDr$findElement(using = 'xpath', value = xpath)
7 | # 单击
8 | btn$clickElement()
```

6、双击

双击使用函数doubleclick(buttonId = 0)，用法与单击一样。

7、在文本框输入文本

函数sendKeysToElement(sendKeys)用于在指定元素中输入文本，sendKeys必须是list型数据。在网页输入文本的步骤大致为，首先找到文本框元素的然后输入文本。另外一个函数sendKeysToActiveElement(sendKeys)用于在当前活动元素中输入文本。

```
1 ## 以百度文本框为例
2 # 打开网页
3 url <- 'http://www.baidu.com'
4 remDr$navigate(url)
5 # 定位文本框
6 xpath <- '//*[@id="kw"]'
7 btn <- remDr$findElement(using = 'xpath', value = xpath)
8 # 输入文本
9 text <- list('偷闲阁CSDN')
10 btn$sendKeysToElement(text)
11 # 输入文本并按回车
12 text <- list('偷闲阁CSDN', key = 'enter')
13 btn$sendKeysToElement(text)
14
```

8、窗口最大化

```
remDr$maxWindowSize()
```

9、截图

截图使用screenshot(display = FALSE, useViewer = TRUE, file = NULL)函数，其中file是截图后存放的路径，截的图像是PNG格式。

```
remDr$screenshot()
```

RSelenium包的准备工作请参考[R语言爬取动态网页：使用RSelenium包和Rwebdriver包的前期准备](#)。抓取示例请参考[R语言：RSelenium包爬取动态网](#)

参考文献：

[RSelenium手册](#)

良心记者揭露紫金山蜂蜜市场黑幕！亲赴深山，发现惊人真相！

新钥 · 鸺鹠



想对作者说点什么

R语言利用RSelenium包或者Rwebdriver模拟浏览器爬取异步加载...

阅读数 9576

1.背景介绍：现在很多网页都采用AJAX这种异步加载的网页结构，简单的页面爬取不... [博文](#) 来自：[abc2009...](#)

RSelenium基本用法总结

阅读数 354

有些数据在script标签里面的js动态脚本中，普通请求器如RCurl和httr没有权限办到(少... [博文](#) 来自：[默默耕耘...](#)

selenium用法详解

阅读数 18万+

selenium是一个用于Web应用程序测试的工具。Selenium测试直接运行在浏览器中，... 博文 来自： 天涯笨熊...

为什么世界名校都在教Python，这篇报告告诉你！

Python火爆全球的原因，你一定要知道！

RSelenium包抓取链家网（上：模拟点击与页面抓取）

阅读数 2281

安装RSelenium包install.packages("RSelenium")#直接从CRAN下载RSe... 博文 来自： Joyliness

以后再有人问你selenium是什么，你就把这篇文章给他

阅读数 1万+

写在最前面：目前自动化测试并不属于新鲜的事物，或者说自动化测试的各种方法论... 博文 来自： TestingG...

几款有用的R包

阅读数 1343

1. autoinst 自动安装缺失包，实际测试发现如果使用了包内的方法，只会导致R报... 博文 来自： 天天磨刀

【R 数据科学】R语言进行数据科学整理最有用的包大全

阅读数 4095

一、数据科学工作流程1.1数据导入1.2数据整理1.3反复理解数据1.4数据可视化1.5数... 博文 来自： 赖德发的...

常用R包

阅读数 306

MachineLearning在R... 的包e1071 隐类分析，短时傅立叶变换，模糊聚类，支... 博文 来自： song_110...

千万不要再乱喝蜂蜜了，蜂民对天发誓要揭露背后惊人黑幕！

博森 · 鸛鵒

R 语言数据分析/爬虫

Python怎么学

转型AI人工智能指南

Java学习路线

28 天算法训练营

2019 Python 开发者日

IT培训机构排名

知网查重入口

数据分析的流程主要包括以下几个步骤，而数据清洗作为第一步至关重要，数据... 博文 来自： weixin_40628687

[登录](#)

一次性下载《R语言实战2》全书的R包及常用的R包

阅读数 7793

系统：CentOS6.5R版本：3.2.5source("http://bioconductor.org/biocLite.R")biocLite(c(... 博文 来自： 我爱红金...



Joyliness

22篇文章

[关注](#) 排名:千里之外

beyondma

62篇文章

[关注](#) 排名:千里之外

豆饼

1篇文章

[关注](#) 排名:千里之外

B_Hanan

12篇文章

[关注](#) 排名:千里之外

RSelenium/Rwebdriver/rdom包抓取表格数据

阅读数 1008

安装RSelenium/Rwebdriverinstall.packages("RSelenium")#直... 博文 来自： Joyliness

R语言量化：KDJ指标的计算

阅读数 1504

KDJ指标，中文名随机指标，由乔治·莱恩（GeorgeLane）创立，是目前期货和... 博文 来自： 偷闲阁

简单粗暴彻底解决selenium+chromedriver无法定位各种元素的方法

阅读数 1万+

解决各种webdriver.find_element定位不准的问题 博文 来自： 一个业务...

一个长期喝蜂蜜的人，竟然变成了这样，再忙也要看一下

恒昌 · 鸛鵒

mac系统 selenium配合无界面谷歌浏览器headless+ 模仿手机登...

阅读数 403

fromseleniumimportwebdriverfromselenium.webdriver.chrome.optionsimportOptionsd... 博文 来自： 小米的博客

【R语言 爬虫】Rwebdriver 安装方法

阅读数 840

RSelenium和Rwebdriver个人刚接触不久，除了语法不太一样以为，都是调用的Selen... 博文 来自： 赖德发的...

R爬虫小白实例教程 - 基于rvest包

阅读数 1万+

一入爬虫深似海，从此复制是路人。 博文 来自： 数据分析...

R Selenium学习笔记（一）

阅读数 447

一、介绍R Selenium的目标是让R更容易的连接Selenium Server。R Selenium绑定了R... [博文](#) 来自: [qq_3329...](#)

R语言：R Selenium包爬取动态网页

阅读数 2261

目前很多网站的数据采用异步加载的方式呈现，以进口商品统计月报为例，当我们点... [博文](#) 来自: [偷闲阁](#)

它居然是满脸斑的克星，好多人都不知道，导致斑点越来越严重！

盐边 · 鸛鸛

R Selenium包抓取豆瓣电影（模拟滚动条）

阅读数 1620

关于辅助插件SelectorGadget SelectorGadget插件可以用来显示网页中任意元素的C... [博文](#) 来自: [Joyliness](#)

R语言微博爬虫RWEIBO

11-14

R语言微博爬虫RWEIBOR语言微博爬虫RWEIBOR语言微博爬虫RWEIBO

[下载](#)

一文排除R语言抓取动态网页（selenium）所有的坑

阅读数 102

最近一直在搞R语言结合自然语言处理，分析公司研究报告相关的课题，但是鉴于单... [博文](#) 来自: [微软最有...](#)

R成精系列-Mac os用R+Docker部署Rselenium搭建爬虫系统

阅读数 111

最近在用R做爬虫的时候遇到了部分网页无法爬取。于是转投Selenium的怀抱。本文... [博文](#) 来自: [u010523...](#)

rvest爬虫及案例分析

阅读数 1327

由于最近准备着学习手数据挖掘，想着先从爬取数据开始。现做一下总结，可能需要... [博文](#) 来自: [qq_4058...](#)

良心蜂农揭露蜂蜜市场！亲赴深山，发现惊人真相！

乐爱实业 · 鸛鸛

R语言中利用pacman安装和加载包

阅读数 88

R语言之强大，就在于它可以安装加载不同的包（package）。很多前沿的科研工作者... [博文](#) 来自: [辉的博客](#)

selenium爬取裁判文书网

阅读数 569

一、设定搜索条件from selenium import webdriverimport timeoptions = webdriver.Chrom... [博文](#) 来自: [码农Robi...](#)

基于R Selenium爬取中国裁判文书网文书数据

阅读数 863

传送：R Selenium基本用法总结 Python与R协同完成【中国裁判文书网】文书内容... [博文](#) 来自: [默默耕耘...](#)

Mac版R语言（四）网络数据自动获取

阅读数 1309

R语言爬虫，R语言网络数据自动获取 [博文](#) 来自: [Nicolelov...](#)

R Selenium包抓取链家网（下：数据存储与容错管理）

阅读数 538

HouseinfoFuncfunction(link){destination"UTF-8"}#获取第i条房... [博文](#) 来自: [Joyliness](#)

一个长期喝蜂蜜的人，竟然变成这样，再忙也要看一下！

乐爱实业 · 鸛鸛

R语言基本统计分析方法（包及函数）

阅读数 4173

转载自：http://blog.csdn.net/s04023083/article/details/40344273摘要：目前经典的统... [博文](#) 来自: [qq_3554...](#)

详解介绍Selenium常用API的使用--Java语言（完整版）

阅读数 1万+

参考：http://www.testclass.net/selenium_java/一共分为二十个部分：环境安装之Java... [博文](#) 来自: [qq_2200...](#)

R常用函数参考

07-10

R语言函数参考 中英文 包括常用函数API，基础包函数的中英文参考，R常用函数列表

[下载](#)

RCurl爬虫抓取验证码例子

阅读数 2112

RCurl抓取验证码流程junjun2016年4月20日抓取验证码#1) 抓取验证的流程#①打开... 博文 来自: 军军的专栏

R语言常用函数

R语言常用函数列表，帮助快速查询 R语言常用函数

12-15

下载

“股市三板斧”坚持只看1指标 赢数千万身家！

君银证券 · 顶新

R语言数据抓取实战——RCurl+XML组合与XPath解析

阅读数 1073

杜雨，EasyCharts团队成员，R语言中文社区专栏作者，兴趣方向为：Excel商务图表... 博文 来自: R语言中...

R爬虫豆瓣电影top250

基于rvest包对豆瓣电影top250的R语言爬虫代码加注释，R爬虫入门首选。

11-19

下载

R语言爬取动态网页：使用RSelenium包和Rwebdriver包的前期准备

阅读数 3889

在使用rvest包抓取新浪财经A股交易数据中我们介绍了rvest包的用法，然而rvest... 博文 来自: 偷闲阁

自动化测试工具Selenium入门

阅读数 3万+

自动化测试工具Selenium入门自动化测试工具Selenium入门简介入门环境搭建代码实... 博文 来自: 剑未佩妥...

Selenium介绍及原理解析

阅读数 2万+

一、什么是Selenium？Selenium是ThoughtWorks公司的一个强大的开源Web功能测... 博文 来自: 拓缝的渴望



硅谷深度学习课程，行业专家1对1代码审阅！

立抢限量席位！学习TensorFlow框架等主流技术，成为顶尖人才

R语言爬虫页面时页面内有show more按钮，抓取页面问题

-

问答

R语言爬虫之rvest包初试（一）

阅读数 5595

版权声明：本文为博主原创文章，未经博主允许不得转载。 R语言中的rvest包是一... 博文 来自: 琉璃兔兔...

Selenium 爬取评论数据，就是这么简单！

阅读数 4834

本文来自作者 秦子敬 在 GitChat 上分享「如何利用Selenium爬取评论数据？」，「... 博文 来自: GitChat...

使用selenium模块模拟浏览器爬去网页，并进行点击定位内容笔记

阅读数 1564

针对模拟浏览器采用.webdriver.Firefox()火狐浏览器模拟Windows系统解决办法如下... 博文 来自: IT届的小...

利用selenium实现截图

阅读数 1743

packagecom.dinfo.hrb.util;importjava.awt.image.BufferedImage;importjava.io Buffere... 博文 来自: 小飞侠的...



硅谷深度学习课程，行业专家1对1代码审阅！

立抢限量席位！学习TensorFlow框架等主流技术，成为顶尖人才

Rcurl小应用，爬取京东评论

阅读数 2892

利用Rcurl包做的小爬虫，爬取了京东上电热水器的评论 博文 来自: 黄同学的...

RCurl爬取天猫评论

10-07

详细介绍了如何用R语言进行爬虫，压缩包中有各种注意事项，包含R语言代码


下载

【R语言爬虫】R语言提交get请求抓取城理财数据

阅读数 2680

一、需求分析抓取城理财数据。抓取url:https://www.cclc.co/debts/lctz_all_all二、实... 博文 来自: 赖德发的...

R语言中处理XML/HTML函数详解-爬虫相关	阅读数 356
本篇介绍XML程序包中的函数，因函数较多，侧重于爬虫相关的函数1.xmlTreeParse(...博文来自：默默耕耘...	
一个RCurl抓取美团网信息的实例	阅读数 831
最近忽然想尝试用R做一个简单的网络爬虫，看了不少资料，查了不少技术牛人的代...博文来自：zsk98031...	
股市经典口诀，涨三不追，跌四不压... 赢百万身价 看完恍然大悟	
君银证券 · 顶新	
Window下R语言启动Rselenium	阅读数 450
首先下载Selenium，下载好之后打开命令窗口，将Selenium的jar包放入命令窗口的路...博文来自：Blackros...	
R语言爬虫	阅读数 348
1.简单爬取一个网页url<"http://hz.house.ifeng.com/detail/2014_10_28/500876...博文来自：樊芳芳的...	
caffe安装系列——安装OpenCV	阅读数 2万+
说明 网上关于caffe的安装教程非常多，但是关于每一步是否操作成功，出现了什么样...博文来自：张学志の...	
图像缩放双线性插值算法	阅读数 1万+
插值算法对于缩放比例较小的情况是完全可以接受的，令人信服的。一般的，缩小0.5...博文	
thymeleaf模板实现html5标签的非严格检查	阅读数 2万+
一、概述最近在springboot项目引入thymeleaf模板时，使用非严格标签时，运行会报...博文来自：Luck_ZZ...	
Sphinx 的介绍和原理探索	阅读数 4552
What/Sphinx是什么 定义：Sphinx是一个全文检索引擎。特性：索引和性能优异易于...博文来自：HugeStaff	
Ubuntu16.04下设置静态IP	阅读数 8万+
最近在自己装有win7系统的thinkpad电脑上，给安装了Ubuntu16.04双系统。想在ubu...博文来自：xiaohuozhi...	
ThreadLocal的设计理念与作用	阅读数 4万+
Java中的ThreadLocal类允许我们创建只能被同一个线程读写的变量。因此，如果一...博文来自：u011860...	
linux上安装Docker(非常简单的安装方法)	阅读数 17万+
最近比较有空，大四出来实习几个月了，作为实习狗的我，被叫去研究Docker了，汗...博文来自：我走小路...	
配置简单功能强大的excel工具类搞定excel导入导出工具类(一)	阅读数 3万+
对于J2EE项目导入导出Excel是最普通和实用功能,本工具类使用步骤简单,功能强大,...博文来自：李坤 大米...	
强连通分量及缩点tarjan算法解析	阅读数 51万+
强连通分量：简言之 就是找环（每条边只走一次，两两可达） 孤立的一个点也是一...博文来自：九野的博客	
MATLAB中注释一段程序	阅读数 1万+
在MATLAB中，可以注释一段程序。使用"%{"和"%}"。例如 %{。。。}%}即可。经...博文来自：知识小屋	
科学计算库numpy常用函数 Go语言 Go语言入门学习 Go语言教程 Go语言培训套餐	
c# 应用r包 go语言写爬虫 go语言 hash包 go语言爬虫 图片 go语言实战 爬虫 r语言python速成班 人工智能与r会议课程	



weixin_40628687

关注

原创

20

粉丝

58

喜欢

20

评论

25

等级：

博客 3

访问：

11万+

积分：

947

排名：

7万+

勋章：

恒

最新文章

R语言作图：使用ggplot2绘制人口金字塔

R语言量化：MACD的计算及使用

R语言量化：KDJ指标的计算

R语言：排序、筛选以及分类汇总操作

R语言量化：alpha和beta的用法——所谓“股性”和alpha策略

个人分类

R语言 数据的读写

1篇

R语言 可视化

5篇

R语言 文本处理

1篇

R语言 网络爬虫

5篇

R语言 数据分析

1篇

展开

归档

2018年8月

1篇

2018年5月

1篇

2018年3月

1篇

2018年2月

9篇

2018年1月

4篇

展开

热门文章

R语言作图：坐标轴设置
阅读数 34286

R语言：排序、筛选以及分类汇总操作
阅读数 24109

R语言：填色等值线图及其色标（color bar）设置
阅读数 6987

R语言处理气象数据：NetCDF格式数据的

读写
阅读数 6385

R语言作图：数学公式
阅读数 4435

- 最新评论
- R语言量化：使用WindR下载Wi...
qq_41969000：wind免费版哪里下载

R语言：填色等值线图及其色标（co...
weixin_40628687：[reply]weixin_44541877[/reply]
使用contour函数

R语言：填色等值线图及其色标（co...
weixin_44541877：您好，我想问下如何在填色图上绘制等值线？

R语言爬取动态网页：使用RSele...
weixin_40628687：[reply]weixin_43208142[/reply]
我没有下载selenium-java-3.9.1.zip；似乎seler ...

R语言爬取动态网页：使用RSele...
weixin_43208142：请问大佬，selenium我应该下载哪种版本？ selenium-server-standalone-3.9. ...

VPN That Works in China

Be Free, Truly Anonymous & Secured. High-Speed Guaranteed. Order Now!

ExpressVPN



CSDN学院



CSDN企业招聘

 QQ客服

 kefu@csdn.net

 客服论坛

 400-660-0108

工作时间 8:30-22:00

关于我们 招聘 广告服务 网站地图

 百度提供站内搜索 京ICP证19004658号

©1999-2019 北京创新乐知网络技术有限公司

网络110报警服务 经营性网站备案信息

北京互联网违法和不良信息举报中心

中国互联网举报中心