**SCIHUB下载器说明文档**

目录

[一、实践1:用下载器获取一个pdf 2](#_Toc31828489)

[1.打开获取链接 2](#_Toc31828490)

[2.输入验证码 2](#_Toc31828491)

[3.提取链接 3](#_Toc31828492)

[二、实践2:检查节点连接速度 3](#_Toc31828493)

[1.节点获取 4](#_Toc31828494)

[2.小提示 4](#_Toc31828495)

[3.本地测试 4](#_Toc31828496)

[4.生存测试 5](#_Toc31828497)

[三、实践3:直接下载 6](#_Toc31828498)

[1.下载测试 6](#_Toc31828499)

[2.使用迅雷com接口 6](#_Toc31828500)

[3.使用IDM下载 6](#_Toc31828501)

[4.其他说明 7](#_Toc31828502)

[四、说明1:关于调试信息 7](#_Toc31828503)

[1. 获取调试信息 7](#_Toc31828504)

[2.读取调试信息 7](#_Toc31828505)

[五、实践4:使用蓝灯加速 8](#_Toc31828506)

[1.部署Lantern 8](#_Toc31828507)

[2.写配置 9](#_Toc31828508)

[3.使用 9](#_Toc31828509)

[六、说明2:其他配置项 9](#_Toc31828510)

[八、其他工具 10](#_Toc31828511)

[1. 站点解析 10](#_Toc31828512)

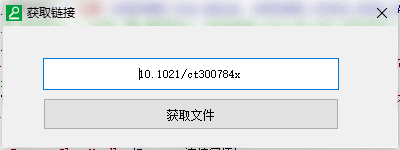
[2. X-mol检索 11](#_Toc31828513)

一、实践1:用下载器获取一个pdf

Q：SCIhub访问很慢，下个pdf简直难于上青天，Sean童鞋你平时怎么下的？

A：确实，所以我为了方便学习就编写了这样可以快速解析下载pdf的小爬虫，也本着SCIhub的精神拿出来造福其他同学，那么现在就开始学习吧

1.打开获取链接



1.1它长者本该长的样子，所以就输入本应该在SCIhub上输的东西

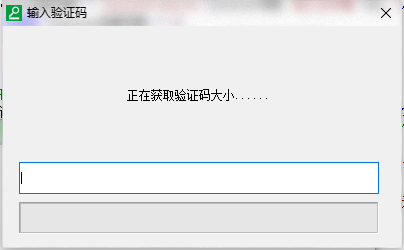
1.2输完点确定

1.3它会检测剪切板，你复制了DOI它识别后会自动填写

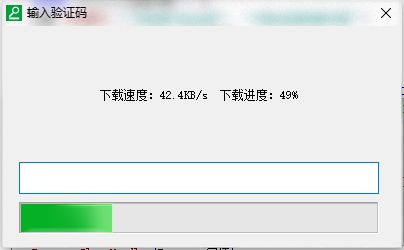
1.4输完让子弹飞一会………

2.输入验证码

2.1如果网站要求输验证码的话就会加载验证码，这个跟本身的网络环境有关，平时用的多的话就会要求输入



2.2因为软件获取验证码是多线程工作的，所以它需要先知道验证码的大小，才能把下载任务分给配不同的处理器去处理，所以你会觉得它获取的速度很快，如果你还是觉得下载的慢的话，请打开IE /黑脸



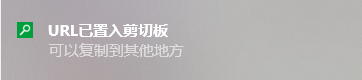
2.3加载完毕，那就输呗



2.4输完按确定，注意验证码是6位，字母之间一般是有联系的，当然，不输六位你还真填不进去 /得意

3.提取链接

然后链接就被复制到剪切板里面了，就可以粘贴到其他地方了，强烈建议用迅雷

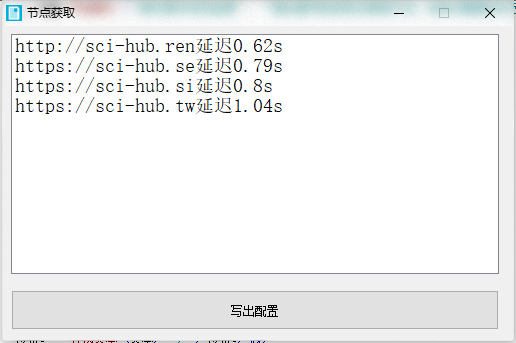


二、实践2:检查节点连接速度

Q：访问的太慢了，而且有的时候就直接提示网络错误，难道软件已经废了吗？

A：只是节点处了问题，SCIhub本来就是非法网站所以接连遭到封杀一直在换域名，你可以手动换节点，而且Sean也写个寻找节点和检测节点的功能，一起看看吧

1.节点获取



1.1原理是直接读取<https://lovescihub.wordpress.com>读取网站的内容

1.2但是据<http://tool.yovisun.com/scihub/>所说只有tw，si和se是官方网站

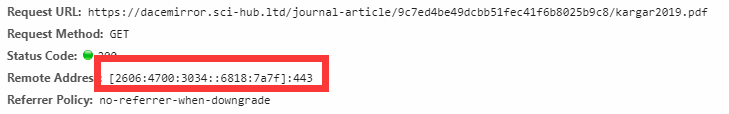
(但是Sean也只是个不明真相的吃瓜群众)

1.3网站给的延迟值比较玄学，这个延迟值也跟ISP、地域、网络环境和实时访问量等都有很强的关系所以参考价值不大

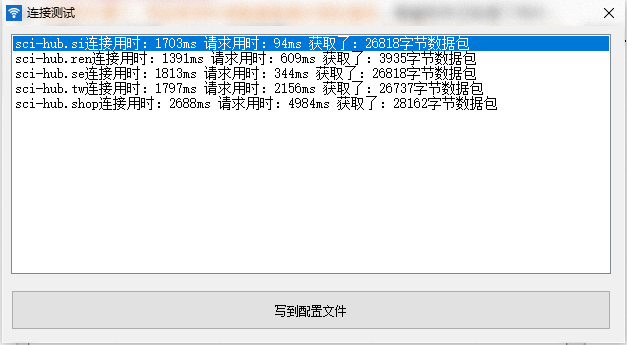
1.4所以，先别慌点那个写到配置文件,本地测试一下一探虚实

2.小提示

如果不挂VPN的话某些节点也是很快的，当然如果网络支持IPV6会更好，其实中国已经普及了IPV6，如果你的电脑不支持就是你路由器或者是光猫的锅了



3.本地测试



2.1测试得到的时间是软件内部真实的操作计时，一定程度上能反应连接的速度，有很大的参考价值，其中连接用时反应了服务器离客户的距离，请求用时反应了服务器的数据响应能力，选中一个写到配置文件就行。

Q：我知道一个更快的节点，它能否可以被软件利用？

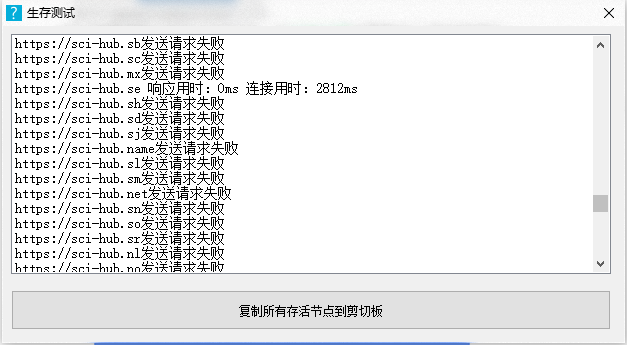
A：打开程序目录里面的host.txt，可以自己添加想要的链接，当然也可以手动删除已废弃的

链接

Q：有些链接是开头http，但是你不让用http开头的，只能填https的，你在搞什么鬼？

A：因为https是加密的http，更安全也更抗封锁一些，区别只是通讯所用端口不一样，http用的80，https用的443端口，两个本质上是访问的都是一个服务器，很多站点都不再支持http(比如Github)，我也没必要费力兼容一个不讨好的东西，当然win XP貌似原生不支持https，再当然，都9102年了你还用xp /黑人问号

4.生存测试



这个就纯属胡闹了，把节点都扫描一遍，节点列表来自[Wadauk](https://github.com/Wadauk/scihub_ck)同学，向他表示感谢

Sean在他的基础上兼容了换行符，其实用lovescihub给的几个就行

三、实践3:直接下载

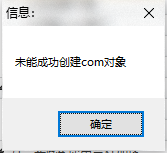
Q：每次都复制粘贴这样好麻烦，能不能直接就把文件下载下来呢？

A：Sean同学是个松鼠党+老司机，熟悉下载的他集成了主流下载接口，让我们来看看吧



1.下载测试

在写入配置文件的时候先测试东西能不能用，选中之后先运行下载测试，以免下载的时候才发现有问题错误，然而有些错误在下载的时候都发现不了的



我们想要调整下载方式的时候，点下载设置下面那几个单选框就行

2.使用迅雷com接口

迅雷作为中国下载界的老大，提供了很多接口，其中com接口还能在网上找到资料

在迅雷接受到投递的消息后做出了激烈的反应，甚至迅雷没有打开就被唤醒了



Q：我每次还是要点立即下载还是好麻烦emmmm，能不能不点就自动下载呢？

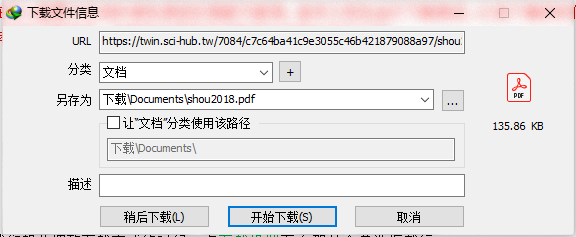
A：迅雷设置里面里面有个一键下载，记得以前叫静默下载？勾上就可以自动下载了

Q：迅雷太卡了，一开各种视频广告，电脑卡爆了，我不想用了

A：虽然是不对的，但是还是推荐[胡萝卜周](http://www.carrotchou.blog/16012.html)同学破解的迅雷X，破解的比较棒而且也没有精简com组件，另外作者也兼容了一款非常牛逼并且轻量的下载软件

3.使用IDM下载

Internet Download Manager，简称IDM，是国外的一款优秀下载工具，该软件同时是一款共享软件，(我就不复制百度百科了)，知名度非常高，网上一大把各种非常好用的注册机和破解版，提供了并不像是用来下载并且外国论坛也找不到的com接口，不过调用命令行接口还是很简单的



4.其他说明

还有一些牛逼的比如说aira2c、curl、网络快车、比特精灵等软件现在已经非常小众，特别是aira2c之前的版本有兼容rpc接口，因为使用不方便，现在已经不再兼容，当然Sean会在其他软件继续兼容rpc接口，喜欢的同学可以找我定制

四、说明1:关于调试信息

Q：能不能告诉我这个软件它让我等了这么久它到底在干什么，等了半天不会又等出个获取链接失败，我好方啊

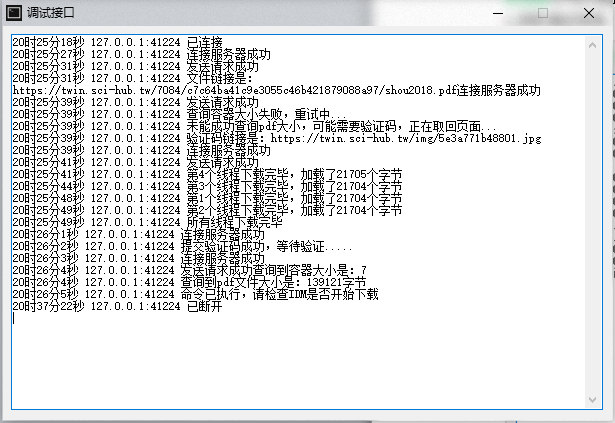
A：软件可以通过调试窗口可以告诉它正在干什么

1. 获取调试信息

1.1打开调试窗口

调试接口并不是什么高端的东西，是我自己写的，如要使用需要打开获取链接之前打开，获取链接才可以找到调试窗口并向它发送信息

2.读取调试信息



2.1这张debug interface完整的演示了文件获取的过程

连接服务器-发送请求-解析返回数据-查询链接信息-猜测是否需要验证码-查询验证码信息-多线程下载验证码-提交验证码-再次查询信息-成功后调用下载接口

Q：你是怎么想到要弄这个反人类的东西的呢？

A：因为连接质量实在是太糟糕了，用户很容易下载失败并且在非调试模式状态出现了错误也不知道哪里出了问题，不利于排查错误，开启调试窗口后下载器的信息通过TCP协议以socket客户端的形式发送到窗口上的socket服务器，这么做也是防止主程序崩溃(编写前期Sean同学被各种各样的bug无语了)后找不到错误信息，这种方案也相比写log有更强的即时性

注意：使用调试接口时请保证2333端口不被占用

五、实践4:使用蓝灯加速

Q：虽然迅雷可以让下载pdf速度变快，使用了最快节点但是软件解析的时候经常慢到无法连接，甚至经常验证码只能读到一半，好气啊

A：如果有VPN（~~very profession network~~）能够快速访问SCIhub的话将对访问连接的稳定性得到很大的帮助，VPN的使用甚至被许多科学研究人员作为必修科目，下面以蓝灯为例演示代理服务器的使用

1.部署Lantern

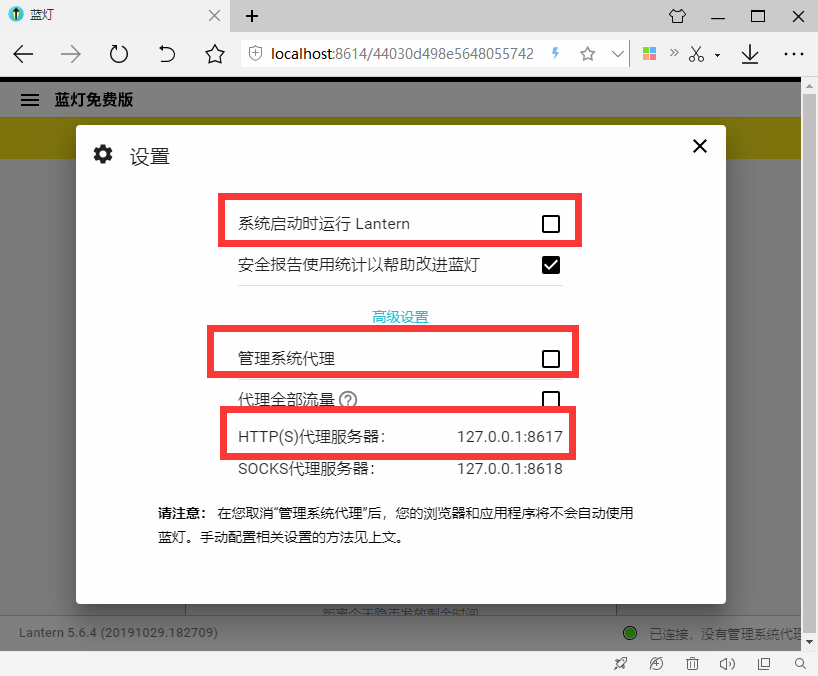
1.1在[Github](https://github.com/getlantern/lantern)上下载蓝灯

1.2安装完了以后打开

1.3右边的菜单按钮-进设置

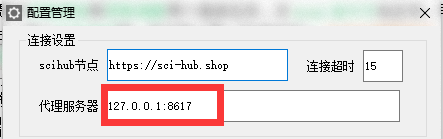


1.4取消管理系统代理（这样可以保证蓝灯只被钦定的软件调用）,取消开机启动



1.5复制https代理服务器地址（注意这个地址可能会变），填入配置管理中的代理服务器编辑框，填错直接就无法访问了，冒号是嘤文的冒号，其他软件(如Shadowsocks和SSR)同理，保存即可

2.写配置



3.使用

写入配置后这个代理是对本地测试、获取链接、简易下载和生存测试几个程序生效，对Aria2也生效，效果跟VPN本身有关，请自行斟酌

很多下载器都支持代理服务器，但不建议使用迅雷使用代理服务器，毕竟迅雷是最没节操的流氓p2p软件

六、说明2:其他配置项

Scihub节点：通过本地测试和获取链接两个软件写配置，对获取链接生效，也可以手输

连接超时：只对本地测试生效，用于定义等待服务器响应时间，单位是秒，通常情况下等待连接时间太长则连接不稳定，过短的话在网络条件不好的情况下可能会全部超时，这个根据自己的网络情况设置，毕竟你只需要在这些host中选一个最快的，默认等待15秒

代理服务器：宁愿不填也一定不要填错，没有不要填，注意冒号是嘤文的冒号

下载器目录：下载器目录只在自定义下载器选项被选中才生效

监控剪切板：系统在剪切板变更时会向软件发送消息，不勾就不监控，仅响应10.开头的DOI

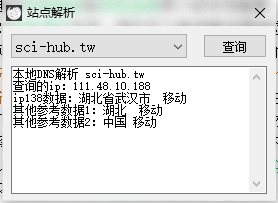
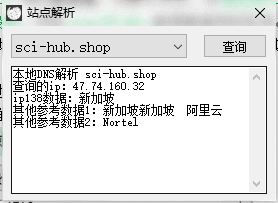
调试端口：可以随便填，只要不冲突，对调试接口和获取链接生效

RPC的设置：仅在Aria2的RPC选项被选中时生效

更新：我把程序的更新放在了配置管理的地方，其实是因为放其他地方不好看

八、其他工具

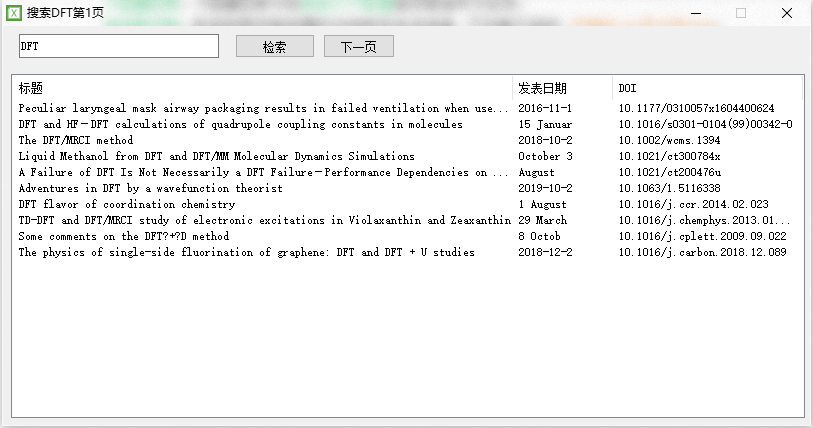
1. 站点解析

站点解析对于Sean来说是个非常有用的东西，可能因为他在襄阳，用的移动(没办法，穷)，所以访问tw就到了武汉的移动节点，在海南会解析到香港节点，shop节点解析到新加坡，在海南可以解析到日本大阪….. 其实这个跟你自己的DNS服务器有关，一般是ISP给的，路由器控制，当然你可以自定义，其实我觉得没多大用



2. X-mol检索



这个也是把X-mol封装进程序，只是方便复制DOI(对，点一下DOI就到剪切板里了)也能避免有时候X-mol网页抽风排版错乱不能使用的情况，本来想做个登录模块，但是有个技术问题搞不定，后续可能会兼容谷歌学术镜像

八、说明

1:关于软件

SCI-hub是一个好网站，但是由于地域和GWF，所以网站连接的成功率很感人，我们可以直接感觉到无论是哪个浏览器加载个pdf都非常吃力，于是为了方便学习我就编写了这样一个软件。

2.关于项目

Github项目主页为<https://github.com/panxiuwen/SCIhub_download>。

有什么好的想法可以与我交流，邮箱是[panxiuwen\_sean@qq.com](mailto:panxiuwen_sean@qq.com)

为了防止被滥用软件不开源

3.软件底层

软件使用易语言编写，不打算使用Python重写。底层使用C函数+win32API实现，网络访问全部调用wininet.dll，Sean在访问上进行优化，理论上连接速度跟系统内的IE（比如win10里的IE12）速度一模一样，但是由于少加载不少东西和渲染，判断上也比浏览器更有正对性，加上下载交给了更专业的IDM和迅雷，大大增加了网站访问的体验性

4.文件依赖

文件包括的执行文件：

主要软件：获取链接

调试软件：调试接口

配置软件：配置管理

站点测试：本地测试、节点获取、站点解析、生存测试

其他工具：X-mol检索

另外：验证码收集计划，希望大家都帮我这个忙

获取链接、站点解析、本地测试需要host.txt

生存测试需要list.txt

krnln.fnr：易语言核心支持库

EThread.fne：多线程支持库

这两个库所有exe都需要

验证码收集计划需要interne.fne互联网支持库

X-mol检索需要iext.fne扩主界面支持库