

信息检索与知识产权 技术分享

分享1：知其所以然

[一道隐形的墙 GFW](#)

[科学上网的原理](#)

[shadowsocks](#)

[全局模式与PAC模式的区别](#)

[Q&A](#)

[Links](#)

分享2：为我所用

[Zotero 软件下载](#)

[推荐理由](#)

[Links](#)

分享3：相见恨晚

[CopyTranslator](#)

[Google Colaboratory](#)

[VisuAlgo](#)

[学术导航](#)

[致谢](#)



扫描二维码在线查看本文

分享1：知其所以然



本文主要是从技术角度来了解的网络限制方式和应对方式。



科学上网的工具很多，八仙过海，各显神通，而且综合了各种技术，所以本文并不做任何科学上网工具的指导和推荐。请具备辨别真善美的能力。。。

最新的科研资讯、科研讨论、科研成果绝大多数都在国外，当你在搜索论文、查找最新资讯、甚至发表论文时都会造成困扰，所以你有必要了解一下科学上网。

科学上网的价值：为了科研，为了科研，为了科研！

一道隐形的墙 GFW

众所周知，目前国内基本无法访问[Google](#)、[Sci-Hub](#)、[Youtube](#)、[Twitter](#)、[Facebook](#)等国外的站点。原因是[长城防火墙 GFW](#)隔离了我们与外界的交流，当然这个隔离并非完全隔离，而是选择性的，因为每一个网络请求都是有数据特征的，不同的协议具

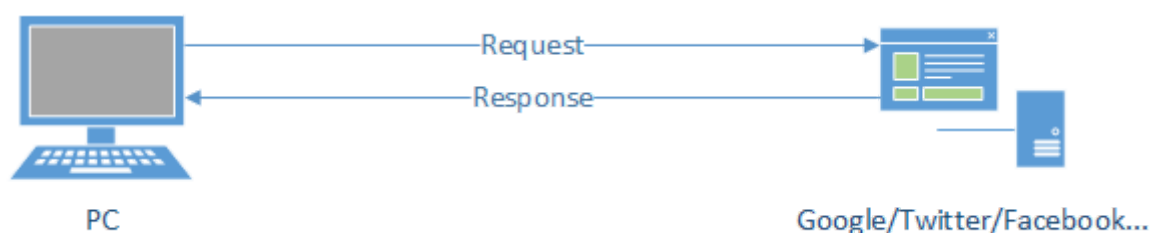
备不同的特征，比如 HTTP/HTTPS 这类请求，会很明确地告诉 GFW 它们要请求哪个域名。所以 GFW 不希望你上的网站就直接阻断无法访问。

长城防火墙 GFW包含多种封锁方式，如域名黑白名单、IP黑白名单、DNS污染、敏感词过滤、数据特征检测等等。其中最容易操作也是最基础的方式便是域名黑白名单，在黑名单内的域名不让通过，IP 黑白名单也是这个道理。如果你有一台国外服务器不在 GFW 的黑名单内，那么你就可以使用中国境内的机器跟这台国外的服务器机器进行通讯。那么一个科学上网的方案就出来了：境内设备与境外机器通讯，境内想看什么网页，就告诉境外的机器，让境外机器代理抓取，然后送回来，我们要做的就是保证境内设备与境外设备通讯时不被 GFW 怀疑和窃听。

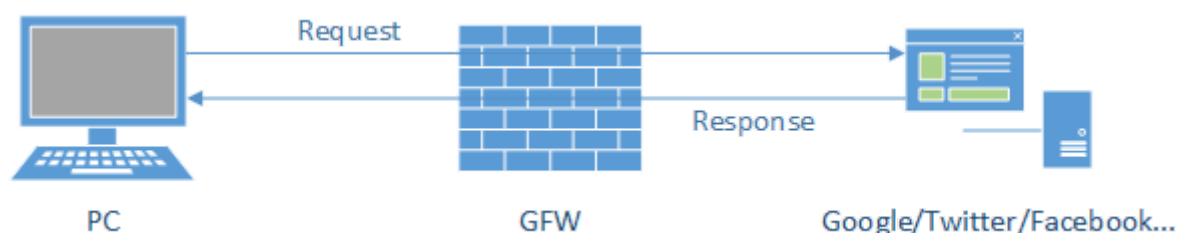
科学上网的原理

下面我们从技术的角度来探讨一下科学上网的原理：

在很久很久以前，我们访问各种网站都是简单而直接的，用户的请求通过互联网发送到服务提供方，服务提供方直接将信息反馈给用户。



后来，GFW 就出现了，他像一个收过路费的强盗一样夹在了在用户和服务之间，每当用户需要获取信息，都经过了 GFW，GFW 将它不喜欢的内容全部过滤掉，于是当客户触发 GFW 的过滤规则时，就会收到 `Connection Reset` 这样的响应内容，而无法接收到正常的内容。

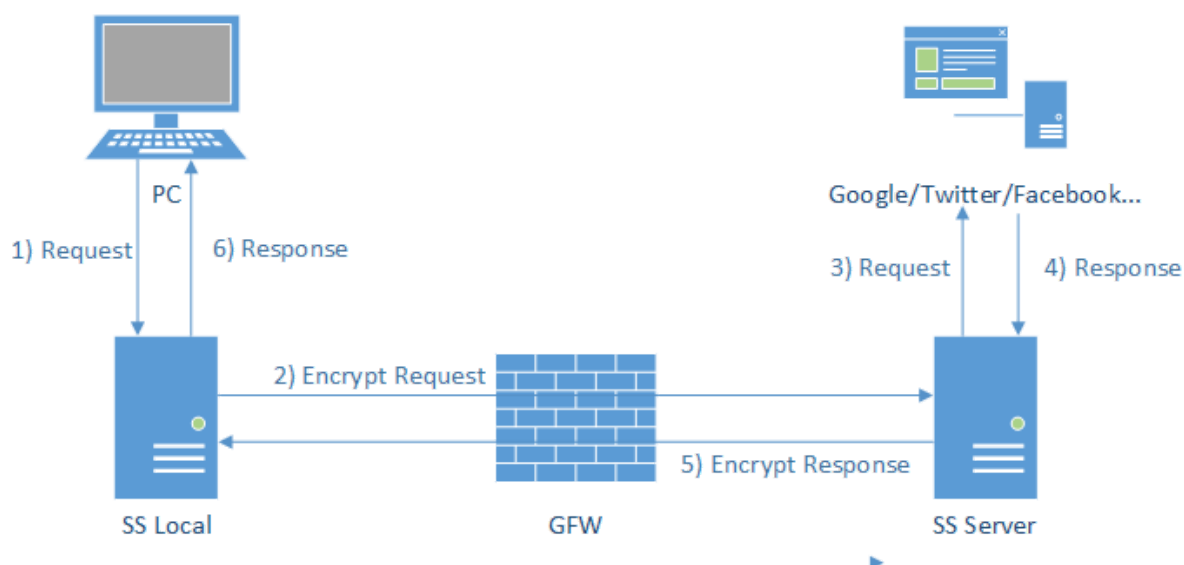


聪明的人们想到了利用境外服务器代理的方法来绕过 GFW 的过滤，其中包含了各种 HTTP代理服务、Socks服务、VPN服务、shadowsocks等。

shadowsocks

shadowsocks 是目前比较流行的用于科学上网的加密传输协议。

shadowsocks 的原理图如下：



1. PC客户端（即你的电脑）发出请求基于Socks5协议跟SS-Local端进行通讯，由于这个SS-Local一般是本机或路由器等局域网的其他机器，不经过GFW，所以解决GFW通过特征分析进行干扰的问题。
2. SS-Local和SS-Server两端通过多种可选的加密方法进行通讯，经过GFW的时候因为是常规的TCP包，没有明显特征码GFW也无法对通讯数据进行解密，因此通讯放行。
3. SS-Server将收到的加密数据进行解密，还原初始请求，再发送到用户需要访问的服务网站，获取响应原路再返回SS-Local，返回途中依然使用了加密，使得流量是普通TCP包，并成功穿过GFW防火墙。

全局模式与PAC模式的区别

简单地说，在全局模式下，所有网站默认走代理。而PAC模式是只有被墙的才会走代理，推荐PAC模式，如果PAC模式无法访问一些网站，就换全局模式试试，一般是因为PAC更新不及时（也可能是GFWList更新不及时）导致的。

Q&A

Q1：翻墙是否违法？

A1：是的，国内明文规定了：计算机信息网络直接进行国际联网，必须使用邮电部国家公用电信网络提供的国际出入口信道，任何单位和个人都不得使用其他信道进行国际联网。违反者由公安机关责令停止联网，给予警告，可以并处15000元以下的罚款。

Q2：中国安全科学上网上网须知

A2：不要在国内外网络散布政治信息；警惕网上所谓“免费VPN”，“免费节点”；不要利用科学上网工具进行牟利。

Q3：我们是否需要长城防火墙？

A3：西方对中国充满敌意，这是这个世界的基本事实。几乎所有西方媒体都以片面报道或完全扭曲任何和中国有关的新闻来获取大批观众。所以我们是需要长城防火墙的。

Q4：长城防火墙的存在对我们造成了影响？

A4：长城防火墙这种管制确实会严重阻碍优质信息的传播，根本上是反生产力的，尽管对许多中国很多网民来说，长城防火墙并不能阻止他们翻墙获取高价值信息，但总有一些年轻国人的视野因此受限，他们会失去很多机会，获取不到成长过程中他们需要的那些随机但却可能改变他们生活的知识。

Q5：为什么你还没能成功科学上网？

A5：你也许对世界并不感兴趣，觉得中国的是最好的，也许你并非这样觉得，而只是感到中国的够用了。的确，有很多东西，中国的确实是最好的，但有很多东西，世界的更好，好很多，你得尝试往外边看它们一眼。很多时候的确中国的够用了，但如果我们让生活停留在够用，就会错过它是多么丰富。

Links

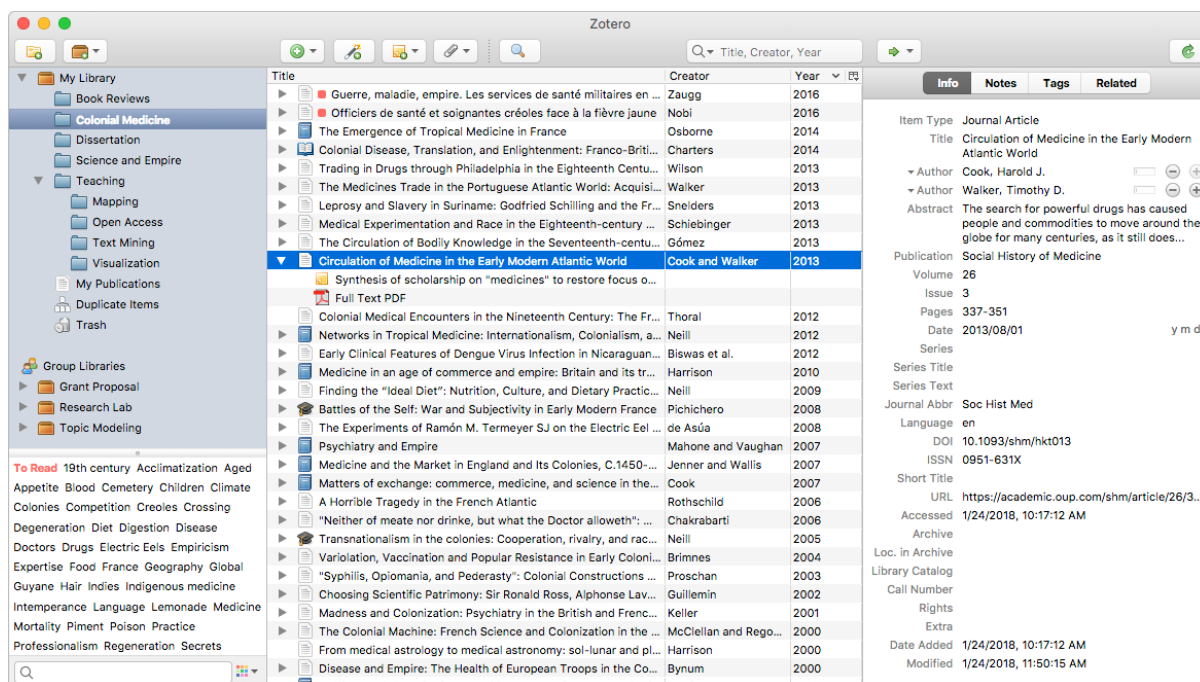
1. 什么是长城防火墙 GFW
2. 科学上网科普好文
3. <https://github.com/getlantern/lantern>

分享2：为我所用

文献是科研工作者最重要的数据资源，阅读文献几乎是科研工作者每天都要做的事情，文献是知识的来源、灵感的来源。搭建最佳的文献生态，服务于我们的科研工作，有助于我们构建知识结构、整理思绪、论文写作。

Zotero 软件下载

Zotero 是一款开源免费、跨平台的知识管理软件，[Zotero官方网站](#)提供下载。



推荐理由

Zotero 几乎满足了我所有的期待，利用它们可以打造最佳的文献管理生态、知识管理生态。以下列举一些Zotero的亮点和特色：

1. 在浏览器端利用Zotero Connector，可以实现对各大主流数据库（Web of Science、Pubmed、ScienceDirect等）或搜索引擎（Google Scholar、百度学术、Research Gate等）具有文献一键抓取能力，甚至支持豆瓣图书的抓取。抓取内容包括文献元数据、PDF、网页链接等。
2. 跨平台实时同步功能，便于随时随地在多平台访问文献。
3. 可与第三方云存储（如坚果云）绑定，达到无限扩展Zotero存储空间的目的，在我看来这是Zotero最具特色的功能了。
4. 作为一款开源软件，Zotero的更新速度很快，保证了它的使用体验。Zotero占用空间很小，运行起来非常流畅，这一点比Endnote有优势。
5. Zotero支持插件扩展，Zotero社区开发出了非常多实用的插件，这使得Zotero更加强大，如PDF文件自动命名功能。
6. Zotero提供分组、标签等功能，方便了对文献的管理。灵活使用标签功能，可以大大提高文献管理的效率。
7. 提供Windows和Mac端的Office插件，方便学术论文写作。Zotero支持非常丰富的期刊参考文献格式，写论文时不用再愁了。
8. Zotero还支持RSS订阅，这一功能使得Zotero能够把学术文献和各种其他资讯集中在一块，可以说融合了科研+生活。

Links

1. [Zotero官网](#)
2. [Zotero系列文章](#)
3. [zotero-shortdoi + Sci-Hub, 让99%的文献都能被免费下载!](#)
4. [Zotero | 用坚果云无限扩展文献存储空间](#)
5. [详解【Zotero+PaperShip+坚果云】文献生态的同步机制, 在iPad上看文献!](#)
6. [Zoo for Zotero, 安卓上阅读Zotero文献的利器!](#)

分享3: 相见恨晚

CopyTranslator

由于论文通常以pdf的形式出现, 而双栏的pdf (单栏的也可能出现类似现象) 在复制的时候又会出现回车(换行)的现象, 这使得我们的复制到各种翻译网页后翻译的效果很差, 需要我们手动删除换行, 才能得到正确的翻译结果。

CopyTranslator诞生时主要是为了解决外文pdf的阅读翻译问题, 但是现在他功能越来越强大, 在网页翻译, 以及其他形式的文本翻译中均表现出色。



官网: <https://copytranslator.github.io/>

Google Colaboratory

Colaboratory 是一个 Google 研究项目, 旨在帮助传播机器学习培训和研究成果。它是一个 Jupyter 笔记本环境, 不需要进行任何设置就可以使用, 并且完全在云端运行。Colaboratory 笔记本存储在 [Google 云端硬盘](#) 中, 并且可以共享, 它支持 Python 2和Python 3, 还包括免费的GPU加速。

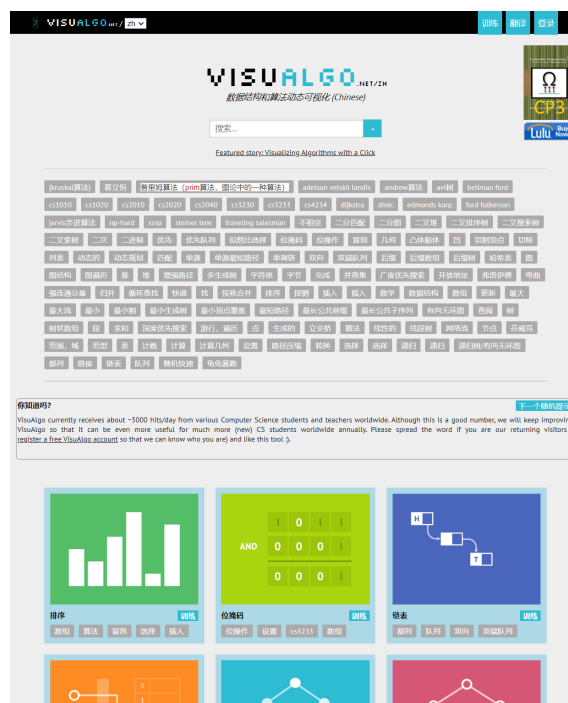


VisuAlgo

推荐一个神奇的数据结构和算法动态可视化学习工具——VisuAlgo。VisuAlgo的用户界面简洁友好，演示生动形象。目的是作为一种自学工具帮助学生更好的理解数据结构与算法。

在功能上它不仅支持暂停、单步、回退等功能，演示算法，查看算法代码的执行过程，更新增加了在线测试功能让大家更加方便测试所学结果。现在的网站还有了中文，更加方便大家学习。

VisuAlgo包含很多高级算法。除了常规的深度优先算法与广度优先算法，我们还设计了深度优先算法的变形以及找割点，连接桥，和有向图中寻找强连接部分的Tarjan's与Kosaraju的深度优先算法。我们甚至包含了2-SAT实例的可视化。



官网：<https://visualgo.net/zh>

学术导航

善用学术导航能降低的资源搜索成本，也降低了使用互联网的门槛，大家根据需要使用以下推荐的学术导航，**请具备辨别真善美的能力**👁️

- [龙猫学术导航](#)
- [大木虫学术导航](#)
- [沃客学术导航](#)
- [虫部落学术导航](#)

致谢

很高兴您能看到最后，希望本文分享能对你有所启发！😊😊😊