

计算机网络 课程实验报告

实验名称	HTTP 代理服务器的设计与实现					
姓名	任永鹏		院系	计算机		
班级	1737102		学号	1173710107		
任课教师	张英涛		指导教师	张英涛		
实验地点	格物 213		实验时间	周六 12 节		
实验课表现	出勤、表现得分(10)		实验报告		实验总分	
	操作结果得分(50)		得分(40)		关 视心力	
教师评语						

实验目的:

本次实验的主要目的。

熟悉并掌握 Socket 网络编程的过程与技术;深入理解 HTTP 协议, 掌握 HTTP 代理服务器的基本工作原理; 掌握 HTTP 代理服务器设计与 编程实现的基本技能

实验内容:

概述本次实验的主要内容,包含的实验项等。

- (1) 设计并实现一个基本 HTTP 代理服务器。
- (2) 设计并实现一个支持 Cache 功能的 HTTP 代理服务器。
- (3) 扩展 HTTP 代理服务器,
- a) 网站过滤: 允许/不允许访问某些网站;
- b) 用户过滤: 支持/不支持某些用户访问外部网站;
- c) 网站引导:将用户对某个网站的访问引导至一个模拟网站(钓鱼)。

实验过程:

(1)将浏览器的局域网设置里将代理服务器设置成127.0.0.1:8080 的ip和端口,从而将浏览器的所有请求发送给自己编写的服务器程序。

首先建立一个ServerSocket,这个ServerSocket是负责监听指定的端口,调用ServerSocket的accept()函数,进入阻塞状态,如果有Sockect连接建立成功,就返回建立成功的Socket对象,命名为Client。

```
ServerSocket server = new ServerSocket(PORT);
Socket client;
while (true) {
   client = server.accept();
```

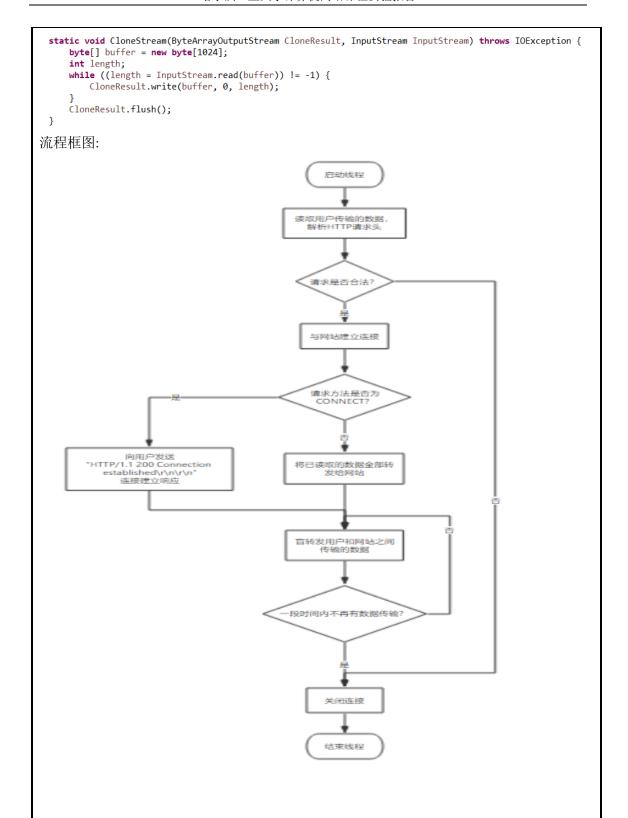
之后通过读取Client 的InputStream输入流,接收到从浏览器发送来的http数据报。将接收的字节流转换成字符串流(即使用BufferReader),然后依据每行进行读写,使用正则表达式Host:\\s*([\\w.-]*)来匹配http报文头部中的host字段,之后建立Server Socket,host为从http协议头部匹配到的host,端口号为http协议服务器固定端口号80。之后将从Client获得的http报文全部转发给Server。

```
BufferedReader ObtainReader = new BufferedReader(new InputStreamReader(new ByteArrayInputStream(ClientCache.
```

这样代理服务器与客户端和目标服务器的两个Socket都建立好了,之后只需要将从目标服务器发来的数据全部转发给客户端,同时将从客户端发送的数据全部转发给目标服务器,直到客户端和目标服务器有一方关闭了连接,则中断双方的通信,结束转发。

```
if (!FireWall.isClientForbidden(client.getInetAddress().getHostName()))
    new Thread(new Client(client, FireWall)).start();
else {
    client.close();
```

在实现的过程中我对inputStream的运行不太清晰,遇到了读取inputStream无法结束的问题。通过查阅资料,发现Socket的inputStream的数据传输与读取文件的不同,因为读取文件时,当读取到文件的结尾时,inputStream会结束,但是Socket的inputSream只有在连接结束时才会结束,所以要想结束读取,我设置了Socket的 outTime,当超时时会抛出SocketTimeoutException异常,通过捕捉这个异常就可以退出inputStream的读取。



HTTPS代理服务器的实现与HTTP代理服务器的实现大体相同,但是在转发过程中使用的是HTTP协议的隧道技术。在一开始Client建立与代理服务器的连接之后,发送Connect HTTP数据报,告知代理服务器目标服务器的 host和port,然后代理服务器需要与制定的host和port建立TCP链接,如果链接成功,就返回给Client一个HTTP Response:HTTP/1.1 200 Connection established\r\n\r\n, 告知Client连接成功,之后同HTTP代理服务器,将Client发送给代理服务器的信息全部交付给目标服务器,将目标服务器发送来的信息全部交付给Client。

在代理服务器收到来自Server的HTTP Response 时,将请求的URL作为HashMap的key,将响应的内容作为value。当客户端再次请求相同的URL时,代理服务器根据URL取出对应的缓存信息,提取出其中的Last-Modified字段的日期,向目标服务器发送一个HTTP请求,使用GET方法,包含If-Modified-Since:日期,如果服务器返回304状态的响应,则代理服务器直接将缓存的副本返回给Client,如果目标服务器返回的是正常的HTTP响应,则代理服务器将新的数据返回给Client,并更新对应的缓存。

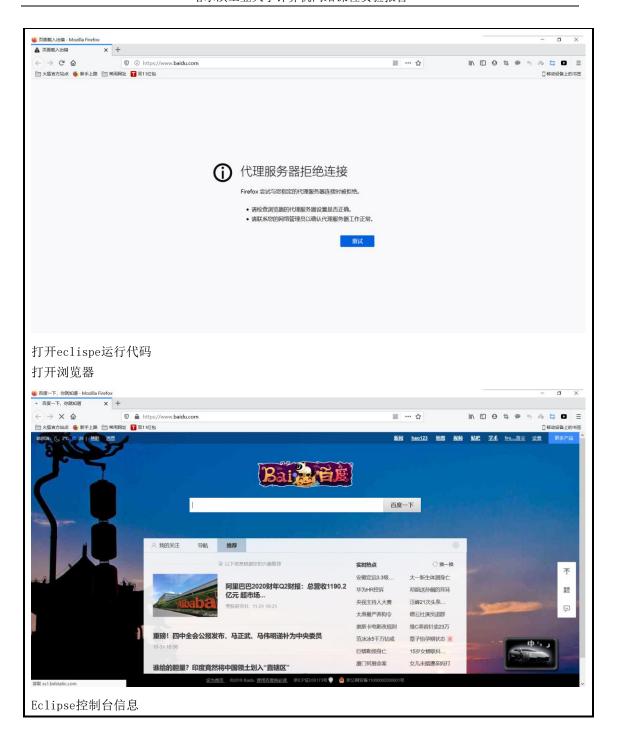
拓展代理服务器功能

- (a) 为了实现屏蔽用户访问某些网站的功能,我建立了一张屏蔽Host的Json格式的文件,然后使用程序读取并解析这个Json文件,从中提取出需要屏蔽的Host,在Client建立链接并发送请求后,解析出请求的Host,判断该Host是否在屏蔽列表中,如果在其中,就立即中断与Client的连接,不再代理转发请求。
- (b) 为了实现屏蔽某些用户访问代理服务器,我在屏蔽Host的列表中建立了一张屏蔽制定用户的列表,处理方式同上,在Client与代理服务器建立链接后,首先检查Client Socket的 Host,如果在屏蔽的列表中,就立即中断与该Client的连接。
- (c) 钓鱼网站的实现:在Client与代理服务器建立连接后,并不将请求转发给目标服务器,而是直接返回一个自己构造好的HTTP响应,然后Client就收到了一个钓鱼网站,而不是真正的网站。

实验结果:

采用演示截图、文字说明等方式,给出本次实验的实验结果。 开启代理:





```
    Problems @ Javadoc   □ Declaration □ Console 
    Console 
    Declaration □ Console 
    Console 

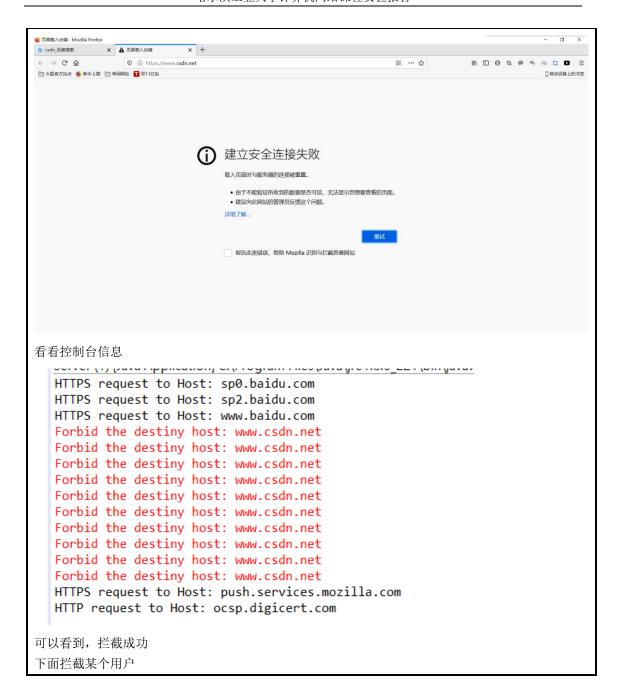
  Server (1) [Java Application] C:\Program Files\Java\jre1.8.0 221\bin\javaw.exe (2019年11月1日下午8:06:58)
  Listening to the port: 8080
  connect: Address is invalid on local machine, or port is not valid on remote machine
  connect: Address is invalid on local machine, or port is not valid on remote machine
  HTTP request to Host: detectportal.firefox.com
  HTTP request to Host: detectportal.firefox.com
  HTTPS request to Host: www.baidu.com
  connect: Address is invalid on local machine, or port is not valid on remote machine
  HTTPS request to Host: ss0.bdstatic.com
  HTTPS request to Host: www.baidu.com
  HTTPS request to Host: ss0.baidu.com
  HTTPS request to Host: ss1.bdstatic.com
  HTTPS request to Host: ss1.bdstatic.com
  HTTPS request to Host: ss1.bdstatic.com
  HTTPS request to Host: sp1.baidu.com
  HTTPS request to Host: ss2.baidu.com
  HTTPS request to Host: ss1.baidu.com
 HTTPS request to Host. ss3 haidu com
屏蔽某些网站的访问:
在config.json中,我添加了
     "forbidden_hosts": [
            "www.csdn.net",
     "forbidden_clients": [
            "#127.0.0.1"
     ]
准备打开csdn
🔞 csdn_百度搜索 - Mozilla Firefo:
  🐕 csdn_百度搜索
 ← → ℃ ☆

    https://www.baidu.com/s?wd=csdn&trsv_spt=1&trsv_iqid=0xd06ce0c3001b4d85&issp=1&tf=8&trsv_h

                                                                                                                                                                                                      II\ □ Θ ≒ ୭ % ≒ □ ≡
 □ 火狐官方站点 🏺 新手上路 🗎 常用网址 📘 双11红包
                                                                                                                                                                                                                                 □移动设备上的书签
 Bai 都百度 csdn
                                                                                                                                                                                                                  百度首页 消息 设置 - fre...落全 -
               网页资讯 视频 图片知道文库 贴吧 采购 地图 更多»
                    CSDN-专业IT技术社区 1977
                    COLON 文 2111以不正位。

在書間機能出品CSDN(ID CSONnews)近日、毕业于加州大学高杉矶分校(UCLA),有着十几年的
芯片宁坦经验的咖啡大下手用高片达、个糠样称、半为自研的PA(PowerAmplifor...

C CSDN技术杠区 = - 百度快丽 - 177条评价
                     CSDN下载
400W+各类资源,技术大牛的分享平台
                                                                 CSDN学院
                                                                  覆盖17+技术领域,近万门IT学习资源,IT
从业者的学习平台!
                     CSDN博客 Csdn问答
专业T技术平台、支持ckedilor及markdown
格式,300万技术博主的共同选择 你的技术难题
                                                                              牛为你答娶解惑,第一时间解决
                   ○ 为参推荐: csdn下载 CSDN官网 csdn官网下载 github中文版 github官网中文 csdn官网管录 csdn官网管页 csdn手机版写博客 csdn是什么好不好
                    CSDN论坛-IT技术交流平台
                    活動性木包収・知道的都不知道推荐 CSDN云计算性目开新 【只送
転割」10天之即 事本乐园 ・NANU-NANA推荐(時勝)22天之前
選水伝図 ・NANU-NAN。
で CSDN様士松図 ・ 高校徳田 - 772条评位
                    中国专业下社区CSDN (Chinese Software Developer Network) 创立于1999年,数力于为中国
软件开发者提供电讯时候属。在线学习、职业发展等全生命周期服务。据下胸有
公司太事记 服务项目 注册方法 紀分規則 上传协议 下载分类 社区平台 更多>>
                                                                                                                                        搜索热点
                                                                                                                                                                                        10-10-C
                                                                                                                                       11章子怡孕期状态 55
                   CSDN下载-IT下载大本营
```





对实验过程中的思考问题进行讨论或回答。

在读取socket的inputStream时,出现了会不断循环的情况。

上网查找相关问题后发现,Socket的inputStream的数据传输只有在连接结束时才会结束,要想结束读取,必须设置计时器。设置了Socket的 outTime,当超时时会抛出SocketTimeoutException异常,通过捕捉这个异常就可以退出inputStream的读取。



心得体会:

结合实验过程和结果给出实验的体会和收获。

实验设计非常好,就是感觉平时看mooc和实验有点脱轨,直接上来编程有点难受

