

# 哈尔滨工业大学

## <<计算机网络>>

### 实验报告

(2016 年度春季学期)

姓名:	易亚玲
学号:	1170300511
学院:	计算机科学与技术学院
教师:	聂兰顺

# 实验一 HTTP 代理服务器的设计与实现

## 一、实验目的

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。

## 二、实验内容

(1) 设计并实现一个基本 HTTP 代理服务器。要求在指定端口（例如 8080）接收来自客户的 HTTP 请求并且根据其中的 URL 地址访问该地址所指向的 HTTP 服务器（原服务器），接收 HTTP 服务器的响应报文，并将响应报文转发给对应的客户进行浏览。

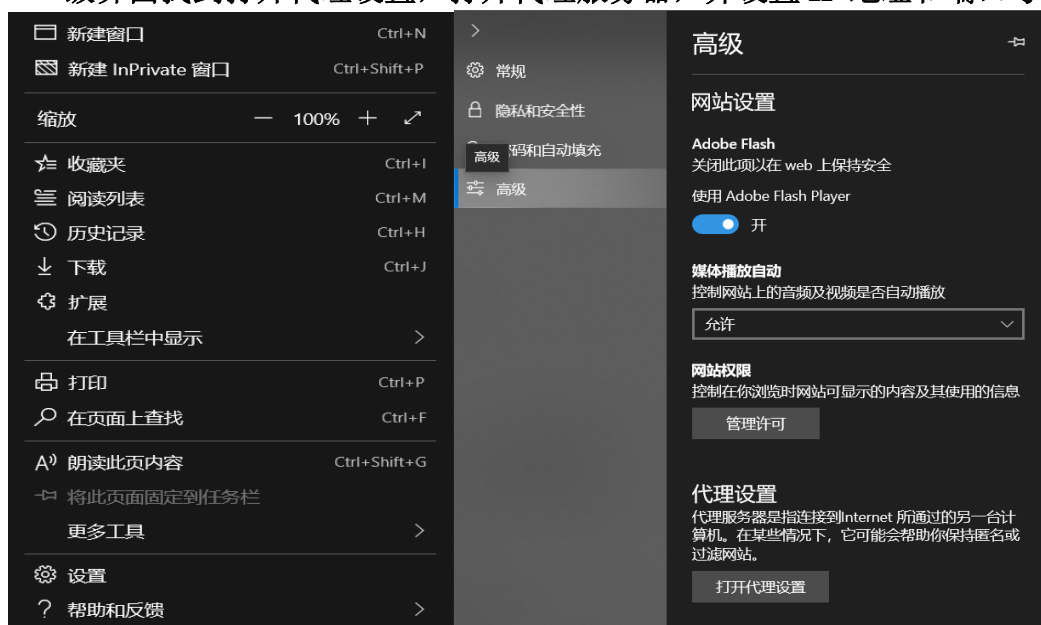
(2) 设计并实现一个支持 Cache 功能的 HTTP 代理服务器。要求能缓存原服务器响应的对象，并能够通过修改请求报文（添加 if-modified-since 头行），向原服务器确认缓存对象是否是最新版本。（选作内容，加分项目，可以当堂完成或课下完成）

(3) 扩展 HTTP 代理服务器，支持如下功能：（选作内容，加分项目，可以当堂完成或课下完成）

- 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）

## 三、实验过程及结果

实验前准备：打开 IE 浏览器，点击右上角的三点，进入设置界面，在高级界面找到打开代理设置，打开代理服务器，并设置 IP 地址和端口号





基本功能：正常的网页浏览，e.g. Http://cs.hit.edu.cn

```
尝试连接: http://cs.hit.edu.cn/_upload/tpl/02/b2/690/template690/images/menu_line.png
CONNECT iecvlist.microsoft.com:443 HTTP/1.0
User-Agent: Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/70.0.3538.102 Safari/537.36 Edge/18.18362
Host: iecvlist.microsoft.com:443
Content-Length: 0
Proxy-Connection: Keep-Alive
Pragma: no-cache
```




打开 cs.hit.edu.cn 后，会在 cache 文件夹里面缓存 cs.hit.edu.cn 返回的消息

```
proxyServer.py x cs.hit.edu.cn_ x filter.json x
1 HTTP/1.1 200 OK
2
3 Date: Thu, 31 Oct 2019 13:50:36 GMT
4
5 Server: Apache
6
7 X-Frame-Options: SAMEORIGIN
8
9 Frame-Options: SAMEORIGIN
10
11 Accept-Ranges: bytes
12
13 Connection: close
14
15 Transfer-Encoding: chunked
16
17 Content-Type: text/html
18
```

过滤网址:

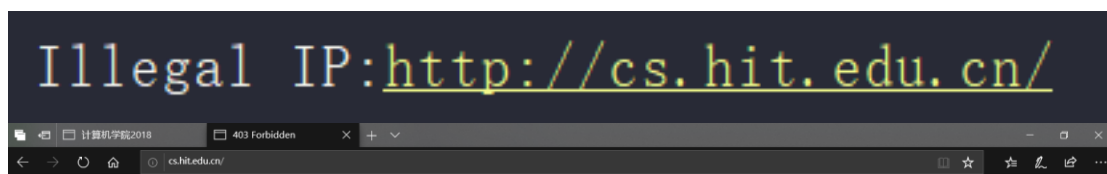
```
Proxy server is ready to receive message:
Denied http://sa.hit.edu.cn/
```



**404 Not Found**

Denied By Root User ZMY

过滤 IP :e.g. 127.0.0.1



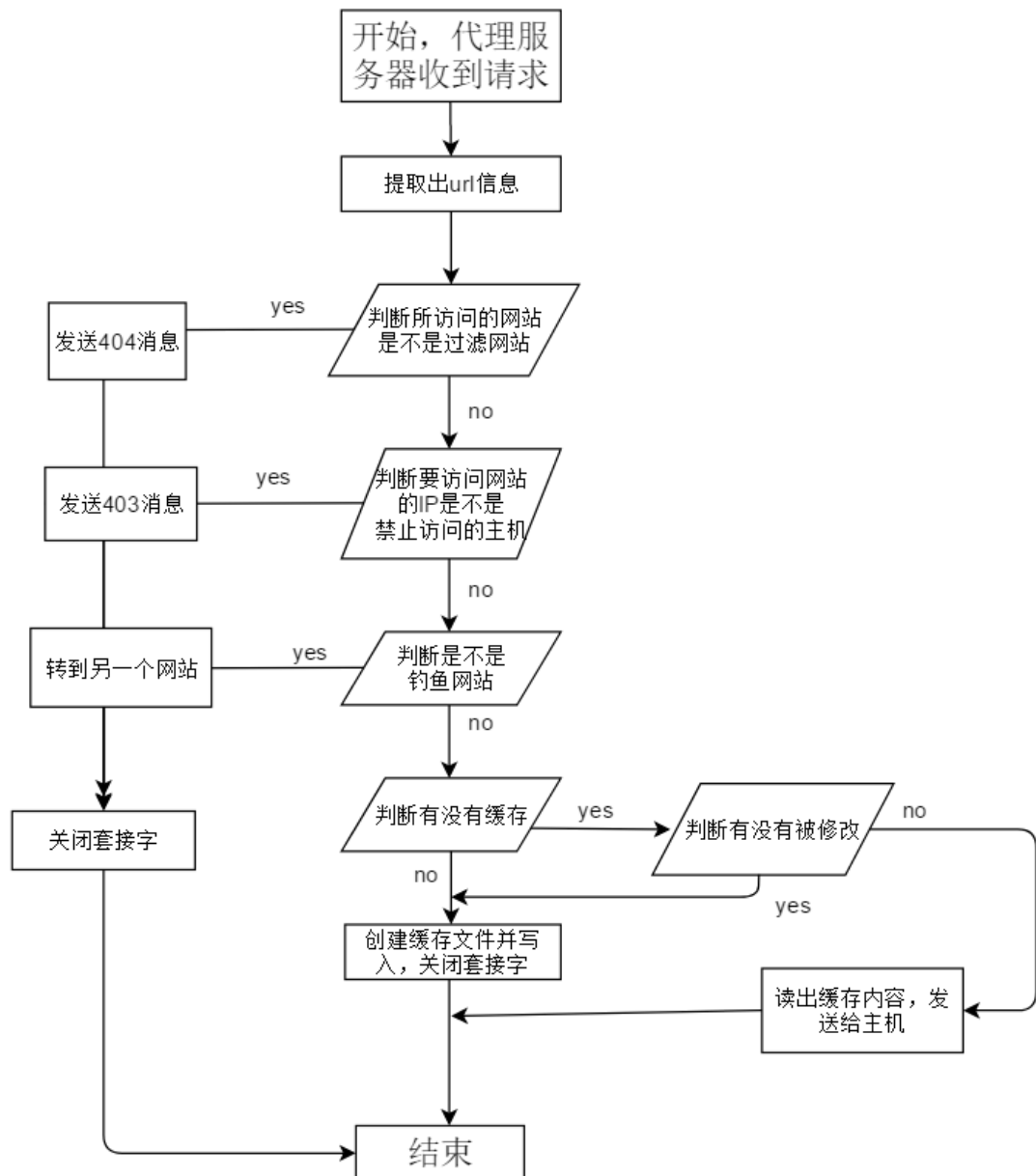
## 403 Forbidden

Denied By Root User YYL yi

钓鱼：访问<http://math.hit.edu.cn>时转到[cs.hit.edu.cn](http://cs.hit.edu.cn)（由于网址服务器的原因，有时候接收不完全）



程序流程图（tcp 连接的函数）：



## 四、实验心得

- 通过本次实验更加深刻地理解了 HTTP 协议下，客户机和服务器之间的交流过程
- 学会用 python 去实现一个简单的代理服务器，实现一些简单的功能比如过滤网站、过滤 IP 以及钓鱼