

README

1. Program:

My folder contains a Makefile, and ProxyServer.java includes the main method.

To start the program, an argument is needed which is the file name.

Like: java ProxyServer config

Main method is in ProxyServer.java, which reads the configure file to get the port and block list. Then start the server on the port above, accept multiple client requests.

HTTPServerThread deals with the processing of requests from clients:

- 1: read input stream from client, save the information in sumInfo;
- 2: judge whether header is http, deny https, ftp, file
- 3: deny POST, HEAD, CONNECTION, PUT, DELETE, TRACE
- 4: when "GET" is got, call the function "get", deal with the GET request
- 5: when "QUIT" is got, terminate the server
- 6: in function "get", get url and potential port number from the input information.

Then send request to server, get response from server, send response to client.

2. What I achieved and not achieved

1. Read configure file to get port and block list
2. Get request from client and request to web server
3. Forward web sites response to client
4. Multiple client connection simultaneously
5. Block web sites which is listed in configure file

3. Screen shot:

Make file:

```
hao@hao-ThinkPad-T440p: ~/Documents/Program/Network/proxy
hao@hao-ThinkPad-T440p:~/Documents/Program/Network/proxy$ make
javac HTTPServerThread.java
javac ProxyServer.java
hao@hao-ThinkPad-T440p:~/Documents/Program/Network/proxy$ java ProxyServer config
port:8888
***server starts!***
```

https is denied:

```
hao@hao-ThinkPad-T440p: ~
hao@hao-ThinkPad-T440p:~$ telnet 127.0.0.1 8888
Trying 127.0.0.1...
Connected to 127.0.0.1.
Escape character is '^]'.
GET https://www.google.com/
Connection: close

SORRY, "https" is not allowed!
```

ftp is denied:

```
hao@hao-ThinkPad-T440p: ~  
hao@hao-ThinkPad-T440p:~$ telnet 127.0.0.1 8888  
Trying 127.0.0.1...  
Connected to 127.0.0.1.  
Escape character is '^]'.  
ftp:/test.txt  
SORRY, "ftp" is not allowed!
```

www.bing.com is blocked in configure, which return 403:

```
hao@hao-ThinkPad-T440p: ~  
hao@hao-ThinkPad-T440p:~$ telnet 127.0.0.1 8888  
Trying 127.0.0.1...  
Connected to 127.0.0.1.  
Escape character is '^]'.  
GET http://www.bing.com/ HTTP/1.1  
Connection: close  
403! Forbidden
```

www.baidu.com is also blocked:

```
hao@hao-ThinkPad-T440p: ~  
hao@hao-ThinkPad-T440p:~$ telnet 127.0.0.1 8888  
Trying 127.0.0.1...  
Connected to 127.0.0.1.  
Escape character is '^]'.  
GET http://www.baidu.com/ HTTP/1.1  
Connection: close  
403! Forbidden
```

CONNECTION is denied:

```
hao@hao-ThinkPad-T440p: ~  
hao@hao-ThinkPad-T440p:~$ telnet 127.0.0.1 8888  
Trying 127.0.0.1...  
Connected to 127.0.0.1.  
Escape character is '^]'.  
CONNECTION: www.baidu.com  
Sorry, only GET method is allowed!
```

POST is denied:

```
hao@hao-ThinkPad-T440p: ~  
hao@hao-ThinkPad-T440p:~$ telnet 127.0.0.1 8888  
Trying 127.0.0.1...  
Connected to 127.0.0.1.  
Escape character is '^]'.  
POST: www.google.com  
Sorry, only GET method is allowed!
```

HEAD is denied:

```
hao@hao-ThinkPad-T440p: ~  
hao@hao-ThinkPad-T440p:~$ telnet 127.0.0.1 8888  
Trying 127.0.0.1...  
Connected to 127.0.0.1.  
Escape character is '^]'.  
HEAD: www.google.com  
Sorry, only GET method is allowed!
```

Successful connect to www.greens.org and get response:

```
hao@hao-ThinkPad-T440p: ~
hao@hao-ThinkPad-T440p:~$ telnet 127.0.0.1 8888
Trying 127.0.0.1...
Connected to 127.0.0.1.
Escape character is '^]'.
GET http://www.greens.org/about/software/editor.txt HTTP/1.1
Connection: close

HTTP/1.1 200 OK
Server: Apache/2.2.22 (Debian)
ETag: "25c8aa-15f5-47ca0b5e2e4c0"
Accept-Ranges: bytes
Vary: Accept-Encoding
Keep-Alive: timeout=5, max=100
Content-Length: 5621
Content-Type: text/plain
Last-Modified: Fri, 08 Jan 2010 05:46:35 GMT
Via: 1.1 ironport.net.rochester.edu:80 (Cisco-WSA/5.0.0)
Connection: close

<title> Text editor?  What?  </title>

<body>
<ul>

<h3> What's text? </h3>
<p>
Computer data are often stored in a simple format called <em>plain text</em>.
In a plain text file, one byte of data (eight bits) represents a character
of text.  The most common computer alphabet began as the American
Standard Code for Information Interchange (<a
href=
"http://web.greens.org/~cls/linux/ascii.7.pdf"
>ASCII</a>) and has expanded to
<a href="http://en.wikipedia.org/wiki/UTF-8">accomodate the rest of the world's
alphabets</a>
as well.
</p>

<p>
Messages between computers are often in plain text, too.
Email messages are plain text.  <a href="editor.txt">HTML Web pages are plain te
xt.</a>
</p>

<h3> What's a text editor? </h3>
A program for creating and changing text files is called a text editor.
A text editor can be used to write email, compose Web pages, create
computer programs, and install the little files that control them.


<p>
Not so long ago, using a text editor was one of the first skills
in computer literacy.
</p>
```

QUIT:

```
hao@hao-ThinkPad-T440p:~/Documents/Program/Network/proxy$ java ProxyServer confi
g
***server starts!***
client number: 1
client number: 2
Client input:QUIT

****SERVER IS TERMINATED BY CLIENT!****
```