暨南大学本科实验报告专用纸

课程名称 《计算机网络实验》 成绩评定

实验项目名称 http协议分析 指导教师 雷小林 、魏林锋

实验项目编号 实验项目类型 综合 实验地点 N117

学生姓名 陈宇 学号 2020101642

学院 信息科学技术学院 系 计算机系 专业 软件工程

实验时间 2022 年 11 月 29 日 下 午～ 11 月 29 日 下 午

1. HTTP 协议简介

HTTP 是超文本传输协议 （Hyper Text Transfer Protocol）的缩写，用于 WWW 服务。

（1）HTTP 的工作原理

HTTP 是一个面向事务的客户服务器协议。尽管 HTTP 使用 TCP 作为底层传输协议，但HTTP 协议是无状态的。也就是说，每个事务都是独立地进行处理。当一个事务开始时，就在万维网客户和服务器之间建立一个 TCP 连接，而当事务结束时就释放这个连接。此外，客户可以使用多个端口和和服务器（80 端口）之间建立多个连接。其工作过程包括以下几阶段。

① 服务器监听 TCP 端口80，以便发现是否有浏览器 （客户进程）向它发出连接请求；

② 一旦监听到连接请求，立即建立连接。

③ 浏览器向服务器发出浏览某个页面的请求，服务器接着返回所请求的页面作为响应。

④ 释放 TCP 连接。

在浏览器和服务器之间的请求和响应的交互，必须遵循 HTTP 规定的格式和规则。

当用户在浏览器的地址栏输入要访问的 HTTP 服务器地址时，浏览器和被访问 HTTP 服务器的工作过程如下：

① 浏览器分析待访问页面的 URL 并向本地 DNS 服务器请求 IP 地解析；

② DNS 服务器解析出该 HTTP 服务器的 IP 地址并将 IP 地址返回给浏览器；

③ 浏览器与 HTTP 服务器建立 TCP 连接，若连接成功，则进入下一步；

④ 浏览器向HTTP服务器发出请求报文（含GET 信息，）请求访问服务器的指定页面；

⑤ 服务器做出响应，将浏览器要访问的页面发送给浏览器，在页面传输过程中，浏览器会打开多个端口，与服务器建立多个连接；

⑥ 释放 TCP 连接；

⑦ 浏览器收到页面并显示给用户。

（2）HTTP 报文格式

HTTP 有两类报文：从客户到服务器的请求报文和从服务器到客户的响应报文。图 1 显示了两种报文的结构。

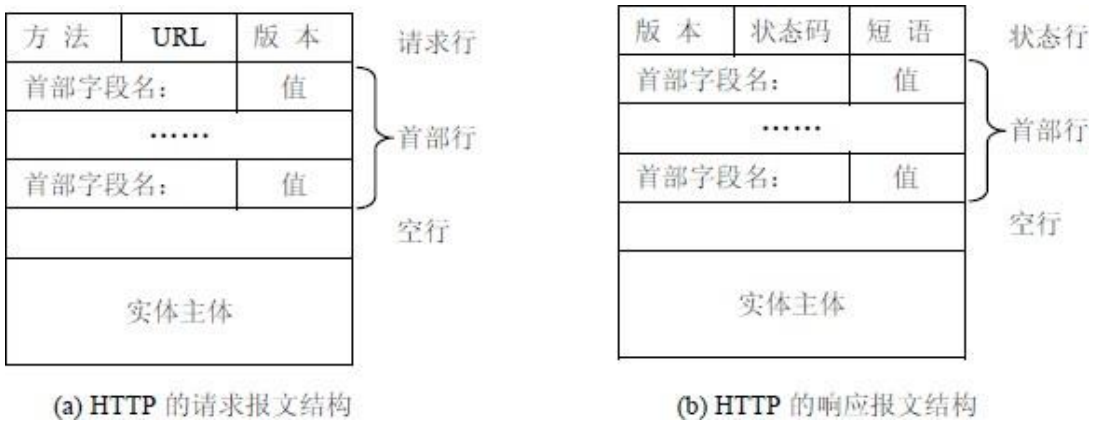


图1 HTTP 的请求报文和响应报文结构

在图 1 中，每个字段之间有空格分隔，每行的行尾有回车换行符。各字段的意义如下：

① 请求行由三个字段组成：

＊ 方法字段，最常用的方法为 “GET”，表示请求读取一个万维网的页面。常用的方法还有 “HEAD（指读取页面的首部）”和“POST（请求接受所附加的信息）

＊ URL 字段为主机上的文件名，这时因为在建立 TCP 连接时已经有了主机名；

＊ 版本字段说明所使用的HTTP 协议的版本，一般为 “HTTP/1.1”。

② 状态行也有三个字段：

＊ 第一个字段等同请求行的第三字段；

＊ 第二个字段一般为 “200”，表示一切正常，状态码共有41 种，常用的有：301 （网站已转移）400（服务器无法理解请求报文）404（服务器没有锁请求的对象）等；

＊ 第三个字段时解释状态码的短语。

③ 根据具体情况，首部行的行数是可变的。请求首部有 Accept 字段，其值表示浏览器可以接受何种类型的媒体；Accept-language，其值表示浏览器使用的语言；User-agent表明可用的浏览器类型。响应首部中有 Date、Server、Content-Type、Content-Length 等字段。在请求首部和响应首部中都有 Connection 字段，其值为 Keep-Alive 或 Close，表示服务器在传送完所请求的对象后是保持连接或关闭连接。

④ 若请求报文中使用 “GET” 方法，首部行后面没有实体主体，当使用 “POST” 方法是，附加的信息被填写在实体主体部分。在响应报文中，实体主体部分为服务器发送给客户的对象。

图2 和图 3 显示了 Wireshark 捕获的 HTTP 请求和响应报文，结合上面的介绍，请自己分析和体会。

```
⊞ Transmission Control Protocol, Src Port: 1068 (1068), Dst Port: 8080 (8080), Seq: 1, Ack: 1, Len: 273
⊟ Hypertext Transfer Protocol
  ⊟ GET /l2_switch.jpg HTTP/1.1\r\n
      Request Method: GET
      Request URI: /l2_switch.jpg
      Request Version: HTTP/1.1
    Accept: */*\r\n
    Referer: http://192.168.1.30:8080/\r\n
    Accept-Language: zh-cn\r\n
    Accept-Encoding: gzip, deflate\r\n
    User-Agent: Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1; SV1; .NET CLR 1.1.4322)\r\n
    Host: 192.168.1.30:8080\r\n
    Connection: Keep-Alive\r\n
    \r\n
```

图2 HTTP 请求报文示例

```
⊞ Transmission Control Protocol, Src Port: 8080 (8080), Dst Port: 1068 (1068), Seq: 7343, Ack: 274, Len: 347
⊞ [Reassembled TCP Segments (7689 bytes): #342(174), #343(512), #345(512), #347(512), #349(512), #350(512), ▾
⊟ Hypertext Transfer Protocol
  ⊟ HTTP/1.0 200 OK\r\n
      Request Version: HTTP/1.0
      Response Code: 200
    Date: Mon, 01 Mar 1993 00:26:11 UTC\r\n
    Server: Start HTTP-Server/1.0\r\n
    Content-Type: image/jpg\r\n
    Content-length: 7515\r\n
    Expires: Thu, 16 Feb 1989 00:00:00 GMT\r\n
    \r\n
⊞ JPEG File Interchange Format
```

图3 HTTP 响应报文示例

2．实验环境与说明

（1）实验目的

在PC 机上访问 RCMS 的Web 页面，截获报文，分析 HTTP 协议的报文格式和 HTTP 协议的工作过程。

（2）实验设备和连接

本地实验室环境，无须设备连接；

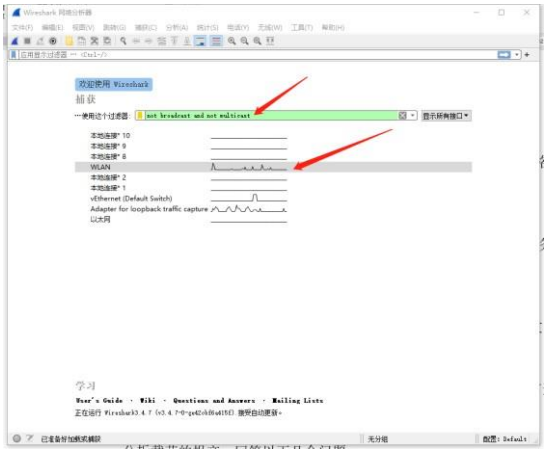
注意：请通过访问可以连接的 WWW 站点或使用 IIS 建立本地 WWW 服务器来进行实验。

（3）实验分组

每六名同学为一组，每人一台计算机独立完成实验。

练习一：捕获校园网认证的用户名和密码

步骤 1：在PC 机上运行 Wireshark，选中对应的网卡，开始截获报文，为了只截获和我们要访问的网站相关的数据报，将截获条件设置为 “not broadcast and not multicast;”



步骤 2：从浏览器上访问暨南大学官方主页 www.jnu.edu.cn，跳转到校园网认证页面，输入用户名和密码，认证成功后关闭网页。（注：若是锐捷客户端认证上网或已经网页认证过的话，先退出锐捷客户端或网页认证，重新进行网页认证）

步骤 3：停止截获报文，将截获的报文命名为 http-学号保存。

分析截获的报文，回答以下几个问题：

1）综合分析截获的报文，查看有几种 HTTP 报文？

四种HTTP报文，有HTTP、GET、POST、HEAD。

2）在截获的 HTTP 报文中，任选一个 HTTP 请求报文和对应的 HTTP 应答报文，仔细分析它们的格式，填写表 1 和表2。

表1 HTTP 请求报文格式

方 法	GET	版 本	HTTP/1.1
URL	http://m.lwvzmdu.cn:16648/		
首部字段名	字段值	字段所表达的信息	
Host	m.lwvzmdu.cn:16648	请求目标主机	

User-Agent	Go-http-client/1.1	用户头
Connection	Upgrade	是否保持固定的HTTP连接
Sec-WebSocket-Key	w2iPrZj3C/bYtCV7WJVQ+Q==	验证服务器是否为Websocket助理
Sec-WebSocket-Version	13	Websocket Draft版本
Upgrade	websocket	使用何种助理

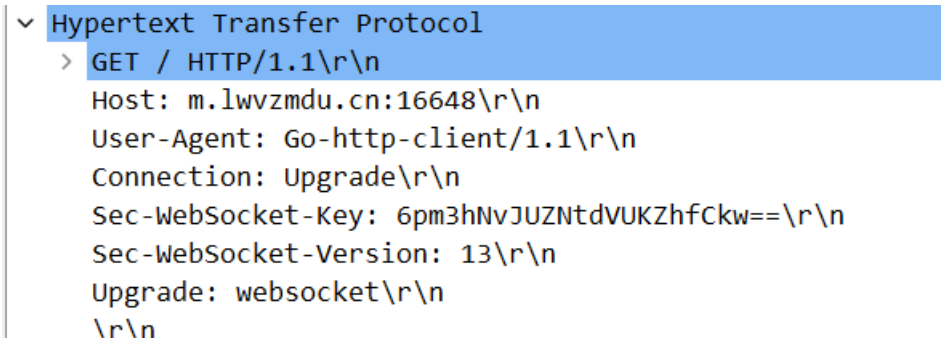
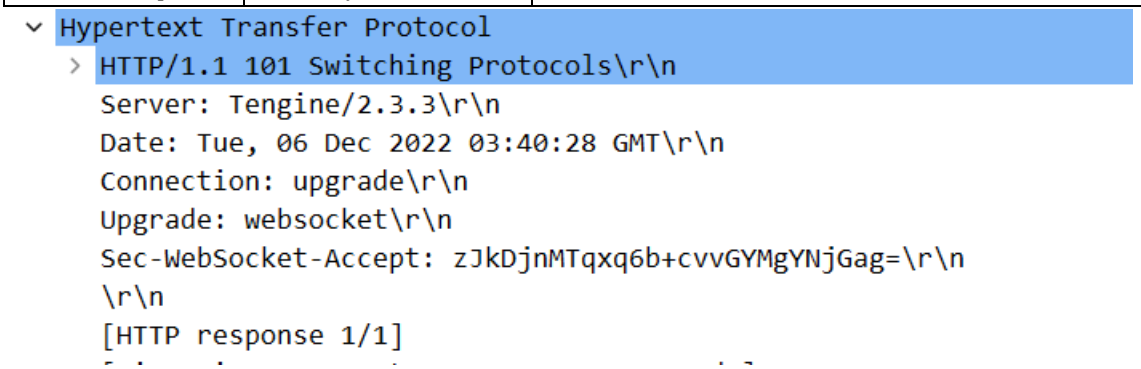


表2 HTTP 应答报文格式

版 本	HTTP/1.1	状态码	101
短 语	Switching Protocols		
首部字段名	字段值	字段所表达的信息	
Server	Tengine/2.3.3	服务器应用程序的信息	
Date	Tue, 06 Dec 2022 03:25:23 GMT	日期信息	
Connection	upgrade	是否保持固定的HTTP连接	
Upgrade	websocket		
Sec-WebSocket-Accept	uRrG14KRyQz37wMWy0yumziF7x8=	对客户端验证的回应	



3）分析在截获的报文中，客户机与服务器建立了几个连接？服务器和客户机分别使用了哪几个端口号？

一个连接：客户端：6948，服务端：16648

4 ）综合分析截获的报文，理解 HTTP 协议的工作过程，将结果填入表 3 中。

表3 HTTP 协议工作过程

HTTP 客户机端口号	HTTP 服务器端口号	所包括的报文号	步骤说明

5 ）在截获的报文中能不能找到上网认证的用户名和密码？如果能找到，请说明原因。

不能。https协议采取了加密。

练习二：捕获 http 网站的登录用户名和密码

步骤 1： 登录一个 http 网站（例如桂林生活网 http://www.guilinlife.com/）点击右上角的登录，随意在账号和密码输入，点击登录，只要有反馈，比如登录失败，就表示已发送数据包。

步骤2： 在 doc 命令行通过 ping www.guilinlife.com方式获取该网站的 IP 地址，然后在wireshark 过滤器上通过该 IP 过滤出该网站所有的数据。

步骤3：停止截获报文，将截获的报文命名为 http-学号-2 保存。分析截获的报文，回答以下几个问题：

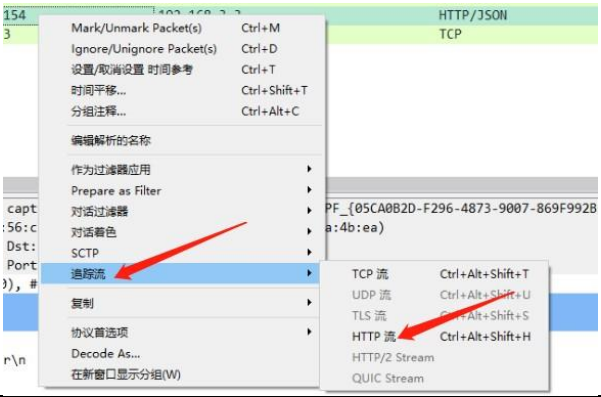
1）在截获的报文中能不能找到登录的用户名和密码？如果能找到，请说明原因。

能，http协议采用的是明文传输用户名和密码。

```
HTML Form URL Encoded: application/x-www-form
  > Form item: "username" = "admin"
  > Form item: "password" = "123456"
  > Form item: "sessionId" = "01CWt674yodh4cl7w
  > Form item: "sig" = "05URXVU-kjM0vFgTcbsPNsS
```

2)在截获的报文上右键菜单选择“跟踪流”->“HTTP流”查看相关信息填写下列表格。

表4 HTTP 请求报文格式



方 法	POST	版 本	HTTP/1.1
URL	http://www.passport.guilinlife.com		
首部字段名	字段值	字段所表达的信息	
User-Agent	Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/107.0.0.0 Safari/537.36 Edg/107.0.1418.62	用户头	
Content-Length	626	消息主体的大小	
Accept	*/*	通知服务器,用户代理能够处理的媒体类型及媒体类	

		型的相对优先级
Accept-Encoding	gzip, deflate	通知服务器,用户代理能够处理的编码方式
Accept-Language	Zh-CN,zh;q=0.9,en;q=0.8,en-GB;q=0.7,en-US;q=0.6	通知服务器,用户代理能够处理的语言
Content-Type	application/x-www-form-urlencoded; charset=UTF-8	客户端能实际处理的内容的内容类型
Cookie	太长了，省略	

```
Host: passport.guilinlife.com
User-Agent: Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/
107.0.0.0 Safari/537.36 Edg/107.0.1418.62
Content-Length: 626
Accept: */*
Accept-Encoding: gzip, deflate
Accept-Language: zh-CN,zh;q=0.9,en;q=0.8,en-GB;q=0.7,en-US;q=0.6
Content-Type: application/x-www-form-urlencoded; charset=UTF-8
Cookie: Hm_lvt_0089713781dae652bbf698e087b5645d=1670074590;
Hm_lpvt_0089713781dae652bbf698e087b5645d=1670074590;
acw_tc=784e2c8f16700745998623405e79a367fe5740f97fff375cffa7833645514e;
PHPSESSID=4q093f1m7uhgpcbb6cvdt12mn7;
u_asec=099%23KAFEAYEKE6YEhYTLEEEEEpEQz0yFZ6NhSXy7Z6VhSXiqW6AHZXJoD6tFDEFETcZdt9TXE7EFD67EE3QTEEx6zIywWE
FE5YwqiGTh%2B1A4kLgBiZWcqmGboSlqcuIQRJfrkg4OViP%2BMJK01wa%2FoZjDfKCypsT2qNGOCYPVWwMQ6%2BCfZZpAIwEsMNLoy
%2F5Qcv64JmDVrCTa3iRhE7EtDqbEwoa3U6wrG0oQ1UWvDs%2FKcgXRwsnj%2FfqZPFbRuz8sDzoSadu4v%2F%2F5rsr33y3h%2FfhS
6dSqbP9W8yXZ9mGTEELP%2F3mABdav6XQTEE9%2F%2F39ZGRa2sEFEpcZdt3illuZdsyaAI9llsUOP%2F3RClllrncZdd%2FRllusos
yaCd%2FllsPAhE7EhsynSt3ll; _uab_collina=167007460026619867654633
Origin: http://passport.guilinlife.com
Referer: http://passport.guilinlife.com/
```

表5 HTTP应答报文格式

版 本	HTTP/1.1	状态码	200
短 语	OK		
首部字段名	字段值	字段所表达的信息	
Date	Sat, 03 Dec 2022 13:37:16 GMT	日期信息	
Content-Type	application/json; charset=utf-8	告诉客户端实际返回的内容的内容类型	
Transfer-Encoding	chunked	传输编码的方式	
Server	nginx	服务器应用程序	
Vary	Accept-Encoding	告诉用户代理选择表现形式（reprentation）的标准	
Access-Control-Allow-Origin	http://passport.guilinlife.com	解决资源跨域的策略	
Access-Control-Allow-Methods	GET	允许客服端使用何种方式获取服务器的服务	
Set-Cookie	Glsh_New_Passport_SSO=deleted; expires=Thu, 01-Jan-1970 00:00:01 GMT; Max-Age=0; path=/; domain=.guilinlif	服务器发送到浏览器或者其他客户端的一些信息	

	e.com; httponly	
Content-Encoding	gzip	服务器能处理的编码方式

```
97HTTP/1.1 200 OK
Date: Sat, 03 Dec 2022 13:37:16 GMT
Content-Type: application/json; charset=utf-8
Transfer-Encoding: chunked
Connection: keep-alive
Server: nginx
Vary: Accept-Encoding
Access-Control-Allow-Origin: http://passport.guilinlife.com
Access-Control-Allow-Methods: GET
Set-Cookie: Glsh_New_Passport_SSO=deleted; expires=Thu, 01-Jan-1970 00:00:01 GMT; Max-Age=0; path=/;
domain=.guilinlife.com; httponly
Content-Encoding: gzip

{"status":102,"msg":"............................"}
```

3）通过上述实验总结 http 协议的安全性。

http协议属于明文传输协议，交互过程以及数据传输都没有进行加密，通信双方也没有进行任何认证，通信过程非常容易遭遇劫持、监听、篡改，严重情况下，会造成恶意的流量劫持等问题，甚至造成个人隐私泄露（比如银行卡卡号和密码泄露）等严重的安全问题。