基于 HTTP 文件下载欺骗系统的设计与实现

王永刚

(福建信息职业技术学院,福建 福州 350003)

摘要 文章提出了一种基于HTTP文件下载欺骗技术的设计方案,能监视局域网用户行为。当目标主机开始执行文件下载时,利用TCP/IP协议没有对源IP地址执行合法性认证的缺陷,创建源IP地址为Web服务器IP地址的HTTP 302重定向数据包。通过重定向目标主机下载的资源文件,和目标主机完成文件下载通信。重定向后的数据文件是后门程序,攻击主机能通过它得到目标主机的管理权限,进而控制目标主机的网络动作。该方案具有很好的隐蔽性。最后进行了测试,测试结果表明该系统的研究思路是正确的。

关键词 HTTP;欺骗系统;文件下载

中图分类号:TP393.08 **文献标志码**:A **文章编号**:1671-0436(2012)02-0037-06

Design of File Download Spoofing System Based on HTTP

WANG Yong-gang

(Fujian Polytechnic of Information Technology, Fuzhou 350003)

Abstract This paper proposes a file download spoofing technology based on HTTP, which can monitor the behavior of users in the local area network. When the target host computer begins the implementation of file download, it creates a source IP address of the Web server IP address HTTP 302 redirect data packets, owing to the lack of authentication mechanisms of the Authenticity of the source IP address. By redirecting the download resource file of the target host computer, it completes the communication of the file download with the target host computer. Data file of redirected is a backdoor program, the attacking host computer can get the authority of the target host computer by it, and then controls its network action. This design successfully spoofs the target host computer in the circumstance of user unknown. The test results indicate the thesis research ideas are correct.

Key words HTTP; spoofing system; file download

当今深入生活的Web服务,是建立在HTTP协议上的全球共享信息库,它是万维网上HTTP服务器的集合。[1]但实际上Web并不安全,网络欺骗和虚假信息在网络上随处可见。在互联网上交易或传送重要数据,用户最忧心的问题就是交易和数据传输是否安全。HTTP协议是万维网的基础,通常使用TCP协议的80端口,目前HTTP协议演化出了很多版本,其中大部分都是向下兼容。[2]

1 HTTP文件下载欺骗系统概述

HTTP文件下载通常都是以资源名称的方式请求下载,与Web服务器建立连接,进行数据通信的。HTTP 302重定向机制能实现对HTTP文

收稿日期:2012-03-26

作者简介:王永刚(1978—),男,硕士,讲师。

件下载进行欺骗。

客户端主机 C 与服务器 S_a 建立 3 次连接后，客户端 C 以资源名称方式向 S_a 请求下载信息，攻击者 A 捕获到此下载请求信息时，伪装成 S_a 给 C 发送 302 重定向数据包，C 接到 302 重定向信息时，用其他端口和重定向 URL（服务器为 S_b）建立连接，发送请求完成攻击者 A 的自定义的欺骗文件的下载。其攻击过程用以下几个步骤表示：

C→S_a：Three hand

C→S_a：GET：filename（EXE/RAR/ZIP） Port：x→80 Flags：ACK/PSH

A（S_a）→C：302：Location（URL） Port：80→x Flags：ACK/PSH

C→S_b：GET：URL（Host：S_b） Port：y→80 Flags：ACK/PSH

S_b→C：200 OK Port：80→y Flags：ACK

只需在合法的响应信息到来之前到达客户端，就能修改客户端下载资源文件。由于攻击者能获得客户端下载资源名称，能把自定义的文件名称改成客户端的下载资源名称，更容易达到欺骗的目的。[3] 要进行文件下载欺骗，最重要的就是构造 HTTP 302 响应欺骗信息。在客户端发出文件下载请求信息后，快速构造 HTTP 302 重定向数据包，先于合法响应数据包到达客户端，就能达到欺骗客户端的目的，把客户端下载的文件修改为自定义的文件。

2 HTTP 文件下载欺骗系统结构设计

2.1 系统设计的运行环境

HTTP 文件下载欺骗系统在监控主机（攻击主机）上运行，是基于 Linux 系统开发的，需安装 Libpcap 库和搭建 HTTP 服务器。另外监控主机还需要配置 2 块网卡，分别负责接收数据包提供给应用程序处理和发送欺骗应答数据包，这样能加快数据处理的速度，也加大了欺骗成功的几率。系统的网络环境是共享式以太网，攻击主机处于旁路监听模式。

2.2 系统总体结构设计

HTTP 文件下载欺骗系统是基于 Linux 系统设计的，它利用 Libpcap 库捕获和过滤数据，在捕获处理数据之前需要从配置文件中读取数据，配置的参数有目标主机 IP、攻击主机 IP 和重定向 URL 等。系统的总体结构设计如图 1 所示。

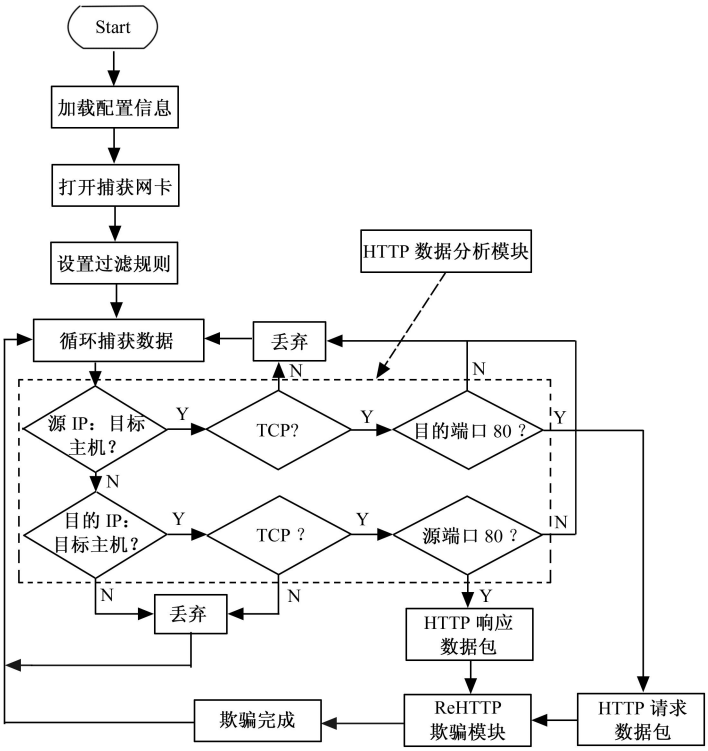

图 1 系统总体结构图

程序利用 Libpcap 库函数捕获网络中的数据包,把网卡设置为混杂模式,先根据以太网首部中的帧类型字段判断协议类型(0X0800 为 IP 数据包),再根据 IP 首部中的 8 位协议字段判断传输层的协议。因为 HTTP 数据包是针对 TCP 的数据包,所以只需要处理 TCP 数据包,其他协议类型的数据包直接丢弃。再从 TCP 数据包中分离出 HTTP 请求报文,即目的端口为 80 的 TCP 数据包,将其送入 ReHTTP 欺骗模块进行处理,继续处理下一个数据包。

3　HTTP 文件下载欺骗系统设计实现

HTTP 重定向欺骗的过程是:监控主机利用旁路监听模式监听局域网目标主机(由配置文件信息确定)的数据包,当发现目标主机有 HTTP 请求数据包(目的端口 80)时,则提取其请求的内容和文件下载特征信息(本文中主要在请求的资源文件为 rar、zip 和 exe 时进行欺骗)作比较,如果匹配则认为目标主机在进行文件下载行为,构造 HTTP 302 重定向数据包,把目标主机下载的文件修改为自定义的文件(可以是后门程序或木马程序);如果不匹配则认为目标主机在浏览网页等,不予处理。

提取 HTTP 请求信息、提取 HTTP 响应信息、是否文件下载判断、四元组匹配模块、构造 HTTP 重定向数据包和消除 ACK 风暴(HTTPResponse 模块)、HTTPRST 模块、发送重定向数据包等功能都是由 ReHTTP 欺骗模块来实现的。ReHTTP 欺骗模块的详细功能设计如图 2 中的虚线框所示。

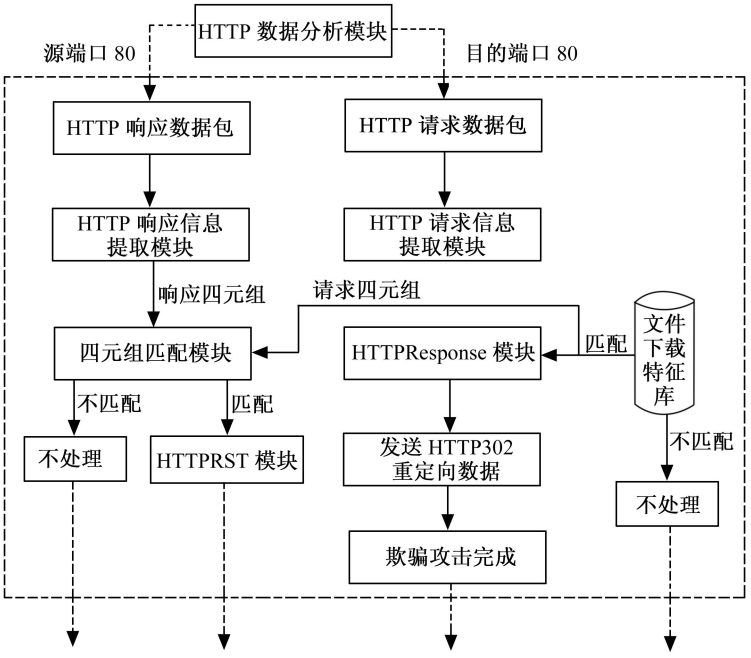

图 2　ReHTTP 欺骗模块详细功能设计图

3.1　循环捕获数据的实现

数据捕获主要利用 Libpcap 库函数实现,主要的函数原型如下:

```
pcap_t pcap_open_live(const char * device,int snaplen,int prosimc,int to_ms,
                      char * ebuf);
int pcap_compile(pcap * p,struct bpf_program * fp,char * str,
                 int optimize,bpf_u_int32 netmask);
int pcap_setfilter(pcap_t * p,struct bpf_program * fp);
int pcap_loop(pcap_t * p,int cnt,pcap_handle callback,u_chat * user);
```

3.2 HTTP 数据分析模块

HTTP 数据分析模块的功能是把 HTTP 请求或响应数据包输入到模块之后,解析 HTTP 请求或响应数据包的各个主要字段,如源 IP、目的 IP、源端口、目的端口以及请求或响应的数据等。其主要的函数原型如下:

```
int read_packet(u_char * httppacket,u_char * etherdst,u_char * ethersrc,
        u_char * ipsrc,u_char * ipdst,u_short tcp_src_port,
        u_short tcp_dst_port,u_int seqnum,u_int acknum,
        int totalLen,u_char * * http_data,int datalen)
```

3.3 HTTP 请求或响应信息提取模块

HTTP 请求或响应信息分别是请求行和响应行的重要信息,在 TCP 首部之后,位于 HTTP 数据的第一行。如请求行为:GET/Path/FileName HTTP/1.0,GET 后面是一个空格,其后紧跟的是要下载的从 WEB 服务器根目录开始的绝对路径的文件,也即请求信息。该路径后又有一个空格,然后是协议名称及协议版本。所以请求信息位于请求行中两个空格之间。

如响应行为:HTTP/1.0 200 OK 第一行(响应行)是协议名称及版本号,空格后面会有一个三位数的数字,即 HTTP 协议的响应状态码,200 表示成功,OK 是对状态码的简短文字描述,即响应信息为状态码也位于响应行中两个空格之间。

2 个信息提取模块的算法一致,所以可以用 1 个接口函数表示。主要的函数接口定义如下:

```
int http_extract_queryorresponse(char * http_queryorrespons,u_char * http_data,int len,int mark);
```

3.4 HTTP Response 模块

HTTP Response 模块的功能就是根据捕获数据包的信息,构造欺骗数据包,计算序列号等重要字段,发送 HTTP 302 重定向数据包到客户端进行欺骗攻击;构造关闭连接数据包,关闭客户端请求端口和服务器的连接。

主要的接口函数定义如下:

1)HTTP 重定向数据包构造函数

```
int RehttpResponse(u_char * pPacket,u_char * dst_etheraddr,u_char * src_etheraddr,
        u_char * ip_dst_addr,u_char * ip_src_addr,u_short tcp_dst_port,u_int seqnum,u_int acknum,
        u_char * ReHttpdata,int redatalen,int querydatalen);
```

函数返回 0 表示发送 HTTP 302 重定向数据成功,返回 -1 表示发送重定向数据失败。参数 ReHttpdata 为 HTTP 302 重定向数据。参数 redatalen 为 HTTP 302 重定向数据的长度。参数 querydatalen 为 HTTP 资源请求数据包中 HTTP 数据的长度,也即请求信息提取模块返回 0 时参数 len 的值,它是为了计算重定向数据包中的确认号。

2)关闭客户端请求端口函数

这个函数主要完成客户端请求端口的关闭功能,其工作原理如下:

假设 A 为客户端以资源名称方式请求的数据包,F 为关闭连接数据包。A_SEQ、A_ACK、A_DATALEN 分别表示数据包 A 的序列号、确认号和数据长度;F_SEQ、F_ACK 和 F_Flag 分别表示数据包 F 的序列号、确认号和标志位,则有下面的数据关系:

F_SEQ = A_SEQ + A_DATALEN,F_ACK = A_ACK,F_Flags = ACK/RST;

接口函数定义如下:

```
int HTTPFIN(u_char * pPacket,u_char * dst_etheraddr,u_char * src_etheraddr,u_char
        * ip_dst_addr,u_char * ip_src_addr,
        u_short tcp_dst_port,u_int seqnum,u_int acknum)
```

其中 pPacket 为构造的数据,其余参数分别为 HTTP 请求数据包 A 的目的以太网地址、源以太网地址、目的 IP、源 IP、目的端口和源端口。其中序列号和确认号等重要字段的构造就由上面的

数据关系确定。若数据发送成功,则函数返回 0,否则返回 -1。

3.5 四元组匹配模块

四元组匹配模块的功能:当捕获到 HTTP 响应数据包 200 OK(即 HTTP 提取响应信息模块的返回值为 1)时,判断此数据包的四元组是否和 HTTP 请求文件下载数据包的四元组匹配,若匹配则表示此数据包是服务器端返回给客户端的真实 HTTP 响应(200 OK),则把四元组信息以及序列号、确认号等信息传递给 HTTPRST 模块,重置服务器端的连接,避免出现 ACK 风暴。

主要的接口函数定义如下:

```
int IsRealResponse(u_char * ip_dst_addr,u_char * ip_src_addr,u_short
        tcp_dst_port,struct FourElement * pParam);
```

3.6 HTTPRST 模块

本模块主要功能是重置服务器端的连接,避免在服务器和客户端通信的过程中插入数据使得客户端和服务器端由于序列号的不同步而循环发送 ACK 认证数据包,形成 ACK 风暴。

这个模块主要由重置连接函数构成,工作原理如下:

假设 C 为服务器端返回给客户端的真实 HTTP 响应数据包(200 OK),R 为重置连接数据包。C_SEQ、C_ACK 和 C_DATALEN 分别表示数据包 C 的序列号、确认号、数据长度;R_SEQ、R_ACK 和 R_Flag 分别表示数据包 R 的序列号、确认号和标志位,则有下面的数据关系:

R_SEQ = C_ACK + 1,R_ACK = C_SEQ + C_DATALEN,R_Flags = ACK/RST;

主要接口函数定义如下:

```
int HTTPRST(u_char * pPacket,u_char * dst_etheraddr,u_char * src_etheraddr,
        u_char * ip_dst_addr,u_char * ip_src_addr,
        u_short tcp_dst_port,u_int seqnum,u_int acknum);
```

其中,pPacket 为构造的数据,其余的参数分别为真实 HTTP 响应数据包 C 的目的以太网地址、源以太网地址、目的 IP、源 IP、目的端口和源端口。

4 HTTP 文件下载欺骗系统测试与分析

4.1 测试过程

首先在攻击主机上备好文件,当捕获到目标主机以资源名称形式的请求文件下载数据包时,提取请求下载的资源文件名,然后将攻击主机中用于欺骗的相同后缀名的文件修改为请求下载的资源文件名,发送重定向欺骗数据包,目标主机下载的资源文件修改为自定义的文件。本系统是基于 Linux 开发的,监控机(攻击主机)安装有 Linux 操作系统和双网卡,网卡分别用于捕获目标主机的文件下载数据包和发送文件下载欺骗数据包。

本次测试的参数设置如下:攻击主机(双网卡主机):192.168.3.231(捕获数据)和 192.168.3.111(发送欺骗数据和作为 HTTP 服务器),目标主机为 192.168.3.100。借助科来网络分析系统监听数据包,设置过滤规则为仅捕获目标主机的 HTTP 数据包。

4.2 测试实例

本测试以某网站的 exe 文件下载为例,使用 Windows IE7.0 浏览器和 Fire Fox 浏览器,下载提示窗口如图 3。

从图 3 可以看出:通过 IE 直接下载和利用 Fire Fox 浏览器下载以资源名称形式请求的 exe 文件,HTTP 文件下载欺骗系统都成功地重定向到监控机的 HTTP 服务器上(192.168.3.111),达到了欺骗的目的。

科来网络分析系统捕捉到 exe 文件下载欺骗的过程,如图 4 所示。数据包 5、6、7 为目标主机与服务器 3 次握手建立连接的过程,数据包 8 为 exe 文件请求数据包,数据包 9 为 302 重定向欺骗数据包,如图 5 所示。数据包 10、11、12 为目标主机与监控主机的 HTTP 服务器(192.168.3.111)建立连接的过程,数据包 15 为下载请求成功信息(200 OK)。

(a)IE7.0 中 exe 文件下载提示窗口　　(b)Fire Fox 中 exe 文件下载提示窗口

图 3　文件下载提示窗口

编.	绝对时间	源	目标	TCP:标志	概要
5	19:26:43.042785	192.168.3.100:3692	220.113.41.126:www-http	..00 0010	序列号=1124267748,确认号=0000000000,标志=....S.,长度=.
6	19:26:43.096492	220.113.41.126:www-http	192.168.3.100:3692	..01 0010	序列号=1692017451,确认号=1124267749,标志=.A..S.,长度=.
7	19:26:43.096563	192.168.3.100:3692	220.113.41.126:www-http	..01 0000	序列号=1124267749,确认号=1692017452,标志=.A....,长度=.
8	19:26:43.099956	192.168.3.100:3692	220.113.41.126:www-http	..01 1000	C: GET /wmxzsetup.exe HTTP/1.1
9	19:26:43.101211	220.113.41.126:www-http	192.168.3.100:3692	..01 1000	S: HTTP/1.1 302 object moved
10	19:26:43.125949	192.168.3.100:3694	192.168.3.111:www-http	..00 0010	序列号=1946997305,确认号=0000000000,标志=....S.,长度=.
11	19:26:43.126424	192.168.3.111:www-http	192.168.3.100:3694	..01 0010	序列号=1145883170,确认号=1946997306,标志=.A..S.,长度=.
12	19:26:43.126466	192.168.3.100:3694	192.168.3.111:www-http	..01 0000	序列号=1946997306,确认号=1145883171,标志=.A....,长度=.
13	19:26:43.128705	192.168.3.100:3694	192.168.3.111:www-http	..01 1000	C: GET /wmxzsetup.exe HTTP/1.1
14	19:26:43.129200	192.168.3.111:www-http	192.168.3.100:3694	..01 0000	序列号=1145883171,确认号=1946997536,标志=.A....,长度=.
15	19:26:43.131328	192.168.3.111:www-http	192.168.3.100:3694	..01 0000	S: HTTP/1.1 200 OK
16	19:26:43.132558	192.168.3.111:www-http	192.168.3.100:3694	..01 0000	S: 继续或非HTTP通信, 1460 字节的二进制数据
17	19:26:43.132586	192.168.3.100:3694	192.168.3.111:www-http	..01 0000	序列号=1946997536,确认号=1145886091,标志=.A....,长度=.
18	19:26:43.134254	192.168.3.111:www-http	192.168.3.100:3694	..01 0000	S: 继续或非HTTP通信, 1460 字节的二进制数据

图 4　exe 文件下载欺骗过程

```
0000  00 E0 11 02 90 03 00 50 BA 69 4F 23 08 00 45 00 01 92 7D 58 00 00 76 06 FC   .......P.iO#..E...}X..v..
0019  11 DC 71 29 7E C0 A8 03 64 00 50 0E 6C 64 DA 23 2C 43 02 F7 CD 50 18 FD E7   ..q)~...d.P.ld.#,C...P...
0032  D4 84 00 00 48 54 54 50 2F 31 2E 31 20 33 30 32 20 6F 62 6A 65 63 74 20 6D   ....HTTP/1.1 302 object m
004B  6F 76 65 64 0D 0A 44 61 74 65 3A 20 54 75 65 2C 20 31 39 20 46 65 62 20 32   oved..Date: Tue, 19 Feb 2
0064  30 30 38 20 30 35 3A 30 31 3A 30 39 20 47 4D 54 0D 0A 53 65 72 76 65 72 3A   008 05:01:09 GMT..Server:
007D  20 4D 69 63 72 6F 73 6F 66 74 2D 49 49 53 2F 36 2E 30 0D 0A 4C 6F 63 61 74    Microsoft-IIS/6.0..Locat
0096  69 6F 6E 3A 20 68 74 74 70 3A 2F 2F 31 39 32 2E 31 36 38 2E 33 2E 31 31 31   ion: http://192.168.3.111
00AF  2F 77 6D 78 7A 73 65 74 75 70 2E 65 78 65 0D 0A 43 6F 6E 74 65 6E 74 2D 4C   /wmxzsetup.exe..Content-L
00C8  65 6E 67 74 68 3A 20 31 34 37 0D 0A 43 6F 6E 74 65 6E 74 2D 54 79 70 65 3A   ength: 147..Content-Type:
00E1  20 74 65 78 74 5C 68 74 6D 6C 0D 0A 43 61 63 68 65 2D 63 6F 6E 74 72 6F 6C    text\html..Cache-control
00FA  3A 20 70 72 69 76 61 74 65 0D 0A 0D 0A 3C 68 65 61 64 3E 3C 74 69 74 6C 65   : private....<head><title
0113  3E 6F 62 6A 65 63 74 20 6D 6F 76 65 64 3C 2F 74 69 74 6C 65 3E 3C 2F 68 65   >object moved</title></he
012C  61 64 3E 0A 3C 62 6F 64 79 3E 3C 68 31 3E 4F 62 6A 65 63 74 20 4D 6F 76 65   ad>.<body><h1>Object Move
0145  64 3C 2F 68 31 3E 54 68 69 73 20 6F 62 6A 65 63 74 20 6D 61 79 20 62 65 20   d</h1>This object may be
015E  66 6F 75 6E 64 20 3C 61 20 48 52 45 46 3D 22 68 74 74 70 3A 2F 2F 31 39 32   found <a HREF="http://192
0177  2E 31 36 38 2E 33 2E 31 31 31 2F 77 6D 78 7A 73 65 74 75 70 2E 65 78 65 22   .168.3.111/wmxzsetup.exe"
0190  3E 68 65 72 65 3C 2F 61 3E 2E 3C 2F 62 6F 64 79                              >here</a>.</body
```

图 5　exe 文件的 302 重定向欺骗数据包

从上述测试分析,可以知道不论客户端是利用 Windows IE7.0 浏览器,还是利用 Fire Fox 浏览器文件下载,HTTP 文件下载欺骗系统都成功将目标主机下载的 exe 文件修改为自定义文件。经测试,其他类型文件也同样能获得上述测试效果,达到了设计的预期目的。

5　结语

本文主要研究了 HTTP 文件下载欺骗技术,研究并提出了一种内部网络用户 HTTP 文件下载行为的监控欺骗系统,并结合局域网环境,根据系统的总体设计结构对系统中的各个模块进行了详细的设计和实现。该系统设计在用户不知情的情况下成功地对目标主机进行了重定向欺骗,具有良好的隐蔽性,最后给出了测试结果。测试结果表明系统设计思路的正确。

[参考文献]

[1]黄晓鸣. WWW 之论[J]. 科技资讯,2007(20):110 - 111.

[2]冯登国. 安全协议——理论与实战[M]. 北京:清华大学出版社,2011:163 - 164.

[3](美)雅各布森. 网络安全基础——网络攻防、协议与安全[M]. 仰礼友,译. 北京:电子工业出版社,2011:263.

责任编辑:唐海燕