# UltraSurf 软件的运行机制分析

张 磊， 谷大武， 陆海宁， 陈 帆
（上海交通大学 信息安全工程学院，上海 200240）

【摘 要】UltraSurf 软件是在互联网客户端运行的应用程序。它使用远程代理服务器和自定义的加密协议，可以突破传统的网络过滤，实现对远程信息的透明访问。文中采用动态反汇编技术，对该软件的通信手段和加密方式等进行了逆向分析，确定了软件的工作流程、加密算法以及网络拓扑结构，还原了软件内部的秘密信息，并对分析结果进行了测试验证。目前，该软件是互联网上具有代表性的动态代理保密通信软件之一，因此文中的研究结果也为其它同类软件的分析提供了借鉴。

【关键词】UltraSulf 软件；逆向分析；动态代理；加密通信
【中图分类号】TP309.7　　【文献标识码】A　　【文章编号】1002-0802(2008)09-0161-04

# Analysis on Runtime Mechanism of UltraSurf Software

ZHANG Lei， GU Da-wu，LU Hai-ning， CHEN Fan
(School of Information Security Engineering, SJTU, Shanghai 200240, China)

【Abstract】UltraSurf is a well-known client application on the Internet. With the help of its private communication protocols and remoting servers as agents, it can be used to penetrate through the network control available, so as to make it accessible to remote information. By utilizing reverse engineering techniques, its internal communication, encryption and processing mechanism ia analyzed. Based on this work, its software structure, working flowchart, encryption algorithms, and network topology is defined. More significantly, all the sensitive information hidden in this software is obtained. The experimental test shows that this analysis is proved to be correct and valid. As the UltraSurf software is nowadays one of the most typical dynamic agent encrypted communication software on the Internet, this effort also provides a reference to other similar softwares.

【Key words】UltraSurf; reverse analysis; dynamic agent; encrypted communication

## 0 引言

目前，UltraSurf（无界浏览）软件是互联网上具有代表性的保密通信软件之一。它通过使用远程代理服务器和自定义的加密协议，可以突破传统的网络过滤，实现对远程信息的透明访问。该软件的内部结构和源代码一直没有公开，文中的主要工作是利用动态反汇编工具 Ollydbg 对无界浏览软件 8.8 版本（文中所讨论的无界浏览软件均指该软件版本）进行逆向分析，确定软件的工作流程、加密算法以及网络拓扑结构，还原软件内部的秘密信息，并通过黑盒测试对分析结果进行验证。

## 1 分析手段及工具

对于网络通信类软件，传统的分析方法是基于网络报文的黑盒分析，通过对软件和网络进行交互的数据进行统计分析，并使用模拟软件与服务器进行通信，从而还原出软件的通信流程和数据信息[1]。文中对无界浏览软件采用的分析方法与传统方法有所不同。首先使用反汇编软件对软件的内部结构进行一定的分析，划分出大致的功能模块，其次结合黑盒分析工具细化软件的行为特征，在此基础上对特定功能（网络通信、加密算法）进行代码定位并使用动态调试方法得到其确切的行为逻辑。与传统的分析方法相比，这种分析

收稿日期：2008-05-26。
基金资助：国家高技术研究发展计划(863)（2006AA01Z405）。
作者简介：张 磊（1984-），男，硕士研究生，主要研究方向为密码与网络安全；谷大武（1970-），男，博士生导师，主要研究方向为密码理论与技术、计算机系统安全；陆海宁（1979-），男，助理工程师，主要研究方向为密码学与网络安全；陈 帆（1984-），男，硕士研究生，主要研究方向为密码与网络安全。

技术具有分析速度快、效率高的特点，并具有很高的准确性和完备性。

文中所用的分析工具包括调试器 Ollydbg1.10[2]、PE 文件信息分析工具 PEiD、网络报文分析工具 Ethereal 等。对无界浏览软件的分析结果表明，该软件是使用 Visual C++ 6.0 开发的 MFC 应用程序，使用了可执行压缩软件 UPX 进行保护[3]。软件基于 MFC Dialog 编程模型开发，使用了 Winsock 作为网络接口，使用了多线程机制进行网络收发，软件通过特定代理提供的服务进行通信。下面将详细介绍笔者的分析结果。

## 2 无界浏览软件的通信过程

图 1 是还原的无界浏览软件的通信流程。该软件启动后，首先检查运行环境，生成特定数据，然后通过多种途径的查询方式获得代理服务器的信息。之后，利用 HTTPS 进行加密通信。软件退出后，本次通信获得的代理服务器信息和 DNS 查询信息将被保存在本地计算机的缓存文件中。

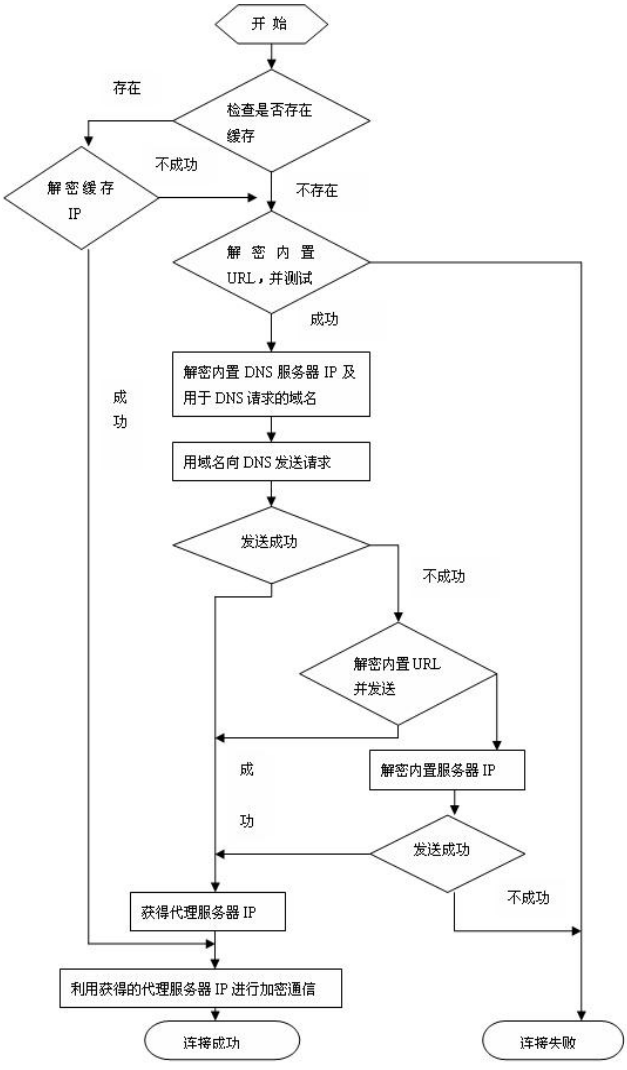


图 1 无界浏览工作流程

### 2.1 软件启动阶段

无界浏览的启动阶段包括两个部分：1)对运行环境的检测和运行选项的设置；2)对本机缓存的查找。对运行环境的检测主要是针对操作系统版本和浏览器类型的检测，无界浏览软件只支持 Windows 2000/XP/2003 操作系统、IE 浏览器 5.0/6.0/7.0 版本下的加密通信代理自动设置。对于运行在 Windows 操作系统下的非 IE 浏览器，软件只支持用户手动设置代理地址的运行方式。

在完成本机环境的检测之后，软件会对本机缓存进行查找。软件在计算机上成功运行后会将使用的代理服务器和查询服务器内容存储在计算机上，这样下次运行时可以直接利用缓存信息进行通信。缓存所处的位置为 Windows 所在的分区\“Documents and Settings\当前用户名\Local Settings\Temp 下的一个临时文件。缓存文件的名称在不同的计算机上是不相同的。名称根据所在磁盘的卷序列号进行一个固定的变换后生成。具体算法见算法 1。

定义：

Para　　32 位整数

Vol　　32 位整数，硬盘卷序列号

Num　　密码表数组，元素类型 char，长度为 6，程序内置

Filename　缓存文件名数组，　元素类型 char，长度为 8

具体算法：

```
Para <= 2
Vol <= Volume Serial
Vol <= Vol ⊕ Para * 32
Vol <= Vol ⊕ 0x81 + Para
If Vol mod 2 == 1 then
    Filename[0] <= 0x7E
Else
    Filename[0] <= ( Vol / 2 ) mod 0x1A + 0x41
End If
For i <= 1 to 6
    Filename[i] <= ( Vol mod Num[i-1] ) mod 0x1A + 0x61
End For
Filename[7] <= ( Vol mod Num[1] ) mod 0x1A + 0x61
Output Filename
```

算法 1 缓存文件名生成算法

缓存信息以加密的二进制数值方式进行保护。算法本身并没有采用标准的对称算法，而且密钥本身也作为密文的一部分进行存储，所以本质上是没有加密保护的。具体算法见算法 2。

定义：

Cipher　　加密的数据　　char 型数组

Cipher-Len　加密数据的长度

Table　　密码表　　char 型数组，长度为 1 2，密码表紧跟在 Cipher 数组后

Dynamic-Table 动态生成的密码表　char 型数组，　长度为 8

Plain　　解密后的数据　char 型数组

具体算法：

```
Real-Len <= (Table[8] | Table[9] | Table[10]|Table[11]) ⊕ 0xFABEBABE
 If Cipher-Len != Real-Len then
    Return “Invalid”
 End if
 Para1 <= Table[0] | Table[1] | Table[2] | Table[3]
 Para2 <= Table[8] | Table[9] | Table[10] | Table[11]
 Para3 <= 0x3F6CB254
 ParaA <= Para1 ⊕ Para2 ⊕ Para3
 Para1 <= Table[4] | Table[5] | Table[6] | Table[7]
 Para2 <= 0xAE985D36
 ParaB <= Para1 ⊕ Para2
 If ParaA == 0 and ParaB == 0 then
    ParaA <= 0x3DCF578A
    ParaB <= 0x78B4FEAE
 End if
 Dynamic-Table <= ParaA | ParaB
index <= 0
 For i <= 0 to Cipher-Len
    Dynamic-Table[index] <=Dynamic-Table[index] ⊕ (i/16)|((i/16)*16)
    Plain[index] <= Dynamic-Table[index] ⊕ Cipher[i]
    Index = (Cipher[i] mod 7) ⊕ index
 End For
Output Plain
```

算法 2 内置数据及缓存文件解密算法

## 2.2 软件代理查询

无界浏览软件的通信过程是基于代理服务器寻找代理，然后使用代理的过程。一共使用了 3 种查询代理的方式：基于 DNS 的查询、基于 gdoc 个人空间的查询以及基于某些特定 IP 信息服务的查询。这些查询方式的目的都是为了获得代理服务器的 IP 信息。当用户在一台计算机上第一次使用无界浏览软件或者缓存文件中的代理服务器信息失效时，软件就会采用上面介绍的三种方式进行代理服务器的搜索。

### 2.2.1 基于 DNS 的查询

DNS 代理查询方式是无界浏览最重要的代理查询方式。它的工作原理是通过向特定 DNS 服务器发送请求，得到的回复中就包含了代理服务器的 IP 地址，然后利用 SSL 和代理服务器进行加密通信，这种查询方式具有一定的隐蔽性。

无界浏览软件内置了一组利用算法 2 加密的特定 DNS 服务器信息：共有 351 个 IP，199 个特殊的查询域名。

其中部分 IP 地址列表如下：

63.251.129.1
12.11.148.11
130.36.31.4
169.144.68.7
208.255.120.34

部分域名列表如下：

www.CNPHOTO.info
www.ZXTGD.info
www.YCWEB.info
www.OA16.info
www.CONOCONPHILLIPS.info

同时软件内部还存有 336 个字符串。在进行 DNS 查询之前会任意选取几个字符串向特定 DNS 服务器发送，软件内部称为“fake query”，然后才会随机选择特殊域名发送给随机挑选的数个内置 DNS 服务器，收到查询返回的信息后经过变换，软件就取得了代理服务器的 IP 地址。发送 DNS 请求的通信方式和正常 DNS 请求一样。虽然无界浏览软件内置了很多 DNS 服务器信息，然而对于特定的计算机，每次选择的 DNS 服务器信息是固定的。因为随机值是根据硬盘和网卡变换得到的。

### 2.2.2 基于 gdoc 个人空间的代理查询

无界浏览软件引入了一种新的隐蔽查询方式，即通过目前网络上一些门户网站提供的个人服务项目作为载体，存放其代理服务器的信息。通过反汇编发现软件内置了 google doc 的域名为 https://docs.google.com/View?docid=dd4gbd38_6c8fpk2。如果 DNS 代理查询方式得不到可用的代理服务器信息，软件将使用 google doc 查询方式。首先连接到 http://docs.google.com，然后将上述域名信息中的最后一部分个人身份标识信息随机拆分成数段进行发送，全部发送完毕后等待 HTTP 应答。收到应答后，将其中的非字母字符全部去掉，然后以第 1-8 个字符与第 13－20 个字符为参数，进行变换即可得到真实的代理服务器 IP 地址。

### 2.2.3 基于某些特定 IP 的代理查询

当软件使用前两种查询方式都无法获得代理服务器信息时，会使用最后一种方式，即利用软件内置的 5 个特定 IP 地址进行信息查询，这些地址是：

211.74.78.17
66.245.217.9
66.245.217.227
66.245.196.247
118.168.50.105

分析过程中发现，软件直接对上述地址发起访问并获得信息，然后将这些信息转化为可用的代理服务器信息。

无界浏览软件通过上述 3 种方式获得代理服务器的 IP 后就开始加密通信的过程。在开始通信之前，软件会修改本地计算机 IE 浏览器的代理设置。默然情况下，软件将 IE 的本地代理设置为 127.0.0.1，端口为 9666，监听 127.0.0.1:9666 上的数据请求并将其转发至搜索到的代理服务器。如果用户已知代理服务器的 IP 地址，可以直接修改此设置，令本地数据请求直接与已知的代理服务器通信。本地代理设置好后就可利用标准的 HTTPS 开始和服务器端建立加密通信。

## 3 无界浏览软件的网络结构和安全性分析

图 2 为软件的网络拓扑结构。从图 2 中可以看到，在多种网络环境下，软件均可以利用 DNS 进行查询；通过门户网站个人空间服务等手段进行秘密通信使得防火墙无法屏蔽。代理服务器查找技术以及根据计算机环境不同而变化的选择查询服务器机制均可以逃避网络流量的分析，然而使用逆向分析手段，这些信息都可以被分析者得到[4]。

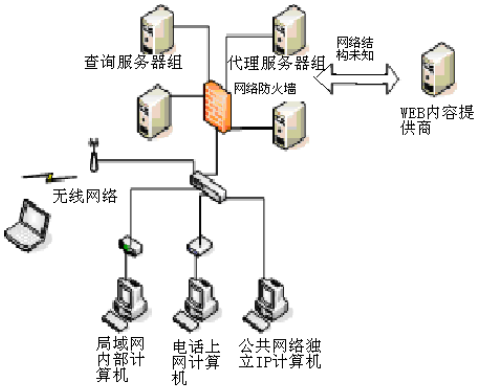


图 2 无界浏览网络拓扑结构

软件的通信过程利用了 HTTPS 的加密传输，这种解密方式可以绕过使用关键字匹配的分析机制。加密通信使用了具有一定强度的加密算法，然而其密钥附着于加密通信数据之中，因此其加密的安全性不能得到保证，通过通信侦听等手段可以对软件的加密通信进行解密。

## 4 验证测试

通过两个方面的测试证明了上述分析过程的正确性。第一，利用抓包软件及防火墙对软件进行黑盒分析。第二，将加解密算法代码实现，观察解密后的缓存内容是否为合法的 IP。

实验环境：一台安装了 Windows 操作系统的计算机(Client)，一台安装了 Linux 操作系统的计算机(Server)，在 Server 上安装有抓包工具 Ethereal[5]及防火墙 Iptables[6]。Client 直连到 Server 端，即所有 Client 网络通信都要通过 Server。

实验步骤：

（1）将所有分析所得 IP 列入 Iptables 封锁名单，并开启防火墙。在 Client 端，清除缓存并运行软件，会发现软件不能正常工作。其他网络通信活动均能正常进行。这证明笔者分析出的 IP 地址的正确性。

（2）关闭防火墙，开启抓包程序。在 Client 端，重新开启无界浏览，发现这次能够正常运行。关闭软件后会发现已经生成了缓存文件。

（3）利用解密算法将缓存文件中的 IP 解密出来。然后再重新开启防火墙及抓包程序，会发现软件仍能正常工作。通过分析抓包信息，能够看到软件是利用缓存解密出的 IP 信息进行的通信。

（4）清除缓存文件，将仅将 DNS 的 IP 信息列入封锁名单，重新开启防火墙及抓包程序。运行无界浏览后分析发现，软件在发送 DNS 请求失败后继而请求 gdoc 通信方式。

（5）清除缓存文件，并将 gdoc 的 URL 及相关的字符串信息列入封锁名单，重新开启防火墙及抓包程序。运行无界浏览，发现软件是通过特定 IP 建立连接的。

上述结论与分析所得结论相符，由此可判断笔者的分析是正确性的。

## 5 结语

无界浏览软件能够成为当前比较流行的自动化网络通信加密软件，其主要优势是具有广泛的查询服务器节点和多样的查询手段，且这些查询手段借助了当前广为使用的网络服务技术，同时其通信还具有一定的加密强度和隐蔽性。

为无界浏览软件提供查询和代理服务的服务器端的运行机制尚不可知，唯一可以了解的信息是它们返回的数据特点，通过观察查询返回的 DNS 结果及 gdoc 的内容会发现这些内容是不断更新的，因此猜测服务器端的数据信息也是处于不断调整之中。

### 参考文献

1 Baset S A, Schulzrinne H G. An Analysis of the Skype Peer-to-Peer Internet Telephony Protocol[C]. INFOCOM 2006. 25th IEEE International Conference on Computer Communications, Barcelona, Spain:Proceedings. April 2006:1-11.

2 段钢. 加密与解密[M].第 2 版，北京:电子工业出版社，2003.

3 飞天诚信,软件加密原理与应用[M].北京:电子工业出版社，2004.

4 看雪学院.软件加密技术内幕[M].北京:电子工业出版社，2006.

5 Sharpe Richard, Warnicke Ed, and Lamping Ulf, Ethereal User's Guide V2.00 for Ethereal 0.10.5[EB/OL].http://www.rootsecure.net/ content/downloads/pdf/ethereal_guide.pdf, 2004.

6 Andreasson Oskar, Iptables Tutorial 1.2.2[EB/OL].http://iptables- tutorial.frozentux.net/iptables-tutorial.html, 2006.