文章编号：1673－9868(2012)11－0132－05

# 一种利用三次握手进行信息隐藏的方法①

胡大辉，　杜治国

西南大学 信息管理系，重庆 荣昌 402460

**摘要**：传统的网络信息隐藏主要利用 TCP 报文头部的冗余位和保留位，但在传输过程中易被截取和破译. TCP 使用三次握手来建立连接，握手过程中产生的数据报序号是随机的，利用随机序号可以传输隐藏信息. 为提高隐藏信息的安全性和减小密文被破译的风险，采用一次一密的方式加密隐藏信息. 实验表明，此方法隐藏效果好，传输信息不易被破获，能穿透大多数防火墙并正常通过入侵检测系统.

**关　键　词**：信息隐藏；三次握手；传输控制协议；防火墙；入侵检测系统

**中图分类号**：TP309　　**文献标志码**：A

传统信息隐藏的对象是静态数据，而网络信息隐藏的是动态数据，两者的隐藏原理完全不一样. 传统信息隐藏主要依赖人类感觉器官的不敏感性，网络信息隐藏基于网络协议在语法或语义上的冗余[1].

目前，多数网络使用的是 IPv4 版本，出于多方面的考虑，该版本在数据报的首部预留了一些冗余和可选字节. 通过精心设计和构造，可以利用这些字节进行信息隐藏传输. 这种方法不增加额外的带宽，可以达到隐藏传输的目的. 例如：TCP 头部的 101-106bit 就从未使用过，可以把需要隐藏的信息分批次放入该位置，从而达到信息隐蔽传输的目的，TCP 报文头部结构如图 1 所示.

大多数的研究者利用保留字、不常用的紧急指针(URG)等进行信息隐藏，这些方法经实验证实是可行的，但由于协议简单，可用位数较少且报文格式已知，因此在实际应用中很容易被破译者获取[2]. 本文利用 TCP 建立连接时三次握手过程中通信发起方发出的数据报的随机序号来隐藏信息，从而达到隐蔽传输数据的目的.

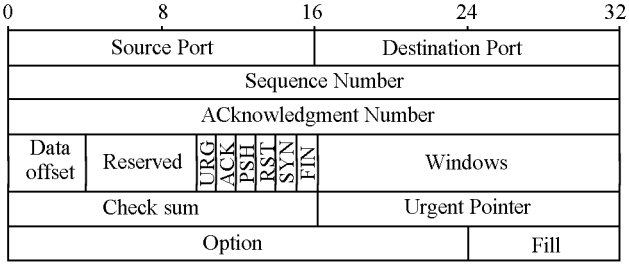


**图 1　TCP 报文段的首部格式**

## 1　三次握手

### 1.1　工作原理

由于 TCP 是可靠的有连接服务，因此 TCP 通信的过程中，A、B 主机之间必须建立连接. TCP 连接过程看似简单，其实很复杂. 为解决此问题，TomLinson 引入了三次握手(Three-way handshake)过程，三次握手的工作过程如图 2 所示[3]. 通过三次握手，通信双方建立可靠连接并传送数据. 由于建立连接的过程并不要求双方以同样的序号开始发送数据，所以它可以与一些不要求全局同步的方法一起使用. TCP 序列号是一个 32 位的字段，用以标识从 TCP 发送端向 TCP 接收端发送的数据字节流，它保证了传输的可靠

① 收稿日期：2011－05－13

作者简介：胡大辉(1977－)，女，重庆大足人，讲师，硕士，主要从事计算机信息安全方面的研究.

性，正常的 TCP 连接建立分 3 步完成.

① 首先 B 打开传输控制模块(TCB)，等待对方的连接请求. TCB 主要包含连接中的一些重要信息，如 TCP 连接表等.

② A 的 TCP 进程也创建 TCB，然后向 B 发出连接请求报文段，这时 TCP 报文中的同步位 SYN 置 1，同时任意选择一个 SEQ 作为发送序号，例如 $x$.

③ B 收到连接请求报文后，同意建立连接，则向 A 发送 ACK，在确认的报文中把 SYN 和 ACK 位同时置 1，由于收到的序号是 $x$，则 ACK 的确认号是 $x+1$，B 同时自己随机选择一个序号 $y$.

④ A 收到 B 的确认报文后，继续向 B 确认. 发送的报文 ACK 置 1，确认号为 $y+1$，A 自己的序号为 $x+1$，这时 TCP 连接已经建立，A 进入可连接状态，B 收到确认报文后，也进入可连接状态.

### 1.2 利用三次握手隐藏信息的方法

由于三次握手过程中双方的 SEQ 序号是随机的，因此可以利用随机的序号来隐藏信号，把需要隐藏信息的 ASC 码以序号的形式发送出去，其工作过程如图 3 所示[4].

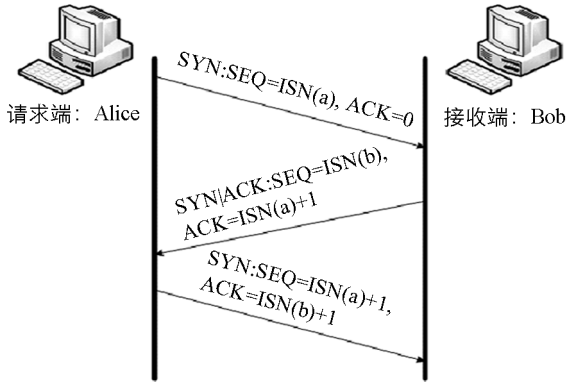


图 2 TCP 的三次握手过程

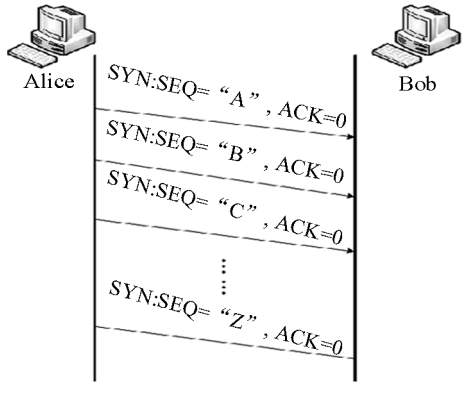


图 3 利用三次握手进行信息隐藏

由于在工作过程中，A 不断发送连接请求报文，B 必须设定相应的规则，否则 B 会认为 A 在对其进行泛洪攻击(SYN Flood). B 设置的规则中最少包含两个方面：

① 可以不间断的接收 A 发来的报文且不认为 A 是在利用泛洪攻击；

② B 不能也不需要对 A 发送 ACK 确认报文.

A 在发送多个报文时必须有一定的间隙时长 $\Delta\tau$，这是为了防止到达 B 时产生数据报序号混乱的情况，经测试可知，$\Delta\tau>=1$ s 时，基本不会产生序号混乱现象.

但是在实际的应用中，必须对 SEQ 序号中隐藏的信息进行加密处理，否则很容易被攻击者破译.

## 2 隐藏信息的加密

为了保证隐藏信息不被攻击者利用，必须采用安全可靠的方式对传输信息加密，本文使用的加密方法依据下列原理.

**定理 1** 设试验 $E$ 的样本空间为 $S$，$A$ 为 $E$ 的事件，$B_1,B_2,\cdots,B_n$ 为 $S$ 的一个划分，且 $P(B_i)>0(i=1,2,\cdots,n)$[5]，则

$$P(A)=P(A\mid B_1)P(B_1)+P(A\mid B_2)P(B_2)+\cdots+P(A\mid B_n)P(B_n)$$

**定理 2** 设试验 $E$ 的样本空间为 $S$，$A$ 为 $E$ 的事件，$B_1,B_2,\cdots,B_n$ 为 $S$ 的一个划分，且 $P(A)>0$，$P(B_i)>0(i=1,2,\cdots,n)$，则

$$P(B_i\mid A)=\frac{P(A\mid B_i)P(B_i)}{\sum_{j=1}^{n}P(A\mid B_j)P(B_j)}$$

其中，$i=1,2,\cdots n$[6].

**定理 3** 设$(P,C,K,E,D)$代表一个加密系统，其中$|K|=|C|=|P|$，$|*|$代表 $*$ 的个数，系

统保密的充要条件是每个密钥 $k \in K$ 被等概率 $1/|K|$ 使用，且对每一个 $x \in P$ 与 $y \in C$，必须有唯一的 $k$ 与之对应，使 $C_k(x)=y$[7].

**推论**　明文 $x \in P$ 发生的概率记作 $P_p(x)$ 称为先验概率，密钥 $k \in K$ 被选择用来加密 $x$ 的概率为 $P_k(K)$，密文 $y \in C$ 的概率分布应为所有密钥 $k$ 与明文 $x$ 加密能得到密文 $y$ 的概率之和[8]. 即

$$P_c(y)=\sum_{\{k:\ y\in C(k)\}} p_K(k)P_p(d_k(y)) \tag{1}$$

其中 $d_k(y)$ 表示用密钥 $k$ 对密文 $y$ 解密得到明文 $x$. 求和范围是所有产生密文 $y$ 的密钥 $k$ 的集合，具有加密功能的 $k$ 的概率和为：

$$Pc(y\mid x)=\sum_{\{k:\ x=d_k(y)\}} P_K(k) \tag{2}$$

**应用**　设有集合 $P=C=K=Z_{26}$（$Z_{26}$ 即字母表），表示明文、密文与密钥都来至 26 个字母组成的字母表. 加密算法为 $e_k(x)=x+k \bmod 26(x \in P=Z_{26})$[9].

由公式(1) 计算密文概率 $P_c(y)$，如果 26 字母等概率选择，则

$$P_c(y)=\sum_{k\in \mathbf{Z}_{26}} p_k(k)p_p(d_k(y))$$

即

$$p_c(y)=\sum_{k\in \mathbf{Z}_{26}} \frac{1}{26}p_p(y-k) \tag{3}$$

其中 $P_p(y-k)$ 表示明文 $x=y-k$ 的概率，对一个不确定的 $y$ 而言，实际是指 $k$ 的概率. 而 $k \in \mathbf{Z}_{26}$，即 $k$ 的概率和为 1，所以由(3) 式可知

$$p_c(y)=\frac{1}{26}\sum_{k\in \mathbf{Z}_{26}} p_p(y-k)=\frac{1}{26}$$

假设 $k$ 等概率均匀分布，由式(2) 则有：

$$P_c(y\mid x)=P_K(k)=P_K(y-x \bmod 26)=\frac{1}{26}$$

由定理 2 可知，

$$p(x\mid y)=\frac{p(x)p(y\mid x)}{p(y)}$$

将其与定理 3 结合使用，可知：

$$P_p(x\mid y)=\frac{P_p(x)\sum\limits_{\{k:\ x=d_k(y)\}} p_k(k)}{\sum\limits_{\{k:\ y\in C(k)\}} p_k(k)p_p(d_k(y))} \tag{4}$$

把公式(4) 应用本例中，即

$$p_p(x\mid y)=\frac{p_p(x)\dfrac{1}{26}}{\dfrac{1}{26}}=p_p(x)$$

由此可知，$x$ 与 $y$ 是相互独立统计的，由此可知这种方法是理想的密码体系.

因为密钥取自 26 个字母的所有排列，就有 26！种方法，如果不知道密钥，采用计算机暴力破译方法破解，1 μs 尝试 1 个解，则需要 26！ $*1\ \mu s=4*10^{26}\ \mu s=10^{13}$ 年即 $10^5$ 亿年才能破解，换句话说，也就是不可破译.

假设 $a$，$b$，$c$，…，$z$ 用 0，1，2…，25 表示，则明文“southwest”用密钥“university”加密过程为：

$y_1=(18+20) \bmod 26=12$

$y_2=(14+13) \bmod 26=1$

$y_3=(20+8) \bmod 26=2$

$y_4=(19+21) \bmod 26=14$

$y_5=(7+4) \bmod 26=11$

$y_6=(22+17)\bmod 26=13$

$y_7=(4+18)\bmod 26=22$

$y_8=(18+8)\bmod 26=0$

$y_9=(19+24)\bmod 26=17$

即加密后的密文为：mbcolnwar.

## 3 实现与测试

实验在局域网环境中进行，交换机使用100Mb/s的以太网交换机，两台主机地址分别是10.240.37.1(发送端)和10.240.37.2(接收端)，两台主机都使用Windows XP SP3操作系统，开发的工具软件是Microsoft Visual C++和WinPcap.

### 3.1 隐藏信息的实现

为防止网络中出现相同数据报，利用TCP数据报头部的6个保留位进行标识设置. 将其置为101010时表示正在传送隐藏信息；将其置为101011时表示隐藏信息传送完毕[10].

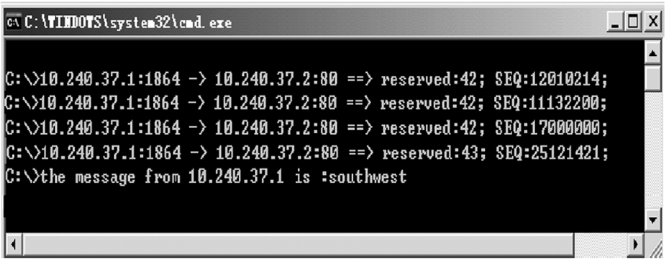


**图4 隐藏信息提取过程**

把明文“southwest”用密钥“university”加密后，捕获网络中传输的数据包，对其首部进行分析结果如图4所示. 由于双方知道加密的方式，根据密钥很快就能得到明文.

在图4中，reserved=42，即是reserved字段值是101010，表示开始正常传输数据；reserved=43，即是reserved字段值是101011，表示数据传输完毕，这时的SEQ序号包含的信息没有任何价值，仅仅是一个随机序号. 由于SEQ序号是32位的，所以每次只能传送4位字母. 在本例中，一共发送9个字母，则第三次SEQ序号只有一个有效字符，其余都以0来填充.

### 3.2 防火墙穿透测试

为测试本方法通过防火墙的效果，在接收端主机上分别安装如表1所示的防火墙，设置不同的安全等级，最后发送端发出的数据报穿透防火墙的结果如表1所示(“√”表示通过，“×”表示未通过). 实验证明，本方法产生的数据报可以通过绝大多数的防火墙.

**表1 防火墙穿透测试结果**

| 防火墙名称 | 低级 | 中级 | 高级 | 防火墙名称 | 低级 | 中级 | 高级 |
|---|---|---|---|---|---|---|---|
| Rising | √ | √ | √ | VRV | √ | √ | × |
| Kaspersky | √ | √ | √ | KILL | √ | √ | √ |
| Norton | √ | √ | × | 360 | √ | √ | √ |
| KingSoft | √ | √ | √ | JIANGMIN | √ | √ | × |

### 3.3 入侵检测系统测试

使用入侵检测系统来测试本方法产生的数据报，在网络中安装如表2所列举的入侵检测系统，最后的测试结果如表2所示. 实验结果表明，入侵监检系统不会对本方法产生的数据报做任何的动作与反应.

**表2 入侵检测系统测试结果**

| 入侵检测系统名 | 检测结果 | 备注 | 入侵检测系统名 | 检测结果 | 备注 |
|---|---|---|---|---|---|
| DIDS | Pass | | AAFID | Pass | |
| CSM | Pass | | GrIDS | Pass | |
| EMERALD | Pass | | MAIDS | Pass | |

## 4 结 论

利用TCP协议进行信息隐藏的方法有多种，本文利用TCP建立连接时三次握手的机会来传输隐藏信

息，采用一次一密的方式对信息进行加密，加密算法在理论上不可破译. 实验表明，本方法具有一定的隐蔽性和不可破译性，产生的数据报能穿透大多数的防火墙，能通过所有的入侵检测系统.

本方法也存在两个需要改进的地方，一是加密算法中，密钥长度必须大于等于明文长度；二是在发送 TCP 报文时没有采用 ACK 确认报文，有较小的丢失数据报的概率.

**参考文献：**

[1] ANENBAUM A S. Computer Networks(Forth Edition) [M]. Peking: Tsinghua University Press, 2004.
[2] 杨　榆. 信息隐藏与数字水印实验教程 [M]. 北京：国防工业出版社，2010.
[3] 谢希仁. 计算机网络 [M]. 北京：电子工业出版社，2010：4.
[4] 杨智丹，刘克胜，李　丽. IPv6 中的网络隐隐蔽通道技术研究 [J]. 东南大学学报：自然科学版，2007，37(9)：141－148.
[5] 田丽华. 编码理论 [M]. 西安：西安电子科技大学出版社，2003.
[6] 戴善荣. 信息论与编码基础 [M]. 北京：机械工业出版社，2005.
[7] 盛　骤，谢式千，潘承毅. 概率论与数理统计 [M]. 北京：高等教育出版社，2001.
[8] 彭　静，候祥勇，马　燕. 一种自适应图像灰度水印算法 [J]. 西南大学学报：自然科学版，2009，31(7_：171－175.
[9] 李元东. 基于模糊信息的群体多维偏好分析决策模型 [J]. 西南师范大学学报：自然科学版，2009，34(10)：82－87.
[10] 陈园园，朱孝成，叶甬渝. 一种改进的 DCF 信息隐藏算法 [J]. 重庆理工大学学报：自然科学版，2011(12)：100－105.

# A Method of Steganogrhpy by Three-Way Handshake

HU Da-hui,　DU Zhi-guo

*Department of Information Management, Southwest University, Rongchang, Chongqing 402460, China*

**Abstract**: The traditional method for network steganogrhpy mainly makes use of the redundant bits and reserved bits at the head of TCP Datagram; however, it is vulnerable to information stealing and attacks. When establishing connection, TCP conducts a three-way handshake, during which a stochastic sequence number is produced. This stochastic number can be used for hidden information transmission. In order to improve the security of hidden information and reduce the risk of being deciphered, this study uses the way of one-time pad to encrypt the hidden information. The experiment results show that this method of hiding information is good in hiding information and not easy to be cracked, and can normally penetrate most firewalls and intrusion detection systems.

**Key words**: steganography; three-way handshake; TCP; firewall; intrusion detection system

**责任编辑**　汤振金