**SOCKS5通讯协议分析研究和改进**

**qwesdfok**

**qwesdfok@vip.qq.com**

**SOCKS5通讯协议分析研究和改进**

**摘 要**

由于校园网、防火墙、数据保护等原因，某些网络资源可能会出现不可访问的原因，必须通过局域网、内网等方式才能进行资源的访问。人们可以添加一些既可以在公网中能被访问，也可在内网中去访问网络资源的计算机，用以连接外网和内网，同时需要对传输的数据进行加密，以保证安全性。这类计算机被称为代理计算机。因此需要使用新的通讯协议，将外网的访问请求通过代理计算机转发到内网服务器。SOCKS5协议便是一种代理协议，用来解决如何通过代理计算机进行资源的访问的问题。SOCKS5协议定义了一系列的协议内容，用以满足外网节点与服务器交互的不同情况。但该协议并未涉及数据加密功能，因此论文将先介绍SOCKS5协议并进行分析，然后提出如何进行改进，以增加安全性。

**关键词** SOCKS5协议 代理 内网资源访问 加密

1. **协议介绍和分析**

通信协议是一种用来解决计算机之间信息交流时所遇到的问题的协议。计算机通讯时会产生许多的问题，比如因为某些原因无法访问互联网资源的计算机，将如何通过一台可访问的计算机，访问互联网资源。一个常见的实例便是在公司内，为了防止公司服务器关键资料外泄，需要对外网隐藏服务器的存在，但又要满足员工或者领导不在公司局域网内时也可以访问服务器，因此，需要通过“跳板”节点，实现外网对内网中服务器的访问。“跳板”计算机节点被称为代理节点，外界与代理节点之间变需要一个通讯协议进行数据的交互，而SOCKS5协议便是解决了这个问题。



图 1 SOCKS5协议网络拓扑结构

Sock5是一种基于TCP和UDP协议的未加密的不安全协议，其解决的问题是类似于简化的路由问题，即数据如何在一台对外或对内网可达的计算机进行转发的问题。其网络拓扑模型见图 1。

图中圆圈表示计算机或者一个发起访问的节点，比如PC机、Android或IOS只能手机或者某个需要访问网络的进程；有向箭头表示连接请求发送的路径，指向服务器节点。并且服务器节点拥有域名为http://www.test.com的网站，但由于正在进行测试，只对内部IP段开放，而代理节点拥有2张网卡，可以通过路由转发功能，转发外界的数据包。下文将使用图中的请求节点、代理节点和服务器节点，来指代协议中的三种实体。

SOCKS5协议定义了三种请求情况：CONNECT、BIND和UDP-ASSOCIATE。CONNECT请求是用来满足请求节点主动发起请求，对服务器节点进行连接的情况。BIND请求用来满足服务器需要向请求节点发起连接的情况，UDP-ASSOCIATE用来满足使用UDP协议的场景。

SOCKS5需要请求节点发送2个报文。当一个请求节点发起上述的三种情况的一种时，首先需要发送一个验证是否是SOCKS5协议的报文。该报文格式见表 1。

表 1 连接发起报文格式

|  |  |  |
| --- | --- | --- |
| VER | NMETHODS | METHODS |
| 1字节 | 1字节 | 1-255字节 |

其中VER表示SOCKS的协议版本号，SOCKS5协议中的VER字段值恒为0x05；NMETHODS为代理节点中的支持的METHOD的数量，数值上等于METHODS的字节数量；METHODS表示代理节点支持的协议认证方式，常用的有0x00无需认证、0x02用户名-密码方式认证和0xFF无可接收方法。一般常见的为0x00，即无需认证，任何人都可以进行连接的建立。

代理节点接收到请求节点发送的报文后，选择METHODS字段中的一种认证方式，并发回给请求节点。返回报文的格式如下表格所示。

表 2 连接响应报文格式

|  |  |
| --- | --- |
| VER | METHOD |
| 1字节 | 1字节 |

其中VER同样表示协议版本号，0x05；METHOD为代理选择的认证方式，一般为0x00。

请求节点收到连接响应报文后，即可知道代理节点工作正常，可以进行请求的代理服务。因此，请求节点将发送代理请求报文，用以通过代理节点访问服务器资源。其格式如下：

表 3 代理请求报文格式

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| VER | CMD | RSV | ATYPE | DST-ADDR | DST-PORT |
| 1字节 | 1字节 | 1字节 | 1字节 | 不定长n | 2字节 |

其中VER为0x05；RSV为保留字段，为0x00；CMD可选值为0x01，0x02和0x03，分别表示CONNECT请求、BIND请求和UDP-ASSOCIATE请求；ATYPE表示后续IP地址的类型，0x01表示为IPv4地址，0x03表示为字符串域名，需要通过DNS解析，0x04表示IPv6地址；DST-ADDR表示所请求的服务器的地址；DST-PORT表示端口。

以图 1为例，CMD为0x01，表示主动访问，ATYPE可以是0x01，DST-ADDR则为IPv4的二进制表示格式0x0A010101或者ATYPE为0x03，则DST-ADDR的第一字节表示后续的DST-ADDR的长度，字符串http://www.test.com共有20字符，则第一个字节为0x14，后续为字符串的ANSCII编码；最后的DST-PORT为0x0050，即80端口。

代理服务器收到该报文后，会根据请求类型进行处理。若为CONNECT请求，则直接与DST-ADDR的DST-PORT建立连接，然后将原先的与请求节点的连接与跟服务器建立的连接进行数据转发。若为BIND请求，则代理节点新监听一个端口，等待服务器的连接请求，然后与CONNECT一样，转发原先的与请求节点的连接的数据。若为UDP-ASSOCIATE，与CONNNECT类似，不过将TCP变成了UDP。

处理完请求后，代理节点需要向请求节点反馈连接的结果。反馈报文格式如下：

表 4 反馈报文格式

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| VER | REP | RSV | ATYPE | BND-ADDR | BND-PORT |
| 1字节 | 1字节 | 1字节 | 1字节 | 不定长n | 2字节 |

其中VER为0x05；REP表示连接建立的结果，常见的有0x00成功建立连接，0x03网络不可达，0x04主机不可达，0x05连接被拒绝；RSV保留字段，0x00,；ATYPE与后续的BND-ADDR和BND-PORT与之前的ATYPE、DST-ADDR和DST-PORT类似。

1. 若先前的为CONNECT请求，则BND-ADDR和PORT为代理节点的IP地址和PORT，若进行了负载均衡，即有多个代理节点，则该值可能会不同于请求节点所请求的IP和PORT。
2. 若为BIND请求，则会发送2次反馈报文，第一次的报文的BND-ADDR和PORT为代理节点所监听的端口，等待服务器节点的连接；第二次报文表明服务器是否成功建立了连接，这里的ADDR和PORT表示服务器节点连接的IP和PORT。
3. 若为UDP-ASSOCIATE请求，则BND-ADDR和PORT表示请求节点后续UDP包所需要发送的地址。
4. **协议安全性研究与改进**

从协议内容中可以看出，SOCKS5协议并没有对所传输的内容进行加密，容易被第三方监听、篡改。因此，在实际应用中，可以修改协议的拓扑结构，与HTTPS中的SSL层类似，添加加密模块。加密模块既可以是对称的，也可以是非对称的，本文采用对称加密的AES加密方式进行数据的加密。修改后的拓扑结构见图 2。



图 2 新拓扑结构

原先的结构被进一步细分，其中，请求节点被细分为App和Client，App值使用代理的程序，比如浏览器，Client表示提供SOCKS5协议的接入口，App将SOCKS5协议数据通过Client发送给代理节点。代理节点中运行的是Server进程，该进程提供解密和解析SOCKS5协议的服务，解析服务是上述所说的SOCKS5协议内容。服务器节点中的Target是指请求节点所要访问的内容，在例子中是指HTTP网页。

此时，若以数据的视角来观测新的拓扑结构，则如下图所示：



图 3 数据视角的新拓扑结构

其中IOStream表示输入输出流，EncryptStream表示加密输出流，DecryptStream表示解密输出流，箭头方向表示数据流动的方向。

由于添加了加密模块，因此需要新的加密协议，用以封装原有的SOCKS5协议。加密协议的数据包格式如下：

表 5 加密协议格式

|  |  |  |  |
| --- | --- | --- | --- |
| VERSION | MAGIC | SIZE | DATA |
| 1字节 | 1字节 | 4字节 | N字节 |

其中VERSION表示版本，为0x01；MAGIC表示用于验证协议的常数，为0xAD；SIZE表示后续的数据的大小；DATA表示加密后的数据。

通过加密协议，并且双方通过预先设置，共享密钥数据，然后通过IO流发送数据。

1. **结论**

本论文主要讲述了SOCKS5通讯协议的内容，介绍了该协议通过定义3种主要的请求类型的方法，解决了请求节点如何通过代理节点进行资源的访问的问题，并且分析了SOCKS5协议不能加密数据而容易导致数据泄露的缺点，最后提出了一种简单有效的改进方法。

SOCKS5协议很好的解决了代理问题，能够让外界的请求节点访问内网数据，因此现在很多应用程序支持SOCKS5协议，比如Chrome浏览器、QQ等应用程序。而使用加密协议后，SOCKS5将变得更加安全和可靠，可以从一定程度上防止数据泄露。