

文章编号: 1000-5013(2007)03-0268-04

一种基于 SOCKS 5 的 Web 安全代理技术

喻小光, 陈维斌, 潘孝铭

(华侨大学 信息科学与工程学院, 福建 泉州 362021)

摘要: 综合公钥密码算法(RSA)和数据加密标准(DES)的优势,用 DES 方法加密待传送的 Web 数据,使用 RSA 方法对 DES 密钥进行加密,提出一种基于 SOCKS 5 的, RSA 和 DES 相结合的 Web 安全代理方案. 通过性能测试和安全性分析表明,使用 RSA 传递 DES 密钥可保证 DES 每次加密都使用新的密钥,杜绝黑客通过分析明文/密文获得密钥,从而能防御网络数据包被恶意截取后造成泄密,安全代理既保证 Web 数据的加密速度,又保证 DES 密钥的安全性和可管理性.

关键词: 网络安全; 代理; SOCKS 5; 公钥密码算法; 数据加密标准

中图分类号: TP 393.08

文献标识码: A

由于 Internet 是一个基于 TCP/IP 的开放系统,几乎所有的数据都是以明文方式在网络间传递,因此数据传输的安全性较差,直接影响 Web 应用的商业前景. 目前解决这个问题主要方法大致分为两类:安全协议法和代理服务器法. 安全协议法实施过程比较简单,但其安全性受到协议本身的限制,如加密算法和密钥强度的选择等,使得该方法不够灵活;而代理服务器法在这些方面具有优势. 本文提出一种基于 SOCKS 5 的 Web 安全代理方案,用以保障 Web 服务器与 Web 客户端之间的数据传输安全.

1 Web 安全代理框架

Web 安全代理模型,如图 1 所示. 图中,CP 为客户代理,SP 为服务器代理,虚线框内为客户代理集合,它们与一个服务器代理共同构成安全代理系统. 客户代理和服务器代理分别负责 Web 客户机和

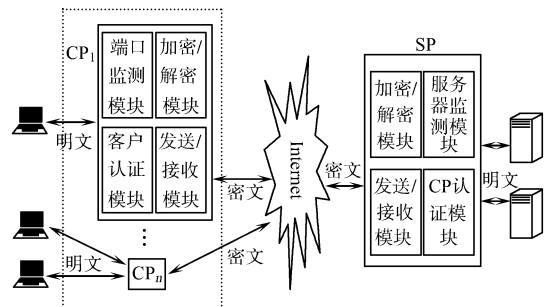


图 1 Web 安全代理模型

Fig. 1 Model of Web security proxy

Web 服务器发送、接收数据的加密/解密处理,两者必须成对出现. 服务器代理对一个或多个 Web 服务器进行监控,负责其数据传输安全;客户代理对一个或多个 Web 客户机进行监控,负责其数据传输安全. 安全代理只对注册(在安装安全代理时进行注册,或运行过程中由管理员注册)的 Web 服务器和 Web 客户机之间的数据进行保密传送. 服务器代理只受理客户代理发送过来的加密过的 Web 请求,客户代理也只受理注册的 Web 客户机发送给注册 Web 服务器的请求. 这样,分别由服务器代理和客户代理把守住 Web 服务器和 Web 客户机的数据出入口,对传出数据进行加密以保证安全.

2 Web 客户代理与服务器代理

2.1 代理工作协议的选择

SOCKS 5 是一个基于客户机/服务器(C/S)模式网络环境的代理服务器实现的规范协议,它独立于应用层协议,并可以应用于多种不同的服务,因此,选择 SOCKS 5 作为代理的工作协议. 其原理^[1]是:

收稿日期: 2006-11-07

作者简介: 喻小光(1976-),男,讲师,主要从事数据库技术及网络安全的研究. E-mail: xiaoyu@hqu.edu.cn.

基金项目: 福建省青年科技人才创新项目(2002J011); 华侨大学科研基金资助项目(04HZR17)

SOCKS 5 代理服务器和客户建立联系后,客户的数据包都以 TCP(Transfer Control Protocol)数据包的形式发送到 SOCKS 5 代理服务器,然后由其转发,并将接收的数据分发给相应的客户。

2.2 客户代理

对客户机 Web 通信的接管,可以通过对客户机浏览器的“代理服务器”选项进行设置来完成,加密则由基于公钥密码算法(RSA)和数据加密标准(DES)的加密方案来实现。同时,通过一个本地的客户信息文件(Client Info)来验证用户是否合法(验证用户名和口令),以及判断收到的请求是否指向受保护的 Web 服务器。客户信息文件使用 XML(Extensible Markup Language)编写,并经过加密处理,其文档格式如下:

```
< Client Info >
  < ServerProxy >
    < ProxyIP > </ ProxyIP >
    < ProxyPort > </ ProxyPort >
  </ ServerProxy >
  < WebServerList >
    < ServerIP > </ ServerIP >
    < ServerPort > </ ServerPort >
    ...
  </ WebServerList >
  < UserList >
    < UserName > </ UserName >
    < UserPwd > </ UserPwd >
    < UserUnit > </ UserUnit >
    ...
  </ UserList >
</ Client Info >
```

图中,ServerProxy 元素存储该客户代理对应的服务器代理的相关信息,WebServerList 元素存储所有 Web 服务器信息,UserList 元素存储所有用户信息。客户代理的工作流程,如图 2 中虚线上部所示。

2.3 服务器代理

对 Web 服务器 HTTP 服务通信的接管,是通过对 HTTP 服务器端口的监听实现,加密也由基于 RSA 和 DES 的加密方案来实现。服务器代理通过一个本地的服务器信息文件(ServerInfo)来验证客户代理是否合法(验证 IP 地址和端口)。其格式如下:

```
< ServerInfo >
  < WebServerList >
    < ServerIP > </ ServerIP >
    < ServerPort > </ ServerPort >
    ...
  </ WebServerList >
  < ClientList >
    < Client ID > </ Client ID >
    < Client IP > </ Client IP >
    < Client Port > </ Client Port >
  </ ClientList >
  < Client Proxy >
    ...
  </ ClientProxy >
</ ServerInfo >
```

其中,ClientList 元素存储该服务器代理对应的所有客户代理的相关信息,WebServerList 元素存储所有被监管的 Web 服务器信息。服务器代理的主要工作流程,如图 2 中虚线下部所示。

2.4 客户代理与服务器代理之间的身份认证

为了防止入侵者伪装成安全代理的一方发出请求或作出应答,安全代理客户端和服务端在通信

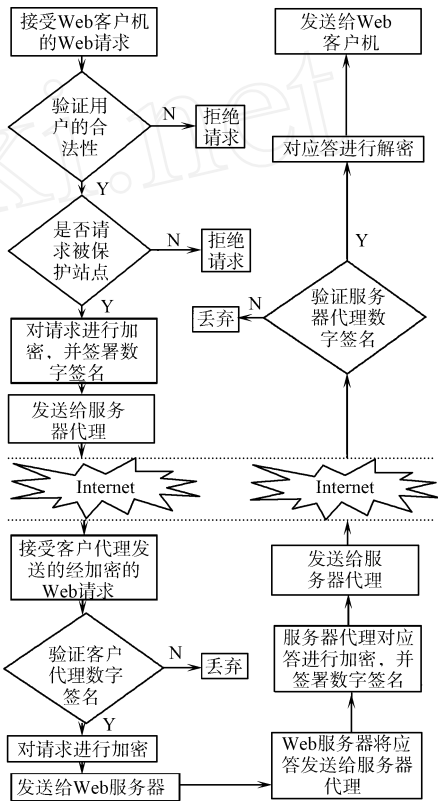


图 2 安全代理工作流程
Fig. 2 Work flow of security proxy

时必须互相进行身份认证。(1) 客户代理只向服务器代理发送请求,并且只接受服务器代理返回的应答;(2) 服务器代理也只接受客户代理发送的请求并作出应答。系统使用 RSA 公钥技术实现身份认证。服务器代理和客户代理在传输数据前对数据进行数字签名;在接收到数据后,首先检验数字签名,验证后方可进行后续操作。

3 对传输数据进行选择性加密

设发送方为 A(加密密钥为 K_{ea} ,解密密钥为 K_{da}),接收方为 B(加密密钥为 K_{eb} ,解密密钥为 K_{db}),加密方案的有如下 5 个具体实现步骤。(1) 发送方首先生成用于 DES 加密的密钥 K ,为了提高数据的安全性,每一个密钥 K 只用一次。(2) 发送方从密钥服务器中获取接收方的 RSA 的公开加密密钥 K_{eb} ,并用 K_{eb} 加密 DES 的密钥 K 形成密文 C_k 。(3) 发送方用 K 加密明文和签名的信息,然后连同 C_k 一起形成密文 C 发往接收方。(4) 接收方接收到 C 后,先用自己的解密密钥 K_{db} 解密出 C 中的 DES 密钥 K ,再利用 K 解密出明文。(5) 发送、接收双方均删除 DES 密钥 K 。

对所有的 Web 通信数据都进行加密会导致 Web 客户和服务器之间的通信效率严重降低,并且使得安全代理成为 Web 通信的瓶颈。事实上,不是所有的数据都需要保密传输。例如客户请求浏览一些非保密信息,就无需对服务器回应的数据进行加密。因此,可以设计一个提高 Web 安全代理工作效率的方案,即对客户机代理与服务器代理之间传送的数据进行分析,如果其中包含的敏感数据达到一定的阈值(敏感度),则判定这些数据应该加密;否则,用明文传送。计算待传送数据敏感度的流程,如图 3 所示。

方案中主要的技术难点在于如何确定待传送数据的敏感度。出于实用性的考虑,只对英文和中文数据的敏感度进行计算,其他语言的数据一律进行数据加密。英文是以词为单位组织的语言,可以根据数据中的标点符号和空格等进行分词处理。而中文是以字为单位组织的语言,无法直接根据数据中的标点符号和空格等进行分词处理^[4],还须辅以中文分词技术。

分词完毕后,将其与预定义的敏感词库中词汇进行匹配,计算出敏感词汇个数,进而计算出待传送数据的敏感度。敏感度 = 敏感词汇个数 / 传送数据词汇个数。敏感词库初始时由系统预先设定,包含常见敏感词汇,如用户名、密码、账号、机密、保密等,用户可以根据自己的实际情况增减敏感词库中的词汇。本方案中采用了最大匹配法,它具有简单高效、分词精度较高的特点。基于词典的传统最大匹配法是一种依据长词优先的机械分词方法,需要最少的语言资源,算法的基本思想是在被处理文档中取出一行字符串,将其中长度为 n 的字符串序列作为待匹配字段,借助分词词典,若词典中存在这样一个 n 字词,则匹配成功,将该词输出;否则,从匹配字符序列中去掉最后一个字重新进行匹配,如此进行直到成功。完成一次匹配后,再取下 n 个字符串序列进行上述匹配,直到文档扫描结束。

选择性加密方案的效率,依赖于敏感度设置及分词算法的效率和准确率。通过对大量 Web 数据传输进行测试,表明选择性加密是可行的,且对提高安全代理的效率有明显作用。使用选择性加密对保密性要求较低的应用尤为有效,但对于大多数通信数据都需要保密的应用,反而会降低安全代理的效率。

4 性能测试与分析

4.1 加密算法对传输数据量的影响

图 4 为密文与明文对传输数据量的影响。图 4 中, C 为需要加密(或传输)的文件大小, R_D 为因加密增加的数据占明文数据量的比例。从图 4 可知,当明文达到一个较大值(5 kB)后,加密造成的影响变得非常稳定,而且增加的数据量也非常有限,基本上不会给传输造成影响。

4.2 加密/解密时间对传输数据量的影响

加密/解密时间对传输数据量的影响,如图 5 所示。图 5 中, C 为需要加密(或传输)的文件大小, R_t 为加密/解密时间在系统时间中所占比率。数据测算时的系统工作环境:网络传输速度 $1\,178\text{ kbit} \cdot \text{s}^{-1}$ 。

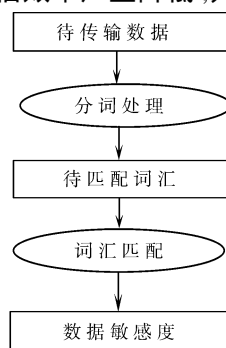


图 3 敏感度计算流程图

Fig. 3 Sensitivity algorithm

从图 5 可知,为传输系统引入安全代理而增加的数据处理时间非常有限,不会对系统正常运行造成明显

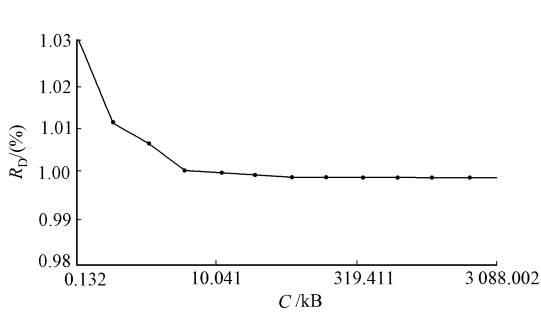


图 4 密文与明文大小比较

Fig. 4 Compare of encryption & plain text

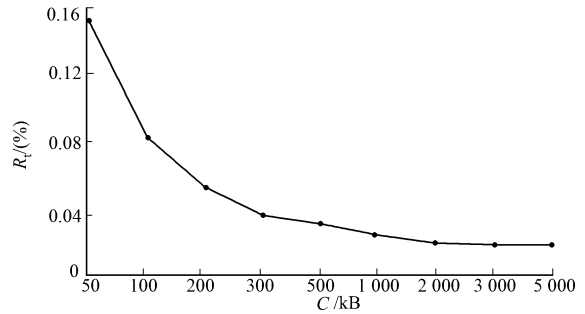


图 5 加密/解密时间在系统时间中的比率

Fig. 5 Encryption and decryption time ratio

影响.综合以上两点,安全代理的效率是有保障的.

4.3 安全性分析

本安全代理使用 DES 加密数据,能防御网络数据包被恶意截取后造成泄密.使用 RSA 传递 DES 密钥,并保证 DES 每次加密都使用新的密钥,使黑客试图通过分析明文/密文从而得出密钥的做法基本无法完成.因此,安全代理保证了数据的安全传输.

5 结束语

基于 SOCKS 5 的 Web 安全代理方案,是将 SOCKS 5 技术、代理服务器技术、RSA 和 DES 加密技术,以及中文分词技术有机地结合在一起,是一个行之有效的 Web 数据安全传输解决方案.进一步的工作设想是在服务器端引入多代理技术,从而实现服务器端代理负载平衡,减轻单个服务器端代理的压力,进而提高服务器端代理的工作效率.

参考文献:

- [1] 陈兵,王立松. SOCKS v5 服务器的研究与实现[J]. 数据采集与处理, 2002, 17(3): 349.
- [2] RONLD R. Adi shamir and leonard adleman: A method for obtaining digital signatures and public-key cryptosystems[J]. Communications of ACM, 1978, 21(2): 120-126.
- [3] 张乐星. 基于 RSA 和高级数据加密标准的网络数据加密方案[J]. 科技通报, 2005, 21(2): 198.
- [4] 苏芳仲,林世平. Web 文本挖掘中的一种中文分词算法研究及其实现[J]. 福州大学学报:自然科学版, 2004, 32(1): 67.
- [5] 于源,衣裘. 中文全切分快速分词方法[J]. 大连铁道学院学报:自然科学版, 2005, 26(2): 84.

The Research on the Technique of Web Security Proxy Based on SOCKS 5

YU Xiao-guang, CHEN Wei-bin, PAN Xiao-ming

(College of Information Science and Engineering, Huaqiao University, Quanzhou 362021, China)

Abstract: Synthesizing the advantages of rivest shamir adleman (RSA) and data encryption standard (DES). Using the method of DES to transmit the encrypt Web data and using RSA to encrypt the encryption key of DES. A Web security proxyscheme combined RSA & DES method based on SOCKS 5 was proposed. The performance test and security analysis show that RSA could guarantee a new encryption key used in each DES encryption, and prevent the hacker to get the encryption key by analysis of plain text and cryptograph. Therefore, the scheme ensured that the secrecy would not be lost even the net data package had been intercepted by hackers. The Web security proxy also ensured the encryption speed, as well as the security and administration of encryption key.

Keywords: network security; proxy; SOCKS 5; RSA; DES

(责任编辑:黄仲一)