对Web数据的安全监控研究*)

王旭仁[1,2] 毕学尧[2] 曹爱娟[2] 李雪莹[2,3] 许榕生[2]

(首都师范大学信息工程学院 北京100037)[1] (中科院高能物理所计算中心 北京100039)[2]

(军事医学科学院医学情报研究所网络信息中心 北京100850)[3]

摘 要 Web已成为许多人日常生活、工作及交流的一种主要工具,针对Web数据存在的信息安全问题,本文提出了实时网络安全监控系统(Real-time Network Security Monitoring System,简记为RNSMS),能够对Web数据(基于HTTP协议和SSL协议的数据)进行实时、有效的获取、重建、监控和存储,在百兆网络环境下丢包率<5%。应用表明RNSMS是一种实用、有效的Web数据安全监控工具。

关键词 网络安全,网络监控,HTTP协议,SSL协议

Research of Web Data Security Monitoring

WANG Xu-Ren[1,2] BI Xue-Yao[2] CAO Ai-Juan[2] LI Xue-Ying[2,3] XU Rong-Sheng[2]

(Information Engineering College of Capital Normal University,Beijing100037)[1]

(Computing Center,Institute of High Energy Physics,CAS,Beijing100039)[2]

(Network Information Center,Institute of Medical Information,Academy of Military Medical Sciences,Beijing 100850)[3]

Abstract Web has already become a common tool in people's daily life,work and communication. In order to solve information security problems in Web data,Real-time Network Security Monitoring System has been proposed in this paper,which can efficiently capture,reconstruct,monitor and restore the Web data by real-time. The loss rate of data packets under 100M networks is lower than 5 percent. Tests show that the system is an applicable and efficient tool in monitoring Web data.

Keywords Network security,Network monitoring,HTTP protocols,SSL protocols

1 前言

World Wide Web称为万维网,简称Web。它的基本结构是采用开放式的客户/服务器结构(Client/Server),分成服务器端、客户接收机及通讯协议三个部分。

服务器结构中规定了服务器的传输设定、信息传输格式及服务器本身的基本开放结构。客户机系统称为Web浏览器,用于向服务器发送资源索取请求,并将接收到的信息进行解码和显示。Web浏览器与服务器之间遵循HTTP协议进行通讯传输。HTTP(Hyper Text Transfer Protocol,超文本传输协议)[1]是分布式的Web应用的核心技术协议,在TCP/IP协议栈中属于应用层。它定义了Web浏览器向Web服务器发送索取Web页面请求的格式,以及Web页面在Internet上的传输方式。

Internet及Web已成为许多人日常生活、工作及交流的一种主要工具。Web的安全问题包括以下几个方面:

1)Web技术的安全 Web赖以生存的环境包括计算机硬件、操作系统、计算机网络、许多的网络服务和应用,所有这些都存在着安全隐患,最终威胁到Web的安全。包括:服务器端的安全;客户端软件(即Web浏览器软件)的安全;运行浏览器的计算机设备及其操作系统的安全(主机系统安全);客户端的局域网(LAN)的安全;服务器端的局域网(LAN)安全;运行服务器的计算机设备及操作系统的安全(主机系统的安全);服务器上的Web服务器软件的安全。

2)Web信息内容的安全 由于Web的广泛使用,而Web上的信息充斥着色情、暴力、无用、反动的各种各样的信息,对Web信息进行监控分析是教育部门、公司企业、安全部门的迫切需要。例如父母对孩子浏览网页信息内容的担忧;企业员工浏览网页信息,对员工的工作效率造成影响,甚至会给公司带来法律问题;对于特权部门,例如公安、军队部门

*)该课题得到国家重点基础研究发展规划(973)项目(编号:G1999035806)资助。王旭仁 博士生,从事信息网络安全、数据挖掘研究;毕学尧 博士,从事网络安全研究;曹爱娟 博士生,从事网络安全研究;李雪莹 博士生,从事网络安全研究;许榕生 研究员,博导,从事信息网络安全研究。

来说，需要了解被监控的对象都在浏览哪些网页以及网页的内容。

3）Web 协议的安全　利用 Web 通讯协议（例如利用 Http 协议发动拒绝服务攻击）进行网络攻击，是网络安全中的一种攻击情形。需要利用分析工具对 HTTP 协议包进行分析，提取攻击模式，以便对攻击进行有效的检测。

本文从 Web 应用的角度出发，主要研究第二种 Web 安全问题。本文作者参与研究开发了实时网络安全监控系统，它能够实现对 Web 信息进行实时获取、实时监控、重建 http 数据包和进行存储等功能。

本文结构如下：第一部分是前言；第二部分是介绍实时网络安全监控系统（RNSMS）的设计和实现；第三部分是监控试验和分析；第四部分是关键技术研究探讨；最后一部分是结论。

2　实时网络安全监控系统

作者参与开发设计的实时网络安全监控系统（Real-time Network Security Monitoring System，简记为 RNSMS）就是以宽带网络建设为硬件平台，对网络信息的内容实现实时监控、分析的软件系统。图1是 RNSMS 的框架。

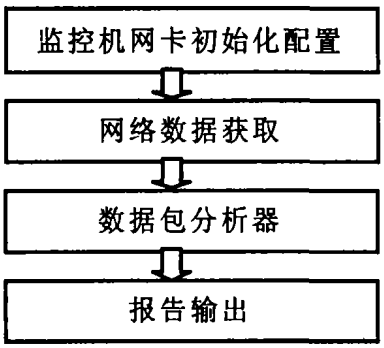

图1　RNSMS 实现框架

2.1　RNSMS 的工作模式

RNSMS 可选的工作模式有三种[2]：基于主机的模式、网络监听模式、网络过滤模式。下面分别加以简单介绍。

基于主机的模式　系统以后台程序的方式运行，这样会占用一定的主机资源。例如主机入侵检测系统（HIDS）使用该模式。

网络过滤模式　系统接收数据包，对数据包处理完毕后再转发，这样会造成一定的网络延迟。例如防火墙、代理网关就使用这一模式。

网络监听模式　系统处于内/外网络中间，但是却独立于两者。不占用网络资源，不会对网络正常通讯产生任何影响。网络入侵检测系统（NIDS）使用该模式。

由于 RNSMS 不仅接收数据包，还要对网络数据包进行处理分析，在网络过滤模式下工作，会成为网络瓶颈，因此 RNSMS 被设计成监听工作方式。

2.2　RNSMS 的功能模块

2.2.1　初始化配置模块　RNSMS 系统运行时，首先进行初始化配置：网卡监听模式的设置使得系统能够接收广播式局域网的数据包，而不影响网络正常运行；设置合适的数据缓冲区大小，缓冲区太小，会造成数据包来不及处理便被丢掉的严重问题；设置对网络数据包的过滤，由于内核空间的过滤不用拷贝数据，因而不占用 CPU 时间，过滤效率较高。

2.2.2　数据获取模块　数据获取模块使用高效的抓包驱动程序 WinPcap，使得整个抓包过程丢包率几乎为零。从网卡链路层获取网络数据包，验证连接是否完整；当一个连接的数据被接收下来后，将数据转给上一层的数据分析模块处理。RNSMS 系统具有功能强大的过滤器，允许用户灵活地监控某个主机、某些主机、某个网段上的所有数据包，并且几乎没有数据包丢失的情况。

2.2.3　数据分析器　实时数据分析器允许用户在抓包和分析的同时查看分析后的网页内容。实时数据分析器又分为若干子功能模块，提供文件重建、统计分析、审计日志等功能，从面到点、从不同的角度分析网络数据。统计分析模块对被监控网络进行流量监控，可以统计网络总流量、分时段流量、各 IP 流量，按照协议不同，分别统计各协议包流量，并绘制动态流量变化图。通过流量统计及流量变化图，管理员可以发现异常流量、突发流量以及网络流量分布，有利于管理人员采取合理的安全措施。生成审计日志模块对获取的网络数据生成审计日志，可供网络管理人员事后分析和取证。重建文件模块将按照 Web 数据使用的协议分别进行重建，在第3节进行详细的讨论。

2.2.4　报告输出模块　将实时数据分析器输出的文件输出到系统窗口，供用户实时察看，同时将文件存档，可以作为取证数据或者进行事后分析。

3　实验分析

3.1　基于 HTTP 协议的 Web 数据的监控和重建

对于基于 HTTP 协议的 Web 数据重建支持多种网页内容的重建（.html，.htm，.asp，.cgi，.php，.php3，.shtml，.shtm，.ide，.jsp，.fcg，.htx，.txt，.text，.xsp，.xml，rxml，.pl 等文件后缀的网页内容），支持多种图片（.gif，jpg，jepg）格式，支持 flash 动画的重建，支持 ActiveX 控件的重建，支持不需要运用额外应用程序就可以播放的视频和音频（这些视频和音频的播放由操作系统默认安装的软件就可以支持）的重建，对于压缩网页能自动解压缩并进行

恢复(例如新浪网页)。所有 Web 数据包内容的重建完全是实时的。下图是当被监控机浏览 www.263.net 网页时由监控机记录下来的情况。

图2 对基于 HTTP 协议的数据包监控示例

3.2 基于 SSL 协议的 Web 数据的监控和重建

安全套接层协议(SSL,Security Socket Layer)是 Netscape 公司提出的基于 Web 应用的安全协议,它包括:服务器认证、客户认证(可选)、SSL 链路上的数据完整性和 SSL 链路上的数据保密性,主要采用公开密钥体制和 X.509数字证书技术来提供安全性保证。目前,大部分的 Web 服务器及浏览器都广泛支持 SSL 技术,它弥补了 TCP/IP 协议安全性较差的弱点。下图是 SSL 握手过程示意图。

```
客户端                                    服务器端
ClientHello      ------->
                                        ServerHello
                                        Certificate*
                                  ServerKeyExchange*
                                 CertificateRequest*
                        <-------- ServerHelloDone
Certificate*
ClientKeyExchange
CertificateVerify*
[ChangeCipherSpec]
Finished          -------->
                                  [ChangeCipherSpec]
                                  <--------Finished
Application Data  <---------->Application Data
```

图3 SSL 握手过程

对于基于 SSL 协议传输的 HTTP 数据包,由于 SSL 协议在传输数据包时使用了加密机制,无法获得解密数据,但是通过对 SSL 握手阶段的数据分析,可以获得握手阶段所有信息,例如有关证书(如有效期,颁发机构)、加密方式、密钥长度等信息。因此,RNSMS 系统按照客户端和服务器端使用的 SSL 协议版本(SSL 2.0/3.0/TLS)[3,4],对这些信息加以重建还原,作为参考信息提供给系统使用者。下面给出了被监控机器通过 SSL 登录到 https://www.hotmial.com 的同时,监控机记录的 SSL 握手部分的信息片断:

```
HandShake Packet Type :- Certificate
  CERTIFICATE INFORMATION :-
  Validity -- Not After Oct 18 23:59:59 2003 GMT
              Not Before Oct 18 00:00:00 2001 GMT
  Subject Distinguished Name -
/C = US/ST = Washington/L = Redmond/O = Microsoft/
  OU = Hotmail/CN = LC1.LAW13.HOTMAIL.PASSPORT.COM
  Issuer Distinguished Name --
    /C = US/O = RSA Data Security, Inc. /OU = Secure
  Server Certification Authority
  RSA Public key size 1024 bits
HandShake Packet Type :- Server hello done
From Record Header -- Protocol Version:3.0
                      Record Length:132
Received a HANDSHAKE packet ···
HandShake Packet Type :- Client key exchange
Length of RSA Encrypted PreMaster Secret -- 128 bytes
RSA Encrypted PreMaster Secret -
0x68b2200b8f33e2c89faa81eab636c493e3d24a441f2c7d260d7
9b17875f44c8
157b2d038d2dca6802944fd30a4221095b2ef93521fb6
520829e59a8177b6dd1ab953a914c6b9038b43d5cc583b652
c7df39352f1bd82499a96dbc823d9fbf7c3c596deed9db3207828
c2d00fc9
7d1e617d512ab796dd26e5e5f52f32fdfdfdfd
```

4 关键技术

4.1 数据包丢包率的减少

在系统中,网络数据先从网卡拷贝到内存再拷贝到用户缓冲区。虽然数据获取模块使用的是高效的抓包驱动程序 WinPcap,但是由于对数据包进行重建、流量分析、生成审计日志等应用操作,丢包问题严重。在系统中已经使用的解决方法主要从以下几个方面进行:(1)RNSMS 系统运行时必须使用和被监控网络相对应的网卡,例如在百兆网络下使用百兆网卡而不是十兆网卡;(2)使用优化的算法,例如在算法中使用了多线程技术,在端口上进行数据包过滤也降低了丢包率;(3)提高数据缓冲区的大小。在(1)(2)条件不变的情况下,在百兆网络环境下,把缓冲区从1024bytes 提高到10M,丢包率从50%下降到5%左右。

4.2 数据的存储

由于监控系统的运行会重建大量数据包,数据的存储是个严重的问题。系统采用的方法是对历史数据超过7天的,予以删除。但是这样无法解决数据需要长期存储的需要。在将来的工作中,准备对数据的存储采取以下办法:(1)转变数据格式进行存储,例如将数据转换成压缩格式或者 libnids 数据的格式进行存储,减少对空间的要求。(2)使用专门的储设备。

结论 本文提出了一种可以监控 Web 数据的实时网络安全监控系统,它具有以下特点:捕捉局域网上的 IP 数据包,并且几乎没有数据包丢失的情

况；实时数据分析器允许用户在边抓包边分析的同时边查看分析后的网页内容；能够对 HTTP 协议数据包解析和解码，显示和保存重建的文件；支持多种网页内容；重建 SSL 协议握手部分的内容，为使用者提供参考信息；允许用户灵活的监控某个主机、某些主机、某个网段。但是本系统在数据的存储问题上还要加强研究。

参 考 文 献

1 HTTP RFC. http://www.w3.org/Protocols/rfc2616/rfc2616.html

2 张承，蒋东兴，刘启新，石岩. 浅析网络监控系统对网络性能的影响. 小型微型计算机系统，2003，23(9)：1059～1062

3 SSL/TLS. http://www.mozilla.org/projects/security/pki/nss/ssl/，2004

4 SSLv3/TLS Sniffer，http://crypto.stanford.edu/～eujin/sslsniffer，2001

（上接第167页）

如果监控总中心和分中心在同一座城市，则可以直接通过 DDN 专线相连，这样可以在不增加费用的前提下提高通信速度。

监控中心与客户终端的连接是通过 Internet 接入的，只需通用的浏览器软件用户就可以轻松、方便、快捷地同监控中心会话，这样既提高了效率又方便了用户。

监控总中心、监控分中心、SMS 短信中心以及客户终端的连接图如图4所示。

4.3 基于 WebGIS 的监控中心平台的实现

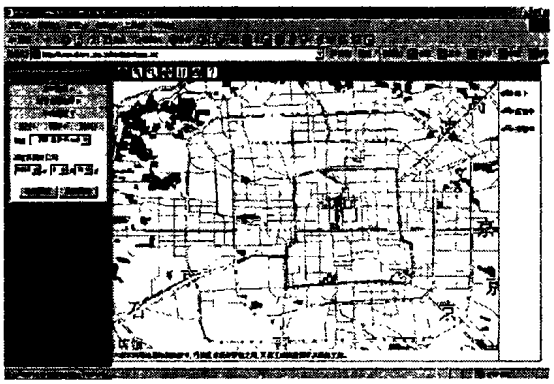

图5 基于 WebGIS 的 B/S 模式的地图显示界面

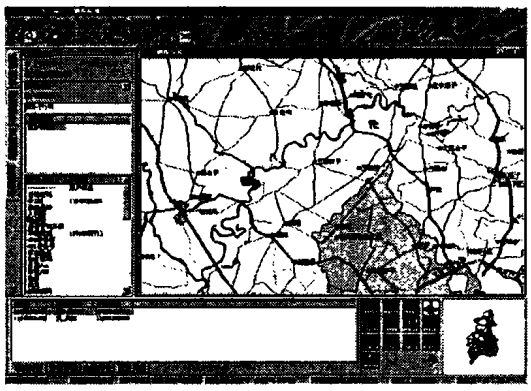

图6 基于 WebGIS 的 C/S 模式的地图显示界面

软件的组件化已成为软件技术发展的潮流；基于 DCOM 的 ActiveX 组件，已成为软件工业的一种标准。基于此，本监控中心平台电子地图部分采用 Intergraph 公司的 GeoMediaWebMap 组件来开发。系统开发周期短，能够简单快速地实现地图编辑、检索、空间分析、专题地图等功能。

在实际开发过程中，充分利用 WebGIS 的特点和组件式的优点来降低开发难度，实现最优的功能和服务。首先创建 WebGIS 组件对象的实例；然后通过该组件的服务器对象来启动 WebGIS 的服务功能，通过监视对象的监视功能来确保 WebGIS 各个对象的实例在各线程中正常运行，这样地图对象才能完成 WebGIS 的地图操作和地图分析的核心功能。在采用 B/S 模式进行开发时，要首先要建立 JSP 应用程序与地图引擎的连接。图5是基于 WebGIS 的 B/S 模式的地图显示界面，图6是基于 WebGIS 的 C/S 模式的地图显示界面。

结束语 “基于 WebGIS 的车辆监控中心平台”的建立给车辆监管部门带来一个实时的、动态的、高精度的、全天候的车辆管理工具；该平台的使用，有利于加强道路安全保障，规范车辆运行情况，提高信息化水平，为解决城市化问题提供最优的方案。该平台在公安、交通运输、公交管理、消防、金融、邮政、农林、水利、医疗救护等领域具有广泛的应用前景。

参 考 文 献

1 李德仁，等. 论空间信息与移动通信的集成应用. 武汉大学学报，2002，27(1)

2 Green D R. Cartography and Internet. The Cartographic Journal，1997，34

3 Getis A. An Introduction to Spatial analysis and GIS. Geograph syst，2001，2：1～3

4 刘基余，等. 全球定位系统原理及其应用. 北京：测绘出版社，1993

5 Derekenaris G，et al. Integrating GIS，GPS and GSM Technologies for the effective management of ambulances. Computers Environment and Urban Systems，2001，25(3)：267～278

6 Bruce E，Bobby M. Clicnt/Server Computing，Architecture，Application and Distributed Systems Management. Norwood：Artech House，1997

7 Paul A L，Michael FG，et al. Geographical Information Systems-Principles and Technical Issues. New York：John Wiley & Sons，Inc.，1999

8 Peter A B，Rachael A M. Principles of Geographical Information Systems. British：Oxford University Press，1998