基于 HTTP 指纹识别的服务器安全对策

徐远纯　郭德先　罗　洁
（景德镇陶瓷学院信息工程学院　景德镇 333001）

摘　要　本文介绍 HTTP 指纹识别的方法，探讨如何抵御 HTTP 服务器识别，提高 HTTP 服务器的安全性。

关键词　超文本传输协议指纹，服务器安全，对策

Safety Measures for Web Server Based on HTTP Fingerprinting

XU Yuan-Chun　GUO De-Xian　LUO Jie
(JingDeZheng Ceramic Institute, Jingdezheng 333001)

Abstract　This paper discusses methodologies for http fingerprinting, suggests possible methods for defending against http-Server fingerprinting, improves the http-server's safety.

Keywords　HTTP fingerprinting, Server safety, Countermeasure

1　引言

随着互联网技术的发展，互联网信息系统应用日益广泛，基于 Web 站点的应用日益普及。服务器提供重要服务，存有大量的数据，一旦崩溃，就会造成数据丢失，服务停止，损失不可估量；若受到恶意攻击，被窃取数据，甚至造成服务器停止，后果非常严重，因而保护 Web 服务器的安全至关重要。由于大部分网站服务器采用 WINDOWS 2000 SERVER 操作系统，Web 服务采用其自带的 IIS 服务，由于众所周知的原因，微软开发的产品先天不足，存在很多漏洞，当然其他公司开发的产品同样也存在这样或者那样的安全漏洞。这就给网络的蓄意破坏者——黑客提供了入侵点。“注意：暴露出服务器软件的版本号，就会使服务器变得更加脆弱，对于那些利用已知的该服务器软件安全漏洞进行的攻击。服务器软件提供商应该使该字段成为一个可配置的选项[1]。”

当前尽管出于安全考虑越来越多的 Web 服务器通过修改源代码或者二进制代码的方法，来改变 Web 服务器的 banner 以伪装自己。比如，在 Apache 服务器的“Server”头域中标记“IIS 5.0”，或者设置头域为“Web Server”，还有的 Web 服务器彻底消除了“Server”头域。但是现在一些工具还是能非常精确地得到服务器的类型和版本号。本文介绍这些常用工具的实现原理，并提出一些抵御措施，增加 HTTP 识别的难度，提高 Web 服务器的安全性。

2　HTTP 识别原理

为了保持各个 Web 服务器软件之间的兼容性，Web 服务器软件提供商必须遵守 HTTP RFCs，但是由于在 HTTP RFCs 中描述 HTTP 特性的实现要求有“must(必须)”，“should(应当)”和“may(可以)”，因而各个服务器软件提供商在实现各自系统的时候就会在符合协议说明文档的程度上有所不同，在具体功能和实现方法上有所差异，这就为服务器识别提供了可能。Http 指纹识别就是要首先收集这些微小的差异形成“指纹”，然后和数据库中的“指纹”进行比较，识别出服务器类型和版本号。

1、主要收集的特征，是从词汇、语义和语法三个方面进行收集

词汇：主要包括所用的单词/短语，字母大小写，标点符号等。

(1)状态代码：对于每个请求，响应都会返回一个状态代码，对于每个状态代码都关联着一个原因短语，例如，对于“404”型号的状态代码，IIS 的原因短语是“Object Not Found”，Apache 原因短语是“Not Found”，而 Netscape 的原因短语是“Not found”。

(2) 字母大小写：尽管在协议规范中要求通用的头部字段必须严格遵守协议，但是在具体实现时仍然会有所变动，例如，一些服务器返回“Content-Length”，而另外的一些则返回“Content-length”。

(3)行结束符：一些服务器用“\n”表示一行的结束，而另外一些服务器利用“\r\n”序列表示一行的结束。

(4)服务器名称：服务器响应的服务器名也可以作为识别的特征，尽管可能是虚假的。

2、语法

HTTP 消息需要用定义好的结构，以便服务器和客户端能够互相理解。然而在请求头的顺序和格式，信息头和它们的内容方面还是有些不同的。

(1)头部字段顺序：HTTP 协议规范建议 HTTP 响应头域顺序如下：“通用头域”，“响应头域”和“实体头域”。但是不同服务器提供商遵循规范的程度不同，如 Apache 服务器把 Date 头放到“Server”头前，而 Netscape 服务器刚好相反，如表 1 所示。

表 1　HTTP 响应头顺序

服务器	基本请求
Apache/1.3.23	Date, Server
Microsoft-IIS/5.0	Server, Date
Netscape-Enterprise/4.1	Server, Date

徐远纯　副教授，硕士，主要研究方向为计算机网络；郭德先　研究生，主要研究方向为计算机网络；罗　洁　研究生，主要研究方向为计算机网络。

(2)格式:一些在头中的元素的格式是可变的或者没有被RFCs指定。例如,Apache/1.3.11返回一个"ETags",用如下的格式:"0-574-38379154;3a5b7811",Jigsaw/2.1.2服务器返回的格式如下:"mvanct:s0jndthg"。

3、语义

当一个服务器收到一个请求的时候,它首先要解释这个请求,然后需要评估是否正确响应了,什么导致了响应失败;并且,服务器必须决定响应什么,在状态行、响应报头、主体中。服务器在解释这些正确和错误的格式的请求和响应上有很大的差别。

(1)一些请求将导致服务器认为请求者是基于HTTP/0.9的客户。对HTTP/0.9的响应只能包括响应主体(不能包括状态行和头)。存在不存在头能够用来推测出服务器如何解释一个服务器请求。

(2)一些专门的头字段:服务器选择包含在响应头中的头字段,一些是协议要求的,很多是可以选择的(如ETag)。例如,对于"501 Method Not Implemented"错误,Apache服务器返回"Allow"头(里面列举了允许的方法),而Jigsaw/2.1.2不返回这个选项。

(3)对于请求和请求对象的长度过长(这个长度通常由服务器来定义),HTTP/1.1说明文档,对请求和请求对象分别提供了"413 Request Entity too Large"和"414 Request-URI Too Long",通过对不同服务器对不同长度请求的响应,如表2,可以看出对于不同长度所引起的错误有明显的区别。

表2 URL长度请求响应

服务器	URL长度	响应
Apache/1.3.12(Win)	1-126	404 Not Found
	217-8176	403 Forbidden
	8177及以上	414Request-URI Too Large
Netscape-FastTrack/4.1	1-4089	404Not found
	4090-8123	500 Server Error
	8124-8176	413 Request Entity Too Large
	8177及以上	400 Bad request

3 常见的Web服务器识别工具

1、UC Davis的HMAP主要以反馈信息在文字、文法上的差异为依据进行服务器映射。例如,对于"404"型号的错误,IIS的原因是"Object Not Found",Apache是"Not Found",而Netscape的是"Not Found"。另外,它还针对不同Web服务器对于超长URL有不同的定义和处理方式的特点,发送了大量不同长度的URL,试探对方的反应,依此判断目标Web服务器的版本。这种方法的缺点是这些反馈信息中特有的语句表示特征易于修改,如果Web管理人员用修改banner的方法修改反馈信息中语句表示,这种映射就会失效。

2、HTTPPRINT

在执行Htpp协议的时候,几乎所有的Http服务器都具有它们独特的方法,如果Htpp请求是合法的并且是规则的,Http服务器返回的应答信息是符合RFC里描述的,但是如果我们发送畸形的Http请求,这些服务器的响应信息就不同了,不同服务器对Http协议行为就是它的基本根据和原理,HTTPPRINT就是通过运用统计学原理,组合模糊的逻辑学方法,利用畸形的请求响应来产生Http签名并保存起来,对一个未知的网站产生的Http签名信息和保存的Http信息进行比较就可以得出服务器的类型。

4 安全对策

由于IIS使用得比较广泛,下面就主要以IIS为例,提出要抵御识别的一些主要范围:

1、首先要把下面一些明显暴露服务器标识的地方,删除或隐藏掉:

(1)要隐藏文件的扩展名。包括URL地址和Content-Location头,扩展名如aspx,asp就明显表明这个服务器用的是微软产品。IIS服务器可以通过Port80 Software公司(在美国加州)开发的PageXchanger工具隐藏文件扩展名。Apache服务器可以在mod_negotiation模块进行配置。

(2)应用不完整的ASP SESSION ID cookie

为了保存客户状态,会话对象会利用ASP SESSION ID cookie:

Set-Cookie: ASPSESSIONIDQGQGGWFC=MGMLNKMDENPEOPIJHPOPEPPB;

这样就暴露出asp信息,我们可以通过ISAPI filter(ISAPI过滤器)来改变ASP SESSION ID cookie的名字。

(3)删除WebDAV

WebDAV(Web分布式创作和版本控制)是HTTP 1.1的扩展,是符合IETF RFC 2518规范的,但是微软WebDAV在响应头部添加了许多信息,因而在不需要这个扩展的时候要把它去掉。

(4)去掉一些头部

一些服务器响应就暴露出其信息,如X-Powered-By和X-AspNet-Version头就明显暴露出你使用的是asp.net,这些响应头应该被除去。

(5)其他方法包括:不要使用集成验证方法,不要把FTP和SMTP服务安装在Web服务器上等

上面都是从应用层保证不暴露出明显的信息,但是底层的协议栈仍然要考虑在内。可以通过防火墙控制。Port80 Software公司为IIS开发的ServerMask模块就能实现上面大部分功能。

2、其次服务器开发者应该提供可配置的特性。例如,配置参数或者提供对源代码编译的环境让系统管理人员控制一个响应中信息的多少和响应什么样的信息。开发者还可以利用识别的原理来纠正一些特性和行为来隐藏服务器标识。下面是一些有效的措施:

通用接口和行为;可配置的选项;利用配置或编译选项,用户可对响应消息的语义,语法和词汇方面有更大的控制权,这些配置项目可以让用户伪装成另一种服务器或者一种不知道的服务器;

可变的输出:使用可变的响应,例如,"File Not Found","Not Found","Not found"可以随机使用在合适的响应中,这样就让HTTP服务器识别工具很难形成该服务器的指纹。

3、默认的消息、页面和脚本也是Http识别的来源,特别是错误页面。对于IIS服务器,Port80 Software公司的CustomError[6]工具可以把错误页面等重定向到自定义到页面。对于Apache服务器也可以通过配置把默认错误页面重定向到自定义的页面。

结束语 对电脑的安全保护一般分为:保护,发现和作出反应。系统管理员和Http服务器工作在保护水平,而入侵

面向 RP 的工业 CT 切片数据格式转换软件开发*)

段黎明[1] 宋 军[2] 聂 璇[1]

(重庆大学 ICT 研究中心 重庆 400030)[1] (重庆交通大学计算机学院 重庆 400074)[2]

摘 要 针对原始的工业 CT 切片数据,应用矢量化软件提取工件封闭轮廓点数据,并通过轮廓配准、数据精简、三角网格划分、端面处理对轮廓点数据进行处理,实现了面向 RP 的工业 CT 切片数据格式转换。为使构成 STL 文件的三角网格更加优化,文中提出了平均点距值法数据精简与 Delaunay 三角化的网格精简相结合的数据精简算法,实例验证了该方法的正确性。

关键词 数据格式转换,工业 CT 图像,STL,RP

Data Transform Software Development Study for RP Based on Industiral CT Slicing Image

DUAN Li-Ming[1] SONG Jun[2] NIE Xuan[1]

(ICT Research Center, Chongqing University, Chongqing 400030)[1]

(College of Computer Science, Chongqing Jiaotong University, Chongqing 400074)[2]

Abstract In this paper the outline point data of the part are extracted based on its Industrial CT slicing images, then the data transform software Development for RP based on Industiral CT slicing image is presented by this schedule of contour Correspondence, data reduction, Delaunay triangulation, dealing with the surface and so on. Essentially the data reduction method on average distance method and Delaunay triangulation is presented, this method improve the triangle mesh quality that structed STL file.

Keywords Data transform, Industrial CT image, STL file, Rapid prototyping manufacture

1 引言

快速原形制造技术(Rapid Prototyping,RP)是一种基于离散堆积成形思想的新型成形技术。目前,针对接触式的三坐标测量数据以及激光三维扫描数据,面向 RP 的数据格式转换软件已经较为普遍,但是针对工业 CT 数据进行格式转换的软件却很少看到,本文研究以位图格式存放的工业 CT 图像数据转化为面向 RP 的数据格式转换软件的开发。

2 面向 RP 数据转换的文件格式概述

本系统所采用的面向 RP 的数据转换文件格式应用 STL 格式。作为面向目前 CAD 实体模型与快速成型设备进行数据交换的准标准,事实上 STL 格式是将实体曲面用一组小平面表示,系统以 SET 格式写这些小平面为 ASCII 文本文件或 Binary 文件,而 Binary 和 ASCII 码两种形式是按照数据储存形式的不同来区别的,本系统应用的是 ASCII 码格式。

ASCII 码格式文件中以 solid, facet, normal, outerloop, vertex, endloop, endfacet, endsolid 等关键词区分各种数据的意义[1]。每个实体的信息以 solid 开头,以 endsolid 结束;每个三角形的信息以 facet 开头,以 endfacet 结束;法向矢量以 normal 引导;顶点信息以 outerloop 开头,以 endloop 结束;每个顶点的坐标以 vertex 引导,每个顶点有 x, y, z 三个坐标值。它的一大优点是可以借助文本编辑软件直接阅读和改动,信息表达直观明了。

3 数据格式转换软件生成

3.1 数据格式转换软件开发思路

针对面向 RP 的数据格式转换软件,采用如下思路进行开发:首先将原始的工业 CT 数据应用已开发的矢量化软件进行边缘提取、轮廓跟踪、细化连接、直线矢量化等前期处理,

*)基金项目:重庆市自然科学基金(CSTC,2006BB2413)。段黎明 副教授,工学博士,研究方向为工业 CT 技术及应用,逆向工程;宋 军 副教授,工学博士,研究方向为嵌入式系统,宽带网络技术。

检测系统工作在发现水平。攻击者掌握的信息比防御者要多得多。隐藏 HTTP 服务器的标识对一个老练的和知识丰富的攻击者不是一个简单的事情。像 HMAP 工具,可以模仿客户端来得到一些响应并且自动地分析这些响应来产生指纹和识别一个服务器。这对攻击者有很大的帮助,识别一个服务器就可以利用它的脆弱性进行攻击。然而这也可以帮助如何加强服务器的安全,通过分析这些工具是如何工作的,可以相应地采取对应措施,在一定程度增加识别的难度,进而提高系统的安全性。

本文作者创新点:针对 HTTP 指纹识别原理,分析并提出相应的服务器安全策略。

参 考 文 献

1 section 14.39. http://www.ietf.org/rfc/rfc2068.txt
2 王新榕.威胁 Web 站点安全的威胁及对策.计算机与网络,2004,5:42
3 刘全.网络控制系统的网络安全研究[J].微计算机信息,2006,9(3):71~73
4 谢希仁.计算机网络(第四版),2005,7
5 黄仲祥.TCP/IP 协议分析和网络攻击方法.通信对抗,2005,4
6 仆胜贤,李鹰.Web 网站安全技术研究.微机发展,2004,5:14
7 Turn dead links, 404s and other site errors into good traffic (& set up IIS error pages easily). http://www.port80software.com/products/customerror/? vid=3491283