基于容侵技术的 Web 服务器设计

王宁波，王先培
(武汉大学 电子信息学院，湖北 武汉 430072)

摘　要：分析了常用安全防护技术的局限性,提出了用容侵技术来构建安全的 Web 服务器,使得服务器在被入侵的情况下仍能对外提供有效服务。

关键词：网络安全；容侵技术；Web 服务器
中图法分类号：TP393.08　　文献标识码：A　　文章编号：1001-3695(2005)06-0123-02

Design of Web Server Based on Intrusion Tolerant Technique

WANG Ning-bo, WANG Xian-pei
(School of Electronic Information, Wuhan University, Wuhan Hubei 430072, China)

Abstract: Firstly analyzes the limitations of often used defensive technologies, and then issues an intrusion tolerant technology to design secure Web server, which can still provide effective services in the case of being attacked.

Key words: Network Security; Intrusion Tolerant Technique; Web Server

在计算机网络得到广泛应用的同时,网络安全问题也显得日益突出。黑客攻击和网络病毒侵扰时有发生。网络系统面临着入侵、事故、失效等危机。网络信息会受到破坏、窃取、窜改等各种威胁,造成信息无法使用。无论是有意的攻击,还是无意的误操作,都将会给系统带来不可估量的损失。为了保护 Web 服务器的安全,一些传统的诸如防火墙、入侵检测、加密技术、认证技术等安全措施得到了广泛使用,在一定程度上把黑客拒之于门外。但这些技术主要侧重于防范,对于黑客攻入系统内部的情形,并没有切实可行的应对之策,一旦黑客攻入,就只能任凭其肆意破坏或使系统被迫停止运行,造成不可估计的损失。本文将在分析此现状的基础上,提出一种容侵技术,它使得即使系统受到了攻击,仍能够对外提供完整的或者是降级的服务。

1　常用安全防护技术的局限性

为了保障 Web 服务器的安全性,许多卓有成效的安全技术得到了应用,防火墙和入侵检测技术是常被使用的有效手段[1],它们使得采用已知攻击方式的黑客行为能够得到及时的发现和屏蔽,但防火墙和入侵检测系统本身具有许多局限性。

防火墙不能防范不经由防火墙的攻击;不能防止受到病毒感染的软件或文件的传输;不能防止数据驱动式攻击;对来自内部的攻击缺少足够的保护,由于最初防火墙的设计思想是以本地私有网络安全可信为前提的,要防范只是来自外部可能的攻击。

入侵检测系统无法弥补安全防御系统中的安全缺陷和漏洞。基于知识的入侵检测系统很难检测到未知的攻击行为,检测具有一定的滞后性,而对已知的攻击,一些没有明显特征的攻击行为也很难检测到,或需要付出提高误报警率的代价才能够正确检测;入侵检测系统的主动防御功能和联动防御功能会对网络的行为产生影响,同样也会成为攻击者的目标。实现以入侵检测系统过敏自主防御为基础的攻击;入侵检测系统主要是对网络行为进行分析检测,不能修正信息资源中存在的安全问题[2]。因此,传统的安全防护技术只重在防范,而没有考虑到系统信息被破坏后的生存性问题。因而设计一个具有容侵能力的服务器能较好地解决这一难题。

收稿日期：2004-06-05；修返日期：2004-07-22

2　容侵的基本理论

2.1　容侵的概念

所谓容侵,就是入侵容忍(Intrusion Tolerance),也就是当一个网络系统遭受入侵,而一些安全技术都失效或者不能完全排除入侵所造成的影响时,容侵就可以作为系统的最后一道防线,即使系统的某些组件遭受攻击者的破坏,但整个系统仍能提供全部或者降级服务。如图 1 所示,它形象地说明了保护、检测、容侵三者之间的关系。

2.2　容侵的策略[3]

容侵系统的策略是指导容侵系统的设计,并决定其运行效果的关键。它与以下因素有关:

(1) 容侵的程度。容侵的程度是指在应用系统受到入侵的时候,容侵系统需要将应用系统保护到的程度。例如,只需要保护服务器上数据的完整性,或必须保护数据的秘密性以及还要提供降级服务,或者还要有自动重组/修复能力等。

(2) 容侵的代价。要实现上述任何一种容侵的程度,都是需要付出相应的代价。这种代价包括硬件设备和软件设计方面的代价,还有效率上的和控制粒度上的代价,还可能包括其他方面的代价。这需要在重要性和代价上进行权衡。

2.3　容侵的方法

容侵系统是建立在入侵检测、容错理论和密码理论基础上

的系统。它所运用的方法大多来源于以下几方面:

(1) 检测与追踪。容侵系统中最前线的是入侵检测系统。但容侵系统的检测模块与独立的入侵检测系统不同在于:前者的行为往往与容侵系统的其他部件是联动的、紧密耦合的。容侵系统对前者在实时性、准确性方面提出了更高的要求。因此,前者在结构、方法等方面有别于后者。而后者更注重于学习性、自适应性等方面。

(2) 冗余与多样性。冗余可以防止系统由于单个部件的受损而导致整个系统即刻崩溃。多样性能够确保同样的攻击不能摧毁其他部件。

(3) 相关算法。由于容侵系统是建立在容错理论和密码学理论的基础上的,因此容侵系统中将会使用到许多与容错和密码相关的算法。例如:冗余/屏蔽算法、检错/纠错算法、门限密码算法、分布式密钥生成算法及其使用方法等。

3 使用容侵技术构建 Web 服务器

3.1 系统总体结构

如图 2 所示,我们采用了一个具有容侵能力的系统结构。防火墙内侧网段上放有一个网络入侵检测系统,通过捕获网络上的数据包来分析、检测明显的攻击,它与防火墙是联动的,一旦发现有攻击模式,就通知防火墙,封堵攻击源。用具有容侵的服务器作为代理服务器,位于两个应用服务器(受保护的 Web 服务器)前面。它是主/备式的,如果代理服务器本身受到了攻击,由灾难评估模块来评估灾难大小,如果不能在短时间内恢复,则启用另一个备用代理服务器,使之成为主代理服务器,而受损代理服务器处于恢复阶段,一旦恢复,使之成为备用代理服务器。它主要是传递客户和服务器之间的请求和响应。两个应用服务器拥有内容相同而形式不同的服务以提供冗余和多样性,而且是同时运行。每个应用服务器内有一个代理模块,它作为服务器的一个进程而存在,主要功能有计算服务器的 MD5 校验和,并将结果送给代理服务器。在服务器内容受损后恢复内容,接收来自代理服务器的重组请求并执行服务器的重组。

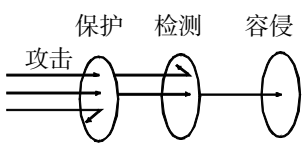


图 1 容侵示意图

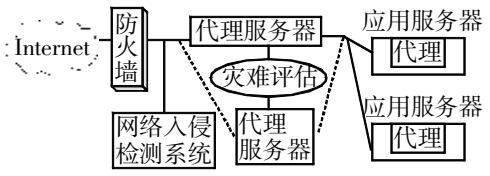


图 2 容侵系统总体结构

3.2 代理服务器及各模块功能

图 3 是代理服务器详细结构图。它主要包括过滤模块、验证模块、表决模块、自适应重组模块和日志与监控模块。

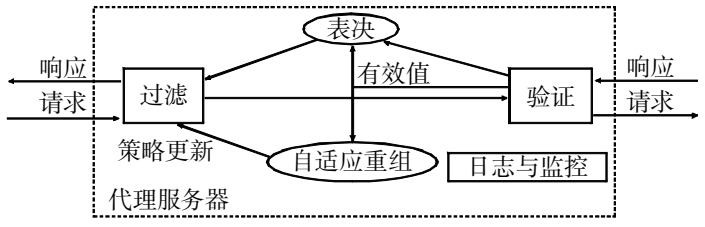


图 3 代理服务器结构框图

(1) 过滤模块。当过滤模块接收到一个客户的请求后,一个新的过滤进程便产生来处理客户的请求。此模块的主要功能是检测和过滤对代理服务器和应用服务器的已知攻击,根据系统的策略把客户的请求传给验证模块以及把服务器的响应传给客户。

(2) 验证模块。验证模块处于过滤模块和服务器之间,对请求和响应进行有效性验证,并传送它们到下一级。有效性验证的目的在于阻止攻击,并把验证有效值传给自适应重组模块以便它能更新过滤的策略。

有效性验证可以被分为两类,即请求有效性验证和响应有效性验证。请求有效性验证程序在过滤模块中执行,主要用于检查攻击模式和过滤对代理服务器和应用服务器的恶意请求,其验证结果被用于执行响应有效性验证,响应有效性验证在验证模块中执行。两种类型的验证值都被送到表决模块和自适应重组模块中。请求有效性验证只能阻止已知攻击。响应有效性验证主要用于检查服务器的响应,它能检测出未知攻击并在以后过滤掉这些攻击。

(3) 表决模块。当客户的请求到达后,一个过滤进程便产生,进而一个表决进程也产生。表决模块的主要功能是收集来自验证模块的有效性的值,把被选择的响应送到过滤模块,把投票的结果送给自适应重组模块。

表决使代理服务器能够提供不间断的服务,即使应用服务器中有一个被攻击,如当一个应用服务器被攻击,如果用户的请求也被传到了另一个应用服务器,表决模块将使用有效性验证的值来选择来自未被攻击服务器的响应。因为同样的攻击对不同的服务器将会有不同的影响,所以通过运行具有相同内容而具有多样性复制的服务器将会提供更强健的服务。另外,内容的复制也使系统有了冗余,即使某一个内容被攻击者改变,通过有效性验证也能选择出正确的内容。内容检查程序对被改变的时间和响应内容的 MD5 哈希值与预存值进行比较。当然这个程序只能用于静态内容的比较而不能用于动态内容的比较。

在表决模块中产生的统计数据将被送到自适应重组模块以便它更新过滤请求的策略。需要收集的统计数据如下:

服务器。每一个服务器的有效值被收集,并计算平均值,服务器的有效值越高,被指定到此服务器的请求就越多。

URL。每一个 URL 的有效值都被收集。

客户。每个客户的有效值都被收集,并被用于自适应重组模块建立“黑名单”。

URL 与服务器。如果一个 URL 被复制到多个服务器并且对此 URL 的请求也被传送到多个服务器,则每个服务器的 URL 的有效值都被收集。

客户与服务器。每个服务器的客户有效值都被收集。

这些被计算的有效值用于自适应重组模块更新过滤请求策略。如果不止一个服务器在运行而且某一个服务器在某个时间段内的平均有效值小于一个阈值,那么此服务器接收请求的数量就减少;反之亦然。

(4) 自适应重组模块。它接收到来自表决模块、验证模块、日志与监控模块的数据,在需要的时候更新过滤策略和执行服务器重组。具体功能包括:①服务器工作模式的改变。服务器工作的模式有两种:全部参与和主/备式。一开始,服务器是全部参与,当攻击增加时,工作模式转变为主/备式。②黑名单。来自表决模块的统计信息被用于去更新策略和黑名单。黑名单里访问某一个 URL 的所有 IP 地址都将被过滤。黑名单里的条目可能是某一个 IP,访问某个 IP 的 URL, (下转第 126 页)

表 2 RC6 加、解密算法

```
算法 1 RC6 加密算法

输入: 128 bits 明文 A, B, C, D
      轮数 r
      轮密钥 S[2r+3]
输出: 密文 A, B, C, D
算法:  = B + S[0]
      D = D + S[1]
      For i = 1 to r do
      {
      t = (B* (2B+1)) <<<5
      u = (D* (2D+1)) <<<5
      A = ((A^t) <<< u) + S[2i]
      C = ((C^u) <<< t) + S[2i+1]
      (A, B, C, D) = (B, C, D, A)
      }
      A = A + S[2r+2]
      C = C + S[2r+3]
```

```
算法 2 RC6 解密算法

输入: 128 bits 密文 A, B, C, D
      轮数 r
      轮密钥 S[2r+3]
输出: 明文 A, B, C, D
算法:  = C - S[2r+3]
      A = A - S[2r+2]
      For i = r down to 1 do
      {
      (A, B, C, D) = (, D, A, B, C)
      u = (D* (2D+1)) <<<5
      t = (B* (2B+1)) <<<5
      C = ((C - S[2i+1]) >>> t) ^u
      A = ((A - S[2i]) >>> u) ^t
      }
      D = D - S[1]
      B = B - S[0]
```

注: 符号“<<<”表示向左循环移位; “>>>”表示向右循环移位。

此外, RC6 还具有相当大的灵活性, 密钥生成算法见表 3 的算法 3。也就是说它可以通过改变密钥生成逻辑中的两个常数 P_{32} 和 Q_{32} 的值使之成为“专用的 RC6”。

表 3 RC6 的密钥生成算法

```
算法 3 RC6 密钥生成算法

输入: 用户密钥 L[c]
      轮数 r
输出: 轮密钥 S[2r+4]
算法:  [0] = P32
      For i = 1 to 2r+3 do
      S[i] = S[i-1] + Q32

      A = B = i = j

      v = 3* max{c, 2r+4}
      for s = 1 to v do
      {
      A = S[iI] = (S[i] + A + B) <<<3
      B = L[jI] = (L[j] + A + B) <<<(A+B)
      i = (i+1) mod(2r+4)
      j = (j+1) mod c
      }
```

由于分组密码多用于高速、实时的数据加解密处理, 所以其实现速度非常重要。这一点从分组密码的设计原则也可以看出。如果采用工程实践中常用的 C 语言编写的程序不能充分利用 TMS320C6201 的并行特性, 无法满足实用中对 RC6 算法实现速度的要求。本文为了达到最佳的实现效果, 经过深入分析、比较, 决定采用手工直接编写线性汇编程序的方法。在程序编写过程中, 我们充分利用了 TMS320C6201 芯片的八个功能单元, 即在一个时钟周期内实现尽可能多的指令。经过反复优化核心代码, 达到了理想的效果: 在 RC6 的多分组加密算法中, 轮变换部分每个时钟周期平均执行 5.0 条指令; 与用 C 语言编写的代码经 TI 提供的汇编优化器优化后的代码执行效率相比, 直接用线性汇编语言编写的代码效率提高了近一个数量级, 速度指标超过了相同环境下文献[2] 中提出的指标。

3 结论

本文中, 包括加、解密算法、测试程序和密钥生成逻辑在内的整个代码总量为 10KB。在实现中, 我们针对 RC6 算法进行分析、综合, 并充分利用 DSP 的特点提高运行的速度。达到的运算能力如表 4 所示。

表 4 RC6 加、解密速度

类型	单分组加密	单分组解密	多分组加密	多分组解密
处理速度(Mbps)	106.7	116.4	157.1	154.2

注: RC6 为 128bit/分组。

在深入分析 RC6 算法的基础上, 我们充分利用 DSP 的特点, 达到了数据处理的理想速度。从表 1 数据可以看出, 实现 RC6 算法的速度超过了国外同行[2,3] 在相同环境下的实现速度。

参考文献:

[1] ES3 Website [DB/OL]. http://csrc.nist.gov/CryptoToolkit/aes/round2/conf3/papers/12-twollinger.pdf.

[2] TEXAS INSTRUMENTS. TMS320C6X Assembly Language Tools User's Guide[R]. Literature Number: SPRU186C, 1998.

[3] Daniel V Bailey, William Cammark, *et al*. Cryptography in Modern Communication Systems[C]. TI DSPS FEST'99, 1999.3-6.

作者简介:

郭艳(1971-), 女, 博士, 主要研究方向为密码、自适应信号处理; 李宁(1967-), 男, 博士研究生, 研究方向为密码、宽带网络。

(上接第 124 页)或者是某个 URL。当某个 URL 有错误或者有漏洞, 对外界不可访问时, 此 URL 将被注册到黑名单。

(5) 日志与监控模块。日志与监控模块在代理服务器中被作为一个进程来监控服务器和其他模块的安全状况, 并保存代理服务器的日志。监控的周期是随着代理服务器状况而动态地改变。其他功能包括: 周期性地检查应用服务器的哈希值, 如果发现应用服务器内有任何错误, 提醒应用服务器内的代理模块以便让其恢复。如果发现代理服务器内任何模块有错误或者性能下降, 通知自适应重组模块进行重组或适应性变化。而且通知灾难评估模块, 灾难评估模块对其进行分析评估, 如果损伤超过某一个阈值, 则启用备用的代理服务器, 使受损的代理服务器处于恢复之中, 一旦恢复, 使其成为备用。

4 结束语

本文构建了一个具有容侵能力的 Web 服务器, 通过使用冗余和多样性的方法既能处理已知攻击, 也能处理未知攻击, 这使得服务器即使在受到攻击的情况下仍能对外提供有效服务。

参考文献:

[1] 王先培, 熊平, 李文武. 防火墙和入侵检测系统在电力企业信息网络中的应用[J]. 电力系统自动化, 2002, (3): 60-63.

[2] 胡道元, 闵京华. 网络安全[M]. 北京: 清华大学出版社, 2004.

[3] StavridouV, Dutertre B, Riemenschneider R A, *et al*. Intrusin Tolerant Software Architectures[C]. DARPA Information Survivability Conference & Exposition Ⅱ, 2001, DISCEX'01. Proceedings, 2001, (2): 230-241.

作者简介:

王宁波(1976-), 男, 四川广元人, 硕士研究生, 主要研究方向为网络安全; 王先培(1963-), 男, 湖北天门人, 教授, 博士, 主要研究方向为网络安全。