SYN flood 攻击检测技术综述

石利平
(广东女子职业技术学院 应用技术系,广东 广州 510450)

摘 要：分析 SYN flood attack 攻击原理,并在此基础上研究几种典型的 SYN flood 攻击检测方法。对主要技术进行分析和比较，这些技术均有各自的优点和局限性，多种方法有机融合互补将成为 SYN flood 攻击检测研究的重点。

关键词：SYN flood 攻击;TCP 协议;检测
中图分类号：TP393.08 **文献标识码**：A **文章编号**：1674-7720(2011)14-0048-03

Overview of SYN flood attack detection techniques

Shi Liping
(Department of Applied Technology, Guangdong Women's Polytechnic College, Guangzhou 510450, China)

Abstract: This paper analyses the principle of SYN flood attack, and some typical detecting methods are researched based on it. This paper summarizes and analyzes each of these schemes. These schemes have their own advantages and limitations, so organically combining several methods will be the focus of detecting SYN flood attack.

Key words: SYN flood attack; TCP protocol; Detect

SYN flood 攻击是一种当前流行的 DoS 与 DDoS 的方式之一，主要是利用 TCP 协议的三次握手的缺陷,造成服务器上 TCP 连接表溢出[1],而攻击服务器响应 TCP 连接请求的能力。TCP 是 TCP/IP 体系中的运输层协议,是面向连接的，其三次握手过程数据包都通过 IP 协议传输。而 IP 协议是一种不可靠、尽力而为的网络协议,缺乏认证和保密措施,因此为否认、拒绝等欺瞒行为开了方便之门[2]。目前,DDos 攻击中约有 90%是 SYN flood 攻击[3],研究 SYN flood 攻击检测技术对网络安全意义重大。

1 SYN flood 原理

TCP 连接的建立是通过三次握手来初始化[4]。三次握手建立 TCP 连接如图 1 所示。

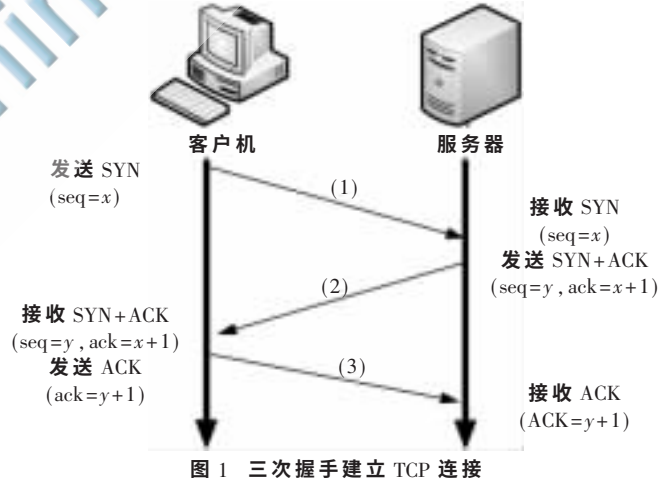

图 1 三次握手建立 TCP 连接

在 TCP 连接的三次握手中,若用户向服务器发送了 SYN 报文后突然死机或掉线，则服务器在发出 SYN+ACK 应答报文后将无法收到客户端的 ACK 报文，导致第三次握手无法完成。这时,服务器端的 TCP/IP 栈中会保留这个半连接状态，然后再次发送 SYN+ACK 给客户端,并等待一段时间后丢弃这个未完成的连接。这段时间称为 SYN Timeout(一般约为 30 s~2 min)[4]。但如有攻击者大量模拟这种情况,产生大量的半连接,服务器将为维护这些半连接而消耗大量的资源,进而导致服务器的 TCP/IP 堆栈溢出，使服务器无法响应正常用户的请求。SYN flood 攻击就是利用这种行为发动攻击。

正常情况下,连接超时后,服务器会清理出 TCP 连接表中已超时连接的相关信息。但攻击者不断给服务器发送大量的欺骗的 TCP 连接请求，使 TCP 连接表一直处于被填满状态,导致服务器无法响应大多数合法的连接请求。

根据攻击源的IP地址，SYN flood攻击可分为以下三种：(1)使用本机的IP发送SYN欺骗包；(2)假冒本网段中未在线的多台主机IP发送SYN欺骗包；(3)随机产生伪造的IP地址发送SYN欺骗包。为增加攻击的力度、隐蔽性和效率，使用随机源地址进行SYN flood攻击是攻击者常使用的方法[5]。

2 SYN flood攻击的检测技术

根据检测的位置SYN flood攻击检测主要分为4种：被攻击的服务器端、攻击源端、攻击源和被攻击服务器间以及正常的主机端。最受关注的是服务器端的检测和防护[6]；

2.1 基于TCP协议SYN-FIN(RST)行为对特征的检测

这种方法检测的是连接终端主机的路由器端，是基于TCP协议SYN-FIN(RST)包对特征和连续变化点的检测[6]。首先收集IP数据包，然后分类数据包。根据IP访问包的TCP头部来识别TCP SYN、FIN和RST。关键是找出IP数据包片偏移中TCP数据位，计算依据公式为IPoffset=Hdr_LengthIP+TCPoffset。接着利用非参数的CUSUM(Cumulative Sum)算法分析SYN-FIN(RST)行为对。CUSUM算法公式如下：

$$y_n=s_n-\min_{1\leqslant k\leqslant n}s_k \tag{1}$$

$$y_n=(y_{n-1}+\tilde{x}_n)^+,\ y_0=0 \tag{2}$$

x_n为一个随机序列模型。如果测试统计y_n大于攻击门限N，则表示有SYN flood攻击发生，否则网络正常运作。该检测方法的优点是准确性高，在线检测速度较快，消耗系统资源少；其缺点是检测在攻击发生后才起作用，攻击对系统已造成一定的危害，且易产生误报警。但作为较早的SYN flood攻击检测方法，其检测思路还是值得借鉴的。

2.2 基于TCP半连接的检测

此检测是在服务器端[7]，主要思路是检测TCP半开连接数是否正常，如异常，则检查半开连接数是否超过服务器容纳最大的半开连接数目，如超出则发出报警。该检测方法的主要步骤为：(1)分析因正常网络拥塞和因DDoS攻击造成的半连接的区别。这种区别是第2步中DARB(延迟探测)算法的基础。(2)利用DARB估计客户端与服务器之间的延迟时长。DARB算法类似于Tracer-oute技术，是在IP层发送带有TTL(time-to-live)字段的数据包到目标网站，测得客户端和服务器之间的延迟时长。算法中有两个重要量：far-hop是探测到最远的路由的TTL值；near_hop是探测到最近的路由的TTL值。(3)概率估计。首先计算一个T时间段内平均延迟$\bar{x}=\frac{\sum_{i\in S}x_i}{|S|}$，其中$x_i$是DARB中第$i$个半连接的延迟时间，$S$是时间$T$内在服务器上收集的半连接样本；然后再用函数$f(x)$来估算检测的延迟概率。

$$f(x)=\begin{cases}\frac{1}{\beta}e^{-\frac{x}{\beta}} & x>0\\ 0 & x\leqslant 0\end{cases} \tag{3}$$

其中$\beta=\bar{x}$，x是随机值，表示DARB算法中检测的一个半连接延迟值。(4)$f(x)$与预定阈值T比较，如$f(x)>T$，则认为半开连接是异常的，否则合法。当异常的半开连接数超过服务器容纳的最大半开连接数目N时，则发出报警。此检测优点是不依赖其他网络设备，且是一种主动积极的方法，系统开销小，能在SYN洪水攻击发生早期进行检测。其缺点是检测算法中参数T和N的值是根据经验而设，如设置不合适，将会影响检测的准确率。

2.3 基于Patricia树的检测

基于Patricia树的检测是基于SYN flood攻击的最为本质的两点特征[8]：(1)短时期内一个服务(一组IP、端口号)接收到大量的SYN包。(2)这些SYN包发起的连接处于半开状态。以Patricia树作为流量统计手段，利用Patricia树记录所有的半开连接及其数量，每个叶节点对应一个发起连接的IP源地址。Patricia树是一棵满二叉树，Patricia树的每一个节点的关键字是二进制字符串。算法主要是维护服务器Patricia树的半开连接表，周期性地进行检查和刷新。如果有攻击产生，Patricia树中将会明显地显示出发动攻击的主机IP地址。此方法的优点是消耗系统资源少，检效率高，但是一种被动检测。

2.4 基于源端网络的双粒度的检测

该检测方法的主要思想是以出入网络的TCP业务量的平衡性为判断依据来检测攻击[9]。此检测分两步：(1)用粗粒度检测引擎对出/入网络的TCP业务进行监控，发现异常就预报警，预警启动细粒度检测引擎、关闭粗粒度引擎。(2)由细粒度检测引擎监控终端网络向外发送的SYN包和接收的SYN/ACK包的数量差，当两种包的数量差迅速增大时，认为有SYN flood攻击，报警启动边界路由器中的过滤引擎，封锁攻击源发送的包。细粒度检测算法主要计算公式为：

$$\bar{\Delta}n=\frac{\Delta n}{\bar{k}} \tag{4}$$

其中：Δn为第n个抽样间隔内终端网络向外发送的SYN包与接收到的SYN/ACK包的数量差。随机序列$\{\Delta n,n=0,1\}$的均值依赖于网络的规模，并随抽样时刻变化；$\bar{k}$是抽样间隔内SYN/ACK包数量的平均值，可进行实时估计和周期性更新，$\bar{k}$的递归估计定义如下：

$$\bar{k}(n)=a\bar{k}(n-1)+(1-a)\mathrm{SYN/ACK}(n)\quad(0<a<1) \tag{5}$$

此检测方法与参考文献[6]的检测思想类似，都是以TCP协议的SYN/ACK包数量为检测基础，均需对IP数据包进行分类。优点是两个粒度检测引擎采用不同的检测机制，检测高效、高速、虚警率低。其缺点主要有：检测效果会随着攻击组规模的增大而衰减。当网络中攻击流与网络正常业务相差不大时，也不易检测出来。

2.5 基于重尾特性的检测

基于重尾特性的检测方法是基于网络流量的检测[10]。主要思想是将统计阈值(以 SYN/TCP 的值为参数)和流量的重尾特性相结合。研究表明,当 SYN/TCP 的值 $\beta \leqslant 0.05$ 时,网络正常;当 $\beta \geqslant 0.15$ 时,网络存在 SYN flood 攻击;当 $\beta \in (0.05, 0.15)$ 时,可能是正常流量,也可能是 SYN flood 攻击。这时对网络流量进行重尾特征检测,当参数 $a \leqslant 2$ 时,认为网络流量为正常流量;当 $a>2$ 时,认为网络发生了 SYN 洪流攻击。a 可从重重尾分布中计算得到。一般正常网络流量呈重尾分布特性,若将浏览器描述成 on/off 信息源,则其生成的数据刚好符合 Pareto 分布。

Pareto 分布的概率密度函数(PDF)为:

$$p(x)=ak^a x^{-a-1}, 0<k \leqslant x \tag{6}$$

它的和余累积分布函数(CDF)为:

$$F(x)=P[X \leqslant x]=1-(k/x)^a \tag{7}$$

其中,X 是一个随机变量,常数 k 表示随机变量最小的取值。

$$\lim \frac{\mathrm{dlg}F(x)}{\mathrm{dlg}(x)}=-a \tag{8}$$

此检测方法的优点是在网络有强 SYN 洪流时检测效率高,但在发生弱 SYN Flood 攻击时,其漏报率稍高,约为 9%。

2.6 基于边界路由器的检测

这种检测方法是在路由器边缘检测 SYN flood 攻击,是基于正常的网络流量下路由器实际转发的 SYN、SYN/ACK 和 ACK 报文的数量进行检测[11]。主要包括两部分:存储模块和检查模块。(1)存储部有两个表:映射表和存储表。表结构如表 1 和表 2 所示。存储表记录 SYN 包源和目的地址,通过构建的 HASH 函数($HASH_a=\sum_{i=1}^{k} hash(i)$,$a$ 表示地址)计算映射,为了防止哈希冲突,为连续的哈希函数分配不同的权重 w_i。(2)检查模块检测和处理两个模块组成。在检测模块通过检查映射表是否有异常情况,方法是选择数据库中的源地址作为嗅探数据包的目的地地址,选择边界路由器的 IP 地址为嗅探包的源地址。然后这个嗅探数据包被发送到目标主机。如果有应答包返回,则说明无 SYN flood 攻击,否则说明有 SYN flood 攻击。通过检测结果,如推断有一个无辜的主机,处理模块发送 RST 包使受害者服务器释放受攻击产生的半开连接。此检测方法的优点是存储保证数据的持久性和容错性。如系统出现任何故障,当再次进入系统时,数据仍然存在,可以直接恢复正常连接。检查过程是主动的,不依赖于其他网络设备。不仅能检测 SYN flood 攻击,还有一定的消除 SYN flood 攻击的功能。

表 1 映射表结构

hash 值	计数器(SYN)	计数器(SYN/ACK)	计数器(ACK)	探测标记 T	攻击标记 E

表 2 存储表结构

hash 值	源地址	目标地址	可疑标记 s	正确标记 A

3 现有技术比较

表 3 对上面介绍的 SYN flood 攻击检测的防御所采用的方法进行简单总结。通过分析和比较已有的 SYN flood 攻击检测和防御方法发现,对于 SYN flood 攻击检测和防御目前还没有完美的解决方案。常规的 SYN flood 防御方法因实现容易,已得到广泛应用,但防御力度不强,不能从根本上避免 SYN flood 攻击的发生。SYN-cookie 技术已得到厂商的支持,已在一些防火墙中使用。输入过滤器在理论上可以彻底避免 SYN flood 攻击的发生,但因其会使路由器处理数据量加大,影响路由器的性能,实施难度较大,还需进一步研究。

表 3 SYN flood 攻击检测方法比较

技术	检测对象	安装位置	优点	缺点
基于 TCP 协议 SYN-FIN(RST)行为对特征	TCP 协议 SYN-FIN(RST)	路由器端	准确性高,在线检测速度较快,消耗系统资源少	检测须在攻击发生前开始,误报率高
基于 TCP 半连接	TCP 半连接	服务器端	不依赖其他网络设备,主动,系统开销小,检测及时	算法中参数 T 和 N 的值需根据经验而设,有一定的局限性
基于 Patricia 树	半开连接及其数量	服务器端	能检测出发动攻击的主机 IP 地址,消耗系统资源少,检效率高	被动检测
双粒度的检测	网络的 TCP 业务量	服务器端	检测效率高、误报率低	当攻击规模的增大时检测效率衰减。攻击流与正常业务相差不大时,检测效果差
重尾特性	网络流量	服务器端	攻击强度大时检测效率高	被动的检测,在攻击强度小时,误报率稍大
基于边界路由器的检测	SYN、SYN/ACK 和 ACK 报文的数量	边界路由器	主动的检测,检测效率高,有一定消除 SYN flood 攻击的能力	占用系统资源稍多

本文分析了 TCP 的三次握手协议,并针对 SYN flood 攻击的原理,对其检测方法进行了探讨。未来的检测和防御,应趋向于如何综合应用上述几种检测方法,取长补短,以达到互相补充和弥补,提高检测精度和效率。

参考文献

[1] 黄发文,徐济仁,陈家松.计算机网络安全技术初探[J].计算机应用研究,2002,19(5):46-48.

[2] 贾春福,刘春波,高敏芬,等.计算机安全原理与实践[M].北京:机械工业出版社,2008.7.

[3] 李海伟,张大方,刘俊,等.一种基于主动探测机制的 SYNFlood 攻击检测方法[J].计算机科学,2010,37(3):117-120.

[4] SYN Flood 攻击的基本原理及防御[EB/OL].http://www.bitscn.com/network/hack/200705/102673.html,2007.
[5] 谢希仁.计算机网络[M].北京:电子工业出版社,2003.
[6] WANG H,ZHANG D,SHIN K G.Detecting SYN flood attacks[J/OL].Proceedings of the Annual Joint Conference of the IEEE Computer Society and Communications Society (INFOCOM),New York,NY,USA,2002,3:1530-1539.
[7] Xiao Bin,Chen Wei,He Yanxiang,et al.An active detecting method against SYN flood attack[J/OL].The 11th International Conference on Parallel and Distributed Systems (ICPADS'05),Volume I,709-715,Fukuoka,Japan,2005.07.
[8] 陈杰,薛质,单蓉胜.一种基于 Patricia 树的检测 Syn Flood 攻击的方法[J].计算机工程,2004,30(13):26-28.
[9] 林白,李鸥,赵桦.基于源端网络的 SYN Flood 攻击双粒度检测[J].计算机工程,2005,31(10):132-134.
[10] 许晓东,杨海量,朱士瑞.基于重尾特性的 SYN 洪流检测方法[J].计算机工程,2008,34(22):179-181.
[11] Ling Yun,Gu Ye,Wei Guiyi.Detect SYN flood attack in edge routers[J].International Journal of Security and its Applications,2009,3(1):32-45.

(收稿日期:2011-02-16)

作者简介:
石利平,女,1971年生,硕士,讲师,主要研究方向:计算机应用,计算机安全等。