文章编号: 1008-1542(2008)01-0033-03

基于小波变换和 CUSUM 算法的 PDoS 攻击检测方法

岳 猛, 姜恩勇, 马润杰, 张文涛

(中国民航大学电子信息工程学院, 天津 300300)

摘 要: Pulsing denial-of-service(PDoS)攻击是一种新型的 DoS 攻击, 它主要利用 TCP 超时重传机制的弱点, 使用周期性的短脉冲向受害者发起攻击, 由于平均攻击速率较低, 传统的检测机制对其无法检测。笔者对这种攻击采用一种基于 Mallat 小波变换和 CUSUM 算法的两步检测法。首先基于 Mallat 小波变换和小波系数提取出攻击的特征, 然后用 CUSUM 算法可将小偏移量进行累加, 提高检测灵敏度, 最后针对 CUSUM 算法的缺陷, 提出了改进的 CUSUM 算法来降低误报率。两步检测法可以准确高效地检测出 PDoS 攻击。

关键词: PDoS; 超时重传; 小波变换; CUSUM 算法

中图分类号: TP393.08 文献标识码: A

Detection of PDoS attack based on wavelet transform and CUSUM algorithms

YUE Meng, JIANG En-yong, MA Run-jie, ZHANG Wen-tao

(College of Electronics and Information Engineering, Civil Aviation University of China, Tianjin 300300, China)

Abstract: Pulsing denial-of-service(PDoS) attack is a new class of DoS attack, which exploits the deficiencies of the minimum retransmission time out (RTO) of TCP to send out short bursts of attack packets to a victim. Because of its low rate, it can elude traditional high rate detection mechanism. This paper proposes a two-step detection approach based on Mallat wavelet transform and CUSUM (Cumulative Sum) algorithms. First, it extracts the characteristic of PDoS attack based on Mallat wavelet transform. Then, the CUSUM algorithms can cumulate the small offsets to enhance the detection sensitivity. In the end, for the defect of the CUSUM algorithms, it puts forward improved CUSUM algorithms to decrease the ratio of false alarm. The two-step detection approach can detect PDoS attack accurately and effectively.

Key words: PDoS; retransmission time out (RTO); wavelet transform; CUSUM algorithms

TCP 协议提供了一种可靠的面向连接的字节流运输层服务。当对端用户可靠的时候, TCP 超时重传机制可以使网络的吞吐量达到最大; 而恶意用户也可以利用超时重传机制的弱点产生拒绝服务攻击, 严重降低受害机器的吞吐量。

PDoS 攻击不同于以往的洪水攻击, 它通过估计合法 TCP 流的 RTO(retransmission time out)作为攻击的周期 T, 发送周期性的短脉冲, 如果攻击流量在 RTT(round trip times)时刻内足够导致数据包丢失, 那么合法 TCP 流将进入超时状态。PDoS 攻击最显著的特点是攻击者使用更少的攻击包就可以使受害主机的

收稿日期: 2007-10-18; 修回日期: 2007-11-30; 责任编辑: 陈书欣

作者简介: 岳 猛(1984-), 男, 河北沧州人, 硕士, 主要从事网络与信息安全方面的研究。

吞吐量大幅度降低, 因此它更加隐蔽; 同时可以灵活地调整参数, 使受害者的受害程度有所不同[1]。

因为 PDoS 攻击具有低速率特性, 所以传统的方法无法有效检测这种攻击[2]。笔者采用一种基于小波分析和自动调整的 CUSUM 算法的两步检测法来检测 PDoS 攻击。

1 PDoS攻击的模型分析

一个 PDoS 攻击脉冲序列可以用一个四元组表示为 $A(T_{\text{Extent}}, S_{\text{Extent}}, T_{\text{Space}}, N)$, 见图 1。其中: T_{Extent}是脉冲攻击长度; S_{Extent}是脉冲幅度; T_{Space}表示 2 个脉冲之间的时间间隔; N 是一次攻击发出的脉冲总数。

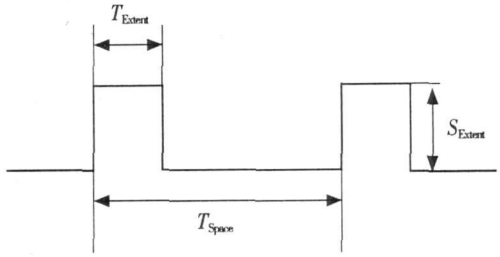

图 1 PDoS 攻击模型

2 两步检测法

要对 PDoS 攻击进行检测, 就要研究攻击发生时网络流量的特性。监测进入受害端的流量, 对到达路由器的数据流量(byte)每隔 0.25 s 进行取样, 取样时间为 900 s。为便于比较, 文中所有图的横轴范围统一为 0~900 s。由图 2 可以看出攻击发生期间(181~720 s) 流量下降且波动更剧烈, 这是因为正常流与攻击流混合造成的。

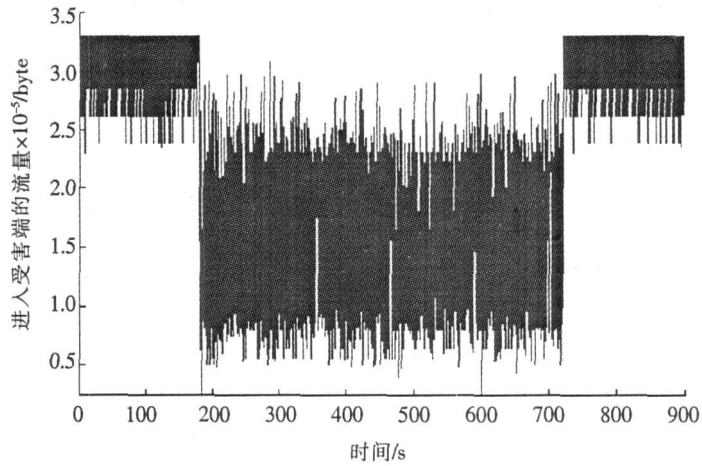

图 2 进入受害端的流量

Fig. 2 Traffic incoming victim

2.1 小波分析

基于上面的分析, 首先用离散小波变换分析流量的异常变化。对于一个信号 $f(t) \in L^2(R)$, 可以用尺度函数 $\Phi_{j,k}(t)$ 和小波函数 $\varphi_{j,k}(t)$ 将其表示为

$$f(t) = \sum_k c_{j_0}(k)\,\Phi_{j_0,k}(t) + \sum_k \sum_{j=j_0} d_j(k)\,\varphi_{j,k}(t), \quad (1)$$

其中, $c_j(k)$ 和 $d_j(k)$ 分别是尺度系数和小波系数, 它们可以通过 Mallat 算法求出[3]:

$$c_j(k) = \sum_m h_0(m-2k)\,c_{j-1}(m), \quad (2)$$

$$d_j(k) = \sum_m h_1(m-2k)\,c_{j-1}(m), \quad (3)$$

式中, h_0 和 h_1 分别是低通滤波器系数和高通滤波器系数, 因为小波系数(尺度 $j=1$) 有高通特性, 所以它可以捕获流量的波动特性。如图 3 所示, 横轴表示取样时间, 纵轴表示小波系数。可以看出在攻击期间流量具有剧烈波动特性。

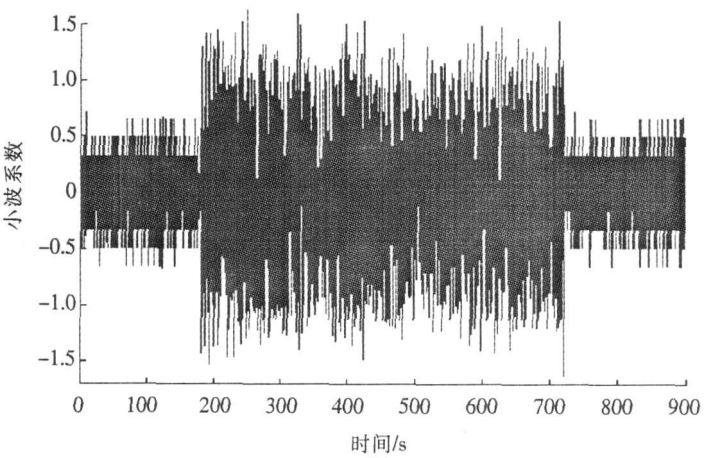

图 3 小波系数

Fig. 3 Wavelet coefficient

为了量化流量的波动特性, 基于信号能量的概念, 定义统计量 E_H。在给出 E_H 的表达式之前, 对取样数据加一个窗口 G, 即每 G 个取样点进行一次小波分析。这里取 $G=120$, 每个检测周期为 0.25 s × 120 = 30 s。这样第 n 个窗口的 $E_H(n)$ 就定义为

$$E_H(n) = \frac{1}{G}\sum_k |\,d^{\ln}_{1,k}\,|^2, \quad (4)$$

它反映了波动的程度，式中 $d_{1,k}^{\ln}$ 是小波系数。

2.2 CUSUM 算法检测

在上面小波分析的基础上，为了能准确及时地检测出 $E_H(n)$ 的突变点，使用 CUSUM 算法。CUSUM 算法是在统计过程控制中常用的算法，该算法是对信息进行累加，将过程中的小偏移量累加起来，达到放大的效果，以便提高检测灵敏度。CUSUM 算法在检测均值的小偏移时比较有效，而且根据点的倾斜程度的改变，可以方便、直观地检测到变化。它一般定义为

$$y(n) = (y(n-1) + x(n))^+, y(0) \equiv 0, n = 1,2,\cdots, \tag{5}$$

式中：$x(n)$ 为观察序列；$y(n)$ 为累计值；Δ 表示为 $y(n-1)+x(n)$，则 $\Delta^+ = \begin{cases} \Delta & (\Delta > 0) \\ 0 & (\Delta \leqslant 0) \end{cases}$。当 $x(n)$ 从负数变到正数时，累计值 $y(n)$ 就会从零变化到高于某个门限的值[4]，这时就检测出了 $x(n)$ 的变化。为了使 $E_H(n)$ 从负到正的变化，对其做如下变换：

$$Z_H(n) = E_H(n) - \beta_H。 \tag{6}$$

式中：β_H 是常数，一般取无攻击情况下 $E_H(n)$ 的上界。图 4 反映出统计量 $Z_H(n)$ 随时间的变化。

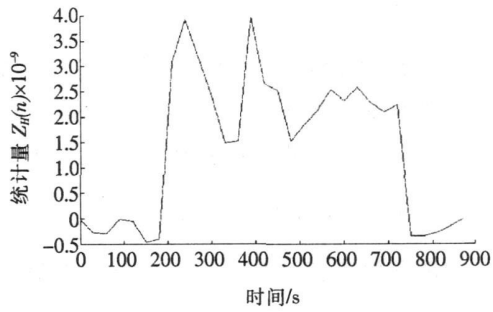

图 4　统计量 $Z_H(n)$

Fig. 4　Statistic $Z_H(n)$

对 $Z_H(n)$ 使用 CUSUM 算法进行检测。在这里 CUSUM 算法的递归定义为

$$y_{Z_H}(n) = (y_{Z_H}(n-1) + Z_H(n))^+,$$
$$n = 1,2,\cdots, \quad y_{Z_H}(0) = 0。 \tag{7}$$

图 5 是 $y_{Z_H}(n)$ 随时间变化的结果，其中 C_{CUSUM} 是门限，当 $y_{Z_H}(n)$ 大于 C_{CUSUM} 时即判定发生攻击。

3　自调整 CUSUM 算法检测

虽然 CUSUM 算法能快速有效地检测出攻击，但是攻击结束后仍会连续报警。针对这个缺陷，下面引入 CUSUM 算法的自调整机制。

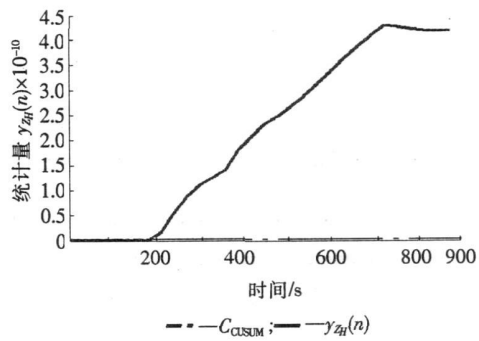

图 5　统计量 $y_{Z_H}(n)$

Fig. 5　Statistic $y_{Z_H}(n)$

在自调整算法里，一旦出现 CUSUM 算法累积值 $y_{Z_H}(n)$ 下降的情况，就在下一个检测周期开始对采样得到的原始流量数据进行检验。把一个检测周期分为 m 个相同的时间间隔小段，用 t 检验法来判定攻击是否已经结束。需要检验：

$$H_0: \mu \geqslant \mu_0 = M_k/m,$$
$$H_1: \mu < M_k/m。$$

式中：M_k 是在正常情况下通过学习得到的一个流量的数学期望；H_0 表示当前时间段流量大于或等于正常流量；H_1 表示当前时间段流量小于正常流量。则拒绝域为

$$t = \frac{F_{\text{avg}} - \mu_0}{s/\sqrt{m}} < t_\alpha(m-1), \tag{8}$$

式中 F_{avg} 为这一段时间流量的平均值。如果 H_1 可以接受，则认为攻击还没有结束；如果 H_1 被拒绝，则认为攻击已经结束，需要调整 $y_{Z_H}(n) = 0$ 以减少累积造成的影响。图 6 就是采用自调整 CUSUM 算法后的检测结果，测试环境与上面相同。与图 5 最显著的区别就是在攻击结束之后，$y_{Z_H}(n)$ 的值将自动调整为 0，消除了累积的负面影响，达到了减少误报的目的。

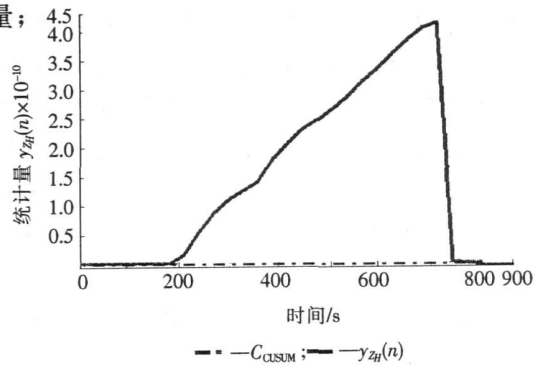

图 6　自调整 CUSUM 算法

Fig. 6　Self-adaptive CUSUM algorithms

（下转第 39 页）

PID 控制算法, 系统的输出 $u(k)$ 可以表示为

$$u(k) = K_P \times e(k) + K_I \times \sum_{j=0}^{k} e(k) + K_D \times [e(k) - e(k-1)]。 \quad (3)$$

系统每 30 s 对 pH 值进行检测, 与设定值比较后, 得到系统偏差 e 和偏差变化率 e_c, 如果 $|e| > 2$, 偏差较大, 系统输出最大, 即 $u(k) = 100$; 如果 $|e| \leqslant 2$, 经模糊化后, 查询模糊控制规则表, 并计算得到 K_P, K_I, K_D 值进行 PID 控制算法从而得到系统输出 $u(k)$。

基于 51 单片机的智能模糊自调整 PID 控制算法程序流程步骤如下。1) 连续 AD 采样 10 次, 去除最大和最小算术平均值滤波算法。2) 计算系统偏差, 和偏差变化率, 查控制规则表, 确定 K_P, K_I, K_D 值。3) 进行 PID 程序计算。4) 执行系统输出, 将 PID 程序计算结果转化为蠕动泵的通断时间。设 T 为程序执行周期, T_1 为蠕动泵最大开启时间, t 为一个基本输出时间单位, $t = T_1/100$, 则系统执行时间 $T_1 = u(k) \times t$, 由于 T_1 较大, 采用时间中断方法, 即用单片机 T_0 产生中断, 在中断程序中每间隔 t 时间对 $u(k)$ 进行递减, 当 $u(k)$ 减为 0 时, 即 T_1 时间到, 系统停止输出。5) 等待下一个程序执行周期, 重复 1)。

T_1 的选取在不同的时候、对不同的对象各不相同。选取过小, 调节阀反应过于迅速, 容易引起系统波动, 阀体经常性动作, 降低使用寿命; 选取过大, 调节阀反应迟钝, 使整个系统的调节时间加长。因此必须根据实际工况, 权衡相互之间的关系, 选取一个恰当的控制周期, 这对系统的整体性能有很大的影响[6]。

系统实际运行输出曲线见图 3, 图中横坐标单位为 min。

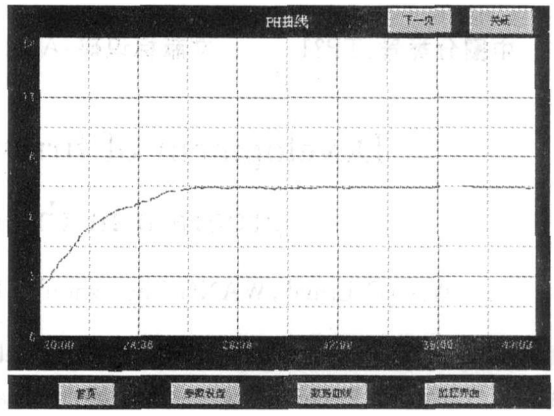

图 3　pH 值控制实验效果图

Fig. 3　Sketch map of fermentation pH control

4　结　语

试验证明, 将智能模糊自调整 PID 控制器应用在发酵过程 pH 值控制系统中, 既提高了系统控制精度, 缩短了系统调节时间, 又保证了系统不会出现超调和失调, 实际运行中, 发酵液中的实际 pH 值缓慢接近设定值, 本系统控制精度为 ±0.05。

参考文献:

[1] 于广平, 张　平, 魏松岩, 等. 仿人智能 PID 控制及在污水处理溶解氧控制中的应用[J]. 微计算机信息, 2006, 22(2): 13-15.
[2] 陈　众, 严　伟, 李祖枢. 基于 HSIC 的非线性 PID 控制器[J]. 控制与决策, 2003, 18(6): 695-696.
[3] 李丙才, 姜映红, 叶碧成. 啤酒发酵过程模糊智能 PID 控制[J]. 兰州理工大学学报, 2006, 32(6): 68-70.
[4] 孟　宇, 彭晓华, 张　浩. 模糊自适应整定 PID 控制及其仿真研究[J]. 机械工程与自动化, 2006, (6): 92-93.
[5] 南忠良, 严新忠. 微生物发酵过程环境控制系统的研究[J]. 天津轻工业学院学报, 2002, (3): 5-6.
[6] 吴小许, 孙　瑜, 黄　敏. 数字量调节阀的控制算法研究与应用[J]. 化工自动化及仪表, 2004, 31(2): 61-63.

(上接第 35 页)

4　结　语

笔者针对 PDoS 攻击提出了一种两步检测法, 此检测法的出发点是基于异常流量剧烈波动的特性。首先用小波分析来提取流量的特性, 然后用 CUSUM 算法来检测突变点。通过实验证明此方法可以实时有效地检测出 PDoS 攻击, 同时通过改进的自调整 CUSUM 算法消除了累积造成的负面影响, 从而降低了误报率。

参考文献:

[1] DELIO M. New Breed of Attack Zombies Lurk[EB/OL]. http://www.acm.org/technews/articles/2001-3/0514m.html, 2001-05-14.
[2] KUZMANOVIC A, KNIGHTLY E. Low-rate TCP-targeted denial of service attacks[A]. ACM SIGCOMM[C]. Karlsruhe: ACM, 2003.
[3] BURRUS C, GOPINATH R, GUO H. Introduction to Wavelets and Wavelet Transforms: A Primer[M]. NJ: Prentice Hall, 1998.
[4] BRODSKY B, DARKHOVSKY B. Non-Parametric Statistical Diagnosis Problems and Methods[M]. Netherlands: Kluwer Academic Publishers, 2000.