基于信号互相关的低速率拒绝服务攻击检测方法

吴志军,李 光,岳 猛
(中国民航大学电子信息工程学院天津市高级信号处理重点实验室,天津 300300)

摘 要: 低速率拒绝服务 LDoS(Low-rate Denial of Service)攻击是一种基于 TCP/IP 协议漏洞,采用密集型周期性脉冲的攻击方式.本文针对分布式 LDoS 攻击脉冲到达目标端的时序关系,提出基于互相关的 LDoS 攻击检测方法.该方法通过计算构造的检测序列与采样得到的网络流量序列的相关性,得到相关序列,采用基于循环卷积的互相关算法来计算攻击脉冲经过不同传输通道在特定的攻击目标端的精确时间,利用无周期单脉冲预测技术估计 LDoS 攻击的周期参数,提取 LDoS 攻击的脉冲持续时间的相关性特征,并设计判决门限规则.实验结果表明基于信号互相关的 LDoS 攻击检测方法具有较好的检测性能.

关键词: 低速率拒绝服务攻击; 互相关函数; 循环卷积; 时序; 检测
中图分类号: TP393.4 **文献标识码:** A **文章编号:** 0372-2112 (2014)09-1760-07
电子学报 URL: http://www.ejournal.org.cn **DOI:** 10.3969/j.issn.0372-2112.2014.09.015

Detecting Low-Rate DoS Attacks Based on Signal Cross-Correlation

WU Zhi-jun, LI Guang, YUE Meng
(*Tianjin Key Laboratory for Advanced Signal Processing*, *Civil Aviation University of China*, *Tianjin* 300300, *China*)

Abstract: Low-rate Denial of Service (LDoS) attack is TCP-targeted attack, which attempts to deny bandwidth of TCP flows. LDoS attacks send intensive periodic pulses at sufficiently low average rate to elude detection of DoS defense system. Based on the sequence relation between the distributed LDoS attack pulses arriving at the destination, a cross-correlation LDoS attack detection method is proposed by using cyclic convolution. This method builds a detection sequence for the purpose of exploring the timing relationship for distributed LDoS attack pulses arriving at the specific destination. Through computing the relation between the constructed detection sequence and sampled network flow sequence, the cross sequence is obtained. The cyclic convolution cross-relation algorithm is utilized to compute the precise time that the attack pulses arriving at the specific destination through different transferring channels. With nonperiodic monopulse prediction technology, the periodic parameters of LDoS attack are estimated, the relation characteristic of the pulse durations of LDoS attacks is extracted, and the threshold rules are designed. Experimental results show that the proposed algorithm of LDoS attack detection based on signal correlation achieves good detection performance.

Key words: low-rate denial of service (LDoS); cross-correlation; circular convolution; time sequence; detection

1 引言

拒绝服务 DoS(Denial of Service)攻击是网络攻击者最常采用的攻击手段之一.根据流量大小,DoS 攻击分为高速率洪水(flood)攻击[1],低速率地鼠(shrew)攻击[2].“地鼠”攻击,又称为 LDoS(Low-rate Denial of Service),它的流量仅占正常流量的 10 ~ 20%[3].由于该攻击在时域的平均速率低,流量占用的带宽小,具有出色的躲避检测能力,是网络犯罪者最钟爱的攻击方式之一.

一个 LDoS 攻击可以用一个三元参数组(T, L, R)来描述.其中:T 为攻击周期,它是两次连续的攻击脉冲之间的时间间隔,是通过来自可信赖源端估计的 TCP 超时重传 RTO(Retransmission Time-Out)计时器执行情况来

收稿日期:2013-08-12;修回日期:2014-01-09;责任编辑:覃怀银
基金项目:国家自然科学基金面上项目(No.61170328, No.U1333116);天津市应用基础与前沿技术研究计划(自然科学基金重点项目)(No.12JCZDJC20900);2013 年民航科技引导资金项目;中央高校基本科研业务费(No.3122013P007, No.3122013D007, No.3122013D003);中国民航大学科研建设平台项目;中国民航大学研究生课程建设项目

计算的;L 为脉冲宽度,它描述了攻击者持续发包的时间段;R 为脉冲幅度,它显示了攻击流的最高速率[4].

低速率分布式拒绝服务 LDDoS 攻击可以估计 RTO 时间来调整其攻击周期,周期性地发送攻击脉冲当脉冲达到最高速率 R 时,使得合法的 TCP 流试图发送数据包到目的主机的网络链路被严重的阻塞.如果 R 和 L 超过一定的值就存在被检测机制发现的可能.因此,攻击采用分布式的方式,来自不同信道的小攻击脉冲在特定位置汇聚成一个大的攻击脉冲.这就需要精确的时间控制,这些小脉冲必然在周期 T 和脉冲长度 L 上具有某种必然关系,即这些小脉冲相互之间有很高的相关性.针对这一特点,本文提出了基于信号互相关的检测算法,通过预先构造攻击模型,计算攻击模型与采集的网络流量间的相关性,从相关序列中提取攻击的各种特征以达到检测的目的.

2　相关工作

由于 LDoS 攻击的平均流量很小,在时域采用统计分析的方法很难检测到 LDoS 攻击.很多学者提出了在频率域利用频谱分析的方法检测 LDoS 攻击.因此,目前 LDoS 攻击的检测方法可以分为时域和频域检测两大类.

频域检测方法采用信号处理技术与网络流量数据处理技术相结合,把经典的信号检测理论和滤波器理论应用到 LDoS 攻击流量的检测和过滤方法中,成为 LDoS 攻击检测新的发展趋势.例如:南加州大学洛杉矶分校的 Chen Yu 和 Huang Kai[3]等提出了采用归一化功率谱密度 PSD,利用攻击流量和正常流量之间归一化功率谱密度 PSD 的最大距离作为判定 LDoS 攻击存在的依据;威斯康辛大学的 Paul 和 Jeffer[4],及国内武汉大学的何炎祥[5]等提出了采用小波处理的思想,利用离散小波变换 DWT(Discrete Wavelet Transform)技术将网络流量变换成为高、中和低 3 个频率分量,以便查找攻击流量.

在时域内,Gabriel 和 Pedro[6]给出了 LDoS 攻击的数学模型,这对 LDoS 检测算法的研究帮助很大.Xiang Yang 和 Li Ke[7]在 2011 年提出了用信息指标检测 LDoS 攻击,给出了广义熵和信息距离两种检测方法,并提出了一种基于信息距离的 IP 追踪方法能够在段时间内追踪到攻击源.

针对相关检测,国内已经有学者已经把互相关算法用于检测 DDoS 攻击,通过计算不同流量间的相关系数,设计判决门限以检测 DDoS 攻击[8],对 DDoS 攻击的时域检测做了很大贡献.

现有研究成果存在一些不足:(1)频域分析的方法要进行时域到频域的转化,然后进行频谱分析,需要较大的计算量,而且需要精确的网络流量的统计模型;(2)时域方法需要对网络流量进行大量的采样和统计.

本文针对上述不足之处,提出了基于信号互相关的 LDoS 攻击检测方法.

3　LDoS 攻击的互相关检测

互相关性检测就是利用信号互相关算法从背景流量中检测 LDoS 攻击流量.

3.1　网络流量分析

正常的网络中 80% 的流量为 TCP[3].因此,本文在实验中用 TCP 流量代表正常流量;用 UDP 流量代表攻击流量.

LDoS 的攻击周期 T 一般选取被攻击者的 RTO 加上 2 ~ 3 倍的往返时间 Round-Trip Time(RTT),即 RTO + 2 ~ 3RTT.脉冲持续时间 L 一般选取 1 ~ 2 个 RTT.由于现实网络 TCP 的 minRTO 一般为 1s,RTT 一般不超过 100ms.本实验选取 T 为 1200ms,L 为 200ms.

在实际网络中,攻击流量是隐藏在正常流量中的,混合了攻击流量的网络流量如图 1 所示.

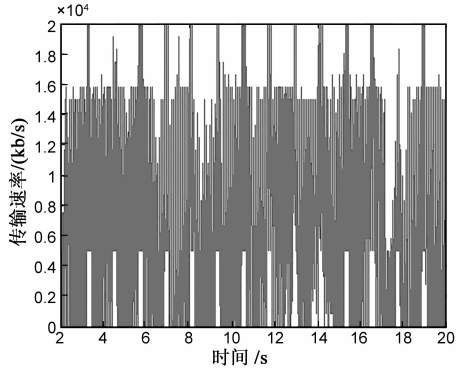

图1　LDoS攻击混合流量(采样间隔为10ms)

从图 1 可以看出,攻击流量几乎完全隐藏在正常流量中.在时域用时间统计平均的方法,很难检测 LDoS 攻击.

隐藏在混合流量中 LDoS 攻击流量仍然具有严格的时序特点.攻击流量中的每个脉冲都是方波脉冲,且具有相同的传输速率 R 和持续时间 L.相邻脉冲间遵循严格的时序关系,所有相邻脉冲的间隔都相同,为周期信号.因此,本文提出了基于信号互相关的算法来检测复杂背景中的周期 LDoS 攻击信号.

3.2　循环卷积的互相关算法

设 X、Y 为两个相同的 LDoS 攻击分布,$x(n)$,$y(n)$,$n = 1,2,\cdots,N-1$ 为分别属于这两个分布的序列.则这两个序列的相关系数 r 定义为:

$$r_{xy} = \frac{\mathrm{cov}(X,Y)}{\sqrt{D_X}\sqrt{D_Y}} = \frac{\sum_n [(x(n) - m_x) * (y(n) - m_y)]}{\sqrt{\sum_n (x(n) - m_x)^2}\sqrt{\sum_n (y(n) - m_y)^2}} \tag{1}$$

其中，m_x 和 m_y 分别是序列 $x(n)$和 $y(n)$的均值.

考虑到 LDoS 攻击序列在传输过程中经过不同的传输路径，造成的延时不同，两个序列的相关系数 r 并不能完全体现出序列间的相关程度.因此，在考虑不同时延时，两个时延为 d 的序列的相关系数 $r(d)$定义为：

$$r_{xy}(d)=\frac{\sum_n[(x(n)-m_x)*(y(n-d)-m_y)]}{\sqrt{\sum_n(x(n)-m_x)^2}\sqrt{\sum_n(y(n-d)-m_y)^2}} \tag{2}$$

其中，$d=0,\pm1,\pm2,\cdots,\pm(N-1)$[9].

式(2)的分子部分进行的是线性卷积计算.

在线性卷积计算中，随着 $y(n)$相对 $x(n)$的位移 d 的增大，$y(n)$与 $x(n)$叠加的部分会减少.为了解决这个问题本文用循环卷积来代替线性卷积.对两个长度为 N 的有限序列 $x(n)$和 $y(n)$，它们的循环卷积 $h(d)$定义为：

$$h(d)=x(n)\otimes y(n)=[\sum_n x(n)*\tilde{y}(n-d)]R_N(d) \tag{3}$$

其中，$\tilde{y}(n-d)$为 $y(n-d)$以 N 为周期的周期延拓.先计算长度为 N 的有限序列 $x(n)$和以周期为 N 的无限序列 $\tilde{y}(n-d)$的卷积，得到周期为 N 的无限序列 $\tilde{h}(d)$，再根据相对位移 d 的取值范围得到 $\tilde{h}(d)$的主值序列 $h(d)$.

基于循环卷积的互相关算法可以定义为：

$$r_{xy}(d)=\frac{\sum_n[(x(n)-m_x)*(\tilde{y}(n-d)-m_y)](R_N(d)+R_N(-d-1))}{\sqrt{\sum_n(x(n)-m_x)^2}\sqrt{\sum_n(y(n-d)-m_y)^2}} \tag{4}$$

对于理想的 LDoS 攻击序列 $x(n)$和 $y(n)$，求得基于循环卷积的 $r(d)$，如图 2 所示.

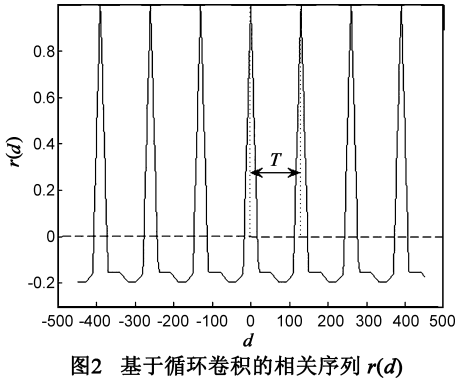

图2 基于循环卷积的相关序列 $r(d)$

从图 2 可以明显看出，应用基于循环卷积的互相关算法得出的 $r(d)$是一个以 T 为周期的周期序列.在实际网络环境中，攻击流量隐藏在正常流量中.因此，采用互相关算法在背景流量中检测 LDoS 攻击，就是从 $r(d)$序列中提取出 LDoS 攻击脉冲的 2 个相关特征 T 和 L.

3.3 基于循环卷积互相关的 LDoS 攻击检测

设含有攻击脉冲的混合流量为 $g(n)$，则：

$$g(n)=x(n)+h(n) \tag{5}$$

式中，$x(n)$为 LDoS 攻击脉冲序列；$h(n)$为背景流量(TCP 流量)，其概率分布符合 TCP 流量一般规律.

检测 LDoS 攻击就是检测 $g(n)$序列中是否隐藏有 $x(n)$序列.由于直接从 $g(n)$中分离 $x(n)$是十分困难的.因此，需要预先构造一个与 $x(n)$属于同分布的序列 $y(n)$，对混合流量序列 $g(n)$进行互相关检测.

令：X 表示序列 $x(n)$的概率分布函数；Y 表示序列 $y(n)$的概率分布函数；G 表示混合流量序列 $g(n)$的概率分布函数；D_X 为攻击序列的方差；D_Y 为构造的检测序列的方差；D_G 为混合流量的方差.则：

$$r_{gy}=\frac{\mathrm{cov}(G,Y)}{\sqrt{D_G}\sqrt{D_Y}} \tag{6}$$

因为，$G=X+H$，且认为 X 与 H 相互独立.所以：

$$\mathrm{cov}(G,Y)=\mathrm{cov}(X,Y)+\mathrm{cov}(H,Y) \tag{7}$$

$$D_G=D_{(X+H)}=D_X+D_H+2\mathrm{cov}(X,H) \tag{8}$$

则，

$$r_{gy}=\frac{\mathrm{cov}(X,Y)+\mathrm{cov}(H,Y)}{\sqrt{D_X+D_H+2\mathrm{cov}(X,H)}\sqrt{D_Y}}=\frac{\sqrt{D_X}}{\sqrt{D_X+D_H+2\mathrm{cov}(X,H)}}\frac{\mathrm{cov}(X,Y)+\mathrm{cov}(H,Y)}{\sqrt{D_X}\sqrt{D_Y}} \tag{9}$$

因此，

$$r_{gy}=\frac{\sqrt{D_X}}{\sqrt{D_G}}r_{xy}+\frac{\sqrt{D_H}}{\sqrt{D_G}}r_{hy} \tag{10}$$

再由于 H 与 X、Y 相关程度很低.所以，

$$r_{hy}\approx0 \tag{11}$$

即，$h(n)$与 $y(n)$的相关序列 $r(d)$近似为零.则，

$$r_{gy}\approx\frac{\sqrt{D_X}}{\sqrt{D_G}}r_{xy} \tag{12}$$

$$r_{gy}(d)=kr_{xy}(d)+\sigma \tag{13}$$

式中，$k\approx\frac{\sqrt{D_X}}{\sqrt{D_G}}$，$\sigma\approx r_{hy}$，

$$r_{xy}(d)\approx\frac{1}{k}r_{gy}(d) \tag{14}$$

其中，混合流量的方差 D_G 可对混合流量采样后计算得到；攻击序列的方差 D_X 可近似地认为等于构造出的检测序列 $y(n)$的方差 D_Y.

根据式(14)，可以利用混合流量 $g(n)$和预先构造的检测序列 $y(n)$计算出 $r_{xy}(d)$.进而从中提取出 LDoS 的相关特征，检测出攻击.关于 $y(n)$的构造会在后面部分介绍.

在求出 $r_{xy}(d)$之后,首先需要设计判决规则来判定混合流量中是否混有 LDoS 攻击脉冲.这里,定义判决的参数有:①采样窗口:一定长的时间窗口,保证可以容纳足够多的异常点;②敏感系数:当相关系数超过这一值,表示可能出现异常;③计数器:用来统计异常范围;④判决门限:用来判定是否有攻击.

互相关检测算法的步骤如下:

(1)在受害端的上一跳路由监测流量,每隔 t 秒的间隔对流量进行取样,一个取样周期为 $t * N$ 秒,得到长度为 N 的序列 $x(n)$;

(2)构造周期估计序列 $y'(n)$,用互相关算法估计周期;

(3)根据估计的周期构造检测序列 $y(n)$;

(4)利用互相关算法求互相关序列 $r_{xy}(d)$;

(5)求出 $r_{xy}(d)$, $d=0,\pm1,\pm2,\cdots,\pm(N-1)$后,对于 d,从 $d=-(N-1)$依次检测对应 $r_{xy}(d)$的值,如果 $r_{xy}(d)$的值大于敏感系数,则计数器加 1;如果 $r_{xy}(d)$的值小于敏感系数,且计数器 >0,则计数器减 1.如果计数器大于判决门限,说明 $r_{xy}(d)$连续大于敏感系数,出现了波峰,则判定攻击存在.

3.4 检测序列的构造

为了保证检测效果,在构造检测序列 $y(n)$的时候需要精确预估计攻击序列的 R, L 和 T 的值.作为检测方,攻击脉冲的相关参数是不可预知的.因此,估计参数与实际参数之间的误差直接影响到检测性能,必须保证各参数的估计值的误差对检测结果的影响缩小在可接受的范围.

3.4.1 参数 R 的预估计

R 的取值只影响 $y(n)$的大小,设 $Z=nY$,则:

$$r_{xz}=\frac{\mathrm{cov}(X,Z)}{\sqrt{D_X}\sqrt{D_Z}}=\frac{\mathrm{cov}(X,nY)}{\sqrt{D_X}\sqrt{n^2D_Y}}=\frac{n\mathrm{cov}(X,Y)}{\sqrt{D_X}*n\sqrt{D_Y}}=r_{xy} \tag{15}$$

根据式(15)得知对 R 值估计的误差 $\delta(\delta=\hat{R}-R)$对检测结果没有任何影响.因此,在构造检测序列 $y(n)$时,R 一般取链路瓶颈的大小.

3.4.2 参数 L 的预估计

参数 L 估计的误差对 $r(d)$序列的周期没有影响.它主要影响每个波峰的形状,当估计的脉冲持续时间 $\hat{L}$ 过大时波峰形状由三角波变为梯形波.下面分析估计误差对 $r(d)$峰值的影响.

首先给出 LDoS 攻击序列的数学模型

$$x(n)=\begin{cases}R, & kT\leqslant n<kT+L\\ 0, & kT+L\leqslant n<(k+1)T\end{cases} \tag{16}$$

其中,$k=0,1,2,\cdots,n$.则,两个参数(R、L、T)相同 LDoS 序列间的互相关序列 $r(d)$为:

$$\begin{aligned}r(d)&=\frac{\sum_{n=0}^{N}(x(n)-m_x)(y(n-d)-m_y)}{\sqrt{\sum_{n=0}^{N}(x(n)-m_x)^2}\sqrt{\sum_{n=0}^{N}(y(n-d)-m_y)^2}}\\&=\frac{a\sum_{n=0}^{T}(x(n)-m_x)(y(n-d)-m_y)}{\sqrt{a\sum_{n=0}^{T}(x(n)-m_x)^2}\sqrt{a\sum_{n=0}^{T}(y(n-d)-m_y)^2}}\\&=\frac{\sum_{n=0}^{T}(x(n)-m_x)(y(n-d)-m_y)}{\sqrt{\sum_{n=0}^{T}(x(n)-m_x)^2}\sqrt{\sum_{n=0}^{T}(y(n-d)-m_y)^2}}\end{aligned} \tag{17}$$

其中,N 为采样窗口大小;a 为采样窗口内脉冲个数.

由于均值 m_x, m_y 对两个序列相关性影响不大,为了简化计算,可以把 $x(n)$、$y(n)$的均值 m_x, m_y 都设为 0.则 $r(d)$的峰值为:

$$\max r(d)=\frac{L(R-0)^2}{\sqrt{L(R-0)^2}\sqrt{L(R-0)^2}}=1 \tag{18}$$

而在构造检测序列 $y(n)$时对其参数 L 的估计会出现一些偏差.设 $\hat{L}$ 为估计值,则当 $\hat{L}>L$ 时,

$$\max r(d)=\frac{LR^2}{\sqrt{LR^2}\sqrt{\hat{L}R^2}}=\frac{L}{\sqrt{L\hat{L}}} \tag{19}$$

当 $\hat{L}<L$ 时,

$$\max r(d)=\frac{\hat{L}R^2}{\sqrt{LR^2}\sqrt{\hat{L}R^2}}=\frac{\hat{L}}{\sqrt{L\hat{L}}} \tag{20}$$

则对参数 L 估计的误差 $\delta(\delta=\hat{L}-L)$对相关序列峰值的影响为:

$$\max r(d)=\begin{cases}\dfrac{\delta+L}{\sqrt{L(L+\delta)}}, & \text{当 } \delta<0 \text{ 时}\\ \dfrac{L}{\sqrt{L(L+\delta)}}, & \text{当 } \delta>0 \text{ 时}\end{cases} \tag{21}$$

根据式(21),当 L 为 200ms 时 max $r(d)$与 δ 之间的关系如图 3 所示.

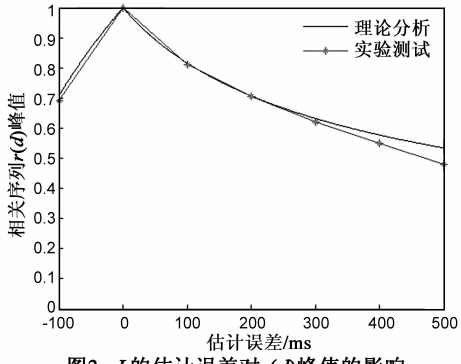

图3　L的估计误差对$r(d)$峰值的影响

在图 3 中,实线为理论分析值;星型实线表示实验测试值.图 3 说明理论分析结果和实际测试结果基本吻合.

在实际情况下,LDoS 攻击脉冲的持续时间 L 一般为 2~3 个 RTT,正常网络 RTT 一般不超过 100ms.过大的 L 会使 LDoS 攻击沦为 DDoS 攻击.因此,在构造检测序列 $y(n)$时,可以根据被攻击网络环境取 1~3 个 RTT 大小做为 L 的取值,估计误差不会超过 1~2 个 RTT.L 的估计误差 δ 在一定范围内,对相关序列 $r(d)$的影响很小,不影响检测结果.

3.4.3 参数 T 的预估计

周期性是 LDoS 攻击最显著的特征.因此,在构造检测序列 $y(n)$时,对 T 的预估计必须十分精确.对于 LDoS 攻击,当其周期固定,相邻脉冲的时间间隔相同时,对其周期估计的误差对检测结果影响很大.因此,在构造 $y(n)$之前,先构造一个 $y'(n)$.$y'(n)$是一个单脉冲的序列,它本身是不具备周期性的,它的作用是检测 $x(n)$的周期性.构造的 $y'(n)$以及它与 $x(n)$的相关序列 $r(d)$如图 4 所示.

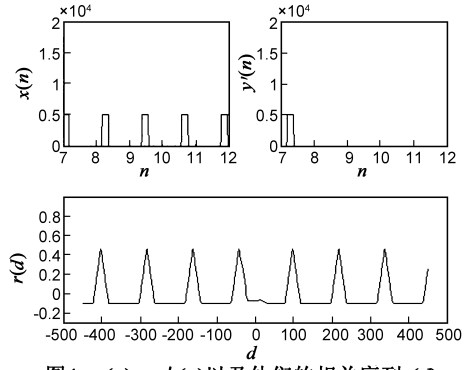

图4 $x(n)$, $y'(n)$以及他们的相关序列$r(d)$

虽然 $x(n)$与 $y'(n)$之间的相关程度不高,但是相关序列 $r(d)$完整的展现出了 $x(n)$的周期性.而且每个波峰的峰值大约为 0.5 左右,明显大于背景流量 $h(n)$与 $y(n)$的相关程度,表明能够用 $y'(n)$从混合流量中提取出 $x(n)$的周期特性.只要计算相邻波峰间的距离,就可以估计出攻击的周期,进而完成对 $y(n)$的构造.

由于估计出的周期 $\hat{T}$ 和实际周期 T 间仍可能出现误差.误差较小时对相关序列 $r(d)$的影响主要体现在对 $r(d)$的峰值的影响,而对相关序列 $r(d)$的周期性的影响不明显.

设 $x(n)$参数 $T=1200$ms,$L=200$ms,采样窗口大小 $N=3000$ms,窗口内脉冲个数为 3 个,估计的周期为 $\hat{T}$,构造的检测序列 $y(n)$以 $\hat{T}$ 为周期,其他参数与 $x(n)$相同.相关序列 $r(d)$取得峰值

$$\max r(d)=\frac{LR^2+2(L-|\delta|)R^2}{\sqrt{3LR^2}\sqrt{3LR^2}}=1-\frac{2|\delta|}{3L} \quad (22)$$

得到 $\max r(d)$与 δ 之间的关系如图 5 所示.

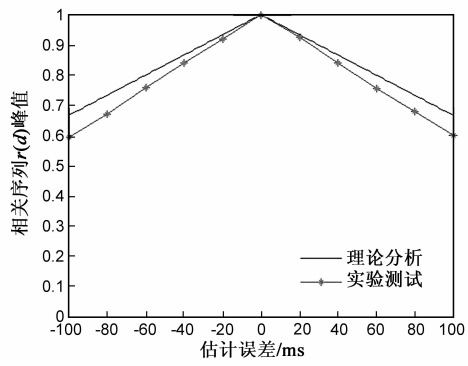

图5 T的估计误差对$r(d)$峰值的影响

在图 5 中,实线为理论分析值,星型实线表示实验测试值.图 5 表明理论分析结果和实际测试结果基本吻合.周期 T 的估计误差 δ 在可以控制的较小范围内对相关序列峰值的影响较小,不影响检测结果.估计误差的范围在下面具体试验中求得.

4 实验及结果分析

本文利用网络仿真软件 *NS*-2 搭建网络环境测试基于信号互相关的 LDoS 攻击检测算法.

4.1 实验环境

测试网络拓扑是一个变形的哑铃形状,如图 6 所示.

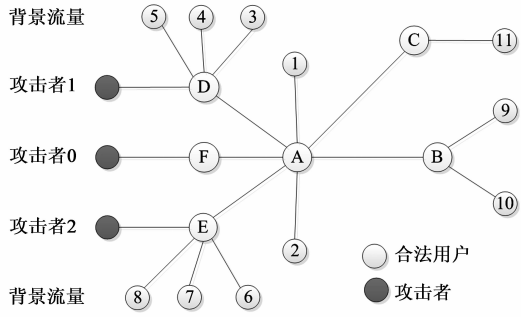

图6 实验环境

合法用户 1 为被攻击目标,它经过路由器 A 和 B 与用户 9、10 建立正常的 TCP 连接.路由器 A 与 B 的链路瓶颈为 15Mbps,攻击者在路由 A 处完成流量汇聚,形成攻击速率 15Mbps、攻击脉宽 200ms、攻击周期 1.2s 的攻击脉冲,然后对路由 A、B 发起攻击使经过 A~B 的通信中断.其中,攻击者 1、2 混合了背景流量,使其更贴近真实流量以增加检测的难度.攻击者 0 未混合背景流量,用来做比较.合法用户 2 用来做检测的干扰项,它经过路由 A 和 C 与用户 11 建立正常的 TCP 连接,它不受攻击,因此,用户 2 的流量为正常的无攻击网络流量.将用户 3~8 设为背景流量,它们避开被攻击的路径 A~B,经过路由 A 和 C 与用户 11 建立正常的 TCP 连接,使得背景流量不受攻击的影响,保证背景流量与攻击流

量是相互独立的.

4.2 实验及结果分析

将检测算法应用于测试网络中的路由器 A 处,在 A 处对与 A 相连路径上的流量进行采样(采样间隔为 10ms),得到的 2000 个点的混合流量序列 $g(n)$.首先构造 $y'(n)$,其中,R = 5Mbps、L = 200ms.对 $g(n)$进行周期估计,统计相关序列中每个峰值对应的延迟 d 和相邻峰值间距见表 1.

表 1　周期估计结果

d	−407	−284	−165	−45	93	210	335	455
Δd	~	123	119	120	118	117	125	120

对其值求平均得到 120.3.采样间隔 d = 10ms,得到估计出的周期为 T = 1203ms.而实际周期为 1200ms,误差为 3ms,其值小于一个采样间隔.因此,用 $y'(n)$估计出的周期比较准确.

重复上述实验 100 次,测得周期的平均值为 1201ms.其中,偏差最大的值为 1217ms,平均误差为 5ms,最大误差为 17ms.由此可以得出用单脉冲互相关方法检测出的周期精确,误差较小,对检测结果影响较小.

根据估计出的周期 T 构造 $y(n)$,并计算 $D_X = D_Y = 2.8790 * 10^6$ 和 $D_G = 2.2788 * 10^7$,得到 $k = \sqrt{D_X}/\sqrt{D_G} = 0.3554$.然后,计算 $g(n)$和 $y(n)$的互相关序列 $r_{gy}(d)$,并根据式(14)和 k 值求出 $r_{xy}(d)$,如图 7 所示.

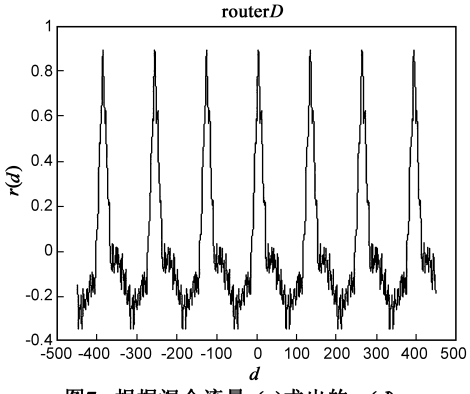

图7　根据混合流量$y(n)$求出的$r_{xy}(d)$

图 7 为混合流量 $g(n)$与 $y(n)$互相关计算后求出的 $r_{xy}(d)$,它与图 3 非常接近,具有明显的周期性,其周期与攻击脉冲的周期相同,并且每个周期内的峰值在 0.8 以上.这表示混合流量序列 $g(n)$中隐藏有与检测序列 $y(n)$相关程度很高的攻击脉冲序列.利用互相关算法能够从混合流量中提取出攻击脉冲的相关特征,达到检测目的.

由中心极限定理可知,当实验次数足够大时相关序列峰值的分布近似服从正态分布.设无攻击的相关序列的峰值分布为正态分布 $N(\mu_0,\sigma_0^2)$,概率密度函数为:

$$f_0(x)=\frac{1}{\sqrt{2\pi}\sigma_0}\exp\left\{-\frac{(x-\mu_0)^2}{2\sigma_0^2}\right\} \tag{23}$$

有攻击的相关序列的峰值分布为正态分布 $N(\mu_1,\sigma_1^2)$,,概率密度函数为:

$$f_1(x)=\frac{1}{\sqrt{2\pi}\sigma_1}\exp\left\{-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right\} \tag{24}$$

通过 100 次实验求出无攻击正常流量的相关序列峰值的均值 μ_0 为 0.2306,标准差 σ_0 为 0.2235;有攻击混合流量的相关序列峰值均值 μ_1 为 0.8705,标准差 σ_1 为 0.1286.

首先给出 3 个检测算法的性能指标:

(a)正确检测(攻击存在并准确检测出攻击)概率;

$$P_D=\int_{\gamma}^{\infty}f_1(x)\mathrm{d}x$$

(b)漏报(攻击存在但未检测出攻击)概率;

$$P_{FN}=\int_{-\infty}^{\gamma}f_1(x)\mathrm{d}x$$

(c)误报(攻击不存在误将正常流量当作攻击)概率.

$$P_{FP}=\int_{\gamma}^{\infty}f_0(x)\mathrm{d}x$$

在不同 γ 取值下求出的 LDoS 攻击检测性能结果见表 2.表 2 说明敏感系数在取值为 0.6 ~ 0.7 时有较好的检测性能.

表 2　不同 γ 取值下异步攻击检测性能

敏感系数 γ	检测率 P_D	漏报率 P_{FN}	误报率 P_{FP}
0.55	99.38%	0.62%	8.62%
0.6	98.35%	1.65%	5.61%
0.65	95.98%	4.02%	3.5%
0.7	91.34%	8.66%	2.09%
0.75	83.5%	16.5%	1.19%

比较图 7,可以看出有攻击的混合流量的相关序列的形状是三角波峰.根据这一特点,在实际检测时采用敏感系数和计数器双门限的检测方法,可以使检测性能进一步提高.

将本文方法与相关领域典型的研究成果进行比较,结果如表 3 所示.

表 3　不同检测方法检测性能比较

检测方法	检测率 P_D	漏报率 P_{FN}	误报率 P_{FP}
归一化累计功率谱密度 NCPSD[3]	88%	12%	16.7%
基于互相关的 DDoS 检测方法[8]	99.82%	0.18%	0.1%
本文方法	98.35%	1.65%	5.61%

本文方法与归一化累计功率谱密度 NCPSD[3]的方法相比较,在检测性能方面,本文方法的检测率提高了 7% ~10%,误检率减少了 10% ~13%.而与基于互相关的 DDoS 检测方法[8]的检测性能十分接近.

5 总结

本文分析了 LDoS 攻击流量的周期性的特点,研究了存在和不存在 LDoS 攻击两种情况下的正常流量和有攻击的混合流量在时域和频域的差异,改进了传统的信号互相关计算方法,提出了基于循环卷积的 LDoS 攻击互相关检测算法.设计了 LDoS 攻击模型的构造方法,预先构造检测序列,设计了判决规则,再计算检测序列与采样得到的网络流量序列的相关序列,并提出了用无周期单脉冲预测估计 LDoS 攻击的周期、脉宽和脉高三个参数的方法从相关序列中提取出 LDoS 攻击参数.实验结果表明本文提出的检测算法具有较好的检测性能.

在 LDoS 攻击发生时,不仅不同路径上的 LDoS 攻击脉冲具有相关性,由攻击造成的一些特征,比如:攻击下的队列平均报文长度的变化和 RTT 的变化等,也具有很强的相关性.本文的研究仅利用了 LDoS 攻击流量上的相关性.在未来的研究中,可以充分考虑到 LDoS 攻击下的网络所表现出的各种特征的相关性,可以进一步提高互相关检测算法的检测性能.

参考文献

[1] 孙长华,刘斌.分布式拒绝服务攻击研究新进展综述[J].电子学报,2009,37(7):1562 – 1570.
SUN Chang-hua, LIU Bin. Survey on new solutions against distributed denial of service attacks[J]. Acta Electronica Sinica, 2009,37(7):1562 – 1570.(in Chinese)

[2] Kuzmanovic A, Knightly E W. Low-rate TCP-targeted denial of service attacks-The shrew vs the mice and elephants[A]. Proc ACM Sigcomm 2003[C]. Karlsruhe, Germany: ACM, 2003. 25 – 29.

[3] Yu CHEN, Kai HWANG, et al. Collaborative Defense Against Periodic Shrew DDoS Attacks in Frequency Domain[R]. CA, USA: USC Internet and Grid Computing Lab, 2005.

[4] Barford P, Kline J, et al. A signal analysis of network traffic anomalies[A]. Proc ACM Sigcomm Internet Measurement Workshop[C]. Marseille, France: ACM, 2002. 71 – 82.

[5] 何炎祥,曹强,等.一种基于小波特征提取的低速率 DoS 检测方法[J].软件学报,2009,20(4):930 – 941.
HE Yan-Xiang, CAO Qiang, et al. A low-rate DoS detection method based on feature extraction using wavelet transform[J]. Journal of Software, 2009,20(4):930 – 941.(in Chinese)

[6] Gabriel M F, Jesús E D V, Pedro G T. Mathematical model for low-rate DoS attacks against application servers[J]. IEEE Transactions on Information Forensics and Security, 2009, 4(3):519 – 529.

[7] Yang X, Ke L, et al. Low-Rate DDoS attacks detection and traceback by using new information metrics[J]. IEEE Transactions on Information Forensics and Security, 2011, 6(2):26 – 437.

[8] WEI W, FENG C, et al. A rank correlation based detection against distributed reflection DoS attacks[J]. IEEE Communications Letters, 2013, 17(1):173 – 175.

[9] WU Z J, MA L, et al. Research on time synchronization and flow aggregation in LDDoS attack based on cross-correlation [A]. Proc IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communication, TrustCom 2012 [C]. Liverpool, United Kingdom: IEEE Computer Society, 2012. 25 – 32.

作者简介

吴志军(通信作者) 男,1965 年出生于新疆库尔勒,现为中国民航大学教授、博士生导师,主要研究方向为网络与信息安全.
E-mail:zjwu@cauc.edu.cn

李 光 男,1988 年出生于天津宝坻,现为中国民航大学通信与信息系统专业研究生,主要研究方向为网络与信息安全.
E-mail:liguang0059@163.com

岳 猛 男,1984 年出生于河北沧州,现为中国民航大学助教,主要研究方向为网络与信息安全.
E-mail:myue@cauc.edu.cn