软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software,2015,26(Suppl.(2)):90−99

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

RoQ 攻击的特征提取和检测*

文 坤[1,2], 杨家海[1,2], 李晨曦[1,2], 程凤娟[3], 尹 辉[3]

[1](清华大学 网络科学与网络空间研究院,北京 100084)

[2](清华信息科学与技术国家实验室(筹)(清华大学),北京 100084)

[3](河南工业大学 信息科学与工程学院,河南 郑州 450001)

通讯作者: 杨家海, E-mail: yang@cernet.edu.cn

摘 要: 降质攻击(RoQ)是一种非典型拒绝服务攻击,具有很强的隐蔽性,大多数传统的基于 DoS 攻击的检测方法不再适用.迄今为止,有不少学者提出了许多新的方法,但这些检测方法在不同程度上存在误报率较高的情况.为此,提出了一种改进的检测方法,它在分析和提取异常突变特征的基础上,对异常突变的局部流量进行了二次频谱分析,提取了攻击的周期特征,从而提高了检测的精确度.模拟实验及对比分析结果表明,该检测方法的检测精度高,其误报率和漏报率都很低.

关键词: 网络安全;RoQ 攻击;异常检测;小波分析;倒频谱

中文引用格式: 文坤,杨家海,李晨曦,程凤娟,尹辉.RoQ 攻击的特征提取和检测.软件学报,2015,26(Suppl.(2)):90−99. http://www.jos.org.cn/1000-9825/15019.htm

英文引用格式: Wen K, Yang JH, Li CX, Cheng FJ, Yin H. Characteristics extraction and detection of RoQ attack. Ruan Jian Xue Bao/Journal of Software, 2015,26(Suppl.(2)):90−99 (in Chinese). http://www.jos.org.cn/1000-9825/15019.htm

Characteristics Extraction and Detection of RoQ Attack

WEN Kun[1,2], YANG Jia-Hai[1,2], LI Chen-Xi[1,2], CHENG Feng-Juan[3], YIN Hui[3]

[1](Institute for the Network Sciences and Cyberspace, Tsinghua University, Beijing 100084, China)

[2](Tsinghua National Laboratory for Information Science and Technology(TNList) (Tsinghua University), Beijing 100084, China)

[3](College of Information Science and Engineering, He'nan University of Technology, Zhengzhou 450001, China)

Abstract: Reduction of quality (RoQ) attack is an atypical denial of service (DoS) attack, which has a strong concealment. Consequently, most traditional methods of detection are no longer applicable. There are a number of new methods developed recently. However, most of these methods have higher false positive rate in varying degree. In this paper, a novel method is proposed based on the principle of time-frequency analysis with Wavelet multi-resolution and Cepstral technique. First, according to different time-domain characteristics, the potential anomaly is detected and the abrupt change point is located. Secondly, the local traffic around the abrupt change point is analyzed by cepstrum. The potential characteristics of attack periodicity is extracted. By the two-stage detection, this new method ultimately can confirm whether the network is affected by the attack. Results of simulations and real network experiments demonstrate that the presented algorithm can detect RoQ attacks accurately with very low false positive rate and false negative rate.

Key words: network security; RoQ attack; anomaly detection; wavelet analysis; cepstrum

拒绝服务(DoS)攻击导致把目标计算机的网络资源及系统资源耗尽,使之无法为请求的用户提供正常的服务.黑客通过僵尸网络就可以发动大规模 DDoS 攻击,在短时间内就能使受害端的网络瘫痪或系统崩溃.目前,在互联网上发生的网络攻击中,DDoS 攻击发生的次数最多、产生的攻击流量最大[1].尽管危害很大,但是,在受害端,它的攻击效果会表现得非常明显,因此,只要在受害端的网络或系统中部署相应的防御措施,就能够进行有

* 基金项目: 国家自然科学基金(61170211)
收稿时间: 2014-05-02; 定稿时间: 2014-08-22

效的防范[2].可是,随着拒绝服务攻击的演变,产生了一种更加隐蔽的攻击方式——降质(reduction of quality,简称 RoQ)攻击.在 2003 年举行的 SIGCOMM 会议上,Kuzmanovic 首次阐述了这种新型的拒绝服务攻击,形象地称之为 Shrew 攻击[3].之后不久,Luo 等人进一步评估了该攻击的不同攻击方式:同步和异步攻击[4,5].他们认为该攻击的最大特征是流量具有周期性脉冲特征,因此,称之为脉冲式拒绝服务(pulsing denial-of-service,简称 PDoS)攻击.Guirguis 等人则称之为降质攻击[6,7],他们认为,在许多网络协议或系统的自适应机制中,都有稳定状态被瞬间改变的时候,而一旦系统长期处于非平稳状态,系统的服务质量将极大降低.尽管这种攻击有多种称呼,但是大家对该攻击的原理已经达成共识,我们称为 RoQ 攻击.这种攻击的目标是 TCP 协议,它利用了 TCP 拥塞控制的自适应机制,用相对较少的攻击成本,却显著地减少或抑制了正常的 TCP 流量,从而降低网络应用中有关 TCP 应用的服务质量.对于被 TCP 流量占去一半以上的互联网而言,这是一个巨大的潜在威胁,已经引起众多研究者的关注.

与传统的 DoS 攻击相比,RoQ 攻击主要有以下不同:第一,攻击目标是网络中的瓶颈链路,而不再是明确具体的网络终端.第二,攻击成本相对较少,攻击时会发送瞬时的高速脉冲流量,而不再是发送持续超量的攻击流(比如,UDP flooding 攻击和 TCP SYN flooding 攻击等).第三,攻击的效果也不再是受害端的网络瘫痪或系统崩溃,而是显著地降低或抑制 TCP 服务质量.同时,当攻击发生时,受害端的服务资源耗用情况不升反降.由此可以看出,RoQ 攻击的特征非常独特,且具有很强的隐蔽性,传统的 DoS 攻击的检测方法并不适用于 RoQ 攻击[8].

目前,研究者们提出了不少新的检测和防御方法,比如,Kuzmanovic[3]、Luo[5]、Sarat 和 Terzis[9]、Sun[10]和 Chen[11]等人都先后提出了不同的方法.近年来,Chen 等人在文献[12]中提出了使用频谱特征分析的方法,通过攻击前后流量在时域和频域的不同表现,从而实现攻击的检测.Yang 等在文献[13]中利用广义熵和信息距离在分离度上的表现,提出了一种基于这两种信息量度的检测算法.上述检测方法都不约而同地针对该攻击对网络流量产生的各种变化情况进行特征分析和检测,通过各种参数的调整,都可以实现很好的检测效果,漏报率可以很低.但是,实际网络中存在大量合法的流量拥塞现象,而且,还有许多网络应用(比如视频点播等)的流量和 RoQ 攻击流量的特征非常相似,对于这些因素造成的网络流量变化,上述检测方法会很容易产生误判,从而在检测时会出现漏报率很低,误报率却较高的糟糕效果.

基于此,我们提出了一种改进的特征提取和检测方法.它的主要贡献在于:在分析和提取异常突变特征的基础上,还对异常突变的局部流量进行了二次频谱分析,提取了攻击的周期特征,从而提高了检测的精确度.实验表明,该检测方法的误报率和漏报率都很低.

本文第 1 节给出攻击原理及流量特征分析.第 2 节给出攻击特征提取和检测算法的具体实现.第 3 节是实验及分析.第 4 节对全文进行总结.

1 攻击原理及流量特征分析

1.1 攻击原理

这种攻击的目标是 TCP 协议,它利用了 TCP 拥塞控制的自适应机制,用相对较少的攻击成本,却显著地减少或抑制了正常的 TCP 流量,从而降低网络应用中有关 TCP 应用的服务质量.如图 1 所示,当网络出现拥塞时,将会出现正常的拥塞调整和控制.当处于 *knee* 状态时,网络会处于一个良好运行状态.但是,当出现 *cliff* 状态时,将不得不进行网络拥塞控制,此时,由于超时等机制的启动,发送端将会把发送速度调为最低[14].

如果频繁出现 *cliff* 状态,那么网络中的 TCP 流量将会显著减少.从这个意义上说,它成为了 RoQ 攻击利用的一个安全漏洞.

我们用 3 种参数描述 RoQ 攻击,如图 2 所示.T 为攻击周期,它与 TCP 机制的时间参数相关;t 为攻击持续时间;R 为攻击脉冲的速率,其大小需要达到能使网络出现短时拥塞.

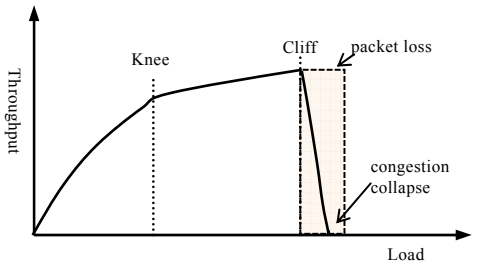


图 1 拥塞控制

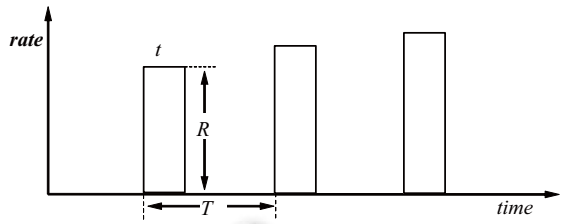


图 2 攻击脉冲

为了定量分析攻击的效果,我们定义了攻击效果γ和攻击成本 C 如下:

定义 1. 攻击效果γ

$$\gamma = 1 - \frac{M_{attack}}{M_{normal}} \tag{1}$$

这里, M_{attack} 和 M_{normal} 分别代表在相同的时间内,发送端在攻击前后成功地发送的 TCP 流量的大小(bytes).

定义 2. 攻击成本 C

$$C = \frac{R \times t}{T} \tag{2}$$

攻击成本 C 代表 RoQ 攻击的平均攻击速率,当γ达到一定阈值时,我们认为这是一次有效的攻击.

1.2 攻击流量特征

1.2.1 突变特征

用户访问网络是随机事件,这也决定了网络流量的产生具有随机性,同时,由于网络用户数量超大,因此,访问的流量也是巨大的,瞬间的流量突变也是很合理的现象.因此,突变性是网络流量的一个基本特征.作为互联网中所占比重最大的 TCP 流量而言,除了同样具有上述特征之外,它的流量特性还取决于自身独有的自适应拥塞控制机制.

TCP 流量的突变生成的原因主要有两种:一种是随机的用户访问,会在访问开始和结束时发生流量突变;二是当网络处于 *cliff* 状态时,TCP 的自适应机制也会导致流量突变.这是因为 TCP 流量的 RTO 和 AIMD 机制会在网络处于 *cliff* 状态时进行非常大的流量大小调整,如图 3 和图 4 所示,详细算法见文献[15].

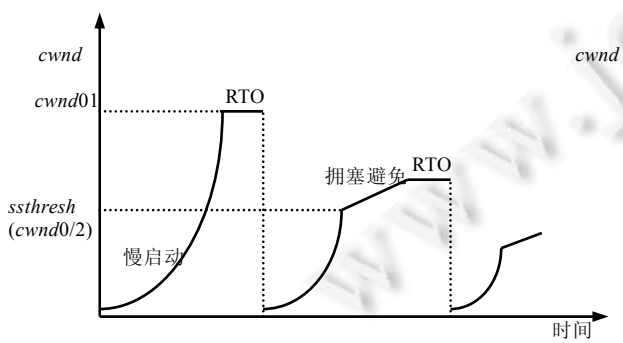


图 3 RTO 机制的状态调整

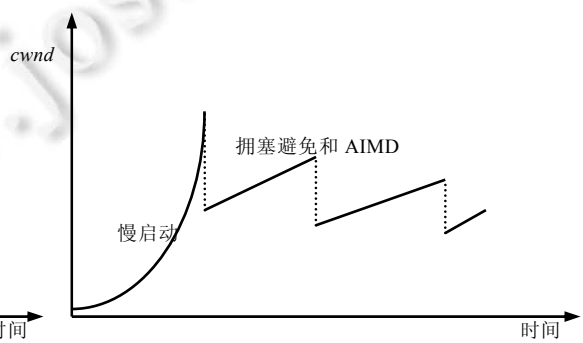


图 4 AIMD 机制的状态调整

RoQ 攻击也正是利用了 TCP 流量的自适应调整机制,使得本是正常的拥塞调整,因为突变过于频繁而显著降低 TCP 流量.

1.2.2 周期特征

根据 RoQ 攻击原理可知,周期性特征也是它的一个重要的攻击特征[3−5].当攻击发生时,如果只发送一次的

突发流,也能使关键链路拥塞并降低 TCP 流量,但是,流量会在瞬间得到恢复.如果想要达到持续的攻击效果(能够长时间地降低 TCP 流量,同时又要避免过高的攻击成本),攻击端就需要周期性地发送攻击流.

当然,这种周期性的攻击必然会引起相关网络流量发生周期性的变化.正是因为这种攻击特征非常独特,所以,周期性特征也是我们主要检测的一个攻击特征.

由此可见,尽管 RoQ 攻击具有很强的隐蔽性,但是,与原有的拒绝服务攻击相比,该攻击仍然会有自身显著的攻击特征,以及因其攻击而引起的相关流量特性的显著变化.

在文中,我们将逐步提取和分析这两个主要攻击特征,从而精确地检测 RoQ 攻击.

2 攻击特征提取和检测算法

2.1 攻击特征提取

我们利用小波的局部时-频分析能力,提取了攻击流量自身及其引起的相关流量的突变特征.在此基础上,我们进一步对突变的局部流量进行二次频谱分析,提取了流量的周期特征.检测算法将对这两种特征进行分析,并最终判断当前网络链路是否存在 RoQ 攻击.整个特征的提取流程如图 5 所示.

2.1.1 动态时间序列

为了达到增加检测速度和减少存储容量,同时还能保持数据之间的连续性的目的,我们引入滑动窗口对信号进行分析.本文使用的滑动窗口模型为 $W(S,\omega,\beta)$,其中 S 是基于滑动窗口的数据流,ω是滑动窗口的宽度(window width),β是滑动窗口的滑动步长(sliding step), ω和β均以一个采样点作为单位长度.

其中 $S(i,j),i<j,i=1,2,3,\ldots;j=1,2,3$ 表示第 i~j 个采样点所组成的数据流,$S(i)$表示第 i 个采样点.工作时,数据流进入大小为ω的当前滑动窗口内,假设当前滑动窗口第 1 个数据是$S(i)$,则当 ω 被填满时窗口内的数据流为 $S(i,i+\omega-1)$,滑动窗口向前滑动β个单位长度,当前滑动窗口被更新为$S(i+\beta,i+\beta+\omega-1)$,随后采样点流入新的滑动窗口直至其被填满后再一次滑动.

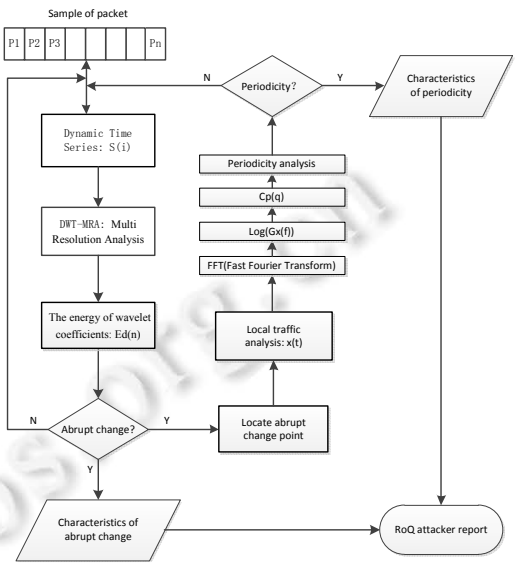


图 5 特征提取流程图

2.1.2 异常突变特征提取

首先,我们需要将采样序列转化为动态的时间序列 $\{S(i)\}$, $S(i)$代表从 i 到 i+ω–1 的维向量.接着,对 $S(i)$进行正交离散小波变换,利用小波多分辨率 MRA 良好的时-频局部化分析流量特性[16],我们就可以聚焦分析信号的局部时域和频域特征.

小波多分辨分析将信号分解成为一个近似的粗糙部分和一系列细节部分.这里的近似粗糙部分对应于信号的低频部分,这里的细节部分对应于信号的高频部分,这些高频部分是分层次的,是在不同的分辨率下逐步产生的.

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 $\psi_{j,k}(t)$ 将其表示为

$$f(t)=\sum_{k=-\infty}^{\infty}c_{J,k}\phi_{j,k}(t)+\sum_{-\infty}^{J}\sum_{k=-\infty}^{\infty}d_{j,k}\psi_{j,k}(t) \tag{3}$$

其中,第 1 部分是信号 $f(t)$的近似逼近部分(coarse approximation),第 2 部分是函数的细节部分(detail),$c_j(k)$和$d_j(k)$分别代表第j层尺度系数和小波系数,前者反映信号在宏观上的变化趋势,后者反映信号微观上的变化特征.在实际应用中,它们可以通过 Mallat 算法快速计算双尺度式(4)和式(5)得到:

$$c_{j-1,k}=\sum_{m=-\infty}^{\infty}h^{*}(m-2k)c_{j,m} \tag{4}$$

$$d_{j-1,k}=\sum_{m=-\infty}^{\infty}g^{*}(m-2k)c_{j,m} \tag{5}$$

根据帕塞伐尔定理[17],为了量化原始数据的波动情况,我们定义统计量 $E_d(n)$ 表示小波系数的能量以反映信号的变化细节.同时,我们使用滑动窗口 $W(d_j(k),\omega,\beta)$来聚集小波系数$\{d_j(k)\}$,见公式(6).

$$E_d(n)=\frac{1}{W}\sum_{k}\left|d_j(k)\right|^2,n=1,2,3\ldots \tag{6}$$

为了观察流量的变化幅度,我们引入公式(7),

$$\delta=\frac{\left|E_d(n+1)-E_d(n)\right|}{E_d(n)} \tag{7}$$

如果δ大于某一设定阈值 th,则认为流量出现异常突变,可能遭受了网络攻击;否则,认为网络流量是正常的.

当检测到异常突变时,需要通过公式(8)定位异常突变点.然后,取异常突变点局部的网络流量,用于下一阶段的流量特征分析和提取.值得注意的是,因为滑动窗口的引入,导致$E_d(n)$的时间尺度发生变化,若 i 代表$\{E_d(n)\}$序列中第 i 个点的横坐标,我们需要明确该点对应的真实时间 t_i 是多少.我们可以将产生该聚合点 t_i 的滑动窗口的中点作为 t_i 的大小.设检测系统的工作开始时间为 t_0,可以通过公式(8)得到第 i 点对应的实际时间.

$$t_i=t_0+\left(\beta\times(i-1)+\frac{\omega}{2}\right)\times\varepsilon \tag{8}$$

文中通过设置阈值,我们利用信号的高频部分提取流量突变特征,而低频部分用于下一步提取流量的周期性特征.

2.1.3 局部的周期特征提取

经过上述异常突变特征分析和提取,我们将对发生异常突变的局部流量进行二次频谱分析.

在信号分析时,时域信号 $x(t)$经过傅里叶变换变为频域函数 $x(f)$或功率谱密度函数 $G_x(f)$.当频谱图上呈现出复杂的周期结构时,如果再进行一次对数的功率谱密度函数傅立叶变换并取平方,则得到倒频谱函数 $C_f(\tau)$[18].其数学表达式为

$$C_f(\tau)=\left|F\left\{\log G_x(f)\right\}\right|^2 \tag{9}$$

由倒频谱的定义可以看出,尽管不同频段的信号共同作用在一起是复杂的,但其对数功率谱可将不同频段的功率谱表现为叠加性,其数学表达式为

$$\log G_y(f)=\log G_x(f)+\log G_n(f) \tag{10}$$

可见当取对数之后,我们将信号的相乘转化为了信号的相加,这将非常有利于识别各频率的组成分量,便于提取我们所关心的信息.因此倒频谱的实质是对功率谱取对数,然后进行频谱分析得到频谱中的周期成分[19].$\log G_x(f)$是源信号,具有明显的周期特征,若对式(10)再进一步作傅里叶变换,可得幅值倒频谱:

$$F\left\{\log G_y(f)\right\}=F\left\{\log G_x(f)\right\}+F\left\{\log G_n(f)\right\} \tag{11}$$

接下来,再作一次傅里叶变换,即做倒频谱,还将保持不同频段表现的叠加性.所以,倒频谱可以实现不同频段表现的分离.由于倒频变换的变换因子为

$$e^{i\omega\tau}=\cos\omega\tau-i\sin\omega\tau \tag{12}$$

τ值大小对应谐波频率的变化,在原信号中既含有低频率变动的信号又含有高频率变动的信号时,$C_f(\tau)$能有效地将这两部分加以区分和隔离.在实际应用中,低频表现的峰值与高频表现的峰值之间的距离 T 反映了这两类信号的相关距离,也就是原信号的时域周期,时域周期与频域周期的关系为 $T=1/\omega\tau$.

通过二次频谱分析,我们利用倒频谱分析的优势,不仅可以将不同频段的细节显现出来,而且可以较好地分离和提取其周期特征.

如果提取出的周期值T'和攻击周期 T 吻合,则认为在异常突变的局部流量中提取到了攻击的周期特征.根据攻击原理可知,攻击周期和拥塞控制机制的时间设置直接相关[3−5],不同的拥塞控制算法的时间设置是不同的,具体参数见文献[15].实验中,我们使用 TCP Reno 协议,其超时时间缺省为 1s,根据攻击原理,同时考虑到网络延迟等因素,我们给出判断是否与攻击周期吻合的周期值范围 $T\in[0.2s,1.3s]$.

2.2 检测算法

在文中,检测算法包括特征提取和攻击识别两个阶段.在特征提取阶段,我们设置布尔参数 E 和 P 分别表示两种特征提取状态,缺省为 False.当两个同时为 True 时,则表示完成特征提取,从而可以进行最后的攻击识别和判断.

我们对第 1 层小波系数$\{d_j(k)\}$通过滑动窗口 $W(d_j(k),\omega,\beta)$进行动态聚合得到统计量$\{E_d(n)\}$,当$\delta<th$ 时,表示当前网络流量状况正常,继续对新的动态序列进行重新分析,当$\delta\geqslant th$ 时,表示网络流量出现异常突变,则 E 设置为 True.此时,根据公式(9),定位异常突变点位置 ,并提取局部流量序列$\{S(i)\}$.

接下来,对新的动态序列进行二次频谱分析,使用倒频谱分离并提取周期特征 T',如果存在显著周期特征,则 P 为 True,否则 P 为 False.

在攻击识别时,如果此时 E 和 P 同时为 True,且 T'与 T 吻合,即可最终判定网络链路一定受到 RoQ 攻击.具体的检测算法见算法 1.

算法 1. RoQ 攻击检测算法.

```
1    Start to collect sampled data
2    Slide window to get sampled series S[i]
3    For every S[i]
4    [c,l] = wavedec(S[i],q,‘db5’)      //use wavelet basis ‘db1’ to decompose S[i] to level m
5         = wrcoef(‘h’,c,l,‘db5’,q)      //reconstruct the high frequency signals
6    Choose at finest scale 1,
7         =energyDensityFunction( )   //use energy density to get energy series
8    If ( th< δ )
9    Continue to next sampled series S[i]
10   Else
11   Abrupt change point exists, E=True.
12   Get local traffic around the abrupt change point, the signals is presented by x(t)
13   Calculate T′ and P by CepstrumAnalysis()
14   If (P==True && T′==T)
15   Attack confirmed
16   Else
17   Continue to next sampled series S[i]
```

3 实 验

首先,我们利用 NS-2 进行简单的网络模拟实验,通过模拟实验,验证了算法的有效性.然后,我们在真实网络环境中做了实验,在真实网络环境中对检测算法进行了验证和对比分析.

3.1 模拟实验实例

3.1.1 背景流量和攻击流量参数

在模拟实验过程中,首先对经过瓶颈链路的背景流量进行配置,如图 6 所示,分别引入了双向的 TCP 流量和 UDP 流量,通过观察发送窗口(cwnd)大小,保持瓶颈链路处于轻度拥塞状态.

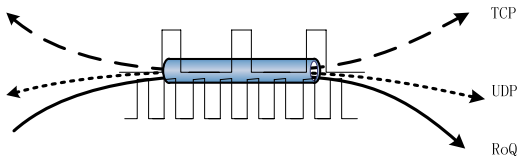


图 6 瓶颈链路流量

然后,我们引入了两种 RoQ 攻击流 C1 和 C3,其参数配置见表 1.其中,T_s 表示攻击开始时间,T_c 表示攻击持续时间, T_p 表示攻击周期,C 表示攻击成本,γ表示攻击效果,阈值 th 设为 0.1.

表 1 两种攻击脉冲参数

No.	T_s (ms)	T_c (ms)	T_p (ms)	C (Mbps)	γ(%)
C1	1 000	100	200	60	26
C3	1 000	200	400	60	19

3.1.2 检测过程

首先,我们选用 db5 正交小波函数将实验中采集到的时间序列信号分解到不同的频率通道上.由于对信号进行 6 层小波分解,则信号被分布在 7 个频带(如图 7 所示).

从图 7 可以看出,攻击开始后,流量发生明显突变.这里,我们把改变明显的高频信号带 d1 用于提取异常突变特征.通过公式(6),建立{$Ed(n)$}动态序列,并计算出δ=0.3,显然超出阈值 th,因此,认为网络流量出现异常突变,此时 E=True.

通过公式(8),定位异常突变点,并对局部流量进行频谱分析.我们选择两个低频信号带 a6 和 d6,重构后进行频谱分析.对比重构前后的 FFT 谱图(如图 8 所示),我们发现在重构信号 FFT 谱图中高频处变得非常干净,信号在高频处几乎没有,这时,我们只保留了低频处的信息.因此,通过小波分解之后再重构,可以消除噪声的影响,从而突出了攻击发生的频带.

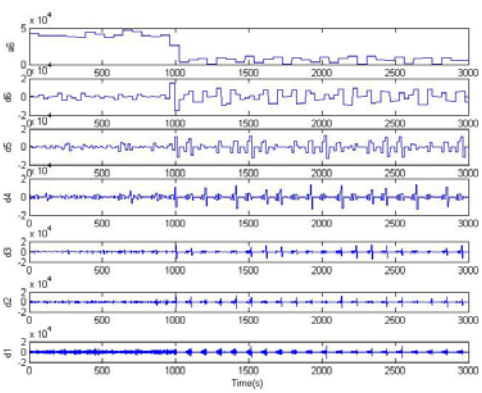


图 7 小波 6 层分解

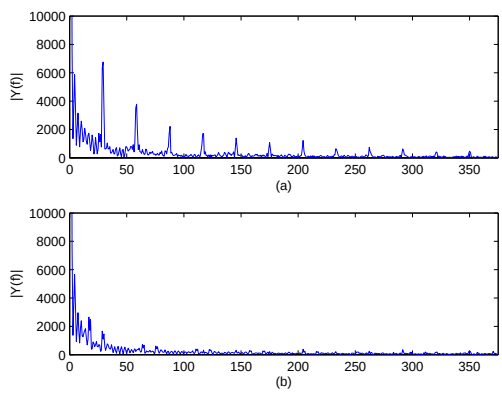


图 8 重构前后的 FFT 谱图

经过对重构信号解调,我们期望识别出与攻击频率有关的信号.尽管突出了攻击频谱,但仍然无法识别出攻击频率.为此,我们在此基础上对信号进行二次频谱分析,它的倒频谱如图 9 所示.

图 9 中,分别在 191ms 和 401ms 处有峰值出现,分别与攻击周期 0.2s 和 0.4s 近似相等,此时,P 设为 True,且周期特征值属于[0.2s,1.3s],由此可以推断,这是 RoQ 攻击引起的流量变化的周期特征.

通过实例验证,表明我们可以使用倒频谱分析复杂频谱图上的周期结构,并能够有效地检测出 RoQ 攻击.

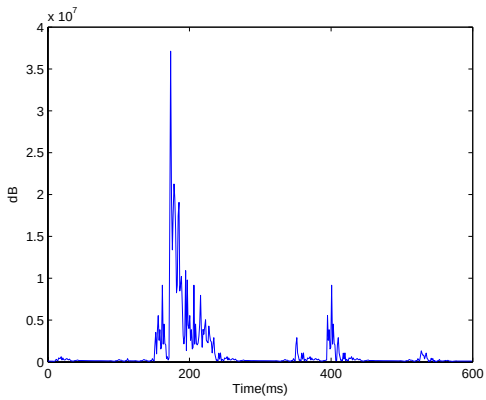


图 9 倒频谱

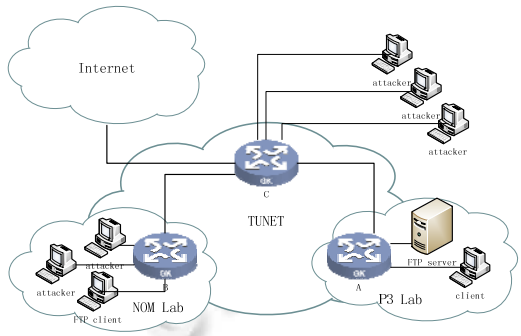


图 10 网络拓扑

3.2 真实网络实验

3.2.1 网络拓扑

为了检验我们的方法在实际网络中的检测效果,我们依托清华大学校园网(TUNET)中的实验室间子网(NOM 实验室和 P3 安全实验室)建立了网络实验平台[20],其网络拓扑如图 10 所示.其中,RoQ 攻击的目标是路由器 C 和 A 之间的瓶颈链路 CA,带宽为 100M,其余链路带宽均为 1G.攻击终端来自一些 NOM 实验室和校园网中的用户终端.

3.2.2 背景流量

实验室的网络间平时的流量有限,而且,与外网的访问流量也有限.因此,为了更好的观察和分析链路中 TCP 流量,我们使用了以下措施配置了背景流量:

1) 安装和配置了 FTP 服务器,提供文件上传和下载服务,用于增加网络中的 TCP 流量.

2) 有目的增多访问互联网提供的各种 Web 服务,比如高清视频点播、软件下载等.

3) 部署 LanTraffic V2.0 流量生成器,合理补充链路间的 TCP 或 UDP 流量.

另外,为了更接近真实的互联网主干流量,我们还引入了 100 次随机的突发流,每个突发流的持续时间在 0.5s~1.0s 之间.

3.2.3 攻击流量

在建立了合法的背景流量后,我们分别引入了 120 种 RoQ 攻击,它包括 10 类大小不同攻击成本,分别从 A_1~A_{12},B_1~B_{12} 到 J_1~J_{12},攻击成本步长为 10Mbps,分别从 40Mbps~140Mbps.每类攻击成本对应 12 种不同攻击周期,周期值大小从 0.2s~1.3s,步长为 0.1s.

3.2.4 数据来源

在实验中,我们配置了路由镜像端口,还使用 wireshark 1.10.2 工具软件全镜像的采集了通过瓶颈链路 CA 的所有数据包.实验时间从 2014 年 3 月 26 日上午 9:00~下午 5:00,在此期间,我们依次注入了 120 种 RoQ 攻击和 100 次突发流.

3.2.5 实验结果

检测算法的精确度通常用误报率(FPR)和漏报率(FNR)两个指标来衡量.我们的检测结果如图 11 所示,其中横坐标为攻击成本,每类攻击成本分别使用 12 种不同周期的 RoQ 攻击.纵坐标为 FPR 或 FNR 大小.

从图 11 可以看出,该检测算法的误报率很低,特别是,随着攻击成本增大,攻击效果越明显时,它的漏报率几乎为 0.当攻击成本较低时,会出现较高的漏报率,这与阈值 th 的大小有直接关系.其主要原因是攻击成本较低时,达不到一定的攻击效果,TCP 流量的减少不明显,往往视为正常的流量变化,这一现象我们在文献[21]中有具体的解释.

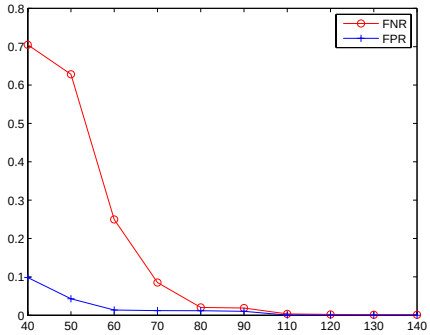


图 11 漏报率和误报率

3.2.6 算法对比

我们与 Hao[12]和 Xiang[13]提出的检测算法做了对比,3 种算法的检测效果如图 12 和图 13 所示.

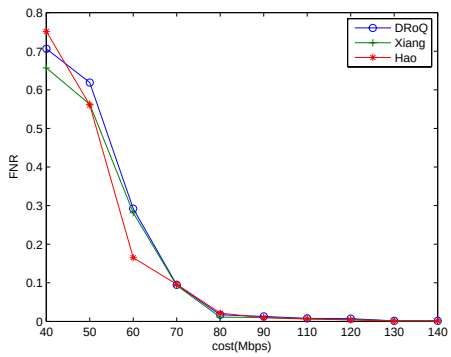


图 12 3 种检测算法的漏报率

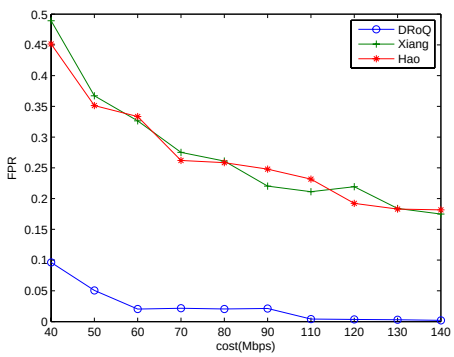


图 13 3 种检测算法的误报率

从以上两图可以看出,3 种算法检测的漏报率比较接近,攻击效果越明显,检测精度会越高;从误报情况看,本文的检测算法的误报率明显低于另外两种算法.主要原因应该是另外两种算法会对许多随机的突发流量产生误判,而我们的算法不仅仅只是根据异常突变特征做出判断的,因此误报率会更低.

4 总 结

本文提出了一种针对 RoQ 攻击的特征提取和检测方法,它利用了小波多分辨率技术良好的时-频分析能力,能够提取流量的异常突变特征.在此基础上,它还进一步对流量突变位置的局部流量进行了二次频谱分析,使用倒频谱提取了局部流量的周期特征.通过两阶段的特征提取和验证,实现了对 RoQ 攻击的有效检测.实验结果表明,该检测方法有很高的精确度,其漏报率和误报率都很低.

致谢 感谢实验室王子玉博士、李福亮博士和其他同学在写作过程中给予的无私支持和帮助.

References:

[1] Worldwide infrastructure security report. Volume VII. Arbor Networks, 2011. http://www.arbornetworks.com/report

[2] Sun CH, Liu B. Survey on new solutions against distributed denial of service attacks. Acta Electronica Sinica, 2009,37(7): 1562−1571 (in Chinese with English abstract).

[3] Kuzmanovic A, Knightly EW. Low-Rate TCP-Targeted denial of service attacks: The shrew vs. the mice and elephants. In: Proc. of the ACM SIGCOMM 2003. Karlsruhe: ACM Press, 2003. 75−86.

[4] Lou XP, Chang RKC. On a new class of pulsing denial-of-service attacks and the defense. In: Proc. of the Network and Distributed System Security Symp. San Diego: The Internet Society, 2005.

[5] Lou XP, Chan EWW, Chang RKC. Vanguard: A new detection scheme for a class of TCP-targeted denial-of-service attacks. In: Proc. of the Network Operations and Management Symp. (NOMS 2006). 2006. 507−518.

[6] Guirguis M, Bestavros A, Matta I. Reduction of quality (RoQ) attacks on Internet end systems. In: Proc. of the 24th IEEE INFOCOM. Miami: IEEE, 2005. 1362−1372.
[7] Guirguis M, Tharp J, Bestavros A, Matta I. Assessment of vulnerability of content adaptation mechanisms to RoQ attacks. In: Proc. of the 8th Int'l Conf. on Networks ICN 2009, 2009. 445−450. [doi: 10.1109/ICN.2009.40]
[8] Bhuyan MH, Kashyap HJ, Bhattacharyya DK, Kalita JK. Detecting distributed denial of service attacks: Methods, tools and future directions. The Computer Journal, 2014,57(4):537−556.
[9] Sarat S, Terzis A. On the effect of router buffer sizes on low-ratedenial of service attacks. In: Proc. of the 14th Int'l Conf. on Computer Communications and Networks (ICCCN 2005). San Diego: IEEE Press, 2005. 281−286.
[10] Sun HB, Lui JCS, Yau DKY. Defending against low-rate TCP attacks: Dynamic detection and protection. In: Proc. of the 12th IEEE Int'l Conf. on Network Protocols (ICNP 2004). Berlin: IEEE COMPUTERSOC, 2004. 196−205.
[11] Chen Y, Hwang K. Spectral analysis of TCP flows for defense against reduction-of-quality attacks. In: Proc. of the IEEE Int'l Conf. on Communications 2007. 2007. 24−28.
[12] Chen H, Chen Y, Summerville DH, Su Z. An optimized design of reconfigurable PSD accelerator for online shrew DDoS attacks detection. In: Proc. of the INFOCOM, 2013. 1780−1787.
[13] Xiang Y, Li K, Zhou W. Low-Rate DDoS attacks detection and traceback by using new information metrics. IEEE Trans. on Information Forensics and Security, 2011,6(2):426−438.
[14] Jain R, Ramakrishnan KK. Congestion avoidance in computer networks with a connectionless network layer: Concepts, goals and methodology. In: Proc. of the Computer Networking Symp., 1988. 134−143.
[15] Paxson V, Allman M. RFC 2988: Computing TCP's retransmission timer. Internet RFCs, 2000. http://rfc.net/rfc2988.html
[16] Chaovalit P, Gangopadhyay A, Karabatis G, Chen Z. Discrete wavelet transform-based time series analysis and mining. ACM Computing Surveys (CSUR). 2011,43(2):37.
[17] Burrus CS, Gopinath RA, Guo H. Introduction to wavelets and wavelet transforms: A primer. Upper Saddle River: Prentice. Hall, 1998.
[18] Duda RO, Hart PB. Pattern Classification and Scene Analysis. New York: Wiley, 1973.
[19] Zhang WQ, He L, Deng Y, Liu J, Johnson MT. Time-Frequency cepstral features and heteroscedastic linear discriminant analysisfor language recognition. IEEE Trans. on Audio, Speech and Language Processing, 2011,19(2).
[20] Yang JH, Wu JP, An CQ. Internet Measurement Theory and Application. Beijing: The People's Posts and Telecommunications Press, 2009 (in Chinese).
[21] Wen K, Yang JH, Cheng FJ, Li CX, Wang ZY, Yin H. Two-Stage detection algorithm for RoQ attack based on localized periodicity analysis of traffic anomaly. In: Proc. of the ICCCN2014. 2014.

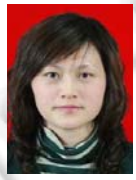
附中文参考文献:

[2] 孙长华,刘斌.分布式拒绝服务攻击研究新进展综述.电子学报,2009,37(7):1562−1571.
[20] 杨家海,吴建平,安常青,互联网络测量理论与应用.北京:人民邮电出版社,2009.

文坤(1976－),男,河南方城人,博士生,主要研究领域为网络测量,网络安全,异常流量检测.

杨家海(1966－),男,博士,教授,博士生导师,主要研究领域为计算机网络,网络管理与测量,网络安全,云计算及安全.

李晨曦(1991－),男,博士生,主要研究领域为网络测量,网络安全,异常流量检测.



程凤娟(1976－),女,副教授,主要研究领域为网络安全,移动自组网.

尹辉(1977－),男,副教授,主要研究领域为智能信息处理,计算机应用.