基于包络延拓和本征波匹配的时变DoS攻击频谱检测

唐赞玉[1] 刘 宏[2]
(吉首大学信息科学与工程学院 吉首 416000)[1] (湖南师范大学数学与计算机学院 长沙 410081)[2]

摘 要 DoS攻击信号具有非平稳时变特性,湮没在色噪声背景的复杂网络环境中,对之难以有效检测。传统方法中采用基于非平稳时变信号处理的Hough变换单谱脉冲响应检测算法,由于二次型时频分布的边缘效应会引起较大包络衰减,检测性能不好。因此提出一种基于包络延拓和本征波匹配的时变DoS攻击信号频谱检测算法来对DoS攻击检测信号进行双曲调频分解,构建信号数学演化模型,得到信号包络和本征波特征提取结果。采用双线性Hough变换法分析频谱特征畸变,进行瞬时频率估计,得到信号的单谱脉冲响幅频响应,在包络时频特征空间优化搜索路径实现包络延拓,基于最小均方误差准则设计本征波匹配滤波器,控制DoS频谱偏移,实现信号频谱检测。仿真结果表明,本算法能在强色噪声背景干扰下提高检测性能,检测概率高于传统算法,且能准确估计参量信息,提高对DoS攻击信号的主动防御能力。

关键词 包络延拓,本征波匹配,DoS攻击,信号检测,网络安全
中图法分类号 TP393 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2015.4.025

Spectrum Detection of Time-varying DoS Attack Signal Based on Envelope Extension and Intrinsic Wave Matching

TANG Zan-yu[1] LIU Hong[2]
(College of Information Science and Engineering, Jishou University, Jishou 416000, China)[1]
(School of Mathematics and Computer, Hunan Normal University, Changsha 410081, China)[2]

Abstract DoS attacks signal has non-stationary and time-varying property. It is lost in the complex network environment with color noise background, and it is difficult to detect. Traditional methods use Hough transform impulse response method to detect the non-stationary signal. Due to the edge effect of frequency distribution, detection performance is not good. A new spectrum detection method of DoS attack signal was proposed based on envelope extension and intrinsic wave matching filtering. The DoS attack signal is processed with hyperbolic frequency modulated signal decomposition, and mathematical evolution model is constructed. Signal envelope intrinsic wave features are extracted. The bilinear Hough transform method is used to analyze the spectrum distortion, instantaneous frequency estimation is obtained, and single pulse response amplitude frequency response is calculated. In time frequency feature space, the envelope extension path search is optimized. Intrinsic wave matching filter is designed based on minimum mean square error criteria. DoS frequency shift is controlled, and the spectrum detection is obtained. Simulation results show that the algorithm can improve the detection performance, and the interference of strong colored noise can be suppressed. The detection probability is higher than traditional methods. It can accurately estimate the parameters, and the active defense ability of network security is improved.

Keywords Envelope extension, Intrinsic wave matching, DoS attack, Signal detection, Network security

1 引言

随着信息安全中对抗与反对抗技术的权衡和发展,网络攻击形式呈多样化,网络攻击的强度和隐蔽性特征逐渐增强,对网络攻击信号的有效准确检测,是保证网络安全、提高网络的生存能力和抵御风险能力的关键[1,2]。网络攻击信号中,以拒绝服务(Denial of Service, DoS)攻击最为常见,通过DoS攻击,使得计算机或网络无法正常运行和提供服务,常见的如网络宽带攻击和网络连通性攻击。DoS攻击信号隐蔽性好,变异性强,在网络周期性控制机制出现的服务间隙发送大量攻击数据,会降低用户的使用性能,导致系统用户崩溃,达到降质服务和拒绝服务的目的。根据相关研究发现,DoS攻击信号本身具有时变性和非平稳性特征,对之难以形成有效的频谱检测算法,因此研究DoS攻击信号的频谱检测算法,对保证网络安全具有重要意义。

DoS攻击信号的出现相对较晚,其变异更新速度较快,对

到稿日期:2014-07-10 返修日期:2014-12-12 本文受湖南省科技计划项目(2012GK3127)资助。
唐赞玉(1978-),女,硕士,讲师,主要研究方向为计算机网络、分布计算与自动控制等;刘 宏(1963-),男,教授,主要研究方向为分布计算与自动控制、人工智能。

DoS攻击信号的检测技术的研究尚处于起步阶段，相关的频谱检测和参量估计技术发展还不成熟。随着现代信号处理技术的快速发展，借鉴水声、雷达、电子信息等领域的现代信号检测技术，将其引入到DoS网络攻击信号检测中，推动了网络攻击信号检测的发展。当前对时变DoS攻击信号检测算法主要分为基于特征提取和包络匹配的信号检测算法和基于线性滤波和异常模式专家系统决策的信号检测算法。其中，文献[3]提出一种基于Rossle混沌平均互信息特征挖掘的攻击检测算法，通过挖掘的互信息这种非线性特征解，实现对具有非线性随机特性的网络攻击信号有效检测，算法在求解具有高斯线性特征分DoS攻击信号频谱时具有有效性，但是当DoS攻击信号是非平稳态时，算法无法有效抑制高斯色噪声干扰，检测性能受限；文献[4]提出一种采用双门限能量检测的协作频谱感知的DoS攻击信号频谱感知技术，即采用能量检测方法，利用瑞利衰落信道进行网络的拓扑控制，其无法满足相应特定要求的拓扑结构，从而影响了检测概率。文献[5]采用高阶谱分析的单谱脉冲响应信号畸变检测方法来检测DoS攻击信号特征，但对高阶谱的求解复杂，算法实时性不好。文献[6]提出的DoS网络攻击信号频谱检测算法建立在基于前馈参数估计和信号动态补偿的基础上，通过设计鉴频器，采用小段接收和信号编码的方法实现对突发衰落信号的跟踪和补偿，从而实现DoS攻击信号的频谱检测。在此基础上，文献[7]对算法进行改进，提出一种任意大频率的微弱信号随机共振检测算法，对实现大频率信号检测具有适用性，但是对多频网络攻击信号检测性能有限。文献[8]采用人工神经网络及粒子滤波的方法实现对DoS攻击信号的频谱分离和聚类，在信号欠定盲分离过程中，检测方法受弱信号幅度和临界阈值的约束。文献[9]受声纳系统使用双曲调频和二次调频信号进行回波定位信号检测的启发，提出一种基于Hough变换的单谱脉冲响应检测，其采用非平稳时变信号处理技术进行DoS攻击信号的频谱分离感知，达到提高信号检测概率的目的。

针对上述问题，本文在综合上述文献研究成果的基础上，特别是针对文献[10]中的算法进行改进，提出一种基于包络延拓和本征波匹配的时变DoS攻击信号频谱检测算法。首先进行非平稳时变DoS攻击数学模型和信号分析，然后进行信号预处理和特征提取，设计本征波匹配滤波器并实现算法改进。仿真实验验证了算法的优越性，并得出可靠性结论，研究成果在网络安全和信号检测等领域具有应用前景。

2 非平稳时变DoS攻击数学模型和信号分析

2.1 非平稳时变DoS攻击数学模型

本文研究和设计基于包络延拓和本征波匹配的时变DoS攻击信号频谱检测算法，需要首先给出在色噪声背景干扰下的DoS攻击信号数学模型，并进行网络安全属性分析，给出相关的问题描述，为下一步构建DoS攻击信号的频谱检测模型奠定数学模型基础[11,12]。

构建DoS攻击系统是一个三维连续的典型自治系统表达[13]，网络入侵特征目标函数表示为：

$$\theta_1(k+1)=\theta_1(k)-\mu\mathrm{Re}[y(k)\varphi^*(k)] \tag{1}$$

式中，$\theta_1(k)$表示初始状态向量，$\theta_1(k+1)$表示二次迭代网络状态向量，μ为DoS攻击系统相空间非线性序列重构收缩系数。

构建DoS网络攻击系统设计中，假设微弱攻击信号幅度为A，对输入信号幅度调整系数为：

$$x(t)=\sum_{i=0}^{p}a(\theta_i)s_i(t)+n(t) \tag{2}$$

设有M个全方向性攻击的DoS信号，一个期望信号A_c和P个干扰信号以$\theta_0,\theta_1,\cdots,\theta_P$的角度输入到DoS攻击检测系统中，对DoS攻击检测系统进行双曲调频分解，得到接收到的信号模型为：

$$\tilde{u}_{e|v,k}=\tilde{u}_{e,k}+\tilde{\Sigma}_{ve,k}^{T}\tilde{\Sigma}_{vv,k}^{-1}(v_k-\tilde{u}_{v,k}) \tag{3}$$

如果将ω_k按照v_k和e_k的组成原则进行分解，则得到多项式相位信号特征分布为：

$$\begin{cases}v_k\sim t_{v_k}(\tilde{u}_{v,k},\tilde{\Sigma}_{vv,k})\\ e_k\sim t_{v_k}(\tilde{u}_{e,k},\tilde{\Sigma}_{ee,k})\end{cases} \tag{4}$$

通常情况下，DoS攻击信号是时变非平稳的，因此采用一个多项式来定义多项式相位情况下的瞬时频率。令q为多项式的阶数，满足的条件是：$q\geqslant p$且为偶数，得到DoS攻击信号的时变瞬时频率估计可以表示为：

$$\hat{f}_{i_q}(t,\tau)=\frac{1}{2\pi\tau}\sum_{k=-q/2}^{q/2}b_k\phi(t+c_k\tau) \tag{5}$$

式中，DoS攻击相位$\phi(t)$为均匀采样的；若c_k为分数，则相位$\phi(t)$为非均匀采样。所以采样间隔的相位差的作用是不同的，其中τ为时间采样步长(相当于Δt)，b_k是控制两个不同相位攻击的DoS信号的相位值差的权系数($b_0=0$)，c_k为相位采样间隔参数，由此得到DoS攻击的数学演化模型的状态转移方程表示为：

$$\begin{aligned}x(n)&=s(n)+v(n)\\&=\omega_{k-1}^{(i)}\frac{p(y_k|X_k^{(i)},Y_{k-1})p(x_k^{(i)}|X_{k-1}^{(i)},Y_{k-1})}{q(x_k^{(i)}|.)}\end{aligned} \tag{6}$$

式中，$s(n)$表示攻击信号，$v(n)$表示色噪声分量，φ_i表示信号的非平稳态瞬时频率估计值。以上述数学模型和状态转移方程为依据，构建DoS攻击信号的数学模型，为后续的DoS攻击频谱检测提供信源基础。

2.2 信号包络和本征波特征提取信号分析预处理

通过演化模型，在模型构建中需要考虑色噪声干扰的影响，相关噪声分布$p(e_k|v_k)$的方差和均值服从如下分布：

$$p(e_k|v_k)\sim t_{(\tilde{v}_k+d_e)}(\tilde{u}_{e|v,k},\tilde{\Sigma}_{e|vk}) \tag{7}$$

此时，本文采用提取信号包络特征和本征波匹配的方法实现DoS攻击信号的频谱检测，对原始攻击信号模型采用双线性Hough变换分析进行信号包络特征和本征波特征提取，采用包络延拓扩展方法求解DoS信号的非高斯函数极限幅频特性，抵消DoS攻击信号的畸变效应，得到瞬时频率的估计为：

$$\hat{f}_i(n)=\frac{1}{2\pi}\sum_{i=0}^{p}ia_in^{i-1} \tag{8}$$

采用双线性Hough变换法分析频谱特征，得到离散时间DoS攻击序列的频谱畸变部分估计为：

$$\frac{1}{2\pi m}\sum_{k=-q/2}^{q/2}b_k\phi(n+c_km)=\hat{f}_{i_q}(n) \tag{9}$$

由于双线性核可以将线性相位变化规律的信号转化为正弦波，因此DoS攻击信号通过双线性Hough变换而得到的信号形式为：

$$s(t)=\underbrace{\sum_{k=1}^{N}p_k\sin(\omega_kn+\Phi_k)}_{u(n)}+\zeta(n) \tag{10}$$

进而得到双线性 Hough 变换单谱脉冲响幅频响应的传输函数为：

$$H(z)=\frac{1+\sin\theta_2}{2}\frac{1+2\sin\theta_1 z^{-1}+z^{-2}}{1+\sin\theta_1(1+\sin\theta_2)z^{-1}+\sin\theta_2 z^{-2}} \quad (11)$$

式中，多项式的零点在单位圆内，幅频响应具有分段常数的特点。然而通过以上分析可知，通过双线性 Hough 变换方法进行包络特征和本征波特征提取，实现对信号的预处理，进行 DoS 攻击信号的频谱检测。由于二次型时频分布的边缘效应，在信号的起始段和终止段会引起较大包络衰减，检测性能不好，故本文将以此算法为基础，引入包络延拓和本征波匹配滤波的思想，实现算法改进。

3 算法改进描述与 DoS 攻击频谱检测实现

3.1 包络延拓算法的提出

本文针对传统方法中由于包络衰减导致频谱检测性能不好的问题，根据 DoS 攻击信号包络特征的双向延拓特性，以 DoS 攻击信号端点或者极值点为对称中心进行延拓，并采用本征滤波匹配方法进行信号抗色噪声干扰处理，实现算法改进。改进算法思想描述如下：采用 Hough 变换单谱脉冲响应检测方法，对时变 DoS 攻击信号的特征和本征波特参数进行特征分解，表达为：

$$\tilde{\Sigma}_{e|v,k}=h_{e|v,k}(\tilde{\Sigma}_{ee,k}-\tilde{\Sigma}_{ve,k}^{\mathrm{T}}\tilde{\Sigma}_{vv,k}^{-1}\tilde{\Sigma}_{ve,k}) \quad (12)$$

$$h_{e|v,k}=\frac{1}{(\tilde{v}_k+d_v)}\times[\tilde{v}_k+(v_k-\tilde{u}_{v,k})^{\mathrm{T}}\tilde{\Sigma}_{vv,k}^{-1}(v_k-\tilde{u}_{v,k})] \quad (13)$$

式中，$\tilde{v}_k=v_k-d+1$ 表示在色噪声背景中，核函数 $k_z^g(t,\tau)$ 取作多项式核时，取信号值的非整数幂之间的乘积，得到单个 DoS 攻击信号在势阱中的正态阈值，DoS 攻击信号检测系统的包络延拓状态方程的线性表达式为：

$$g(x,y)=\begin{cases}0, & |x-y|\leqslant\Delta\\ c(|x-y|-\Delta), & |x-y|>\Delta\end{cases} \quad (14)$$

式中，Δ 的最优选择为两个连续点瞬时频率变化的最大期望值，即对于较小的瞬时频率变化（$|x-y|\leqslant\Delta$）代价函数为 0，以此构成多个窄带信号。包络延拓搜索路径示意图如图 1 所示。

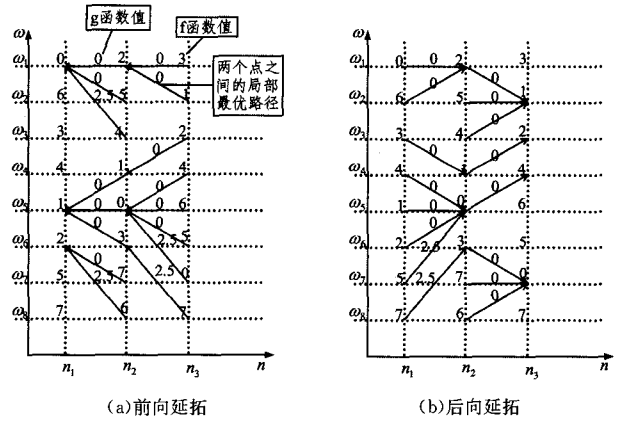

图 1 DoS 攻击信号的包络延拓向前、向后搜索路径

图 1 表示以两个时间点之间的局部最优路径为基准，对 DoS 攻击信号的包络时频特征空间进行向前、向后沿着由确定的局部最优路径得到的包络延拓搜索方案，需要计算整个路径的代价函数值并记录路径每个时间点对应的包络解，表征为 t 分布的位置和尺度参数分别为：

$$\tilde{u}_k=u_k,\tilde{\Sigma}_k=\frac{1+V_{11,k}}{(v_k-d+1)V_{11,k}}\Lambda_k \quad (15)$$

3.2 本征波匹配滤波设计频谱检测算法实现

以上述搜索得出的最优包络延拓特征为基础，对时变 DoS 攻击信号进行频谱检测。本文提出采用本征波匹配滤波方法进行频谱检测，采用最小均方误差准则设计一个能去除多个已知干扰频率成份的本征波匹配滤波器，考虑一种简单的本征波匹配滤波器传输函数：

$$H(z)=\frac{1+az^{-1}+z^{-2}}{1+arz^{-1}+r^2z^{-2}},0<r<1 \quad (16)$$

结合式(11)可见，输入的非平稳时变 DoS 网络攻击信号 $n(k)$ 的实部 $n_r(k)$ 和虚部 $n_i(k)$ 分别为独立的色噪声，通过调解不同的陷波器频率参数 a 和带宽参数 r，可使网络攻击信号的本征波极点稳定在收敛区域内。这样可以得到本征波匹配滤波频率：

$$\omega_0=\arccos(-a/2) \quad (17)$$

当 a 变化时，本征波陷波频率也随之变化；当 $r\to1$ 时，本征波匹配滤波的带宽减小，有利于检测信号频谱，得到频谱检测概率表示为：

$$P_d=1-\prod_{i=1}^{N}[(1-P_{di})(1-P_{ei})+P_{di}P_{ei}] \quad (18)$$

根据频谱检测信道衰落因子，设计色噪声背景下的 DoS 攻击信道频谱融合准则，通过本征波匹配滤波控制调整 DoS 频谱偏移，使检测的攻击信号特征与 f 的组成成分最佳匹配。信号的频谱分解为：

$$\gamma_i=\frac{1}{N-1}*\frac{\sum_{i=1}^{N}(SNR_i)-SNR_i}{\sum_{i=1}^{N}SNR_i} \quad (19)$$

其中，SNR_i 表示信噪比参量，通过估计对传输信噪比分配可信度，可得到信度之和大于 1/2 的攻击信号频谱检测虚警概率为：

$$P_f=\sum_{\sum c_i=1}^{N}\sum_{\gamma_i c_i\geqslant1/2}\prod_{i=1}^{N}(P_{f,i})^{c_i}(1-P_{f,i})^{c_i} \quad (20)$$

通过 k 次分解后，对干扰进行有效过滤，提高非平稳时变 DoS 攻击频谱检测概率。基于本征波匹配滤波的频谱检测原理实现如图 2 所示。

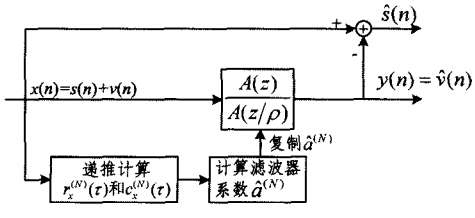

图 2 包络延拓本征波匹配滤波频谱检测原理

4 仿真实验和结果分析

为了测试和验证本文提出的基于包络延拓和本征波匹配滤波算法设计的时变 DoS 攻击频谱检测性能，基于 Matlab 仿真实验平台进行仿真实验。实验仿真环境为：IntelCore3-530 1G 内存，操作系统为 Windows 7，网络 DoS 攻击信号采集于 MIT 林肯实验室 KDD Cup2013 网络病毒数据库，病毒数据库中有大量的 DoS 攻击样本，病毒样本序列的种类和攻击强度具有普适性特征，因此在实验中具有参考意义。网络信息样本数选择为 1024，访问次数为 10332，攻击信号样本数为

998，攻击次数取 135。

假设背景噪声为色噪声模型，网络复杂背景中的噪声频谱 $f=20\text{kHz}$，方位参数 θ 在 $(0,2\pi]$ 之间均匀分布，攻击信号频谱检测系统的采样频率选择和网络环境噪声采集的频率相同，为 196.608kHz。实验信噪比为 -10dB，即信号的幅度 A 约为 0.178。自适应初始步长 μ_0 均选为 $\mu_0=0.001$；本征波匹配滤波器带宽参数选为 $\theta_2=0.46\pi$。在本征波匹配滤波过程中，需要输入的参考信号为两个 LFM 信号，信号由 LFM 信号和正弦调频信号组成，两个信号的瞬时频率没有交点，样本数为 10000，单个节点检测覆盖半径 $c=2.5$，滤波器频谱信道衰落增益为 $\Delta_1=3$，恒虚警概率检测条件下，采用包络延拓本征波匹配融合准则。在仿真实验中，攻击信号频谱检测方法在不同采样点数下进行，其中判决门限 $G_T=20\sigma^2$，恒虚警概率 $p_f=0.1$。得到时变 DoS 攻击信号加噪声的波形如图 3 所示，由图 3 可见，DoS 攻击信号几乎完全淹没在噪声背景中，无法有效拦截和识别攻击 DoS 信号。

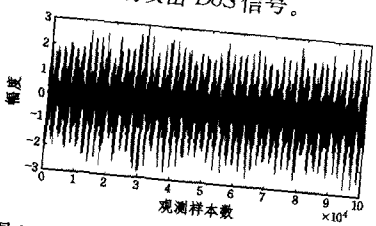

图 3　原始 DoS 信号淹没在噪声中的波形

采用本文算法，首先进行非平稳时变 DoS 攻击数学模型构建，提取信号包络和本征波特征，得到包络特征的幅度响应和延拓结果，如图 4 所示。

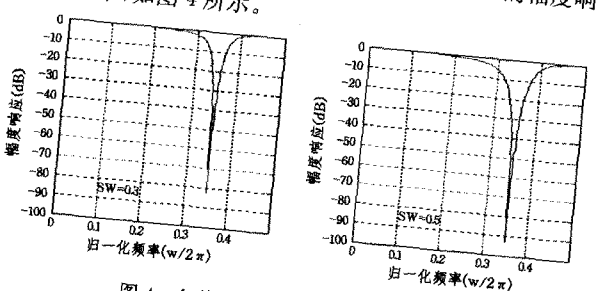

图 4　包络特征幅度响应及延拓结果

以上述信号模型为依据，然后采用本征波匹配滤波算法进行频谱检测实验，得到采用本文方法的非平稳时变攻击信号的频谱检测结果，如图 5 所示。分析图 5 结果可知，本文算法可以很好地抑制色噪声的影响，并估计信号的频率，在起始频率处得到了小于 200Hz 的估计值，而在截止频率处却获得大于 40Hz 的估计值，起始频率和截止频率处只产生了相较小的误差，同时说明改进算法能有效地估计攻击信号的频率等参数，为及早预测和拦截攻击信号提供数据依据。

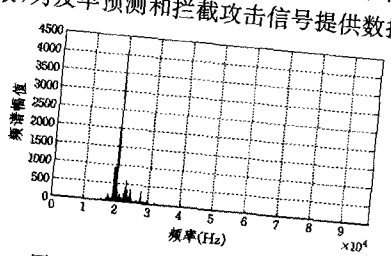

图 5　本文方法进行频谱检测结果

为了对比算法性能，在同等条件下，采用文献[9]中传统 Hough 变换单谱脉冲响应检测算法进行频谱检测，得到的结果如图 6 所示。由图分析可见，采用传统方法时，攻击信号的频谱被完全淹没在噪声中，检测频谱出现多个峰值，无法有效区分，无法实现有效的 DoS 攻击检测，这是因为传统 Hough 变换单谱脉冲响应检测算法二次型时频分布的边缘效应，在信号的起始段和终止段会引起较大包络衰减，降低了频谱检测性能，而本文算法能弥补该缺陷。

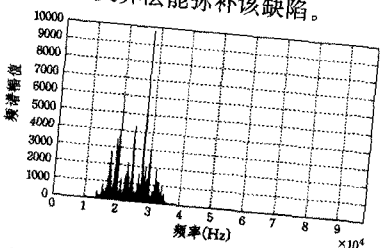

图 6　传统 Hough 变换单谱脉冲响应方法进行频谱检测结果

最后，为获得时变 DoS 攻击信号的频谱检测性能曲线，将本文算法与多种算法进行检测概率的性能对比，采用 Monte-Carlo 方法进行检测性能测试，仿真次数为 10000，得到在背景干扰信噪比为 -11dB、-13.5dB、-15.5dB、-8dB 和 20dB 下的频谱检测性能曲线，如图 7 所示。

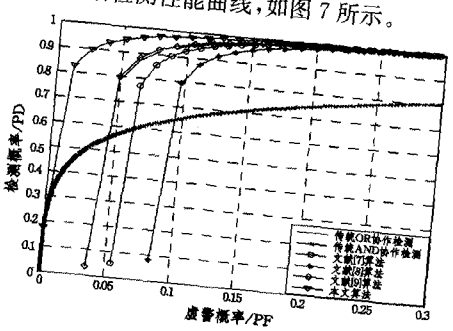

图 7　检测性能曲线

从检测性能曲线分析可知，采用本文算法在低联合虚警概率下的检测概率要优于其它算法，其它算法尤其在低虚警概率 0.1 以下，检测可信度的过程受到起始段和终止段包络衰减的影响严重，而本文算法可以避免此情况发生。本文算法在 SNR 为 -11dB 时正确检测概率在 80%以上，而其它方法几乎不能形成有效的检测输出，从而证明了本文算法在通过包络延拓和本征波匹配滤波后，在抗色噪声干扰和提高检测性能方面具有优越性。

结束语　DoS 攻击信号本身具有时变性和非平稳性特征，对之难以形成有效的检测算法，因此研究 DoS 攻击信号的检测算法，对保证网络安全具有重要意义。本文提出一种基于包络延拓和本征波匹配的时变 DoS 攻击信号频谱检测算法，分析非平稳时变 DoS 攻击数学模型和信号模型构建方法，通过提取信号包络和本征波特征实现频谱检测预处理，优化 DoS 攻击信号的包络延拓的向前、向后搜索路径，设计本征波匹配滤波实现频谱准确检测。研究结果表明，算法对二次型时频分布的边缘效应引起的包络衰减和色噪声干扰都具有较好的抑制性能，频谱检测概率得到了提高，展示了其在网络安全防御和信号检测等领域的优越性能。

参 考 文 献

[1]　刘衍珩，付枫，朱建启，等. 基于活跃熵的 DoS 攻击检测模型[J]. 吉林大学学报：工学版，2011，41(4)：1059-1063

[2]　江先亮，金光，杨建刚，等. 面向自治域的 DoS 攻击流抑制模型[J]. 通信学报，2013，34(9)：132-141

[3]　王进，阳小龙，隆克平. 基于大偏差的……

DDoS检测机制及性能分析[J]. 软件学报,2012,23(5):1272-1280
[4] 张永铮,肖军,云晓春,等. DDoS攻击检测和控制[J]. 软件学报,2012,23(8):2258-2072
[5] 王睿. 一种基于回溯的Web上应用层DDOS检测防范机制[J]. 计算机科学,2013,40(11A):175-177
[6] 夏秦,王志文,卢柯. 入侵检测系统利用信息熵检测网络攻击的方法[J]. 西安交通大学学报,2013,47(2):14-19
[7] 周华,周海军,马建锋. 基于博弈论的入侵容忍系统安全性分析模型[J]. 电子与信息学报,2013,35(8):1933-1939
[8] Bimal K M,Gholam M A. Differential epidemic model of virus and worms in computer network [J]. International Journal of Network Security,2012,14(3):149-155
[9] Zhu Q Y, Yang X F, Yang L X, et al. Optimal control of computer virus under a delayed model [J]. Applied Mathematics and Computation,2012,218(23):11613-11619
[10] 张辉. 自体集网络入侵检测中的高效寻优算法仿真[J]. 计算机仿真,2013,30(8):297-300
[11] 樊爱宛,时合生. 基于特征选择和SVM参数同步优化的网络入侵检测[J]. 北京交通大学学报,2013,37(5):58-61
[12] 饶雨泰,杨凡. 网络入侵搅动下的网络失稳控制方法研究[J]. 科技通报,2014,30(1):185-188
[13] 罗柏文,沈彩耀,于宏毅. 采用余弦调制滤波器组的多径衰落信号子带合成[J]. 信号处理,2013,29(5):537-543

(上接第118页)

有效地约减训练集,可以在保证分类精度基本不变的前提下大大减少训练时间;在Poker-hand数据集上约减后的SVM的分类精度反而提高了,这是由于减样的过程把部分噪声数据也除掉了。

第三,加权K最近邻方法对训练集的约减效果比普通K最近邻方法效果更好。对比WKNN-SVM和KNN-SVM可知,WKNN-SVM比KNN-SVM的分类精度更好,这是一种更有效的减样方法。

5.2 Android数据集仿真实验

使用上文提到的Android数据集构建方法,采集了不同时间段中3个规模不同的数据集,如表4所列。

表4 Android数据集描述

数据集	训练集样本数		测试集样本数
	正类	负类	
Android-1	587	93	170
Android-2	2600	684	821
Android-3	4886	924	1450

本实验将本文提出的FWKN-SVM方法和传统的SVM方法(LIBSVM)进行对比,实验选取相同的参数。准确率的对比结果如图2所示,训练时间的对比结果如图3所示。

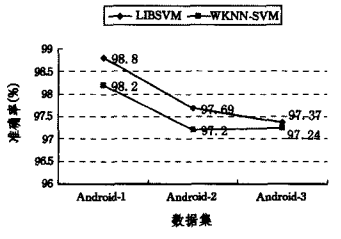

图2 准确率对比图

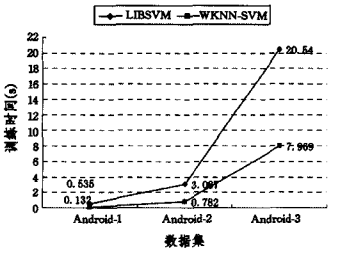

图3 训练时间对比图

上述结果表明:

第一,本文采用的数据集采集和构建方法是有效的,是适用于Android手机SVM入侵检测研究的。因为使用两种方法得到的准确率都比较高,在实时应用中的效果比较好。

第二,WKNN-SVM比SVM更适合于Android入侵检测,在保证准确率和误报率基本不变的前提下大大减少了训练时间,而且,训练样本集越大,训练时间减少得越明显,约减的效果越明显。

结束语 本文提出了一种基于加权K最近邻支持向量机的Android手机入侵检测方法。该方法通过分析恶意软件对系统造成的影响定义了Android入侵检测系统行为,并进行数据集构建;考虑到各个特征值对分类结果的不同影响,使用了基于类内类间距离的特征加权方法求解样本的K最近邻,进而得出训练样本边界向量集;针对得到的边界向量集合进行支持向量机训练,可以在保证分类精度不变的前提下提高训练速度并减少内存占用,适用于实际应用中的大规模样本分类问题。因此,本文的方法为解决Android手机异常入侵检测提供了一种思路。

参 考 文 献

[1] Vapnik V. The nature of statistical learning theory [M]. Springer,2000
[2] 钱权,耿焕同,王煦法. 基于SVM的入侵检测系统[J]. 计算机工程,2006,32(9):136-138
[3] 莫宇祥,俞建銮,王磊,等. 基于角色的Android手机平台木马检测系统[J]. 现代计算机:上半月版,2012(12):51-55
[4] 罗瑜,易文德,王丹琛,等. 大规模数据集下支持向量机训练样本的缩减策略[J]. 计算机科学,2007,34(10):211-213
[5] 孙发圣,肖怀铁. 基于K最近邻的支持向量机快速训练算法[J]. 电光与控制,2008,15(6):44-47
[6] 陈振洲,李磊,姚正安. 基于SVM的特征加权KNN算法[J]. 中山大学学报:自然科学版,2005,44(1):17-20
[7] Burges C J C. Geometry and invariance in kernel based methods [M]// Advances in Kernel Methods-Support Vector Learning. Cambridge,MA:MIT Press,1998:89-116
[8] 乜聚虎. 智能手机异常检测技术研究与实现[D]. 合肥:中国科学技术大学,2011
[9] 周忠军,苏红旗. Android智能手机入侵检测系统设计[J]. 科技资讯,2012(18):30-32