基于多尺度特征融合的异常流量检测方法

陈鸿昶 程国振 伊 鹏
(国家数字交换系统工程技术研究中心 郑州 450002)

摘 要 快速、准确地检测异常是网络安全的重要保证。但是由于网络流量的非线性、非平稳性以及自相似性,异常流量检测存在误报率高、检测率低、不能满足骨干网实时性要求等问题。该方法综合了希尔伯特-黄变换(Hilbert-Huang Transform,HHT)和 Dempster-Shafer 证据理论(D-S evidence theory)评测框架。前者将不同的流特征分别分解为多时间尺度上的固有模态函数(Intrinsic Mode Function,IMF),滤除特征中的非线性、非平稳分量;后者将前者分解得到的多尺度特征作为证据融合并最终做出决策。通过对 KDD CUP 1999 的入侵检测系统(Intrusion Detection System,IDS)基准数据的实验表明,该方法能有效区分突发流量(crowd flow)和拒绝服务攻击(Denail of service,DoS)攻击流,整体上在保证低误报率前提下检测率达到 85.1%。目前该方法已经作为入侵检测的子模块实现,并试用于某骨干网入口处检测异常。

关键词 异常检测,拒绝服务攻击,希尔伯特-黄变换,D-S 证据理论
中图法分类号 TP393 **文献标识码** A

Anomaly Traffic Detection Based on Multi-resolution Feature Fusion

CHEN Hong-chang CHENG Guo-zhen YI Peng
(National Digital Switching System Engineering & Technological R&D Center, Zhengzhou 450002, China)

Abstract Detecting network traffic anomaly quickly and accurately is playing significant roles in guaranteeing network secuirity. But it has high false alarm rate, low detect rate, and can't perform real-time detection in the backbone very well due to its nonlinearity, nonstationarity and self-similarity. For this status quo, we proposed a novel multi-resolution fusion detection method. It combines Hilbert-Huang transform(i. e. , HHT) and Dempster-Shafer(i. e, D-S) theory. The former decomposes traffic features on multi-time scales to intrinsic mode function(IMF), and filters nonlinear, nonstationary ingredients effectively; the latter fuses the multiscale elements and makes a decision. Based on the KDD CUP 1999 intrusion detection system evaluation data set, this detector detects 85.1% attacks at low false alarm rate which is better than related ones, and recognises DoS from burst traffic. At present, this method has been performed as a detector and run in a backbone network.

Keywords Anomaly detection, DoS, HHT, D-S theory

1 引言

异常检测是网络入侵检测系统的重要组成部分。近年来,通过对网络异常流的检测的深入研究,研究人员提出了很多的方法,它们按照所采用的技术分为基于统计方法[6,7,10,11]、基于数据挖掘方法以及基于机器学习方法[2-5,9]。基于这些方法实现了大量已应用的实时网络异常检测系统,包括 SRI International 推出的 NIDES[12]、KDD CUP 1999 的冠军 EMERALD[4]、基于包头的检测系统 PHAD[2]等。这些系统应用于实际网络中,在一定程度上取得了较好的效果,但是仍存在不足。首先,检测率和误报率仍比较高,使用 KDD CUP 1999 的数据集测试,PHAD 的检测率仅为 27%,EMERALD 也只有 50%;其次,不能有效区分正常的突发流量与 DoS 攻击,正常流量具有分形特性,这种特性在受到攻击后会有所变化,这种变化在单一尺度下不能得到体现;最后,虽然大部分系统是多变量的(例如,NIDES),但是它们对于多个变量没有融合起来,或者仅进行简单的组合,不能提炼出不同特征的固有属性。

针对上述问题提出了基于 HHT[9]和 D-S 证据理论[1]的异常检测方法—多尺度特征融合(multi-resolution fusion, MRF)的检测系统,通过实验分析证明了该系统的有效性。MRF 是应用 HHT 把多个网络流特征多尺度分解,将不同种类的异常分散到包含不同频率成分的 IMF 中;然后由 D-S 证据融合引擎融合多尺度分量。HHT 是基于经验的对非线性

到稿日期:2011-03-20 返修日期:2011-06-20 本文受国家高技术研究发展计划(2009AA01A346),国家"十一五"科技支撑计划(2008BAH37B02)资助。

陈鸿昶(1964—),男,博士生导师,主要研究方向为计算机应用、电信网攻防技术、程控交换技术,E-mail:chenhongchang@ndsc.com.cn;程国振(1986—),硕士生,主要研究方向为异常流量检测,E-mail:guozhencheng1986@gmail.com(通信作者);伊 鹏(1975—),男,副教授,主要研究方向为路由交换与调度技术。

和非平稳信号的多尺度分析方法。D-S 证据理论作为不确定推理理论之一，满足比贝叶斯概率论更弱的条件，并给出了不确定性的定量描述，在“正常”与“异常”之间增加了“可疑”的过渡带，其在信息融合中的作用是可对 HHT 分解得到的多个 IMF 分量进行融合，有效减少了信息损失。将使用 KDD CUP 1999 数据集[5]进行的实验与相关异常检测系统做了比较。结果表明，MRF 在较低误报率的条件下，提高了检测率，达到 85.1%，并且在多尺度分解后，有效区分了异常流与正常流，例如正常突发流与 DoS 流。本文第 2 节给出了 HHT 和 D-S 证据理论的背景知识；第 3 节阐述了 MRF 异常检测算法的框架；第 4 节给出了算法实现和实验验证，以及与相关算法的比较；最后给出结论。

2 理论背景

HHT[13]是基于经验的非线性和非平稳信号的多尺度分析方法。HHT 包含了两个步骤：经验模态分解(Empirical Model Decomposition，EMD)和希尔伯特变换(Hilbert transformation)。首先假设任何信号都是由不同波动的简单本征模态组成，采用 EMD 分解将信号逐级分解为不同时间尺度上的波动。波动由本征模态函数(IMF)表示，每个 IMF 分量具有以下两个特点：(1)从整个数据集的角度看，极值点和过零点相等或最多相差一个；(2)在任何点上，极大值包络和极小值包络的平均和为零。与简单谐波函数的固定幅度和频率不同，IMF 表示简单波动模态，幅度和频率随时间可变，其分解过程如下。

(1) 初始化令 $r_0(t)=f(t)$，$i=1$。

(2) 抽取第 i 个 IMF 分量：①初始化使 $h_0=r_i(t)$，$k=1$；②计算 $h_{k-1}(t)$ 的极大值和极小值点；③对 $h_{k-1}(t)$ 的极大值和极小值点分别进行 3 次样条插值，形成上下包络线；④计算上下包络线的均值 $m_{k-1}(t)$；⑤得到 $h_k(t)=h_{k-1}(t)-m_{k-1}(t)$；⑥若满足停止条件，那么 $c_i(t)=h_k(t)$，否则令 $k=k+1$ 并转到②。

(3) 令 $r_i(t)=r_{i-1}(t)-c_i(t)$；

(4) 若 $r_i(t)$ 中的几点多余 2 个，那么转到(2)，否则结束，$r_i(t)$ 是 $f(t)$ 的残余分量，分解结果为：

$$f(t)=\sum_{j=1}^{n}c_j(t)+r_n \tag{1}$$

这样，获得了 n 个 IMF 分量和一个残余分量，IMF 模式分量代表了原始信号中包含的不同时间尺度的特征信号，残余量代表了原始数据中的趋势量信息，对经过 EMD 后的 IMF 进行 Hilbert 变换，得到解析信号为：

$$z(t)=c(t)+jH(c(t))=a(t)e^{j\Phi(t)} \tag{2}$$

式中，$a(t)$，$\Phi(t)$ 分别是幅度函数和相位函数，对相位函数求导可得瞬时频率。

D-S 证据理论是建立在非空有限域 Θ 上的理论，Θ 称为识别框架，表示某一问题的所有可能答案的集合 $\Theta=\{\theta_1,\theta_2,\cdots,\theta_j,\cdots,\theta_N\}$，集合中元素两两互斥。由识别框架 Θ 的所有子集组成的一个集合称为 Θ 的幂集，记作 2^Θ。识别框架 Θ 的任一子集 A 都与问题的一个答案的命题对应，D-S 证据理论的目标就是根据问题的部分观察 $E_1,E_2,\cdots,E_m$(其中 $E_i(i=1,2,\cdots,k)$ 又称为证据)，通过推理得出问题的答案。这些证据是问题答案的不确定表现，证据对答案的支持程度通过基本信任分配函数度量(Basic Probability Assignment，BPA)。

定义 1 设 Θ 为识别框架，基本信任分配函数 m 是一个从集合 2^Θ 到[0,1]的映射，A 表示识别框架 Θ 的任一子集，记作 $A\subseteq\Theta$，且满足：

$$\begin{cases} m(\varnothing)=0 \\ \sum\limits_{A\subseteq\Theta} m(A)=1 \end{cases} \tag{3}$$

式中，$m(A)$ 称为事件 A 的基本信任分配函数，它表示证据对 A 的信任程度，如果 $m(A)>0$，那么称 A 为焦元(local element)。

定义 2 设 $m_1,m_2,\cdots,m_n$ 是同一识别框架 Θ 上的 n 个基本信任分配函数，焦元分别为 $A_i(i=1,2,\cdots,N)$，则 D-S 合成规则为：

$$m(A)=\begin{cases} \dfrac{\sum\limits_{\cap A_i=A}\prod\limits_{1\leqslant i\leqslant N} m_i(A_i)}{1-K}, & A\neq\varnothing \\ 0, & A=\varnothing \end{cases} \tag{4}$$

式中，$K=\sum\limits_{\cap A_i=\varnothing}\prod\limits_{1\leqslant i\leqslant N} m_i(A_i)$。

3 MRF 检测框架

Leland[14]等证明网络流量在不同时间尺度上具有自相似性；网络异常的发生往往同时体现在流量的不同特征上，而在正常情况下，流量的多个特征上同时表现出异常的可能性很小。因此，采用 HHT 对不同流量特征进行时间多尺度分解得到各 IMF 分量，将不同类异常分散到不同的尺度分量上，排除非异常尺度分量的干扰，最后由 D-S 融合引擎将不同证据融合，得到综合信度分配函数。

3.1 方法框架

图 1 给出了框架的基本结构，首先高速设备将骨干网特定的流量转发到系统中，并统计易于区分的流量特征作为基本数据，将不同特征分别进行 EMD 分解得到不同时间尺度的 IMF 分量。本节假设 IMF 分量服从正态分布(第 3.3.2 节将验证该假设)，并由此得出 BPA 函数，然后将其输入到 D-S 融合引擎中做出最终决策。高维流特征经过 EMD 分解后存在 3 种证据融合方案：第一，完全合成法，即将所有的 IMF 分量(不同特征)作为原始证据直接合成得到总的信度函数；第二，内部合成法，首先在同一特征内部将不同 IMF 分量作为原始证据合成得到基本证据，然后再对不同特征下的基本证据合成为总的信度分配函数；第三，交叉合成法，将不同流量特征在相同时间尺度上的 IMF 分量作为原始证据合成得到基本证据，然后将基本证据合成为总的信度分配函数。

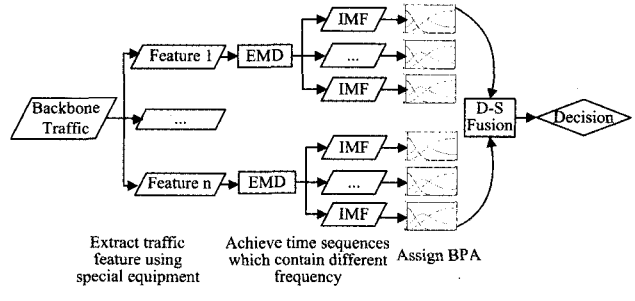


图 1 MRF 异常检测系统框架

3.2 Hilbert-Huang 变换及特征分解

Hilbert-Huang 变换可有效处理非线性非平稳信号，因此，对于待分析数据没有特别的要求，也不需要做特殊处理。分解过程如下：设有 N 个流量特征$\{f_1,f_2,\cdots,f_i,\cdots,f_N\}$，对

任意一特征 f_i 进行 m 层的 HHT 变换得到 m 个 IMF 分量$\{c_{i1}, c_{i2}, \cdots, c_{ij}, \cdots, c_{im}\}$，将 N 个特征的 IMF 分量组成矩阵 C：

$$C=\begin{bmatrix} c_{11} & \cdots & c_{1m} \\ \vdots & \ddots & \vdots \\ c_{n1} & \cdots & c_{nm} \end{bmatrix} \tag{5}$$

式中，每行表示同一个特征的 EMD 分解得到的 IMF 分量，每列表示各个特征在相同时间尺度下的 IMF 分量。

3.3 D-S 证据融合与检测引擎

3.3.1 系统识别框架定义

本文不对异常分类，所以网络异常检测的目的就是判断流量的正常与异常。因此，识别框架 Θ 表示为$\{N,A\}$，N 为正常，A 为异常，并且 $N\cap A=\varnothing$。由定义 1 可知，基本信任分配函数 m:$\{N,A\}\rightarrow[0,1]$，$m(\varnothing)=0$，$m(\Theta)+m(N)+m(A)=1$，其中 $m(\Theta)$表示证据的不确定性，$m(N)$表示证据支持流量正常的信度，$m(A)$表示支持异常的信度。

3.3.2 IMF 分量的分布

本节基于 KDD CUP 1999 的入侵检测系统基准数据的多个特征，采用 Q-Q(Quantile-Quantile)图验证 IMF 分量服从正态分布的假设。Q-Q 图是验证两个分布函数统计上的差异的图形化方法。因此，本节比较了标准正态分布与 IMF 分量。如果这两个分布匹配，那么 Q-Q 图应该近似对角直线。图 2 给出了标准正态分布与 IMF 分量的 Q-Q 曲线图。由图可知，曲线近似为对角直线，仅仅在偏差较大的尾部出现弯曲。这证明 IMF 分量服从中心极限定理。

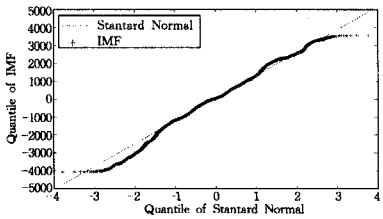


图 2 IMF 的分布与正态分布的 Q-Q 图

由 IMF 的定义可知，IMF 均值为 0，假设方差为 σ^2，借用模糊数学的隶属度函数的概念，得到 IMF 分量到 BPA 的映射。为了避免证据合成时的全冲突悖论，对正态型隶属度函数做了修正，增加了隶属因子 p，可得 $m(N)=pe^{-(x/\sigma)^2}$，其中 x 表示 IMF 值，$0<p<1$ 称为隶属因子，一般 p 取 0.85～0.9。基于上述分析给出了图 3 所示的 BPA 函数，其中 $m(A)$ 可由 $m(\Theta)=1-m(N)-m(A)$求出。区间$[L,H]$的左右移动可调整虚警率和漏报率。由 IMF 分量矩阵式(5)可得 BPA 矩阵式(6)。

$$m^c=\begin{bmatrix} m^c_{11} & \cdots & m^c_{1m} \\ \vdots & \ddots & \vdots \\ m^c_{n1} & \cdots & m^c_{nm} \end{bmatrix} \tag{6}$$

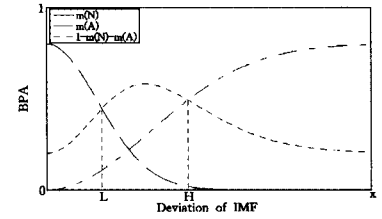


图 3 基于正态分布的基本信任分配函数

3.3.3 多尺度信息的融合

第 3.1 节中提到了对 IMF 的融合存在 3 种方案：完全合成法、内部合成法、交叉合成法。对于 3 种不同的方案包括 3 类不同的证据合成情况：同一特征的不同时间尺度，不同特征的相同时间尺度以及不同特征不同时间尺度的合成。从 BPA 矩阵式(6)的角度分析，3 种方案分别为矩阵所有元素的一次合成，先按行后按列合成以及先按列后按行合成。

方案 1 不同特征不同时间尺度的 IMF 分量的合成可表示为：

$$m^c(D)=\frac{\sum\limits_{\cap D_p=D}\prod\limits_{i=1}^{n}\prod\limits_{j=1}^{m}m_{ij}(D_p)}{\sum\limits_{\cap D_p=\varnothing}\prod\limits_{i=1}^{n}\prod\limits_{j=1}^{m}m_{ij}(D_p)} \tag{7}$$

方案 2 同一特征的不同时间尺度的 IMF 分量的合成可表示为：

$$m^c_i.(D)=\frac{\sum\limits_{\cap D_p=D}\prod\limits_{j=1}^{m}m^c_{ij}(D_p)}{\sum\limits_{\cap D_p=\varnothing}\prod\limits_{j=1}^{m}m^c_{ij}(D_p)} \tag{8}$$

首先在各特征内部完成证据的合成得到 $m^c_i.(D)$。一般地，仅仅考虑某一特征误报率会很高，必须将多个特征融合起来得到总的质量函数 $m^c(D_p)$。在 D-S 证据理论中，证据合成结果应该符合大多数证据，但在本方案中却恰恰相反。假设待分析特征中包含一个异常，例如短时异常经过 EMD 分解后，异常被映射到高频分量上，则在高频 IMF 分量中的基本信任分配函数趋向于支持异常，其他 IMF 分量则趋向于支持正常，证据间存在冲突。但是根据流量多尺度分析的目的(多尺度分析的目的就是将不同的异常分离到不同频率分量上)，这种冲突不包含有效信息。如果某一特征中任意一个 IMF 分量上存在异常，则该特征表现为异常，具有一票否决制，该现象称为 0 信任悖论。通常 D-S 证据理论的证据合成要避免 0 信任悖论，而这里恰恰利用了这个悖论。

方案 3 不同特征的相同时间尺度的 IMF 分量的合成可表示为：

$$m^c._j(D)=\frac{\sum\limits_{\cap D_p=D}\prod\limits_{i=1}^{n}m^c_{ij}(D_p)}{\sum\limits_{\cap D_p=\varnothing}\prod\limits_{i=1}^{n}m^c_{ij}(D_p)} \tag{9}$$

与方案 2 相同，合成也是分两步进行，首先将相同尺度上的 IMF 分量合成，相同尺度上的 IMF 分量包含同类异常，然后再将 $m^c._j(D)$继续合成，最终得到 $m^c(D)$。

文献[14]在证据中各焦元之间互斥的情况下得出合成规则满足结合律，n 个证据合成的算法复杂度为 $O(n)$。结合 D-S 证据合成规则的交换律，可推知高维流特征经 EMD 分解后的 IMF 分量 D-S 融合引擎与顺序无关。

所以上述 3 种方案虽然合成顺序不同，但是结果一致。根据方案 1，多尺度分解后的信息融合算法复杂度为 $O(mn)$，因为 3 种方案结果一致，最终 MRF 的信息融合算法复杂度为 $O(mn)$。m 是原始数据经过 EMD 多尺度分解后的层数，通常不超过 10 层，即 $m<10$，算法复杂度小于 $O(n)$。因此，MRF 符合检测的实时性需求。

从 3 种方案的合成过程可以清楚地看到 MRF 在信息融合中的实际意义。EMD 将每个流特征分解为不同尺度的 IMF 分量。这些 IMF 包含了特征的不同频率成分，即包含了不同种类的异常。在合成过程中若按照方案 2 的特征内部合成考虑，可观察每个特征的融合结果；若按照方案 3 的交叉合成，可观察所有特征在同一尺度上的融合结果。

3.4 MRF 错误概率分析

事实上，异常检测系统在检测某一类异常时，背景流量和其它类异常就可能成为干扰项。它们会影响系统的判决，并最终使误报率增大，例如突发流量就可能被误判为 DoS 攻击行为。MRF 将各特征分解成包含不同频率成分的时间序列，同时也将不同异常分散到不同的时间序列中。这种处理使得 BPA 能更准确地描述特征的状态，从而降低系统误报率。

下面本节从理论上推导 MRF 的误报率低于单纯的 D-S 融合框架。不失一般性，针对一个流特征的多个 IMF 展开证明。

证明：设某一流特征 X 经过 EMD 分解可得 $X=\sum_i C_i+R_n$，重写该式的集合表示形式为 $X=(\bigcup_{i=1}^{n} C_i)\cup R_n$。分类集合 $\Theta=\{N,A\}$，则平均错误概率为：

$$\begin{aligned}P_{e0}&=P(X\in R_A|N)P(N)+P(X\in R_N|A)P(A)\\&=P((\bigcup_{i=1}^{n}C_i)\cup r_n\in R_A|N)P(N)+P((\bigcup_{i=1}^{n}C_i)\cup r_n\in R_N|A)P(A)\\&=P(N)\sum_{i=1}^{n}P(C_i\in R_{A_i}|N)+P(r_n\in R_{A_r}|N)P(N)+P(A)\sum_{i=1}^{n}P(C_i\in R_{N_i}|A)+P(r_n\in R_{N_r}|A)P(A)\end{aligned}$$

根据 EMD 分解定义 r_n 作为趋势分量不包含有效信息，属于干扰噪声。在对其他 IMF 分量融合时，D-S 融合规则会去除一些冲突信息。在此过程中，部分 IMF 分量被强化，部分分量被弱化，例如方案 2 中所举例子，忽略弱化的 IMF 分量，可得

$$\begin{aligned}P_{e0}&\geqslant P(N)\sum_{i=1}^{n}P(C_i\in R_{A_i}|N)+P(A)\sum_{i=1}^{n}P(C_i\in R_{N_i}|A)\\&\geqslant P(N)\sum_{j=1}^{m}P(C_j\in R_{A_j}|N)+P(A)\sum_{j=1}^{m}P(C_j\in R_{N_j}|A)\\&=P_{e1}\end{aligned}$$

证毕。

其中 $m<n$，所以 EMD 分解后再融合可有效降低判决的平均错误概率。

4 实验及结果分析

本节验证 MRF 检测方法的有效性，为了便于与其他方法比较，采用广泛使用的基准评测数据集 KDD CUP 1999 进行实验测试。采用整个数据集测试算法的有效性，并与相关入侵检测系统进行比较。

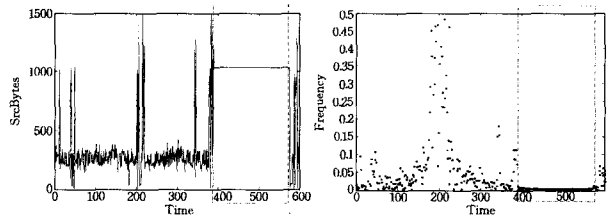
4.1 测试数据集

KDD CUP 1999 包括约 5,000,000 条记录，每条记录由 41 个连接特征组成，其中 34 个数据特征，7 个标识特征，包含 DoS，U2R，R2L，Probe 以及正常流等 5 类数据。第二步实验采用 kddcup. data. corrected. gz 作为测试数据。该数据集记录了 5 周数据，其中前两周用于算法训练，后 3 周用于测试，第一周和第三周为正常数据，第二周包含 43 次攻击(18 种)，最后两周包含了 58 种 201 次攻击，其中 40 种是新的攻击。

4.2 实验结果

MRF 方法可处理任何网络流特征，因此特征的选取没用特殊限制，也无须对数据预处理，但为方便验证，选取 duration，count，srv_count，src_bytes，dst_host_count，dst_host_srv_diff_host_rate，dst_host_rerror_rate 等 12 个特征进行分析。MRF 方法首先对特征进行 EMD 分解得到特征的 IMF 分量，然后再对其进行 D-S 证据合成，最终做出判断。

限于篇幅，这里只给出 src_bytes 的 EMD 分解结果。选取 600 条记录，进行 EMD 分解，原始数据如图 4(a)所示，数据中包括 DoS 攻击流、突发流量和正常流量。其中红色方框内为 DoS 攻击流量，从 397 点开始，574 点处结束。图 4(b)给出了原始流量的时频分布图，其中正常流量的频率成分散落在低中频，突发流量有部分高频成分，而红框内的 DoS 攻击流的频率成分单一，且分布在低频。



(a) 流量时间序列图　　(b) 时频分析图

图 4　流特征 src_bytes 的 600 条记录数据

图 5 给出了各多尺度分解后的 IMF 分量和趋势分量，由图看出，正常流量在 IMF 中几乎没有波动；DoS 攻击流在各 IMF 分量上均有较大的偏离，且偏离的时间跨度较大；突发量的波动相对于 DoS 流较小，波动的时间跨度也较小，并且突发流波动在 IMF 分量波动的贡献不大，在低频分量上没有贡献。因此，原始特征经过 EMD 分解后可以将正常流量、突发流量与 DoS 攻击流区分开，有效降低了误报率。

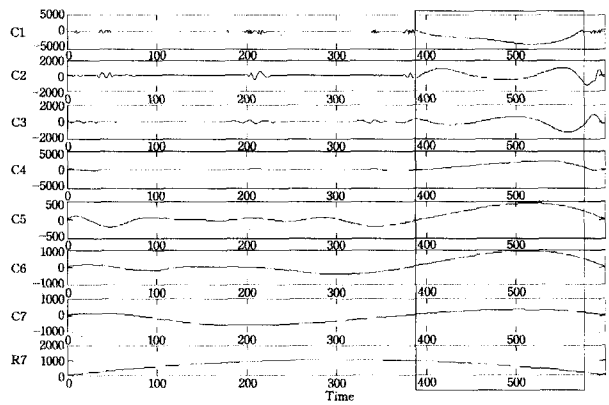


图 5　流特征 src_bytes 的 600 条记录的 EMD 分解

一般异常检测方法每次建立网络正常模式均需要大量的历史数据和较长时间用于训练，降低了实时性。HHT 是针对处理小样本不平稳数据而提出的经验性方法，它将数据多尺度分解，去除不平稳分量，从而有利于 D-S 证据检测。实验表明：在训练阶段完成后，MRF 每次检测时间小于 10s，而硬件上报周期为 20s，在 10G 流量的压力测试下满足实时检测的要求。

MRF 采用第一周的数据进行训练，建立正常轮廓，然后使用第四周和第五周的数据进行测试，得到的测试结果如表 1 所列。

表 1　MRF 系统性能

Category of attacks	Amount of detected attacks/Total attacks	Detection Rate
DoS	65/65	100%
U2R	35/43	81.4%
R2L	40/56	71.4%
Probe	31/37	83.8%
Total	171/201	85.1%

4.3 与相关检测系统的比较

有关网络异常检测的研究很多，这里将采用 KDD CUP 1999 数据集进行测试的几个异常检测系统与 MRF 做了比较，包括 KDD CUP1999 竞赛的冠军 EMERALD 在内，在误报率为 10%的情况下，结果如表 2 所列。

表 2 MRF 与其它算法的比较

Systems	Detection Methods	Amount of detected attacks/Total attacks	Detection Rate
EMERALD	Expert System	85/201	42%
LERAD	Learning goog rules from training set	114/190	60%
PHAD	Packets head anomaly detection	54/201	27%
ALAD	Using well know ports	60/201	30%
NETAD	PHAD and ALAD	132/201	66%
FAD	D-S theory	119/201	59%
MRF	Based on Hilbert-Huang transform and D-S theory	171/201	85.1%

实验结果表明，MRF 性能超过了当年 KDD CUP 的冠军 EMERALD 以及采用动态学习规则的 LERAD。仅仅使用 D-S证据理论的 FAD 检测系统与其它方法比较并没有太大的性能提高，并且与 NETAD 比较性能还有所降低，但是在与 HHT 相结合使用后，D-S 证据理论检测有了较大的提高，因为网络流特征信号是非线性和不平稳的，并且正常流量特征表现为自相似性，HHT 将信号进行不同时间尺度的分解去除了趋势等不平稳分量，并将多尺度分量作为证据，有效区分了突发流和 DoS 流。

结束语 HHT 是一种自适应的时频局部化多尺度分析方法，适合处理非线性、非平稳信号；D-S 证据理论作为不确定推理理论之一，已经广泛用于多传感器信息融合等领域。最近，该理论被引入到网络异常检测中，在多特征融合方面取得了一定效果，但其检测率仍然不理想。本文将 HHT 与其结合，提出了 MRF 异常检测方法，它从多尺度的角度检测异常，并使用 KDD CUP 1999 数据集进行验证。实验表明，该系统在保证一定误报率的情况下提高了检测率。

参 考 文 献

[1] Dempster A. Upper and lower probabilities induced by multivalued mapping[J]. Annals of Mathematical Statistics, 1967, 38(2): 325-339

[2] Mahoney M V, Chan P K. PHAD: Packet Header Anomaly Detection for Identifying Hostile Network Traffic[R]. Melbourne: Department of Computer Science, Florida Institute of Technology, 2001

[3] Mahoney M V, Chan P K. Learning Nonstationary Models of Normal Network Traffic for Detecting Novel Attacks[C]// Proc. of the Eighth International Conference on Knowledge Discovery and Data Mining. Edmonton: ACM, 2002: 376-385

[4] Porras P A, Neumann P G. EMERALD: Event Monitoring Enabling Responses to Anomalous Live Disturbances[C]// Proc. of the 20th National Information Systems Security Conference. Baltimore, 1997: 353-365

[5] Lippmann R, Haines J W, Fried D J, et al. The 1999 DARPA off-line intrusion detection evaluation[J]. Computer Networks: The International Journal of Computer and Telecommunications Networking, 2000, 34(4): 579-595

[6] Moayedi H Z, Masnadi-Shirazi M A. Arima model for network traffic prediction and anomaly detection[C]// Proc. of the International Symposium on Information Technology. 2008

[7] Li Zong-lin, Hu Guang-min, Yao Xing-miao. Detecting Distributed Network Traffic Anomaly with Network-Wide Correlation Analysis[J]. EURASIP Journal on Advances in Signal Processing, 2009

[8] Kline J, Nam S, Barford P, et al. Traffic Anomaly Detection at Fine Time Scales with Bayes Nets[C]// Proc. of the Third International Conference on Internet Monitoring and Protection. Washington: IEEE Computer Society, 2008: 1-10

[9] Huang N E, Shen Z, Long S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis[C]// Proc. of the Royal Society of London. 1998, A454: 903-995

[10] 诸葛建伟，王大为，陈昱，等. 基于 D-S 证据理论的网络异常检测方法[J]. 软件学报，2006，17(3)：463-471

[11] Lakhina A, Crovella M, Diot C. Diagnosing Network-Wide Traffic Anomalies[C]// Proc. of ACM SIGCOMM. Portland: ACM, 2004

[12] Andeson D, Frivold T, Tamaru A. Next-generation intrusion detection expert system(NIDES)[R]. Software Users Manual, Beta-Update release. Menlo Park: Computer Science Laboratory, SRI International, 1994

[13] Mahoney M V, Chan P K. Learning Models of Network Traffic for Detection Novel Attacks[D]. Computer Science Department, Florida Institute of Technology, 2002

[14] Leland W E, Taqqu M S, Willinger W, et al. On the Self-similar Nature of Ethernet Traffic [J]. Transactions on Networking, 1994, 2(1): 1-15

[15] Silveira F, Diot C, Taft N, et al. ASTUTE: Detecting a Different Class of Traffic Anomalies[C]// Proc of SIGCOMM. 2010

(上接第 37 页)

[7] Zhou Heng-min, Zeng D, Zhang Chang-li. Finding Leaders from Opinion Networks[C]// ISI 2009. 2009: 266-268

[8] Zhai Zhong-wu, Xu Hua. Identifying opinion leaders in BBS[C]// IEEE Proceedings of Web Intelligence and Intelligent Agent Technology. 2008

[9] Scott J. Social Network Analysis: A Handbook[M]. Sage Publications, London, 2000

[10] Wu F, Huberman B A. Finding communities in linear time: A Physics approach[J]. Phys. J B, 2003, 38: 331-338

[11] Newman M E, Girvan M. Finding and evaluating community structure in networks[J]. Physical Review E, 2004, 69: 026113

[12] Cover T M, Hart P E. Rates of convergence for nearest neighbor procedure[C]// Proc. HaWaii Int. Conf. on System Science. 1967: 413-415

[13] Yang Shen, Li Shu-chen, Zhen ling. Emotion mining research on micro-bolg[C]// SWS 2009. 2009: 71-75