一种基于超统计理论的非平稳时间序列异常点检测方法研究

杨 越[1,2] 胡汉平[1] 熊 伟[1] 丁 帆[1]
(华中科技大学图像识别与人工智能研究所 武汉 430074)[1]
(中国地质大学机电学院测控系 武汉 430074)[2]

摘 要 从非平稳时间序列的分布函数及其参数入手,主要研究分布函数不变分布参数变化的这一类非平稳的时间序列异常点检测方法,提出了基于超统计的异常检测方法,并将其应用于非平稳网络流量时间序列。从网络流量的非平稳和突发性特点出发,特别考虑到由于攻击流量所引起的流量特性的变化,结合超统计理论,主要研究分布参量的变化。根据超统计的理论,先应建立分布统计模型,研究分布模型不同参数变化对分布的决定性作用,从而将异常网络流量的检测研究转化成对慢变量参数序列的检测研究。该检测方法大大降低了计算的复杂度。通过大量实验表明该方法具有良好的效果。

关键词 时间序列,非平稳,超统计,网络流量
中图法分类号 TP31 **文献标识码** A

Novel Non-stationary Time Series Anomaly Detection Model Based on Superstatistics Theory

YANG Yue[1,2] HU Han-ping[1] XIONG Wei[1] DING Fan[1]
(Institute for Pattern Recognition & AI, Huazhong Univ of Sci & Tech, Wuhan 430074, China)[1]
(School of Electronic Information & Mechanics, China University of Geosciences, Wuhan 430074, China)[2]

Abstract Because of network traffic non-stationary property it can hardly use traditional way to analyze the complicated network traffic. A new detection method of non-stationary network traffic based on superstatistics theory was discussed. According to the superstatistics theory, the complex dynamic system may have a large fluctuation of intensive quantities on large time scales which causes the system to behave as non-stationary which is also the characteristic of network traffic. This new idea provides us with a novel method to partition the non-stationary traffic time series into small stationary segments. We used the slow parameters of the segments as a key determinant factor of the system to describe the network characteristic and analyze the slow parameters with time series theory to detect network anomaly. The result of experiments indicates that this method can be effective.

Keywords Time series, Non-stationary, Superstatistics, Network traffic

1 引言

非平稳时间序列是其分布函数随时间变化的过程[1]。对于同一分布的非平稳过程,主要是分布参数随时间的变化。分布参数变化可分为参数变化是确定函数,还是随机过程。本文研究针对参数是确定函数,即参数是宽平稳过程的非平稳时间序列。目前很多领域都存在这类非平稳时间序列,如异常的网络流量[2]。时间序列异常点的检测方法一般可分为基于统计模型的非参数检验和基于特征量的检测(或参数检测)[3]。前者的优点在于统计模型是根据先验所有数据建立的,具有很强的整体性,包含的信息量大,其缺点在于计算的数据量大,不适合实时处理;相对于前者而言,基于特征量检测方法的优点在于所需的数据量较统计模型少,计算复杂度较低,但它所能提供的信息量较少。本文研究的是分布函数不变、分布参数变化的非平稳时间序列。因分布函数不变,故提出了先离线建立统计模型,再根据分布函数的参数变化进行实时检测,即基于统计模型的参数检验方法。这种方法是传统两种方法优势的互补,建立统计分布模型避免了特征量检测没有充分利用信息量的缺点,又具有参数检测计算量小的特点。而统计模型异常检测,在建立统计模型或分布时,首先要考虑序列的平稳性,只有建立在平稳性至少宽平稳的基础上才能进行统计模型的研究。对于一般的非平稳转化为平稳序列的处理方法,时间序列的非平稳有两种情况:一种是退势非平稳序列,另一种是差分平稳序列[4]。对于前者,直接退势(减去)即可变为平稳序列;对于后者,可以直接差分。大多数非平稳的时间序列一般可通过一次或多次差分的形式变为平稳的。但也有一些时间序列,无论经过多少次差分,都不能变为平稳的。针对的统计分布函数不变,分布参数是随机或者是复杂变化的非平稳这类序列。由于没有考虑到序列的非平稳性和突发性,也没有考虑到统计分布参数是随机或者是

到稿日期:2010-07-16 返修日期:2010-11-14 本文受国家自然基金(60773192)资助。
杨 越(1978—),女,博士生,讲师,主要研究方向为信号处理与信息安全等,E-mail:cugyy@126.com;胡汉平 男,教授,主要研究方向为信息安全等;熊 伟 男,博士生,主要研究方向信息安全;丁 帆 女,硕士生,主要研究方向为信息安全。

复杂变化的，因此不可避免地存在异常检测问题。所以采用微积分的思想——分窗处理，可有效降低其在子窗内的显著性和复杂性，将非平稳序列划分成平稳子段，在根据流量的本身统计特性确定相应的统计模型的基础之上，根据保持统计模型的参数在子窗内的宽平稳性原则进行分窗处理。在平稳子窗内研究分布函数的参数变化。这种研究非平稳时间序列的分布函数的参数变化的方法，即“统计之统计”过程，正是超统计理论处理非平稳时间序列研究方法[5,6]。超统计理论属于物理学的前沿领域，弥补了传统统计方法的不足。超统计理论用于研究多个动力学子系统的复合或复杂过程，这种系统在较大的时间尺度下存在某种强度量的大幅度波动，这种强度量在超统计理论中被称为慢变量。根据超统计理论，研究含有异常的非平稳性时间序列统计分布的慢变量参数序列，即决定系统内在变化的主要因数，从而达到研究原始序列的目的，通过慢变量参数序列的变化寻找出异常点。

2 基于超统计的非平稳时间序列的检测模型

根据数学建模分析时间序列的一般步骤[7]，对于基于超统计的非平稳时间序列异常点检测可分为以下几个步骤：

(1)对于所研究的时间序列进行分析及平稳化处理。

首先需要分析时间序列的平稳性，判断该时间序列是平稳的时间序列，还是非平稳的时间序列。若为非平稳序列，研究序列属于具体哪一类型非平稳的序列，即是否能用传统的非平稳转平稳的方式处理。若为复杂的统计分布不变、统计分布参数变化的这一类非平稳序列，提出合理的平稳化处理方法。本文是根据微积分分窗处理的思想，可以有效降低其在子窗内的显著性和复杂性，采用非平稳序列将其划分成平稳子窗。参考启发式分割算法是一种能将非平稳时间序列按照均值划分为平稳子序列的有效方法（由 Galvan 在 2001 年提出）[8]，但其按照整个序列统一来划分窗口的，不是按照时间序列时间增长的方向分割的。提出改进的启发式分割方法，其主要思想如下：

记待测的时间序列为 $x(t)$，其包含的点数为 N；记初始窗口部分为 Ls，初始滑动窗口为 Lm，$Ls+Lm$ 的区域为当前窗口。设 i 为当前窗口的截止位置，$Ls+Lm$ 的长度为 $N_2(i)$，其 Ls 的长度为 $N1$，分别计算 Ls 和 $(Ls+Lm)$ 部分的平均值 $\mu_1(i)$，$\mu_2(i)$，以及标准偏差 $s_1(i)$ 和 $s_2(i)$，则当前窗口的合并偏差 $S_D(i)$ 为：

$$S_D(i)=\left(\frac{(N_1-1)\times s_1(i)^2+(N_2-1)\times s_2(i)^2}{N_1+N_2-2}\right)^{\frac{1}{2}}\times\left(\frac{1}{N_1}+\frac{1}{N_2}\right) \tag{1}$$

式中，用统计值 $T(i)$ 来量化表示 Ls，$Ls+Lm$ 的差异：

$$T(i)=\left|\frac{\mu_1(i)-\mu_2(i)}{s_D(i)}\right| \tag{2}$$

若 $3G\geqslant T(i)\geqslant G$，其中 G 为设定域值，$G=0.5$，则进入下一步。否则，按照以下规则调整 Lm 的长度：

若 $T(i)<G$，则进一步缩放滑动窗口 $Lm=Lm+Lf$（Lm 初始值为 10，Lf 为滑动增量，范围为大于 0、小于 Lm，初始可以取 3～5），重复上面的步骤。

若 $T(i)>3G$，则进一步缩放滑动窗口 $Lm=Lm-Lf$（Lm 初始值为 10，Lf 为滑动增量，范围为大于 0、小于 Lm，初始可以取 2～3），重复上面的步骤。

根据 Lm 最后一个点所在位置作为第一个子窗分割点 W_1 后，从 W_1 位置按照式(2)计算下一个 $T(i)$，得到第二个子窗分割点 W_2，直到序列的结尾，即可依次求出每个子窗的大小，nn 记作子窗总数，即时间序列由 nn 个子窗组成，记 $x\{(x_1,x_2,\cdots,xw_1),(xw_{1+1},\cdots,xw_2),\cdots,xw_m)\}=x\{xw_1,xw_2,\cdots,xw_m\}$。

(2)根据具体时间序列的研究对象的具体特征提出合理的假设，建立合理的统计分布模型。这里可以根据序列的特点(如序列的直方图等)提出可能的适合模型的假设、模型的假设检验。

(3)根据假设的数学模型，求解模型相应参数。

(4)对(3)求解出来的模型参数分布检验，以确定假设是否合理。

(5)根据计算出来的参数序列，分析并找到系统的慢变量，最后建立综合的突变决策判决模型，找到异常点。

3 基于超统计网络流量时间序列的异常点检测

已有的实验已经证实异常流量是一种复杂的非线性或随机性的变化过程[9,10]，也反映了网络异常流量的突发性(即显著性复杂变化)特点。根据网络异常流量，特别是由于攻击引起的异常流量非平稳和突发性的基本特点(异常流量是一种复杂的非线性或随机性的变化过程)，可将前述的基于超统计理论的异常检测方法用于网络流量时间序列，研究通过监测流量统计参数序列的统计特征进行网络异常流量的实时检测。

实验所采用的网络流量数据是麻省理工林肯实验室的信息系统技术组在美国国防部高级研究项目署(DARPA) 及空军研究室赞助下为计算机网络入侵检测系统评估提供的测试用数据集合[11,12]。实验所用的网络流量时间序列是对 1 秒内到达的数据包数量的累积，测试数据集中共有 5 周，每周 5 天的数据，其中第 4 周和第 5 周的数据存在攻击。在 DARPA99 数据中包含的主要 4 种攻击中，只有拒绝服务攻击(DOS 攻击)会对网络流量产生影响。因此，对网络流量进行异常检测，检测的攻击种类主要为 DOS 攻击。

根据正常网络流量的特性建立适合的统计分布模型，主要的网络流量统计分布模型有 Poisson 分布、Pareto 分布、指数分布、对数正态分布和威布尔分布等分布模型[13,14]。通过假设检验等方法可以选取较优的分布拟合模型。根据具体的分布模型进行相应的参数评估和置信度区间分析，最后还应以不同的分布模型采用相应的方法进行检验，以确定模型的准确性。

首先根据正常网络流量时间序列建立统计分布模型。实验所使用的正常网络流量在局部表现出很强的突发性，并且从其直方图中可以看出实际网络表现出明显的重尾特性。因此基于正常网络流量的上述特点，在本文中选择离散广义 Pareto 分布对局部网络流量进行分析。这里采用图检验方法和 K-S 检验方法进行检验[15]。从图检验的结果直观发现所有点基本上拟合在一条直线上，说明符合离散广义 Pareto 分布的分布图检验。再进一步计算出 K-S 检验的结果。表 1 计算的结果也证明了分布模型建立的准确性。

表 1　对于不同窗口 K-S 分布检验结果

number	KSSTAT	CV	P
160	0.0777	0.0954	0.2745
80	0.0616	0.1339	0.9099
60	0.1086	0.1539	0.4461
60	0.1473	0.1869	0.3116
80	0.0748	0.1339	0.7386
60	0.0887	0.1539	0.7032
60	0.0524	0.1539	0.9952
80	0.0995	0.1339	0.3813
80	0.0995	0.1339	0.8162
180	0.0802	0.0900	0.1862
80	0.0839	0.1339	0.6005
100	0.0853	0.1201	0.4370
120	0.0851	0.1098	0.3317
50	0.1108	0.1869	0.6696
60	0.0881	0.1539	0.7108
50	0.0922	0.1869	0.8602

在广义 Pareto 分布的统计模型基础之上，针对异常流量，对网络流量非平稳序列进行平稳性转化——根据保持统计模型的参数在子窗内的宽平稳性原则进行分窗处理。采用改进的启发式分割方法即可依次求出每个子窗的大小。用 nn 记作子窗总数，即网络流量由 nn 个子窗组成，记 $x\{(x_1,x_2,\cdots,xw_1),(xw_{1+1},\cdots,xw_2),\cdots,(\cdots,xw_m)\}=x\{xw_1,xw_2,\cdots,xw_m\}$。

再计算子窗中统计分布模型参数，根据参数序列建立综合判决模型，检测异常。

这里采用 2001 年 Rasmussen 的 GPD 参数估计方法广义概率权重矩估计(GPWM)[16]。GPWM 方法：

$$\tilde{\alpha}_v=\frac{1}{n}\sum_{j=1}^{n}(1-p_{j:n})^v x_{j:n} \tag{3}$$

式中，$x_{1:n}\leqslant x_{2:n}\leqslant\cdots\leqslant x_{n:n}$)是经过排序的样本，$v$ 可以取任意实数。GPWM 方法通常取 $v_1=1,v_2=1.5$。$p_{j:n}$是样本累计分布函数的 Kaplan-Meier 估计，即经验累计分布函数。尺度参数 b 和形状参数 k 可以通过以下两式计算：

$$k=\frac{\tilde{\alpha}_{v_1}(v_1+1)^2-\tilde{\alpha}_{v_2}(v_2+1)^2}{\tilde{\alpha}_{v_2}(v_2+1)-\tilde{\alpha}_{v_1}(v_1+1)} \tag{4}$$

$$b=\tilde{\alpha}_{v_2}(v_2+1)(v_2+1+k) \tag{5}$$

通过多次叠加运算，即可计算出离散广义 Pareto 分布的形状参数和尺度参数。根据式(4)和式(5)计算每一个子窗的参数值 k 和 b，形成用于检测的参数序列$\{k_1,k_2,\cdots,k_m\}$和$\{b_1,b_2,\cdots,b_m\}$(nn 为子窗总数)。

广义 Pareto 分布的两个参数(尺度参数、形状参数)中尺度参数对流量起缩放作用，而形状参数对广义 GPD 分布起决定性作用。形状参数反映流量的内在特性，其变化反映了流量的变化——正常流量的波动，即为慢变量。故对异常网络流量的检测研究，可转化成对慢变量参数序列的检测研究。网络流量具有非线性、非平稳性和复杂性的特点，网络流量的异常变化是一种突变过程。根据控制变量以及状态变量的个数，就可以选择适当的突变模型来描述网络流量的动态特征。根据前面得到的参数序列$\{k_1,k_2,\cdots,k_m\}$和$\{b_1,b_2,\cdots,b_m\}$，控制变量个数为 2 个(控制变量为其他的个数也可以，只是要采样的突变模型不一样而已)，即离散广义 Pareto 模型的形状参数和尺度参数，状态变量为网络流量 $x\{xw_1,xw_2,\cdots,xw_m\}$，那么就可以选择尖点突变模型。尖点突变模型是由两组状态变量(u,v)来描述系统 x 的一组状态变量(u,v)所组成的参数空间，也称为控制空间，其势函数可以表示如下：

$$F(u,v,x)=x^4+aux^2+vbx \tag{6}$$

式中，a,b 为系数，x 为状态变量即网络流量 $x\{xw_1,xw_2,\cdots,xw_m\}$，$u,v$ 为控制变量(前面计算出来的形状参数$\{k_1,k_2,\cdots,k_m\}$和尺度参数$\{b_1,b_2,\cdots,b_m\}$)。不同控制参数的数值组合会形成不同结构的势函数，要求出这些势函数的稳定解，只需要对其微分，便可以得到它们的稳定曲面：

$$\frac{\partial F}{\partial x}=0$$

$$M_F:\{(u,v,x)\,|\,4x^3+2aux+bv=0\} \tag{7}$$

消去状态变量，可以得到由(u,v)组成的分叉集 S：

$$8a^3u^3+27b^2v^2=0 \tag{8}$$

根据已经建立的突变模型，设定一个阈值 ξ，利用新观测到的网络流量数据，应用相空间重构方法，可以计算出对应的特征量 $P(u,v)$。根据式(8)计算 P 到分叉集 S 的距离 D。若 $D\leqslant\xi$，则检测出网络流量的异常。ξ 阈值可以根据最开始 5 个子窗的特征量 P 到分叉集的距离的平均值进行计算，在实际计算中可以做相应的缩放。

采用常用 ROC (Relative Operating Characteristic)曲线来分析检测方法的效果特性。ROC 曲线的横坐标 Pfa 是误检率，纵坐标 Pd 是检出率。从图 1 可以看出，该方法在较低误检率的情况下，检出率较高，实验效果明显优于其它方法。

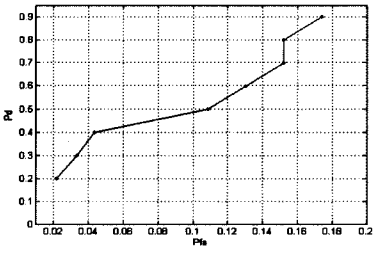

图 1　ROC 曲线

结束语　本文主要针对分布函数不变、分布参数变化这一类非平稳时间序列异常检测，根据其非平稳突变的特性采用超统计理论研究其分布参数的变化。再运用于非平稳的网络流量，根据其特点提出一种新的异常流量检测方法，该方法充分考虑到异常流量所引起的流量突变以及非平稳等特性，通过寻找反映流量变化的参数序列，变换了研究对象，将对复杂的网络流量数据的研究集中到对某个决定系统本质变化的参数序列的研究，既充分又避免了传统有统计模型参数检验计算的复杂性。经过多次实验获得良好的实验效果，并且降低了计算的复杂度。本文选取了慢变量参数序列来研究网流量的异常，取得了较好的效果。因此有理由相信超统计理论也可以应用于其它领域的非平稳时间序列分析。

参 考 文 献

[1] Davison A C. Modeling Excesses Over High Thresholds, With an Application[J]. Statistical Extremes and Applications, 1984: 461-482

[2] Mukherjee B, Heberlein L T, Levitt K N. Network Intrusion Detection[J]. IEEE Network, 1994, 8(3): 26-41

[3] Rasmussen P, Ashkar F, Rosbjerg D, et al. The POT Method for Flood Estimation: A review[M]. Hipel K W, ed. Stochastic and Statistical Methods in Hydrology and Environmental Engineering. Kluwer Academic Publishers, 1994: 15-26

(下转第 117 页)

参 考 文 献

[1] Wang Yao, Wenger S, Wen Jian-tao, et al. Error resilient video coding techniques[J]. IEEE Signal Processing Magazine, 2000, 17(4):61-82

[2] Wang Yao, Zhu Qin-fan. Error control and concealment for video communication: a review[J]. Proceedings of the IEEE, 1998, 86(5):974-997

[3] He Zhi-hai, Cai Jian-fei, Chen Chang-wen. Joint source channel rate-distortion analysis for adaptive mode selection and rate control in wireless video coding[J]. IEEE Transactions on CSVT, 2002, 12(6):511-523

[4] Ru Cong-chong, Yin Liu-guo, Lu Jian-hua, et al. UEP Video Transmission Based on Dynamic Resource Allocation in MIMO OFDM System[C]// Proceedings of EEE Wireless Communications and Networking Conference(WCNC'07). 2007:310-315

[5] Fang Tao, Chau L-P. Efficient content-based resynchronization approach for wireless video[J]. IEEE Transactions on Multimedia, 2005, 6(7):1021-1027

[6] 邱锦波,冯镔,喻莉,等. 基于宏块重要性测度的重同步方法[J]. 计算机科学,2010,37(5):81-83

[7] Lambert P, Neve W D, Dhondt Y, et al. Flexible macroblock ordering in H. 264/AVC [J]. Journal of Visual Communication and Image Representation, 2006, 17(2):358-375

[8] MPEG-4 Standard[S]. ISO IEC 14496-2, 2003

[9] Zhang Rui, Regunathan S L, Rose K. Video Coding with Optimal Inter/Intra-mode Switching for Packet Loss Resilience [J]. IEEE Journal on Selected Areas in Communications, 2000, 18(6):966-976

[10] Cheng Liang, Zarki M E. Perceptual quality feedback based progressive frame-level refreshing for robust video communication [C]// Proceedings of 2004 IEEE Wireless Communications and Networking Conference(WCNC'04). 2004:2047-2052

(上接第 83 页)

[12] Dunbar R I M. Coevolution of neocortical size, group size and language in humans[J]. Behavioral and Brain Sciences, 1993, 16(4):681-735

[13] Fu F, Chen X, Liu L, et al. Social dilemmas in an online social network: the structure and evolution of cooperation[J]. Physics Letters A, 2007, 371(1/2):58-64

[14] Milgram S. The small World problem[J]. Psychology Today, 1967, 2:60-67

[15] Watts D J, Strogatz S H. Collective dynamics of 'small—world' networks[J]. Nature, 1998, 393:440-442

[16] Miklas A G, Gollu K, Chan K, et al. Exploiting Social Interactions in Mobile System[C]// Proc. of 9th Intl. Conference on Ubiquitous Computing. Heidelberg: Springer, 2007:409-428

[17] Chaintreau A, Fraigniaud P, Lebhar E. Opportunistic Spatial Gossip over Mobile Social Networks[C]// Proc. of the 1st ACM SIGCOMM Workshop on Social Networks. New York: ACM, 2008:73-78

[18] Hui P, Crowcroft J, Yoneki E. BUBBLE Rap: Social Based Forwarding in Delay Tolerant Network[C]// Proc. of 9th ACM International Symposium on Mobile Ad-hoc Networking and Computing. New York: ACM, 2008:241-250

[19] Marti S, Ganesan P, Garcia-Molina H. SPROUT: P2P routing with social networks[C]// Proc. of International Workshop on Peer-to-Peer Computing & DataBases. Heidelberg: Springer, 2004:425-435

[20] Mislove A, Gummadi K P, Druschel P. Exploiting social networks for Internet search[C]// Proc. of the 5th Workshop on Hot Topics in Networks. 2006

[21] 郎君,秦兵,宋巍,等. 基于社会网络的人名检索结果重名消解[J]. 计算机学报,2009,32(7):1365-1373

[22] Yu Hai-feng, Gibbons P B, Kaminsky M, et al. SybilLimit: a near-optimal social network defense against Sybil attacks[C]// Proc. of the 2008 IEEE Symposium on Security and Privacy. Picataway, NJ: IEEE, 2007:885-898

[23] 陈端兵,万英,田军伟,等. 一种基于社会网络分析的 P2P 僵尸网络反制策略[J]. 计算机科学,2009,36(6):101-104

[24] Boykin P O, Roychowdhury V P. Leveraging social networks to fight spam[J]. Computer, 2005, 38(4):61-68

[25] Lam H Y, Yeung D Y. A Learning Approach to Spam Detection based on Social Networks[C]// Proc. of the Fourth Conference on Email and AntiSpam. Mountain View. California, USA, 2007

[26] 黄文良,刘勇,钟志强,等. 基于复杂网络的垃圾短信过滤算法[J]. 自动化学报,2009,35(7):990-996

[27] Aharony N, Reed D P, Lippman A. Social Area Networks: Data Networking of the People, by the People, for the People[C]// Proc. of International Conference on Computational Science and Engineering. February. Picataway, NJ: IEEE, 2009:1148-1155

(上接第 95 页)

[4] Pickands J. Statistical Inference Using Extreme Order Statistics [J]. The Annals of Statistics, 1975, 3:119-131

[5] Beck C, Cohen E G D. Supertatistics[J]. Physica A, 2003(332):267-275

[6] Cohen E G D. Supertatistics[J]. Physica D, 2004(193):35-52

[7] Beck C, Cohen E G D, Harry L. Swinner from time series to superstatistics[J]. Phys. Rev., 2005, 72:056133

[8] Mahoney M, Chan P. An analysis of the 1999 Darpa/Lincoln Laboratory Evaluation Data for Network Anomaly Detection[C]// Proceeding of Recent Advances in Intrusion Detection (RAID) Lecture Notes in Computer Science. 2003, 2820:220-237

[9] Ye N, Chen Q. An Anomaly Detection Technique Based on A ChiSquare Statistic for Detecting Intrusion into Information System, Quality and Reliability Engineering International[J]. 2001, 17:105-112

[10] Tan K M C, Maxion R A. Determining the Operational Limits of an Anomaly-based Intrusion Detector[J]. IEEE J. Sel. Areas Commun, 2003, 21(1):96-110

[11] Sommer R, Paxson V. Enhancing Byte-level Network Intrusion Detection Signatures with Context[C]// CCS '03: Proceedings of the 10th ACM Conference on Computer and Communications Security. 2003:262-271

[12] DARPA 99 Intrusion Detection Data Set Attack Documentation [DB/OL]. http://www.ll.mit.edu/IST/ideval/docs/1999/attackDB.html

[13] He Di, Leung H. Network Intrusion Detection Using CFAR Abrupt-Change Detectors[J]. IEEE Transactions on Instrumentation and Measurement, 2008, 57:490-497

[14] Paxson V, Floyd S. Wide Area Traffic: The Failure of Poisson Modeling[J]. IEEE/ACM Trans. Networking, 1995, 3(3):226-244

[15] Dupuis D J. Estimating the Probability of Obtaining Nonfeasible Parameter Estimates of the Generalized Pareto Distribution[J]. Journal of Statistical Computation and Simulation, 1996, 54:197-209

[16] Rasmussen P F. Generalized Probability Weighted Moments: Application to the Generalized Pareto Distribution[J]. Water Resources Research, 2001, 37(6):1745-1751