基于支持向量机的网络流量异常检测

温志贤[1]，李小勇[2]
(1. 天水师范学院 数理与信息科学学院，甘肃 天水 741000；
2. 上海交通大学 计算机科学与工程系，上海 200030)

摘 要：提出了一种基于支持向量机的网络流量异常检测方法. 分析了支持向量机的基本原理，结合网络流量异常检测的特点，讨论了异常检测的特征选择问题；提出了网络流量对称性、TCP 报文 SYN 和 SYN/ACK 对称性以及协议分布等具有鲁棒性的特征参数，描述了数据的预处理方法. 测试结果表明，所选特征参数可有效地检测网络攻击导致的流量异常变化，说明基于支持向量机的检测方法具有较好的泛化能力.

关键词：异常检测；入侵检测；支持向量机；端口扫描；网络安全
中图分类号：TP 393.07　　文献标识码：A　　文章编号：1001-988X(2005)03-0027-05

Network traffic anomaly detection based on support vector machine

WEN Zhi-xian[1]，LI Xiao-yong[2]
(1. School of Mathematics, Physics and Information Science, Tianshui Normal College, Tianshui 741000, Gansu, China;
2. Department of Computer Science and Engineering, Shanghai Jiaotong University, Shanghai 200030, China)

Abstract: A network traffic anomaly detection mechanism is presented based on support vector machine (SVM). Theory of SVM is introduced first, and then feature selection is discussed in depth. Many features, including symmetry of network traffic, symmetry of SYN and SYN/ACK packets, protocol distribution, are introduced in network traffic anomaly detection. And preprocessing of data is explained in detail. Experimental results show that the selected features can be used to detect traffic anomaly incurred by network attacks, and the detection mechanism based on SVM has good capability of generalization.

Key words: anomaly detection; intrusion detection; support vector machine; port scan; network security

随着信息技术的日益发展和计算机网络的普及应用，对网络流量管理提出了越来越高的要求. 网络流量的管理包括网络流量的建模和分析、流量预测、流量异常检测、流量工程等[1]. 经多年研究，对于流量异常检测已提出多种方法，但各种检测方法的不足在于选择的特征参数单一，因此检测网络流量异常的能力受到限制，而且需要采集大量数目的报文才能进行判别，导致检测的延迟较大. 笔者提出了一个基于支持向量机的网络异常检测方法.

1　支持向量机原理[2]

考虑简单的二值分类问题，给定符合某种未知概率分布 $F(x, y)$ 的训练数据集 (x_i, y_i)：

$$\{(x_1, y_1), (x_2, y_2), \cdots, (x_l, y_l)\},$$

其中 $x_i \in \mathbf{R}^N$，$y_i \in \{-1, +1\}$，l 为样本集的数目. 希望能设计一个最优分类器 $f_\alpha: \mathbf{R}^N \longrightarrow \{-1, 1\}$，$\alpha \in \Lambda$，不仅能使 2 类数据正确分开，而且使得分类间隔最大. 记分类面的方程为 $x \cdot w + b = 0$，则

收稿日期：2005-03-01；修改稿收到日期：2005-04-18
作者简介：温志贤 (1979—)，男，甘肃天水人，讲师. 主要研究方向为机算机网络.

$$y_i(w\cdot x+b)-1\geqslant 0,\quad i=1,\cdots,l, \tag{1}$$

此时分类间隔为 $2/\|w\|$. 使间隔最大等价于使 $\|w\|^2$ 最小，即

$$\min\varphi(w)=\frac{1}{2}(w\cdot w). \tag{2}$$

对于线性可分数据集，构造最优分类面的问题可以转化为在(1)式的约束下最小化(2)式，这是一个二次规划问题，该问题的解由以下 Lagrange 泛函的鞍点给出：

$$L(w,b,\lambda)=\frac{1}{2}(w\cdot w)-\sum_{i=1}^{l}\lambda_i(y_i(w\cdot x+b)-1), \tag{3}$$

其中 λ_i 为非负 Lagrange 系数. (3)式是一个二次凸函数，存在唯一的最优解. 在鞍点处，由于 w 和 b 的梯度为 0，所以

$$w=\sum_{i=1}^{l}\lambda_i y_i x_i, \tag{4}$$

$$\sum_{i=1}^{l}\lambda_i y_i=0. \tag{5}$$

将(4)、(5)式代入(3)式，就转化为一个较简单的二次规划问题，即在约束条件(5)、(6)下，最大化(7)式.

$$\lambda_i\geqslant 0,\quad i=1,\cdots,l, \tag{6}$$

$$W(\lambda)=\sum_{i=1}^{l}\lambda_i-\frac{1}{2}\sum_{i,j}\lambda_i\lambda_j y_i y_j(x_i\cdot x_j). \tag{7}$$

在求得 w 和 b 后，根据 $\mathrm{sgn}(w\cdot x+b)$ 即可判定 x 所属的分类.

在输入空间线性不可分的情况下，需在最大间隔分类面和最少错分样本之间取得折中，即构造一个软间隔的分类超平面，在条件(1)中增加一个松弛项 $\xi_i\geqslant 0$，即

$$y_i(w\cdot x_i+b)-1-\xi_i\geqslant 0. \tag{8}$$

相应的目标函数为求 $f(w,\xi)=\frac{1}{2}\|w\|^2+C\left|\sum_{i=1}^{n}\xi_i\right|$ 的最小，其中 $C>0$ 为常数，控制对错分样本的惩罚因子. 求解广义最优分类面的对偶问题与(7)式基本相同，只是条件约束变为 $0\leqslant\lambda_i\leqslant C$.

在输入空间是非线性情形下，统计学习理论引入了一个重要概念即核函数. 通过核函数将输入空间变换到一个高维特征空间，然后在特征空间中构造最优分类面实现分类. 核函数只要满足 Mercer 条件即可. 设核函数为 $K(x_i,x_j)$，则二次规划的目标函数为

$$W(\lambda)=\sum_{i=1}^{l}\lambda_i-\frac{1}{2}\sum_{i,j}\lambda_i\lambda_j y_i y_j K(x_i\cdot x_j), \tag{9}$$

相应的最终决策函数为

$$f(x)=\mathrm{sgn}(w\cdot x+b)=\sum_{\text{Support vector}}\lambda_i y_i K(x_i,x)+b. \tag{10}$$

只要选择合适的核函数 $K(x,y)$，就能确定不同类型的支持向量机.

2 网络流量异常检测的特征参数选择

特征选择对网络异常检测的准确性具有直接影响，选择时应多角度反映正常网络流量的多维特征，从而使系统能够检测多种类型的网络攻击. 选取以下特征用于网络流量异常检测.

1) 流量对称性. Internet 中的大多数网络应用都是基于交互的请求/应答方式，如 ICMP 的各种请求报文、UDP 的域名解析请求、基于 TCP 的 FTP 协议，以及目前应用最为广泛的 Web 访问协议 HTTP 等. 这种交互式的特征，以及 TCP 协议面向连接对接收到的数据进行确认的机理[3]，决定了从一个网络发出的报文数和达到该网络的报文数应基本相同，即网络流量具有对称性. 而当发生 DoS 攻击，大规模网络扫描等攻击时，网络流量的对称性会遭到破坏.

通过对 1999 年 DARPA 评测网络入侵检测系统所使用的数据集前 7 d 正常数据的分析[4]，做出图 1~4. 图 1 为任意一天一段时间内进、出网络的报文数的统计，从中可以看出网络流量的对称性.

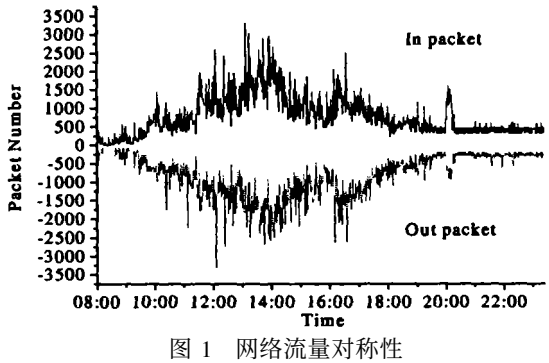


图 1 网络流量对称性

设 P_i, P_o 分别表示方向为 in 和 out 的报文数，则网络流量的对称性 F_n 定义为

$$F_n=\frac{P_o}{P_o+P_i}. \tag{11}$$

2) SYN/ACK 对称性. 由 TCP 协议的“三次握手”机制可知[3]，在正常情况下，当接收到来

自 Client 的连接请求(SYN 报文)时，Server 将发送 SYN/ACK 报文作为应答，即 SYN 报文和 SYN/ACK 报文应一一对应. 因此，在一段时间内，网络中出现的 SYN 和 SYN/ACK 报文数目应基本相同. 图 2 为任意一段时间内 SYN 和 SYN/ACK 报文数的统计，直观地显示了其对称性.

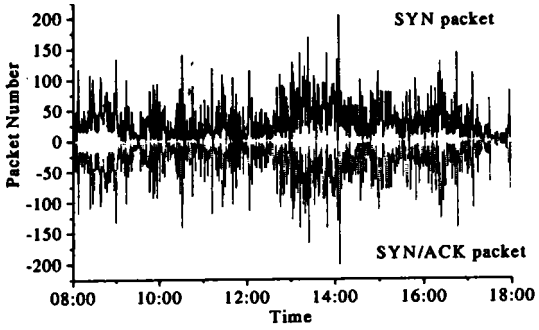


图 2　SYN 和 SYN/ACK 对称性

定义 SYN 和 SYN/ACK 对称性为

$$S_n = \frac{P_{SYN}}{P_{SYN} + P_{SYN/ACK}}, \qquad (12)$$

其中，P_{SYN}，$P_{SYN/ACK}$ 分别为单位时间内 SYN 和 SYN/ACK 报文的统计.

3) 协议分布. 对网络流量的统计分析表明，TCP, UDP 和 ICMP 报文在网络流量中的分布具有很强的规律性，TCP 报文一般在总网络流量中占有较高比例(对于不同的网络，分布可能不同. 如在有 DNS 服务器的子网内，UDP 报文可能会占较大比例). 若发生基于 UDP 或 ICMP 协议的洪流攻击，不同协议报文的分布即会发生明显变化. 图 3 显示了正常情况下某几天 TCP 包在总网络流量中的比例(p_{TCP}为某时刻 TCP 的报文数，p_{IP}为某时刻总报文数)，表明 TCP 报文的分布具有很强的规律性.

4) 异常报文统计. 异常报文如 TCP RST，ICMP 目的主机或端口不可达等，往往是检测网络攻击的重要依据[5]. 对 TCP 报文的统计分析表明，在正常情况下，TCP Flag 只呈几个固定的值，如图 4 所示. 它们与 flag 的对应关系为 2：SYN；16：ACK；17：ACK/FIN；18：ACK/SYN；24：ACK/PUSH. 除以上几种取值外的其他 TCP 报文都将作为异常报文. 在所有报告网络异常的 ICMP 报文中，如目的主机不可达、目标端口不可达、TTL 为 0 等，也作为异常报文记入统计.

5) 流量强度和服务分布. 陈硕等人[6]的研究表明，单位时间内的包数能够较好地表征网络流量

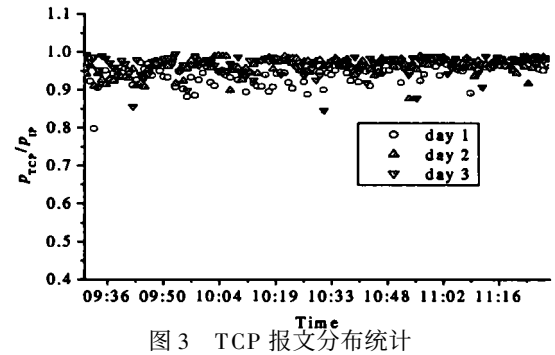


图 3　TCP 报文分布统计

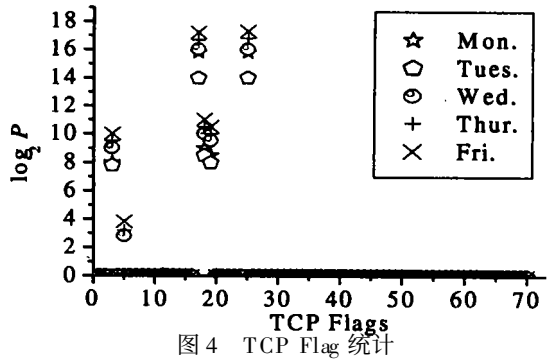


图 4　TCP Flag 统计

强度，而字节数则具有很大的随机性，因此将包数作为网络流量强度的一个特征. 同时，也将“IP 流”的统计作为网络流量强度的特征之一. IP 流定义为“任意 2 个主机间的通信报文的序列”，它只与通信主机的地址有关，与通信内容无关. 当发生 DDoS 攻击、大规模网络拓扑扫描、蠕虫等网络攻击时，都会导致 IP 流急剧增加. 分别对基于 TCP 和 UDP 的各种网络服务(端口范围 1~1024)进行统计，选择了使用最为频繁的多种网络服务(占网络总流量的 80%以上)，统计它们在总网络流量中的比例作为另一个特征.

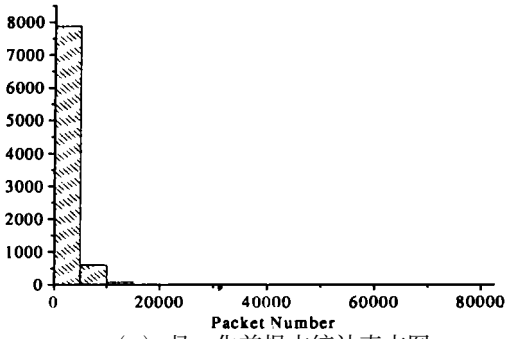
3　数据预处理

在对支持向量机进行训练以前，首先要求对数据进行归一化[7]. 文中所选特征中，流量对称性、协议分布和服务分布的取值范围为(0, 1)，数据的离散度较小；而流量强度、异常报文统计的取值为自然数，数据离散度很高，如图 5(a). 如果使用线性归一化方法，对于离散度较大的属性，在判别中的有效性将会下降. 因此，对 LibSVM 进行了扩展，实现 Sigmoid 归一化方法[8]，其定义为

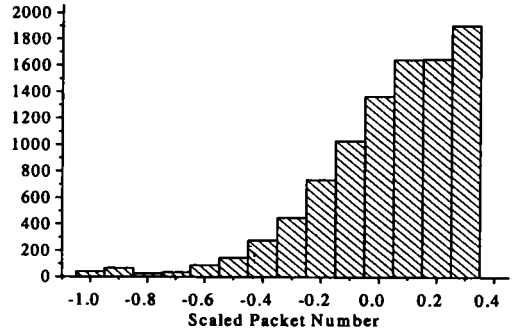
$$v = \frac{v - e^{\alpha}}{1 + e^{\alpha}}, \quad \alpha = \frac{v - m_A}{s_A}, \qquad (13)$$

其中 m_A, s_A 为属性 A 的均值和标准差，v 为归一化后的值. 图 5(b)为使用 Sigmoid 归一化后报文

数统计的直方图. 从中可以看出经过归一化处理后, 其分布更趋于均衡.



(a) 归一化前报文统计直方图



(b) 归一化后报文统计直方图

图5 压缩前后报文数统计分布对比

在协议分布中, 最初共选择了2个特征: TCP报文在IP报文中的比例和UDP报文在IP报文中的比例. 对统计数据的分析表明, 2个属性之间存在强相关性(负相关, 相关系数为0.86). 为消除冗余信息, 提高检测系统运行效率, 删除了UDP报文在IP报文中的比例这一属性.

在网络流量很小时(每分钟报文数小于100), 文中提出的流量对称性、协议分布、服务分布等特性不再成立. 因此, 在进行训练和测试时, 过滤了网络流量很小的统计数据. 绝大多数情况下网络流量都会大于此值, 因此不会对应用造成影响. 在一些统计数据中存在异常值, 如当TCP报文数为0时, SYN和SYN/ACK对称性为无效值, 基于TCP的网络服务的分布也为无效值. 对这些异常值, 用每种属性正常训练数据的平均值替代. 异常报文统计在多数情况下的值都为0, 为了消除对检测算法性能的影响, 将值为0的数据用0.1代替.

4 实验结果

选用1999年DARPA评测网络入侵检测系统所使用的数据集[4], 对基于SVM的网络流量异常检测系统进行训练和测试. SVM工具包使用LibSVM[7]. 测试数据集共由35 d的原始报文组成, 其中第1周、第2周为不含攻击的训练报文, 第3周为含有攻击的训练报文, 第4、第5周的数据为测试报文, 其中含有第2周未曾出现的网络攻击. 测试数据集中的攻击分为4类[4]: DoS: 拒绝服务攻击, 如SYN Flooding, Smurf攻击; R2L: 远程权限获取, 如口令猜测; U2R: 各种权限提升, 如各种本地和远程Buffer Overflow攻击; Probe: 各种端口扫描和漏洞扫描.

在上述攻击中, Flooding类型的DoS攻击, 高强度的端口扫描和漏洞扫描都会导致网络流量特征的变化. 对报文进行解析后, 以60 s为单位进行统计, 在过滤了小流量的统计数据后(每分钟报文数小于100), 共得到约14 000条训练数据记录, 从中随机选取700条作为训练数据, 其中120条为"异常"数据. 对测试数据集进行处理后, 共得到约10 000条测试记录, 按照数据说明对其进行标记. 正常数据标记为"+1", 异常数据标记为"-1".

首先将训练数据和测试数据统一进行预处理, 包括对强度属性和异常报文统计使用Sigmoid进行归一化, 对其他属性进行最大-最小值归一化, 对异常值进行处理. 然后用训练数据对支持向量机进行训练, 用测试数据对其进行测试.

表1 使用不同核函数的测试结果

训练参数		检测结果			
SVM类型	SVM参数	支持向量数	检测率	误报警率	运行时间
Linear	$C=1$, $g=0.2$	92	98.8%	2.30%	0.53 s
Polynomial	$C=50$, $g=0.1$, $d=3$	104	97.2%	1.50%	0.94 s
Radial Basis	$C=1$, $g=0.1$	135	98.2%	0.11%	0.82 s
Sigmoid	$C=1$, $g=0.1$	170	96.3%	0.05%	1.27 s

表2 使用RBF核函数在不同参数下的测试结果

训练参数	支持向量数	检测率	误报警率	运行时间
$C=0.1$, $g=1.5$	133	97.9%	0.76%	0.82 s
$C=1$, $g=0.1$	135	98.2%	0.11%	0.83 s
$C=10$, $g=1$	33	96.1%	2.9%	0.54 s
$C=50$, $g=2$	39	95.7%	3.6%	0.55 s

测试结果如表1、表2所示. 其中表1为使用不同的核函数测试的结果, 表2为使用不同参数对RBF核函数的测试结果.

测试结果显示, 基于支持向量机的网络异常检测方法不仅可以有效地检测Neptune, TCPReset,

Smurf, Email Bomb 等各种 DoS 攻击，和 Queso (Nmap), Satan, Msan, Reset Scan, Port Sweep, SAINT 等各种高强度的扫描行为，同时误报警率较低. 这表明文中选择的特征参数能够有效地检测网络攻击导致的流量异常变化，而且说明基于支持向量机的检测方法具有较好的泛化能力，能够检测到训练中未出现的新攻击.

在不同的核函数中，Sigmoid 误报警率最低，RBF 的综合性能较好. 由于流量的采样周期为 60 s，SVM 方法对每条记录的平均检测时间在毫秒级，说明文中实现的网络流量异常检测方法也具有很高的实时性.

参考文献:

[1] 邹伯贤，李忠诚. 基于 AR 模型的网络异常检测[J]. 微电子学与计算机，2002，(12)：1—6.

[2] Jrisitianini N, Shawe-Taylor J. 支持向量机导论[M]. 李国正，王　猛，曾华军 译. 北京：电子工业出版社，2004. 53—79.

[3] Jon Postel. RFC 793[A]. DARPA. *Transmission Control Protocol—DARPA Internet Program Protocol Specification*[C]. Cacifornia: Information Sciences Institute, 1981. 7—52.

[4] Licoln Laboratory, Massachusetts Institute of Technology. DARPA intrusion detection evaluation [EB/OL]. http://www.ll.mit.edu/IST/ideval/index.html, 2003-09-16.

[5] Tanenbaum A S. 计算机网络[M]. 第 4 版. 潘爱民 译. 北京：清华大学出版社，2004. 437—472.

[6] 陈　硕，安常青，李学农. 分布式入侵检测系统及其认知能力[J]. 软件学报，2001，(2)：225—232.

[7] Chang Chih-Chung, Lin Chih-Jen. LIBSVM: a library for support vector machines [EB/OL]. http://www.csie.ntu.edu.tw/~cjlin/libsvm/, 2004-02-25.

[8] HAN Jia-wei, Kamber M. 数据挖掘：概念与技术[M]. 范　明，孟小峰 译. 北京：机械工业出版社，2001. 209—217.

(责任编辑　惠松骐)

(上接第 26 页)

如果在嵌入原始图像之前对秘密信息进行依靠密钥控制的随机置换，可进一步增加隐藏信息的安全性. 由上述实验结果可见，算法能成功地实现较多数据秘密信息的盲隐藏，在 256×256×8 bit 的公开图像中可以隐藏 128×64 的二值图像，而文献[6]在一幅 512×512×8 bit 图像中嵌入的数据不超过 800 bit.

参考文献:

[1] Wayner P. 隐显密码学[M]. 第 2 版. 杨立平，严　毅，何晓辉 译. 北京：电子工业出版社，2003. 5—20.

[2] QI Dong-xu, ZOU Jian-cheng, HAN Xiao-you. A new class of scrambling transformation and its application in the image information covering[J]. *Sciences in China* Ser E, 2000, **43**(3): 304—312.

[3] 柳葆芳. 基于融合的数据隐藏算法[J]. 电子学报，2001，**29**(11)：1445—1448.

[4] 张贵仓，王让定，章毓晋. 基于迭代混合的数字图像隐藏技术[J]. 计算机学报，2003，**26**(5)：569—574.

[5] LIN S D, Shie S C, CHEN C F. A DCT-based image watermarking with threshold embedding[J]. *International Journal of Computers and Applications*, 2003, **25**(2): 130—135.

[6] WANG Y, LIN S. Wavelet tree quantization copyright protection watermarking [J]. *IEEE Transactions of Image Processing*, 2004, **13**(2): 154—163.

(责任编辑　惠松骐)