文章编号: 1001-0920(2005)03-0251-06

# 基于KPLS的网络入侵特征抽取及检测方法

杨辉华[1,2], 王行愚[1], 王 勇[1,3], 何 倩[3]
(1. 华东理工大学 信息科学与工程学院, 上海 200237; 2. 桂林电子工业学院 计算机系, 广西 桂林 541004; 3. 桂林电子工业学院 网络信息中心, 广西 桂林 541004)

**摘 要**: 从特征抽取的角度研究提高入侵检测性能问题, 提出应用核偏最小二乘(KPLS)进行入侵特征抽取和检测的方法. 其优点在于KPLS能非线性地抽取输入特征的多个正交分量, 并保持与输出类别的相关性, 可同时完成入侵特征抽取和判别. 将该方法应用于基于Linux主机的入侵检测实验, 取得了比SVM和KPCR等方法更好的效果.

**关键词**: 核方法; 核偏最小二乘; 非线性特征抽取; 异常检测; 支持向量机
**中图分类号**: TP18; TP393 **文献标识码**: A

# KPLS approach for network intrusion feature extraction and detection

*YANG Hui-hua*[1,2], *WANG Xing-yu*[1], *WANG Yong*[1,3], *HE Qian*[3]
(1. College of Information Science and Engineering, East China University of Science and Technology, Shanghai 200237, China; 2. Department of Computer, Guilin University of Electronic Technology, Guilin 541004, China; 3. Network Information Center, Guilin University of Electronic Technology, Guilin 541004, China. Correspondent: YANG Hui-hua, E-mail: yang98@ecust.edu.cn)

**Abstract**: A novel KPLS based network intrusion feature extraction and detection approach is put forward, among which KPLS serves as simultaneously a non-linear feature extractor and a decision maker. KPLS approach bears the merits that it can not only extract orthogonal score vectors from explanatory variables, but also remain good correlation with response variables. The feature extraction procedure and decision making procedure can be achieved at one time. The novel method is applied to an up-to-date Linux-hosted IDS experimental system and better performance is attained in comparison to SVM and KPCR etc.

**Key words**: kernel method; kernel partial least squares (KPLS); nonlinear feature extraction; anomaly detection; support vector machines

## 1 引 言

网络空间是虚拟与现实之间的一种特殊空间, 入侵检测和信息安全控制是其中的重要研究课题[1]. 异常检测根据网络入侵状态偏离正常使用这一原理进行检测, 能检测出新的攻击方式. 麻省理工学院(MIT)林肯实验室公布了两套标准测试数据, 其中最著名的是KDD Cup 99 IDS(以下简称KDD 99)数据集. 在入侵检测领域, 现已提出基于神经网络[2]、支持向量机[2,3]、D-S证据理论[4]等检测方法, 这些算法主要使用KDD 99数据集. 然而, 随着网络软硬件环境和攻击手段的快速更新, 该数据集的局限在所难免. 因此, 在新的网络环境中进行入侵检测仿真便显得十分必要.

入侵特征分析可对入侵检测算法的性能产生重要影响. 目前主要有两大类特征分析方法: 特征选择和特征抽取. 特征选择根据某种性能判据, 从所有输

**收稿日期**: 2004-05-09; **修回日期**: 2004-08-02
**基金项目**: 国家重点基础研究发展规划项目(2002CB312200); 国家自然科学基金项目(69974014); 教育部高校博士点基金项目(20040251010).
**作者简介**: 杨辉华(1972—), 男, 湖南澧县人, 讲师, 博士生, 从事智能信息处理、模式识别等研究; 王行愚(1944—), 男, 上海人, 教授, 博士生导师, 从事智能控制、网络信息安全等研究.

入特征中选择重要特征, 并去掉次要特征, 这有利于缩短检测时间, 发现某类攻击的本质特征. 这方面的代表性工作参见文献[2]. 但特征选择方法是一个离散过程(变量或者保留, 或者丢弃), 预测模型常常表现出较高的方差. 特征抽取是一种连续的方法, 不会因为变量多而过多地降低性能[5].

特征抽取并不从输入特征中显式地去掉某些特征, 而是对输入特征进行线性或非线性变换, 从中抽取代表性分量, 用这些分量代替原输入特征, 进一步分析和判别. 当数据产生的机理不明确, 输入维数高或样本数据少时, 特征抽取有一定的优势. 目前, 从特征抽取的角度研究入侵检测问题还少有报道.

本文主要目的在于探讨入侵检测的特征抽取方法对入侵检测性能的影响, 并通过实验进行验证. 为达此目的, 一方面模仿MIT 的入侵检测实验环境, 建立一个基于Linux 主机的IDS 环境, 进行仿真实验; 另一方面探讨将KPLS 方法直接用于入侵特征抽取和判别, 并将结果与 SVM 和 KPCR 等方法进行比较.

## 2 基于PLS 和KPLS 的特征抽取、回归与分类方法

首先介绍线性PLS 方法, 然后通过核技巧引入KPLS. 本文对PLS 和KPLS 均从特征抽取、回归和分类3 个方面进行论述.

### 2.1 线性PLS

#### 2.1.1 PLS 特征抽取

考虑线性PLS 建模中两个数据块之间关系的一般情况. 设 $x \in X \subset R^N$, $y \in Y \subset R^M$ 分别为数据块 $X$ 和 $Y$ 中的 $N$ 维和 $M$ 维行向量. PLS 通过潜变量对两个数据块的关系进行建模, 它将 $n \times N$ 零均值矩阵 $X$ 和 $n \times M$ 零均值矩阵 $Y$ 分解为如下形式:

$$X = TP^{\mathrm{T}} + F, \tag{1}$$

$$Y = UQ^{\mathrm{T}} + G. \tag{2}$$

其中: $T$ 和 $U$ 是由抽取的 $p$ 个得分矢量组成的 $n \times p$ 矩阵, $P$ 和 $Q$ 为负载矩阵, $F$ 和 $G$ 为残差矩阵.

抽取 $T$ 和 $U$ 有多种方法. PLS 的经典形式是基于非线性迭代偏最小二乘(NIPALS) 算法. 下面简要介绍通过矩阵本征值问题求解 $T$ 和 $U$ 的方法. 权矢量 $w$ 对应于如下本征值问题:

$$X^{\mathrm{T}}YY^{\mathrm{T}}Xw = \lambda w \tag{3}$$

第1 个本征矢量[6]. $X$ 的得分矢量 $t$ 可通过计算 $t = Xw$ 而得到.

通过在式(3) 两边同时左乘 $X$, 即可发现得分矢量 $t$ 还可通过求解如下本征值问题:

$$XX^{\mathrm{T}}YY^{\mathrm{T}}t = \lambda t \tag{4}$$

第1 个本征矢量而求得 $Y$ 的得分矢量 $u$ 的估计为

$$u = YY^{\mathrm{T}}t \tag{5}$$

PLS 特征抽取分量的个数最多等于 $X$ 的秩, 即 rank($X$). 对从 $X$ 和 $Y$ 中抽取出的分量 $T$ 和 $U$, 可进一步运用其他方法处理, 如运用 SVM 进行回归、分类等.

#### 2.1.2 PLS 回归与分类

PLS 回归直接利用以上步骤抽取的特征进行回归分析. PLS 回归模型用矩阵形式可表示为

$$\hat{Y}\_{\mathrm{reg}} = XB. \tag{6}$$

其中 $B$ 是一个 $N \times M$ 的回归系数矩阵, 即

$$B = X^{\mathrm{T}}U(T^{\mathrm{T}}XX^{\mathrm{T}}U)^{-1}T^{\mathrm{T}}Y. \tag{7}$$

如果对式(6) 简单地取符号函数 sign, 则得 PLS 模式识别模型

$$\hat{Y}\_{\mathrm{pr}} = \mathrm{sign}(XB). \tag{8}$$

### 2.2 非线性KPLS

基于Mercer 核的SVM 在机器学习领域取得了巨大成功, 引发人们将传统的各种可用内积表达的线性方法"核化", 从而成为非线性方法. Rosipal 和Trejo[7] 将线性PLS 方法推广为非线性KPLS 方法.

#### 2.2.1 KPLS 特征抽取

KPLS 的基本思想是首先选择Mercer 核 $K(\cdot, \cdot)$, 该核隐含的非线性变换 $\varphi: x \mapsto \varphi(x)$, 将输入空间 $X$ 变换到高维Mercer 特征空间 $F$, 并用 $\Phi$ 代表 $X$ 空间的数据映射到 $S$ 维特征空间 $F$ 所得的 $n \times S$ 矩阵($S$ 可以为无穷大); 然后使用核技巧, 得 $K = \Phi\Phi^{\mathrm{T}}$, $K_{jj} = K(x_i, x_j)$ 是 $n \times n$ 的 Gram 矩阵. 类似地, 选择核 $K_1(\cdot, \cdot)$, 对应于映射 $\psi: y \mapsto \psi(y)$, 将输出 $Y$ 映射到特征空间 $F_1$, $\Psi$ 为 $F_1$ 空间 $n \times S_1$ 的矩阵, $K_1 = \Psi\Psi^{\mathrm{T}}$ 是 $n \times n$ 的 Gram 矩阵.

对 PLS 方法核化的关键是将 PLS 问题表达为数据的内积形式, 再用 Gram 矩阵直接代替内积, 即为非线性 KPLS 方法. 观察式(4) 和(5), 由于 $XX^{\mathrm{T}} = \langle X, X\rangle$, $YY^{\mathrm{T}} = \langle Y, Y\rangle$ 是内积形式, 用 $K = \langle\Phi, \Phi\rangle$ 代替 $XX^{\mathrm{T}}$, $K_1 = \langle\Psi, \Psi\rangle$ 代替 $YY^{\mathrm{T}}$, 即可从式(4) 和(5) 推出 KPLS 关于 $t$ 和 $u$ 的估计

$$KK_1t = \lambda t, \quad u = K_1t \tag{9}$$

类似于 PLS 方法, 假定使用零均值的非线性 KPLS 模型, 为此需要中心化特征空间的元素, 方法如下:

$$K = \left(I_n - \frac{1}{n}1_n1_n^{\mathrm{T}}\right)K\left(I_n - \frac{1}{n}1_n1_n^{\mathrm{T}}\right). \tag{10}$$

其中: $I_n$ 是 $n$ 维的单位矩阵, $1_n$ 是元素全为1 的 $n \times 1$ 列向量. $K_1$ 的中心化方法同上.

抽取得分矢量 $t$ 和 $u$ 后, 矩阵 $K$ 和 $K_1$ 减去基于 $t$ 和 $u$ 的秩1 逼近, 再抽取新的 $t$ 和 $u$, 如此下去, 直至达到需要的个数为止. 最多可抽取 rank($K$) 个分量.

$t$ 不同的消去方法对应于不同的 KPLS 算法, 并决定 $t$ 是否正交

**2.2.2　KPLS 回归与分类**

与 PLS 方法类似, KPLS 的回归系数矩阵

$$B = \Phi^{\mathrm{T}} U (T^{\mathrm{T}} K U)^{-1} T^{\mathrm{T}} Y. \quad (11)$$

对训练数据进行回归, 有

$$\hat{Y} = \Phi B = K U (T^{\mathrm{T}} K U)^{-1} T^{\mathrm{T}} Y = T T^{\mathrm{T}} Y. \quad (12)$$

对于测试样本, 有

$$\hat{Y}_t = \Phi_t B = K_t U (T^{\mathrm{T}} K U)^{-1} T^{\mathrm{T}} Y = T_t T^{\mathrm{T}} Y, \quad (13)$$

其中 $K$ 和 $K_t$ 需要经过中心化处理

同理, 对式(12) 和(13) 取符号函数即可用于分类

## 3　基于核正交 PLS 的机器学习方法

下面介绍一种变形的 KPLS 方法 —— 核正交 PLS 算法(仍称为 KPLS)[6], 它用于模式识别问题具有更好的理论解释

Barker 和 Rayens[8] 从统计学的观点分析指出: PLS 方法与 Fisher 的 LDA 和 CCA 之间存在紧密的联系　一个基本事实是 PLS 可看成一种带罚的 CCA, 即

$$[\mathrm{cov}(t, u)]^2 = [\mathrm{cov}(Xw, Yc)]^2 = \mathrm{var}(Xw)[\mathrm{corr}(Xw, Yc)]^2 \mathrm{var}(Y_c), \quad (14)$$

其中 $X$ 和 $Y$ 空间均带 PCA 形式的罚 Barker 和 Rayens 建议在 PLS 判别分析时去掉无意义的 $Y$ 空间罚 $\mathrm{var}(Yc)$, 此时 PLS 变形为正交 PLS 方法　正交 PLS 方法将式(4) 变成如下本征值问题:

$$XX^{\mathrm{T}} Y (Y^{\mathrm{T}} Y)^{-1} Y^{\mathrm{T}} t = \lambda t, \quad (15)$$

而核正交 PLS 将式(4) 变为

$$KY (Y^{\mathrm{T}} Y)^{-1} Y^{\mathrm{T}} t = K \tilde{Y} \tilde{Y}^{\mathrm{T}} t = \lambda t, \quad (16)$$

$$u = \tilde{Y} \tilde{Y}^{\mathrm{T}} t \quad (17)$$

其中 $\tilde{Y} = Y (Y^{\mathrm{T}} Y)^{-1/2}$. 核正交 PLS 能抽取 $\mathrm{rank}(K)$ 个得分矢量 $t$, 并且 $t$ 互相正交

与基于仅在 $F$ 空间中最大化数据变异原则的 KPCA 方法相比, 核正交 PLS 能提供一个理论性更好的降维方法　PLS 与 CCA 及 Fisher 线性判别分析的紧密联系, 促使本文将核正交 PLS 方法应用于分类

对于只有一个响应变量($M = 1$) 的二值分类问题, 以 Jong 提出的 PLS 的快速算法 SIMPLS 为基础, 并参考文献[6], 本文从机器学习的角度, 给出如下核正交 PLS(仍称为 KPLS) 学习算法和测试算法　其伪代码如下:

**KPLS 训练算法**

输入: $X$ (零均值, 标准差为 1)

　　$Y$ (对分类编码为 ± 1; 对回归零均值, 标准差为 1)

　　$p$ (得分矢量的个数)

输出: $Y\_hat\_reg$, $Y\_hat\_pr$, $T$; $C$, $U$, $K$, $mean\_Y$

处理: 计算 $Y$ 的均值 $mean\_Y$, 对 $Y$ 零均值化;

　　根据给定的核及核参数, 求 $X$ 对应的核矩阵 $K$, 对 $K$ 中心化

　　初始化 $K_{res} = K$

　　for $i = 1$ to $p$

　　$t_i = K_{res} Y$

　　$t_i \leftarrow t_i / \| t_i \|$

　　$u_i = Y (Y^{\mathrm{T}} t_i)$

　　$K_{res} \leftarrow K_{res} - t_i (t_i^{\mathrm{T}} K_{res})$

　　$Y \leftarrow Y - t_i (t_i^{\mathrm{T}} Y)$

　　end

　　$T = [t_1, t_2, \cdots, t_p]$, $U = [u_1, u_2, \cdots, u_p]$

　　$C = T^{\mathrm{T}} * Y$

　　计算 $Y\_hat\_reg = T * C + mean\_Y$

　　计算 $Y\_hat\_pr = sign(Y\_hat\_reg)$

**KPLS 测试算法**

输入: $X_t$ (零均值, 标准差为 1)

　　$C$, $T$, $U$, $K$, $mean\_Y$

输出: $Yt\_hat\_reg$, $Yt\_hat\_pr$, $T_t$

处理: 计算中心化的测试样本 Gram 矩阵 $K_t$

　　投影测试样本, 即 $T_t = K_t U (T^{\mathrm{T}} K U)^{-1}$

　　计算 $Yt\_hat\_reg = Tt * C + mean\_Y$

　　计算 $Yt\_hat\_pr = sign(Yt\_hat\_reg)$

对于模式识别问题, 输出取 $Yt\_hat\_pr$; 对于回归问题, 输出取 $Yt\_hat\_reg$; 对于特征抽取问题, 输出为 $T$ 和 $T_t$, 它们是从输入空间 $X$ 中抽取的相互正交的特征, 可作为其他系统(如 ANN, SVM 等) 的输入, 进一步进行分类、回归或其他处理

## 4　基于 KPLS 的网络入侵检测实验系统

### 4.1　实验系统方案设计

为验证 KPLS 特征抽取和分类算法的有效性, 并考虑到原有的攻击方法不再有效, 已发展了多种新的入侵方法, 因此本文模仿 MIT 林肯实验室对入侵检测系统进行测试的著名实验, 建立了一个入侵检测实验环境

本实验参考了 MIT 测试实验的基本思想, 即遵照 MIT 提出的建立 IDS 测试系统的重要原则: 网络拓扑结构相同; 网络攻击策略相同; 参照 MIT 的特征选择方法, 并且充分利用审计数据　本实验与 MIT 实验不同的是: 在特征选取上除网络特征外, 同时提取一部分主机特征; 网络软硬件环境是当今的主流配置; 攻击工具是最新的

设计实验环境如图 1 所示. 在服务器(192.168.67.21)上安装 Red Hat Linux7.3 操作系统, 开启 Web, FTP, SSH, Telnet 等网络服务. 在服务器上进行入侵检测的数据采集, 其他局域网内和局域网外的主机与服务器建立正常的网络连接, 形成服务器正常的基本流量

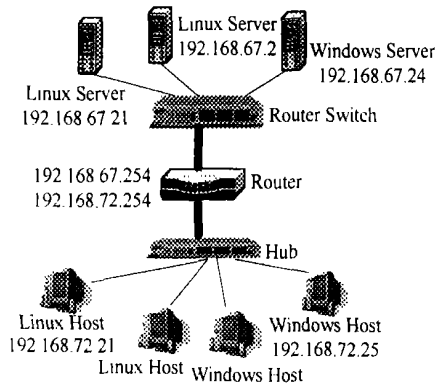


图 1 入侵检测实验环境

### 4.2 网络攻击设计

针对图 1 所示的实验系统, 从网络上收集一些有效的新式攻击工具, 对服务器实施以下攻击: 扫描和远程密码的蛮力攻击; 远程溢出攻击, 试图获取系统管理员权限; 拒绝服务攻击; 本地溢出攻击, 试图提升用户权限; 安装后门, 隐藏入侵痕迹. 在 192.168.72.0 网段选择一台 Linux 主机(192.168.72.32)实施主要攻击, 选择其他机器偶尔进行扫描等攻击. 实施攻击的同时记录攻击时间和攻击状态.

### 4.3 基于 KPLS 的入侵检测方案设计

为实现在上述实验环境下的入侵检测, 基于 KPLS 算法设计了如下入侵检测方案: 数据采集程序从 Linux 目标主机采集原始数据, 经特征获取程序加工产生矢量形式的样本数据(输入特征), 送 KPLS 检测器非线性抽取重要特征并进行判别, 检测结果按 XML 格式提交到事件数据库. 基于 KPLS 的入侵检测器的设计是整个 DS 系统的核心, 具体过程如图 2 所示

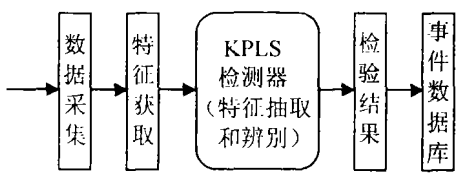


图 2 基于 KPLS 的入侵检测系统结构

### 4.4 入侵检测系统的特征选取

参考林肯实验室对 DS 测试的有关方法[9], 本文以 1 min 为单位进行特征采集并检测. 通过对 Linux 主机的入侵特性进行分析, 并参考 KDD 99 数据集的经验, 选取 5 个类别共 29 个参数作为入侵检测器的输入特征, 各特征含义见表 1.

表 1 入侵检测系统采集的特征

| 特征类别 | 特 征 含 义 | 特征取值 |
|---|---|---|
| 用户级特征 | 本分钟检测的用户登录失败次数 | 整数 |
| | 上一分钟检测用户登录失败次数 | 整数 |
| 系统级特征 | 本分钟交换分区利用率 | 实数 |
| | 本分钟物理内存利用率 | 实数 |
| | 上一分钟物理内存利用率 | 实数 |
| | 本分钟系统级 CPU 利用率 | 实数 |
| | 上一分钟系统级 CPU 利用率 | 实数 |
| | 本分钟用户级 CPU 利用率 | 实数 |
| | 上一分钟用户级 CPU 利用率 | 实数 |
| 进程级特征 | 本分钟系统运行进程总数 | 整数 |
| | 上一分钟系统运行进程总数 | 整数 |
| | 当前系统运行进程总数 | 整数 |
| | FTP 相关进程数 | 整数 |
| | FTP 进程 CPU 资源占用率 | 实数 |
| | FTP 进程内存占用率 | 实数 |
| 端口连接特征 | TCP 开放端口数目 | 整数 |
| | UDP 开放端口数目 | 整数 |
| | RAW 开放端口数目 | 整数 |
| | 非 TELNET, FTP, HTTPD 端口已建立连接数目 | 整数 |
| | TELNET 端口等待连接数目 | 整数 |
| | TELNET 端口已建立连接数目 | 整数 |
| | TELNET 端口处于其他状态数目 | 整数 |
| | FTP 端口等待连接数目 | 整数 |
| | FTP 端口已建立连接数目 | 整数 |
| | FTP 端口处于其他状态数目 | 整数 |
| | HTTPD 端口等待连接数目 | 整数 |
| | HTTPD 端口已建立连接数目 | 整数 |
| | HTTPD 端口处于其他状态数目 | 整数 |
| 网络延迟特征 | 网络延迟 | 实数 |

### 4.5 入侵检测系统的特征采集

要得到上面定义的入侵特征, 需要从 Linux 主机中采集有关原始信息并进一步预处理, 才能作为 KPLS 检测器的输入数据. 实验中设计了 4 个脚本程序共同完成数据采集任务, 每分钟产生一条样本数据, 它是一个 29 维的矢量, 属性或取离散值, 或取连续值, 范围相差很大. 通过进行归一化处理, 使均值为零, 标准差为 1, 从而成为 KPLS 的输入特征.

### 4.6 KPLS 模型选择

如前所述, KPLS 有训练和测试两种基本工作模式. 本文使用 RBF 核, 此时 KPLS 算法有两个重要参数需要确定: 一个是 RBF 核宽度参数 $\sigma^2$, 另一个是抽取的分量个数 $p$. 通过交叉验证法来确定这两个参数, 使用格点搜索法对最优参数进行搜索. 检测前需要调整并训练 KPLS 检测器, 使之工作在最佳点

### 4.7　攻击与检测

如图1 所示, 首先同步全网时钟, Linux 服务器提供多种正常服务, 通过一些主机产生不同的背景流量, 并由多台计算机轮流发起攻击. 为便于对照, 发起攻击时严格记录发起和结束时间、攻击类型及具体攻击工具. 这些信息与从Linux 服务器取得的特征样本进行对比, 从而确定哪些时间对应的样本是攻击, 哪些是正常网络活动. 在不同时段、不同网络条件下进行攻击和数据采集, 累计527 min, 相应取得527 条样本数据. 其中有457 组正常样本和70 组攻击样本, 它们共同构成训练样本. 按4.6 节的方法进行模型选择得到最佳模型, 此时KPLS 即训练完毕, 进入工作状态. 然后实施攻击和检测实验, 攻击间隔进行, 测试过程持续75 min, 其中35 min 没有进行攻击, 另外40 min 发动攻击. 实验结果见表2 和表3.

表2　基于KPLS 的入侵检测器的训练与检测结果

| 攻击类别 | 攻　击　工　具 | 训练过程攻击 | 检测过程攻击 | 正确检测次数 |
|---|---|---|---|---|
| Probe | netcat | 4 | 0 | 0 |
| | Nmap | 8 | 3 | 3 |
| | proft_put_down.c | 7 | 2 | 2 |
| | tn | 0 | 2 | 2 |
| | Brutus | 7 | 1 | 1 |
| | XScan | 6 | 0 | 0 |
| | Hscan | 1 | 5 | 4 |
| | NetScan Tools | 6 | 4 | 3 |
| R2L | httpd.c | 0 | 1 | 1 |
| | mayday_linux.c | 0 | 1 | 1 |
| | mod_mylo_exploit.c | 1 | 0 | 0 |
| | pam_lib.c | 0 | 1 | 1 |
| U2R | hatorihanzo.c | 1 | 3 | 3 |
| | c_marbles.c | 1 | 1 | 0 |
| DOS | Nmap | 3 | 0 | 0 |
| | teardrop.c | 4 | 0 | 0 |
| | portscan.c | 2 | 2 | 2 |
| Roolkit | netcat | 11 | 5 | 5 |
| | wipe | 1 | 4 | 3 |
| | stcpshell.c | 7 | 5 | 5 |
| 合计 | 20 种攻击 | 70 | 40 | 36 |

在训练阶段, 除将一次NetScan Tools 攻击错误识别为正常外, 其余526 个样本均判断正确. 在检测阶段, 经过训练的KPLS 检测系统正确识别出35 种正常模式中的34 种; 对发起的40 组攻击, 正确识别出其中的36 组, 有4 次漏检, 可能的原因是相应的训练样本数太少; 对4 种在训练过程中未出现的新的攻击模式, 系统全部正确识别.

研究表明, 随着抽取的分量个数不同, 分类精度随之变化. 如果固定抽取的分量个数, 再选择最优的核参数, 则可得到图3 的检测结果. 可见, 当抽取的分量达到一定数量时, KPLS 算法在较宽的范围内具有较高的分类精度, 这使模型选择变得容易.

表3　入侵检测算法性能比较

| 算　法 | 检测混淆矩　阵 | 检测精度/% | 虚警率/% | 训练时间/s | 检测时间/s |
|---|---|---|---|---|---|
| C_SVM | N　A←判为<br>35　0　N<br>23　17　A | 69.33 | 0 | 0.12 | 0.03 |
| KRR | N　A←判为<br>35　0　N<br>13　27　A | 82.67 | 0 | 1.28 | 0.03 |
| KPCR | N　A←判为<br>33　2　N<br>6　34　A | 89.33 | 5.56 | 8.05 | 0.33 |
| KPLS | N　A←判为<br>34　1　N<br>4　36　A | 93.33 | 2.70 | 7.52 | 0.34 |

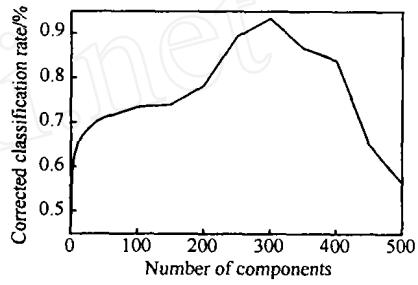


图3　KPLS 分类精度随抽取的分量个数变化

为进一步验证KPLS 入侵特征抽取及检测方法的性能, 对支持向量机(C-SVM), 核岭回归(KRR)[5]和核主分量回归(KPCR)[10]方法进行对比分析. 它们均使用RBF 核函数, 使用交叉验证法确定最优参数, 训练和检测结果见表3. 其中: N 表示正常(Nomal), A 表示攻击(Attack). 可见, 包含非线性特征抽取过程的KPCR 和KPLS 方法, 在检测精度上均优于不进行特征抽取的SVM 和KRR 方法, 这说明有效的特征抽取能提高入侵检测精度. C-SVM 能正确判断所有的正常行为, 但漏检率较高, 通过对错分样本进行分析, 发现主要是那些训练样本较少的攻击.

这一现象一方面可从所研究的特征抽取的角度来解释, 即C-SVM 将输入数据非线性地投影到Mercer 特征空间, 在特征空间进行线性分类, 并没有进行特征抽取; 另一方面可能是RBF 核对应的非线性特征与实际问题的非线性特征不一致, 由此导致精度不高. 而KRR 具有较高的分类精度, 可能的原因是RR 为一种连续的收缩方法[5], KRR 继承了这一特点, 即它们虽然选择了所有特征, 但对主分量

进行收缩, 起到了类似于特征抽取的作用.

从虚警率上看, 各算法都很低或为零. 从时间代价上看, SVM 训练和检测均最快, 而KPCR 和KPLS 因为需要进行特征抽取, 所以耗时最长, 但对异常检测算法而言是可以满足要求的. 从检测精度上看, 文献[3]运用 SVM 方法, 基于 KDD 99 数据集, 对 DOS, U2R 和 Probe 的最优检测精度分别为 86%, 90.01% 和 98.19%, 与 KPLS 方法相当.

## 5 结 论

本文从特征抽取、回归和分类的角度, 对 PLS 和KPLS 进行论述, 并从机器学习的角度出发, 给出核正交 PLS 的学习和测试算法. 其得分矢量相互正交, 可用于非线性特征抽取、回归和模式识别等问题. 建立一个模仿林肯实验室的 IDS 测试实验环境, 给出相应的特征选择方案. 通过多种新式攻击对研究方法进行验证, 并与SVM 和KPCR 等方法进行对比, 实验结果表明本文方法是有效的.

## 参考文献(References)


[1] 王行愚. 在虚拟与现实之间——自动化若干发展方向刍议[J]. *自动化学报*, 2002, 28(Suppl): 77-84.
(Wang X Y. Automatic control: Virtuality vs reality [J]. *Acta Automatica Sinica*, 2002, 28(Suppl): 77-84.)

[2] Andrew H Sung. Identify important features for intrusion detection using support vector machines and neural networks[A]. *IEEE Proc of the 2003 Symp on Application and the Internet* [C]. Orlando: IEEE Computer Society Press, 2003: 209-216.

[3] 李辉, 管晓宏, 咎鑫, 等. 基于支持向量机的网络入侵检测[J]. *计算机研究与发展*, 2003, 40(6): 799-807.
(Li H, Guan X H, Zan X, et al. Network intrusion detection based on support vector machine [J]. *Journal of Computer Research and Development*, 2003, 40(6): 799-807.)

[4] Wang Y, Yang H H, Wang X Y, et al. Distributed intrusion detection system based on data fusion method [A]. *The 5th World Congress on Intelligent Control and Automation* [C]. New Jersey: IEEE Press, 2004: 4331-4334.

[5] Trevor Hastie, Robert Tibshirani, Jerome Friedman. 范明, 等译. *统计学习基础——数据挖掘、推理与预测* [M]. 北京: 电子工业出版社, 2004.

[6] Roman Rosipal, Leonard Trejo, Bryan Matthews. Kernel PLS-SVC for linear and nonlinear classification [A]. *Proc of the 20th Int Conf on Machine Learning* [C]. Washington, 2003: 640-647.

[7] Roman Rosipal, Leonard J Trejo. Kernel partial least squares regression in reproducing kernel Hilbert space [J]. *J of Machine Learning Research*, 2001, 2: 97-123.

[8] Matthews Barker, Williams Rayens. Partial least squares for discrimination [J]. *J of Chemometrics*, 2003, 17: 166-173.

[9] Matthew V Mahoney, Philips K Chan. An analysis of the 1999 DARPA/Lincoln laboratories evaluation data for network anomaly detection [R]. Florida: Florida Institute of Technology, 2003.

[10] Bernhard Scholkopf, Alexander J Smola. *Learning with kernels: Support vector machines, regularization, optimization and beyond* [M]. Cambridge: MIT Press, 2002.


# 下 期 要 目