异常检测中支持向量机最优模型选择方法

张雪芹[1]，顾春华[1]，吴吉义[2]

(1. 华东理工大学信息科学与工程学院 上海 徐汇区 200237; 2. 杭州师范大学电子商务与信息安全重点实验室 杭州 310036)

【摘要】为了构建一个具有良好的学习性能和推广能力的异常检测分类器，在结构风险最小(SRM)原则下讨论了基于支持向量机(SVM)的异常检测分类器的设计准则，提出了SVM分类器模型及其参数快速选择和评估方法，并给出了异常检测分类器训练步骤。针对KDD'99网络入侵检测数据集，实验结果表明，该方法能够有效地缩短入侵检测分类模型建立时间，而且建立的入侵检测分类器检测精度较高。

关 键 词 异常检测； 模型选择； 参数估计； 结构风险； 支持向量机

中图分类号 TP393.08 **文献标识码** A doi:10.3969/j.issn.1001-0548.2011.04.017

Support Vector Machine Based Optimal Model Selection Method in Anomaly Detection

ZHANG Xue-qin[1], GU Chun-hua[1], and Wu Ji-yi[2]

(1. School of Information Science and Engineering, East China University of Science and Technology Xuhui Shanghai 200237;

2. Hangzhou Key Lab of E-Business and Information Security, Hangzhou Normal University Hangzhou 310036)

Abstract In order to construct an anomaly detection classifier which has good learning and generalization ability, under the structural risk minimization (SRM) principle，the design rules of a support vector machines (SVMs) based anomaly detection classifier is discussed. The model and its parameters selection and estimation method of a SVM classifier are proposed. The training steps of a SVM anomaly detection classifier are given. Experiments on KDD'99 network intrusion detection dataset indicate that the proposed methods can speed up the process of constructing an intrusion detection classifier and the classification accuracy is higher.

Key words intrusion detection; model selection; parameters estimation; structure risk; support vector machines

随着网络和计算机安全问题的日益严重，入侵检测作为一种主动防御手段越来越受到重视。入侵检测的本质是通过检测将正常数据和异常数据分开。根据检测方式的不同，入侵检测技术分为异常检测和误用检测两种。由于异常检测方式具有发现未知攻击和新型攻击的能力，已成为研究的热点。目前，用于异常检测的方法主要有神经网络、聚类、支持向量机、数据挖掘和数据融合等[1-5]。

支持向量机(support vector machines，SVMs)是在统计学理论基础上发展起来的一种新的机器学习方法，它通过求取能使两类样本以最大间隔分开的最优分类面建立分类模型。不同于神经网络等基于经验风险最小的传统的机器学习方法，支持向量机使用更为科学的结构风险(structure risk minimum，SRM)最小化原则建模，能使学习机器通过对有限样本的学习，不仅具有适当的经验风险值，而且具有较强的推广能力，即具有对未知数据进行正确分类的能力[6-7]。

由于支持向量机分类模型具有对未知数据进行正确分类的能力，许多研究者就建立基于支持向量机的入侵检测分类模型展开了研究。文献[8]采用支持向量机建立网络入侵检测模型，实验表明支持向量机具有比神经网络更好的分类性能；文献[9]使用MIB和SVM算法实现了对DOS/DDOS和网络蠕虫攻击的高效检测；文献[10]采用无监督支持向量机和聚类进行网络异常检测，实验验证了所提方法能够有效检测未知攻击。上述研究表明，基于支持向量机理论建立的入侵检测模型对识别网络和计算机攻击是有效的。

但是，支持向量机理论对于如何选择合适的核函数和参数来构造分类模型还没有一般性的理论指导。为此，不同于上述文献所述内容，本文将在结

收稿日期：2009－12－04； 修回日期：2010－09－05

基金项目：国家自然科学基金(60773094)

作者简介：张雪芹(1972－)，女，博士，副教授，主要从事信息安全、模式识别方面的研究.

构风险最小的原则下，根据一个具有良好学习性能和推广能力的入侵检测分类器的设计准则，定量给出入侵检测分类器推广性的界和经验风险值的界定方法，进而提出递进式快速模型的参数估计方法和基于SRM的入侵检测分类模型的建立步骤。

1 基于SVM的入侵检测分类器模型

对于给定的入侵检测审计数据$(\boldsymbol{x}_1,y_1)$，$(\boldsymbol{x}_2,y_2)$，…，$(\boldsymbol{x}_N,y_N)$，$\boldsymbol{x}$为特征空间的输入数据，$\boldsymbol{x}\in R^{\mathrm{d}}$，$y$为类标志，$y\in\{+1,-1\}$，$N$为样本数，$d$为输入维数。如果$y$的值为1，表示对应的样本为正常；如果$y$的值为–1，表示对应的样本为异常，即有入侵发生。

根据支持向量机理论，对于待分类样本存在一个超平面，使得两类样本完全分开，如图1所示。图中“X”和“O”分别代表两类样本，虚线为类分隔面(interval plane)，实线为分类超平面(hyperplane)。SVM试图寻找最优超平面(optimal hyperplane)以最大间隔分隔两类样本。其中，带圈的“X”和“O”样本决定类分隔面上，称为支持向量。

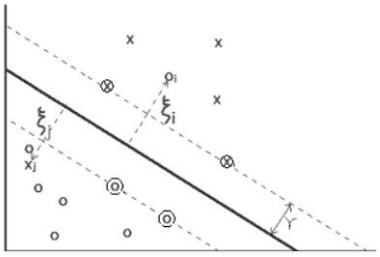

图1 支持向量机的超平面

求解最优超平面可看成解二次规划问题：

$$\min\frac{\|\boldsymbol{w}\|^2}{2}+C\left(\sum_{i=1}^{N}\xi_i\right) \tag{1}$$

$$\text{s.t.}\quad y_i(\boldsymbol{x}\cdot\boldsymbol{w}+b)\geqslant 1-\xi_i\ ,\ i=1,2,\cdots,N$$

式中，ξ为松弛变量；C为惩罚因子，为人工指定的常数，起到控制对错分样本进行惩罚程度的作用。

求解式(1)可转化为求解其对偶问题：

$$\max_{\alpha}\ W(\alpha)=\sum_{i=1}^{N}\alpha_i-\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j K(\boldsymbol{x}_i,\boldsymbol{x}_j) \tag{2}$$

$$\text{s.t.}\quad \sum_{i=1}^{N}y_i\alpha_i=0\quad 0\leqslant\alpha_i\leqslant C,\ i=1,2,\cdots,N$$

式中，α_i为拉格朗日乘子；对应于$\alpha_i>0$的向量称为支持向量；$K(\boldsymbol{x}_i,\boldsymbol{x}_j)=\phi(\boldsymbol{x}_i)^{\mathrm{T}}\phi(\boldsymbol{x}_j)$为核函数。常用的核函数有线性核函数、多项式核函数、高斯核函数等。在支持向量机中，不同的核函数将形成不同的分类模型。决策函数为：

$$f(x)=\operatorname{sgn}\left[\sum_{i=1}^{N}\alpha_i^{*}y_i K(\boldsymbol{x}_i,\boldsymbol{x})+b^{*}\right] \tag{3}$$

由于最终判别函数只包括待分类样本与支持向量的内积求和，因此识别时的复杂度取决于支持向量的个数。

2 异常检测分类器模型及参数的选择方法

2.1 结构风险最小化原则在SVM中的体现

传统的机器学习算法都是以经验风险最小化原则(empirical risk minimization，ERM)为基础的算法，往往会出现过拟合、模型过于复杂等弊端。结构风险最小化原则(structural risk minimization，SRM)是文献[6]提出的一种分类器设计方法[6]，它能有效避免出现过拟合以及由于模型过于复杂而造成分类器推广能力差的问题。

SRM在SVM上的具体体现为：1) 对于线性可分的情况，最优分类面就是在固定经验风险为0的前提下，寻求期望风险的界最小化的子集；2) 而在线性不可分的情况下，软间隔最优分类面是控制错分样本，求期望风险的界的最小值。

在该原则下，SVM分类器的设计通常包含两个任务：1) 选择适当的分类函数子集$f_k(x)$，使之对问题来说有最优的分类能力；2) 从选出的子集中选择一个函数使经验风险最小[11]。

2.2 基于SRM和SVM的异常检测分类器的设计准则

异常检测分类器不仅要求分类精度高、检测速度快，而且要求识别未知攻击的能力强，即一个性能良好的异常检测分类器不仅要求具有小的控制错分样本数(经验风险值小)，同时要求具有强的对未知数据进行正确预测的能力(推广能力强)。

根据结构风险最小原则，一个SVM异常检测分类器的设计包括以下两步：

1) 检测模型选择，即确定支持向量机的核函数形式，得到适当的分类函数子集，保证分类器具有良好的推广能力。

2) 模型参数估计，即在模型确定后，估计模型参数，确定分类函数，保证分类模型经验风险最小。

其中，经验风险最小要求控制错分样本数最小，因此经验风险值可由分类误差Err界定，有：

$$\text{Err}=\frac{\text{错分样本数}}{\text{总训练样本数}}\times 100\% \tag{4}$$

而其推广性能可由以下定理[11]保证。

定理 1 如果一组训练样本能够被一个最优分

类面或广义最优分类面分开，则对于测试样本分类错误率的期望的上界是训练样本中平均支持向量占总训练样本数的比例，即：

$$E(P_{\text{error}}) \leqslant \frac{E[\text{支持向量数}]}{\text{训练样本数}} \tag{5}$$

证明　留一法(leave-one-out, Loo)对测试错误率的估计是无偏估计[12]，在使用留一法进行测试错误率估计时，有：

$$P_{\text{error}}(f) = \frac{1}{N}\sum_{i=1}^{N} L(f^i(x_i), y_i) \tag{6}$$

式中，f^i表示去掉第i个样本后在剩余样本上得到的分类规则；$f^i(x_i)$表示使用该规则对样本x_i进行分类；$L(f^i(x_i), y_i)$表示留一法的分类结果，分类正确取1，反之取0。从SVM和留一法的原理可知，对于非支持向量，在留一法测试时不会产生测试错误，因为非支持向量未被用于构造分类面，当去掉该样本进行留一法训练时，分类面不会改变，不会产生错分，只有去掉支持向量才可能产生测试错误。因此，P_{error}不会大于支持向量对样本数之比。支持向量对样本数之比是对测试样本分类错误率的期望的上界。

通过选择合适的核函数，构造一个支持向量数量相对较少的最优或广义最优分类面，可以得到推广性能较好的入侵检测模型。

目前，常用的SVM模型选择和参数估计方法除证明中提到的留一法外，还有交叉验证法(cross validation，CV)和基于测试的方法。其中，留一法是交叉验证法的特例。交叉验证法和留一法对模型参数的选择准确性高，但存在计算复杂度高、计算时间长的问题，不适用于大样本的入侵检测数据集。为此，本文提出递进式模型参数估计方法和基于SRM的SVM入侵检测分类器构建法。

2.3　递进式模型参数估计方法

SVM异常检测模型参数估计涉及核参数和惩罚因子C，在特定的数据空间中，核参数和惩罚因子的变化都存在一定的规律。如对于惩罚因子C，它的作用是实现训练错误率与模型复杂度间的折衷，C的取值大表示对经验误差的惩罚大，错分样本比例低，经验风险值小，同时，支持向量数目减少，推广性能增强，分类器复杂度增加。但是，当C超过一定值时，分类器的复杂度达到了数据空间允许的最大值，此时C的改变不会再影响分类性能。因此，可以先采用测试方法获得参数变化规律，选定取值区域，然后在选定区域搜寻最优参数。

递进式参数估计方法描述如下：

1) 粗粒度、大步长地选择参数变化区间，在训练集上进行模型训练，在测试集上进行测试，初步获得能使分类器具有良好学习和推广性能的核参数和惩罚因子取值区间。

2) 根据步骤1)确定的参数区间，在该区间内使用交叉验证法细粒度、小步长地寻找最优参数，减少寻优时间。

2.4　基于SRM的异常检测分类器训练方法

基于SRM的SVM异常检测分类器训练过程如图2所示。该方法可描述如下：

1) 模型选择使用测试法训练多个模型，使用$E(P)$值定量界定入侵检测模型推广性的界，当$E(P)$值相近时，并结合考虑分类误差Err，实现模型选择。

2) 参数估计使用递进式参数估计方法，用分类误差Err进一步界定模型经验风险值，并结合考虑该模型下不同参数时的$E(P)$值，实现最优参数的快速寻找。

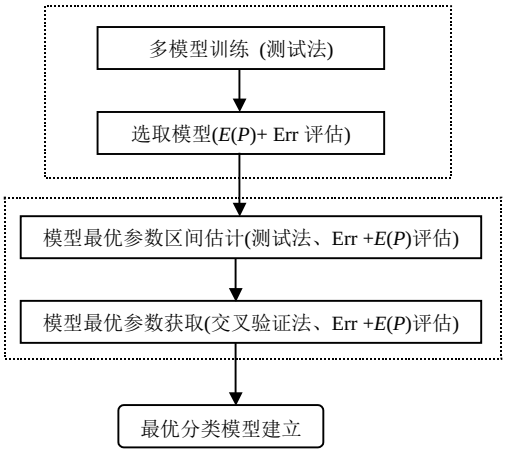

图2　基于SRM的异常检测分类器的设计步骤

3　实验与结果

3.1　数据描述与预处理

实验所用数据来自于KDD'99[13]。KDD'99是1998年美国国防部(DARPA)与麻省理工学院的林肯实验室共同推出的入侵检测评估计划。在该计划中，林肯实验室模仿美国空军的一个局域网进行仿真，采集了9周的网络连接数据，模拟了政府和空军的1 000个主机的100多个用户的正常通讯，同时包含4类38种攻击。

实验从原始的10%数据集中随机选出两个子数据集作为训练集和测试集，分别取名为“1percent”和“2percent”。每个记录集中含有大约50 000条记录，每个记录集中正常记录和异常记录是随机从原始数据集中选取的，但是遵循它们在原始的10%数据集

中的分布比例。

3.2 仿真实验

实验采用LIBSVM实验包[14]，根据第2节所述的基于SRM的SVM异常检测分类器建立过程，首先采用测试法进行模型选择，即核函数选择，然后采用递进式参数估计法寻找最优模型参数，最后建立入侵检测分类器并进行性能评估。

实验中，数据集含有正常记录和多种攻击记录，分类时将各种攻击记录均视为异常类而不作攻击类别区分，即针对网络混合攻击模式建模。

实验1 模型选择

核函数是SVM方法中少数几个能够调整的参数之一。采用不同的核函数将产生不同的支持向量机模型，对同样的数据可产生不同的分类结果。为了获取推广性能良好的入侵检测分类模型，实验采用测试法选择SVM分类模型，以$E(P)$值界定模型推广性能，结合Err值和建模训练时间(Train)对模型进行评估。由于实验中总的训练样本数是一定的，因此实际评估时以支持向量的数目(#SVs)代替$E(P)$值。为了验证方法的有效性，将实验结果与使用交叉验证法得到的结果进行比较。

SVM常用的核函数有线性核、RBF核、多项式核和Sigmoid函数，用测试法和交叉验证法进行模型选择，模型使用默认参数，实验评估结果如表1所示。

表1 两种模型选择方法比较

Kernel	测试法			CV (5-fold)		
	Train/s	Err/(%)	#SVs	Train/s	Err/(%)	#SVs
线性核	27.80	0.68	671	139.6	0.54	548
RBF核	39.94	1.08	977	220.85	0.73	811
多项式核	97.06	3.73	2914	636.65	1.02	2 598
Sigmoid核	53.39	2.28	1 130	268	0.73	948

从表1可以看出：1) 使用线性核建立的分类器支持向量数目最少，分类误差最小，模型训练时间最短，RBF核次之。因此，线性核和RBF核支持向量机模型推广性能较好。2) 使用测试法和交叉验证法所获得的结果趋势是一致的，使用交叉验证法耗时1 265.1 s，而使用测试法模型选择耗时218.19 s，模型选择时间缩短5.8倍。

实验2 参数估计

针对实验1选出的线性核和RBF核模型，实验中需要确定的参数：惩罚因子C和RBF核宽σ。实验采用本文提出的递进式模型参数估计方法寻找最优模型参数。同样，为了验证方法的有效性，将实验结果与使用交叉验证法的结果进行比较。

1) 核宽σ取值区间估计。

取C=1，在RBF核下，按一定步长调节核宽σ进行SVM训练，建立的分类器的性能结果如表2所示。从表2可以看出，在σ为[0.002, 200]范围内，训练时间和支持向量数目都存在由大变小，再由小变大的过程；而分类误差存在由大变小，再由小变大的过程。在$0.02<\sigma<2$区间内，分类器分类误差小，支持向量数目较少，分类器学习时间较短，分类器的学习和推广性能好。因此，可以初步选定核参数σ的取值范围为[0.02,2]。实验中，参数区间估计累计耗时970.99 s。

表2 不同核宽参数σ 时分类器性能评估

σ	Train/s	Err/(%)	#SVs
0.000 2	119.56	3.62	4 638
0.002	49.95	3.23	1 639
0.02	34.69	1.09	1 013
0.2	22.55	2.9	685
2	28.99	3.19	485
20	160.79	21.25	989
200	554.46	21.26	2 691

2) 惩罚因子C 取值区间估计。

按一定变化步长调节C ，进行SVM训练，建立的分类器的性能评估如表3所示。

表3 不同参数C时分类器性能评估

C	Train/s	Err/(%)	#SVs
2^{-2}	43.12	3.34	1 349
2^{-1}	35.61	1.00	871
2^{0}	27.04	0.68	671
2^{1}	25.43	0.65	629
2^{2}	25.28	0.43	587
2^{3}	26.12	0.59	558
2^{4}	26.87	1.22	502
2^{7}	37.64	2.32	410

从表中可以看出，随着C的增大，支持向量数目减少。分类误差和训练时间均存在由大变小，再由小变大的趋势。在$1<C<8$区间内，分类器分类误差低，支持向量数目少，分类器学习时间较短，分类器的学习和推广性能好。因此，可以初步选定惩罚因子C的取值范围为[1,8]。实验中，参数区间估计耗时247.11 s。

3) 最优参数估计。

根据前面的实验结果，σ的搜索范围在[0.02,2]，C的搜索范围在[1,8]，在该两参数空间中使用5-fold交叉验证法，实验结果如表4所示。可以看出，在线性核模型下，耗时0.6 h寻找到最优参数C=8，在RBF

核模型下，耗时1.46 h寻找到最优参数组合C=2，σ=2。使用最优参数建立的SVM线性分类器的分类误差为0.05%，RBF分类器的分类误差为0.59%，分类误差低，而且支持向量数目少，模型推广性能好。而没有使用递进式参数估计的交叉验证法，由于需要在较大的参数空间中搜索最优参数，实验中耗时8 h以上仍未能寻找到最佳参数。表中，ST项指最优参数搜索时间，Param指寻找到的最优参数。

表4　两种参数搜索方法比较

	Kernel	ST/h	Param	Train/s	Err/(%)	#SVs
递进式估计法	RBF	1.46	C=2, σ=2	38.31	0.05	396
CV (5-fold)	Linear	0.6	C=8	26.12	0.59	558
	RBF	>8	—	—	—	—
	Linear	>8	—	—	—	—

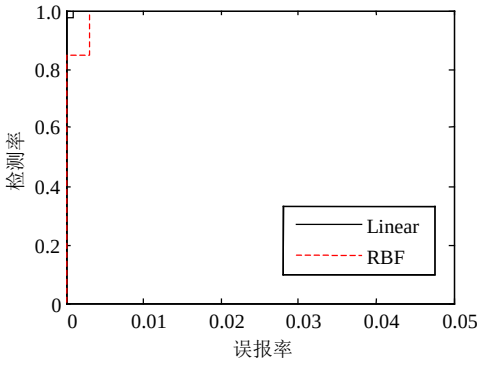

图 3　分类器的 ROC 曲线

实验3　分类器检测性能考察

进一步考察本文所提方法构建的分类器的检测性能，使用实验2中搜索得到的最优参数，分别以线性核(C=8)和RBF核(C=2,σ=2)建立分类器，根据SVM训练结果作出ROC(receiver operator characteristic)曲线，如图3所示。同一ROC曲线上的点代表同一检测方法在阈值不同时的误报率和漏报率。ROC曲线下的面积越大，表明模型的检测性能越好。可以看出，该分类器检测率高，误报率低，其中，线性分类器具有更优良的检测性能。

4　结 束 语

本文针对如何在入侵检测中建立一个具有良好学习性能和推广性能的异常检测分类器做了详细的讨论。在结构风险最小的原则下给出了异常检测分类器的设计准则，以及递进式模型参数快速估计方法和异常检测分类器建立步骤。基于KDD'99数据集的实验验证了所提出方法的快速、有效性。同时，针对DOS、Probing、U2R和R2L共4种单一网络攻击模式，采用所提出的方法构建模型进行实验，同样获得了好的结果。本文提出的方法对于建立基于支持向量机的分类器具有指导意义。

本文工作得到杭州市电子商务与信息安全重点实验室开放基金(HZEB201009)的资助，在此表示感谢！

参　考　文　献

[1] JIANG D B, YANG Y H, XIA M. Research on intrusion detection based on an improved SOM neural network[C]//Fifth International Conference on Information Assurance and Security, IAS'09. Xi'an: IEEE, 2009: 400-403.

[2] JIANG S Y, SONG X Y, WANG H, et al. A clustering-based method for unsupervised intrusion detections[J]. Pattern Recognition Letters, 2006, 27 (7): 802-810.

[3] ZHANG X Q, GU C H. CH-SVM based network anomaly detion[C]//Proceedings the Sixth International Conference on Machine Learning and Cybernetics. Hong Kong: IEEE, 2007: 3261-3266.

[4] WU S Y, YEN E. Data mining-based intrusion detectors[J]. Expert Systems with Applications, 2009, 36(3): 5605-5612.

[5] PARIKH D, CHEN T H. Data fusion and cost minimization for intrusion detection[J]. IEEE Transactions on Information Forensics and Security, 2008, 3(3): 381-389.

[6] VAPNIK V N. The nature of statistical learning theory[M]. 2nd ed. New York: Springer-Verlag, 1999.

[7] VAPNIK V N. Statistical learning theory[M]. New York: John Wiley & Sons, 1998.

[8] HUAND H P, YANG F C. Intrusion detection based on active networks[J]. Journal of Information Science and Engineering, 2009, 25(3): 843-859.

[9] YU J, LEE H. Traffic flooding attack detection with SNMP MIB using SVM[J]. Computer Communications, 2008, 31(17): 4212-4219.

[10] SONG J, TAKAKURA H. Unsupervised anomaly detection based on clustering and multiple one-class SVM[J]. IEICE Transactions on Communications, 2009, E92B(6): 1981-1990.

[11] 边肇祺, 张学工. 模式识别[M]. 第2版. 北京: 清华大学出版社, 2000.
BIAN Zhao-qi, ZHANG Xue-gong. Pattern Reorganization[M]. 2nd ed. Beijing: Tsinghua University Press, 2000.

[12] 董春曦, 杨绍全, 饶鲜, 等. 支持向量机推广能力估计方法比较[J]. 电路与系统学报, 2004, 9(4): 86-91.
DONG Chun-xi, YANG Shao-quan, RAO Xuan, et al. Comparison in generaIization performance of support vector machine algorithms[J]. Journal of Circuits and Systems, 2004, 9(4): 86-91.

[13] DAPPA. KDD'99 dataset[DB/OL]. [2009-12-03]. http://kdd.ics.uci. edu/dataset/kddcup99/ kddcup99.htm.

[14] CHANG C C, LIN C J. LIBSVM: a library for support vector machines[EB/OL]. [2009-12-03]. http://www.csie.ntu.edu.tw/-cjlin/libsvm.

编　辑　税　红