第 34 卷 第 11 期 计 算 机 工 程 2008 年 6 月
Vol.34 No.11 Computer Engineering June 2008
·安全技术· 文章编号：1000—3428(2008)11—0166—02 文献标识码：A 中图分类号：TP391

基于集成学习的入侵检测方法

李 凯，陈 武
(河北大学数学与计算机学院，保定 071002)

摘 要：入侵检测是近年来网络安全研究的热点。利用多分类器技术，研究了基于集成学习的入侵检测方法。应用 Bootstrap 技术生成分类器个体，为了提高分类器的差异性，应用聚类技术对分类器进行聚类，在相应的聚类结果中选取不同的分类器个体，并选择不同的融合方法对分类结果进行融合。针对入侵检测数据的实验表明了该集成技术的有效性。

关键词：集成学习；融合；入侵检测；泛化性能

Intrusion Detection Method Based on Ensemble Learning

LI Kai, CHEN Wu
(School of Mathematics and Computer, Hebei University, Baoding 071002)

【Abstract】 Intrusion detection is a highlighted topic of network security research in recent years. Intrusion detection method based on ensemble learning is studied by using multiple classifiers. Some classifiers are created by Bootstrap technique. To improve their diversity, clustering technique is applied to them for choosing diverse individuals in each cluster. Then different fusion techniques are used to combine different classification results. Experiments are conducted with intrusion detection data set and show that intrusion detection based on ensemble learning is effective.

【Key words】 ensemble learning; fusion; intrusion detection; generalization capability

1 概述

目前，网络安全已经成为一个全球性的重要问题。为了保护计算机网络的安全，通常采用防火墙、防病毒软件、加密技术、用户认证、入侵检测系统等技术。其中入侵检测技术是近年来出现的新型网络安全技术，被认为是防火墙之后的第 2 道安全闸门。

从本质上看，入侵检测属于分类问题，基于这种思想，人们提出了许多入侵检测方法，但是，这些方法却存在不同程度的缺陷，例如误报率高、漏报率高等。为了解决这些问题，研究人员利用融合技术提高入侵检测的性能，例如投票方法、证据理论方法等，并取得了一定的效果。本文主要研究基于集成学习的入侵检测。集成学习属于多分类器系统，主要由两部分组成，即集成个体的选取与分类器分类结果的组合(融合)。对于融合方法，人们提出了许多用于入侵检测的方法，而将集成学习用于入侵检测的研究相对较少。从差异性入手，笔者研究了集成学习算法及提高学习性能的机理，并针对入侵检测数据进行了实验。

2 集成学习的差异性

自从Hansen等人提出神经网络集成后，个体的差异性的概念应运而生。究竟什么是差异性？差异性如何度量？这些成为集成学习需要面对的问题。由于差异性还没有通用的定义，因此人们从不同角度对集成学习的差异性进行研究，例如：在神经网络集成学习中，Hansen等人提出：两个分类器的差异性是指这两个分类器个体对新的输入数据产生的不同分类错误。增加集成个体之间的差异性可以提高集成学习的泛化性能。另外，一旦差异性定义确定，如何度量个体间的差异性呢？为此，人们提出了很多度量方法[1-2]。下一步将研究度量方法与集成的性能的关系。若能得到一种与集成的正确率相关的差异性度量方法，则集成学习的目标就转化为最大化集成个体间的差异性，但研究人员提出的差异性度量与集成正确率间并不存在这种单调关系，在某种程度上都具有一定的局限性。

2.1 使用增加数据方法提高神经网络集成的差异性

当数据集规模较小时，通过学习得到的模型其正确率很难得到保证，另外，模型之间的差异性是很低的。为此，笔者研究了以神经网络分类器作为个体的集成学习算法——DBNNE[3]，该方法通过生成额外的数据试图增加集成学习的差异性。在DBNNE算法中，将集成中分类器的预测与集成预测的不一致性作为差异性度量，更精确地说，若$C_i(x)$是第i个分类器对实例x的类标号的预测，$C^*(x)$是集成的预测，则第i个分类器在实例x上的差异性定义为

$$d_i(x)=\begin{cases}0 & \text{若} C_i(x)=C^*(x)\\ w_i & \text{否则}\end{cases}$$

其中，$w_i\in(0,1]$为第i个分类器的投票权值，本文使用w_i=1。为了得到集成规模为n实例个数为m的差异性，对上面的项取平均值得

$$d_{\text{ensemble}}=\frac{1}{n\cdot m}\sum_{i=1}^{n}\sum_{j=1}^{m}d_i(x_j)$$

使用这种度量用来估计集成中的一个分类器与集成的预测不一致的概率。

基金项目：河北省教育厅基金资助项目(2006406)；河北大学博士基金资助项目(2006075)
作者简介：李 凯(1963－)，男，教授、博士，主研方向：机器学习，模式识别，数据挖掘；陈 武，助理实验师
收稿日期：2007-12-24 **E-mail：**likai@hbu.edu.cn

2.2 使用聚类技术提高神经网络集成的差异性

在基于划分的聚类研究中，通常使用距离度量两个点间的相似性。当这个距离值很小时，则认为这两个点具有一定的相似性，反之，则认为这两个点具有较大的差异。基于这种思想，笔者研究了神经网络模型间的相似程度(也称为差异性)。将衡量模型相似的距离度量定义为 $\rho(M_i(\boldsymbol{x}),M_j(\boldsymbol{x}))$，该定义方式是一种抽象的度量，在不同应用场合中，具有不同的表达形式，例如：为了确定两个神经网络模型的相似性，可以在一个验证集 V 上定义，即利用每个数据点的目标输出与实际输出的偏离程度的绝对值之和作为两个神经网络模型间的距离。此距离越大，则说明两个模型的差异性越大，反之，则差异性很小。在聚类算法中，当两个模型差异性很大时，说明这两个模型有可能在不同的簇中，相反，则在同一个簇中。由于在同一个簇中的模型都是相似的模型(按照上面的距离度量)，因此可以从中选择一个作为簇的代表模型。

3 集成学习算法

在集成学习中，研究人员提出了许多不同的算法，例如 bagging, adaboost。下面主要研究通过聚类技术产生集成个体的方法，该算法是文献[4]方法的一种推广形式，对模型聚类进行了研究[5]。本文将基于聚类技术的集成算法命名为G_CENN。G_CENN算法如下：

Step1 生成一组学习模型集 E。

Step2 对模型集 E 聚类，然后在每个簇中选择一个模型，设通过选择得到的模型为集合 P。

Step3 对集合 P 上的模型融合。

可以看到，G_CENN 算法中的每一步可能有多种选择方法，例如在 Step2 中可以使用不同的聚类算法，例如层次聚类方法、基于划分的聚类方法等；同样在 Step3 中可以使用不同的融合技术，例如投票方法、决策模板方法、朴素贝叶斯方法等。实验选择了层次聚类算法，簇间距离度量分别为最小距离度量、最大距离度量、平均值距离度量与平均距离度量。笔者将簇间距离使用最小、最大、平均值及平均距离的聚类算法分别称为 Hier_min, Hier_max, Hier_mean, Hier_avg。

4 实验研究

为了验证集成技术在入侵检测中的性能，实验选用了 KDD CUP1999 数据，该数据集包含正常连接数据和异常连接数据，共有 41 个属性，可以分为 3 类，即基本属性集，内容属性集，流量属性集。入侵数据有 4 种类型：DoS, R2L, U2R, Probing 攻击。实验选择了与 ftp 服务相关的一个子集。训练集由 4 028 个记录构成，其中 2 000 条记录为正常数据包，剩余 2 028 条记录为入侵数据包；测试集由 5 472 条记录构成，其中 4 172 条记录为正常数据包，1 300 条记录为入侵数据包。另外，为了测试神经网络分类器集成检测新的攻击类型的能力，在测试集中选择了 54 条入侵记录，这些记录的入侵类型并未包括在训练集中。检测率(DR)定义为正确检测到的入侵个数除以测试集中所有入侵数据个数的百分比；误报率(FP)定义为将正常数据错误检测为异常数据的个数除以测试集中所有正常数据个数的百分比。

在神经网络实验中，首先使用单个多层感知器神经网络在 3 个特征空间及整个特征空间中训练，隐层神经元个数都为 15，使用 BP 算法训练神经网络。当均方误差达到 0.001 时停止训练。训练的 4 个神经网络在测试集上的性能见表 1。为了研究集成方法在入侵检测中的性能，使用了 5 种聚类技术用来选择具有差异的神经网络，表 2 给出了集成算法 G_CENN 使用不同聚类技术的检测结果。

表 1 神经网络在测试集上的性能

分类器类型	检测率	误报率
MLP－基本属性	95.38	3.26
MLP－内容属性	96.47	2.28
MLP－流量属性	89.67	24.15
MLP－所有属性	96.89	3.79

表 2 G_CENN 使用不同聚类技术的检测结果

聚类方法	检测率	误报率
Hier_min	97.65	2.87
Hier_max	98.92	2.75
Hier_mean	98.31	3.02
Hier_avg	98.15	3.28
K_means	97.87	2.93

笔者选择了 4 种不同的融合方法，分别是：Majority, Average, Decision template 与 Naive-bayes 方法，针对集成方法 G_CENN, DBNNE, Bagging 与 Adaboost 研究了入侵检测的性能，表 3、表 4 给出了集成技术 G_CENN, DBNNE, Bagging, Adaboost 使用不同融合方法的检测结果。

表 3 4 种集成技术使用不同融合方法的检测结果

集成方法	检测率	误报率
G_CENN	$DR1$＝97.02	$FP1$＝3.04
	$DR2$＝98.92	$FP2$＝2.75
	$DR3$＝99.04	$FP3$＝3.27
	$DR4$＝99.17	$FP4$＝2.56
DBNNE	$DR1$＝97.29	$FP1$＝2.85
	$DR2$＝98.14	$FP2$＝3.16
	$DR3$＝97.58	$FP3$＝3.24
	$DR4$＝98.37	$FP4$＝3.06
Bagging	$DR1$＝97.31	$FP1$＝2.81
	$DR2$＝98.35	$FP2$＝3.07
	$DR3$＝97.18	$FP3$＝3.48
	$DR4$＝99.06	$FP4$＝3.15
Adaboost	$DR1$＝97.46	$FP1$＝2.57
	$DR2$＝98.53	$FP2$＝3.43
	$DR3$＝98.02	$FP3$＝3.16
	$DR4$＝98.95	$FP4$＝2.97

表 4 集成技术使用不同融合方法对已知和未知入侵的检测结果

集成方法	已知攻击	未知攻击
G_CENN	$DR1$＝98.17	$DR1$＝80.36
	$DR2$＝98.75	$DR2$＝81.43
	$DR3$＝99.13	$DR3$＝86.39
	$DR4$＝96.23	$DR4$＝80.45
DBNNE	$DR1$＝98.40	$DR1$＝81.25
	$DR2$＝99.07	$DR2$＝79.64
	$DR3$＝98.49	$DR3$＝83.57
	$DR4$＝99.12	$DR4$＝82.39
Bagging	$DR1$＝98.57	$DR1$＝81.13
	$DR2$＝99.16	$DR2$＝80.05
	$DR3$＝98.32	$DR3$＝82.30
	$DR4$＝99.03	$DR4$＝78.95
Adaboost	$DR1$＝98.91	$DR1$＝82.03
	$DR2$＝99.20	$DR2$＝83.01
	$DR3$＝99.37	$DR3$＝83.35
	$DR4$＝98.57	$DR4$＝78.14

在表 3 中，$DR1$～$DR4$ 分别表示使用 Majority, Average, Decision template, Naive-bayes 方法得到的检测率；$FP1$～$FP4$ 分别表示使用 Majority, Average, Decision template, Naive-bayes 方法得到的误报率。

由实验结果可以看到，使用神经网络分类器的集成技术

（下转第 176 页）