基于 FCM-C4.5 的双过滤入侵检测机制

滕少华 严远驰 刘冬宁 吴 昊
(广东工业大学计算机学院 广东 广州 510006)

摘 要 针对现有入侵检测技术误报率高、未知攻击检测难,而单一检测技术难以检测复杂的网络攻击等问题,提出一种基于 FCM-C4.5 的双过滤入侵检测机制。检测机制分两层对数据进行过滤,第一层采用模糊 C 均值聚类算法 FCM(fuzzy C-means algorithm)初步过滤掉明显的正常数据,从而减少了第二层过滤的数据量;第二层运用决策树 C4.5 算法进行细过滤,从而获得效率与精度的提高。通过 KDD CUP 99 数据集的实验表明,该检测机制既能检测到已知攻击又能检测到未知攻击,且具有较高检测率和较低误报率。

关键词 FCM C4.5 双过滤 入侵检测

中图分类号 TP3 **文献标识码** A DOI:10.3969/j.issn.1000-386x.2016.01.073

A DUAL FILTRATION INTRUSION DETECTION MECHANISM BASED ON FCM AND C4.5

Teng Shaohua Yan Yuanchi Liu Dongning Wu Hao
(*School of Computer Science and Technology, Guangdong University of Technology, Guangzhou 510006, Guangdong, China*)

Abstract Existing intrusion detection technology has high false alarm rate, and is difficult to detect the unknown attacks, while the single detection technology is difficult to detect complicated network attacks. Aiming at these problems, in this paper we propose an FCM and C4.5-based dual filtration intrusion detection mechanism. The detection mechanism is divided into two layers to filter the data, the first layer uses fuzzy c-means clustering (FCM) algorithm to filter out obvious normal data initially so that reduces the data amount to be filtered by second layer; and the second layer uses C4.5 decision tree algorithm to carry out refined filtration so that achieves the improvement in efficiency and accuracy. It is demonstrated by the experiment on the Knowledge Discovery and Data Mining (KDD'99) that the detection mechanism proposed in this paper can detect both known attacks and unknown attacks with higher detection rate and lower false alarm rate.

Keywords FCM C4.5 Dual filtration Intrusion detection

0 引 言

随着互联网的飞速发展和网络环境的日趋复杂,网络安全已经成为一个全球性的问题[1-3]。入侵检测作为一种通过实时监测目标系统来发现入侵攻击行为的安全技术,现已成为网络安全领域中的一个研究热点[3-5]。

现有入侵检测技术主要分为误用检测和异常检测。误用检测是根据对攻击的先验知识建立入侵模型,利用已知的攻击模式来发现攻击,它对已知的攻击检测非常有效,但对未知攻击检测能力弱[6,7]。异常检测是寻找网络中与正常网络行为偏离的行为,利用已经建立的正常行为判别模型来检测入侵攻击行为,它可以检测出未知的攻击行为,但误报率较高[4,8]。

针对以上两种检测技术的不足,本文将误用检测和异常检测相结合,提出一种基于 FCM 和 C4.5 的双过滤入侵检测机制。经实验表明,该检测机制能明显克服单一检测技术的不足,充分发挥 FCM 对未知攻击的检测能力和 C4.5 低误报率的优势,实现对已知攻击和未知攻击的检测,并且具有较高的检测率和较低的误报率。

1 相关工作

误用检测通过预先建立的入侵攻击行为模型来检测入侵行为,由于其能以非常低的误报率检测到已知攻击,在入侵检测领域得到广泛应用。文献[10]提出一种 wrapper 型的特征选择算法来构建轻量级入侵检测系统,检测系统首先删除冗余数据,然后采用 wrapper 特征选择算法选择合适的属性,最后采用 neuro-tree 算法建立检测模型。文献[11]将深度包检测技术应用到木马程序检测上,研究实现了一个基于 DPI 技术并应用 GRETA 正则库匹配攻击的分布式的木马检测系统。文献[12]利用粗糙集理论进行属性约简,并对属性进行离散化和特征选择,构建了一种基于 Q-learning 算法和粗糙集理论的网络入侵检测系统。这些研究对已知攻击具有较好的检测率,但对未知攻击的检测能力较差。

收稿日期:2014-05-22。国家自然科学基金项目(61272067,61104156);教育部重点实验室基金项目(110411)。**滕少华**,教授,主研领域:数据挖掘,网络安全,大数据。**严远驰**,硕士生。**刘冬宁**,副教授。**吴昊**,硕士生。

而异常检测是通过找出网络中与正常网络行为模型偏离的行为来检测入侵的,显然它可以检测到未知攻击。文献[13]将自组织映射和 k-means 聚类算法相结合,提出一种自适应动态聚类的入侵检测模型,检测模型能在没有任何人工干预的情况下根据当前网络数据状态自动进行簇的重建,实现完全无监督的自适应检测。文献[14]提出了两种新的聚类算法:改进的竞争学习网络(ICLN)和有监督的改进的竞争学习网络(SICLN),ICLN 通过一个新的奖惩更新规则克服了标准竞争学习神经网络的不稳定性,SICLN 通过使用带标签数据训练新规则来达到更好的聚类效果。文献[15]将支持向量机回归的分类融合应用到网络异常行为分析中,在 SVM 参数选择时采用交叉验证和深度优先搜索算法进行优化选择,并通过融合证据理论,建立网络异常检测模型。这些研究大都通过算法的改进来提高检测效果,不但增加了算法的复杂度且误报率较高。

基于此,本文提出一种基于 FCM-C4.5 的双过滤入侵检测机制。第一层利用 FCM 进行粗过滤,将大量明显正常数据过滤掉,留下少数正常数据和绝大部分的异常数据送入第二层进行细过滤。第二层利用 C4.5 低误报率的特点过滤掉剩下的正常数据,最终得到一个高检测率、低误报率的检测结果。

2 FCM-C4.5 双过滤检测机制

由于不同网络协议有不同的属性值,同一种协议的数据包相对具有更多的相似性[16]。为更好地提高检测精度和检测效率,本文按网络数据协议类型(TCP/UDP/ICMP)建立了三个检测代理,每个代理分两层过滤检测,各检测代理协同工作形成一个整体检测系统。系统模型如图 1 所示。

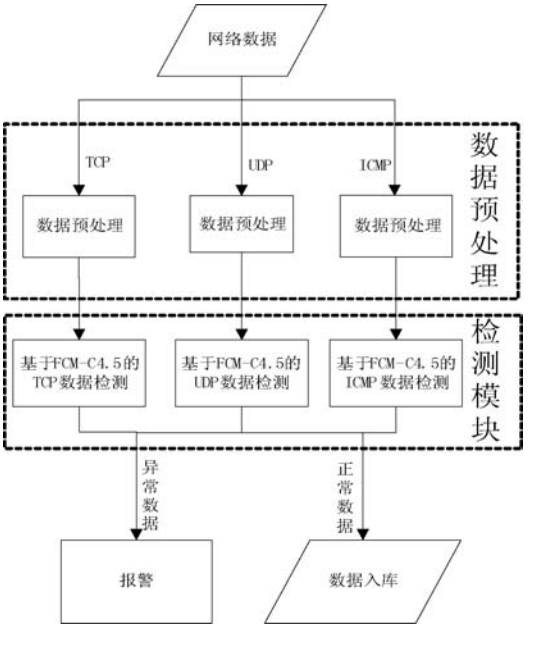

图 1 基于 FCM-C4.5 的双过滤协同入侵检测机制

网络数据按协议类型的不同送入相应的数据预处理模块进行预处理,再将预处理后的数据送入相应的检测模块进行检测。检测模块首先利用 FCM 聚类进行第一层过滤,将数据初步划分为正常数据和异常数据,然后利用决策树模型对异常数据进行二次过滤,得出检测数据的最终检测结果。基于 FCM-C4.5 的 TCP 数据检测流程如图 2 所示,UDP 和 ICMP 两个模块与之类似。

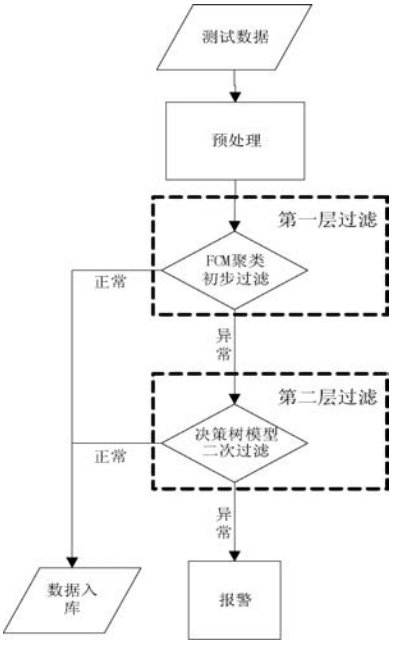

图 2 基于 FCM-C4.5 的 TCP 数据检测模块

2.1 FCM-C4.5 训练算法

FCM-C4.5 训练算法分为 FCM 聚类算法和 C4.5 算法。

FCM 聚类算法建立在两个假设上:1) 正常行为数据的数目远远大于入侵行为数据的数目;2) 入侵数据在某些属性的取值上明显偏离正常取值范围[6]。FCM 算法按照“类内相似度最大化,类间差异性最大化”的原则将数据集划分成若干个类,然后通过设定阈值把数据划分为正常与异常[8,9]。具体算法如下:

算法 1 FCM 聚类算法

1) 通过多次试验,选定迭代终止阈值 ε、N 和聚类个数 k;

2) 用[0,1]间的随机数初始化隶属度 u_{ij},使其满足 $\sum_{i=1}^{c} u_{ij} = 1, 1 \leqslant j \leqslant n$;

3) 训练集 $Train = \{x_1, x_2, \cdots, x_n\}$;

4) 随机选取 k 个样本作为初始聚类中心 $\{v_1, v_2, \cdots, v_k\}$;

5) 选择 $x_i \in Train$, $i = 1, 2, \cdots, n$;

6) 计算 x_i 到各聚类中心的距离,距离最小值记为 d;

7) 将 x_i 划分到距离为 d 的聚类中;

8) $Train = Train - \{x_i\}$;

9) 如果 $Train = \Phi$,转 10),否则转 5);

10) 更新隶属度矩阵 U 和聚类中心 V;

11) 如果聚类中心变化小于阈值 ε 或者迭代次数不小于 N,转 12),否则 $Train = \{x_1, x_2, \cdots, x_n\}$,转 5);

12) 对于每个聚类,如果聚类中数据样本的数目比聚类平均具有的样本数的一半($n/2k$)还小,则将该聚类标记为异常,否则将该聚类标记为正常。

C4.5 算法是由一个个连接记录组成的训练数据出发,在生长树的每一步中通过计算,选择信息增益率最高的属性作为当前节点的测试属性,最后得出可用于判断连接记录所属类型的决策树[10]。具体算法如下:

算法 2 C4.5 算法

1) 对训练集中数据的各项属性进行预处理;

2) 计算各个属性的信息增益率;

3) 挑选信息增益率最大的属性作为决策树的根节点;

4）在剩下的候选属性中选择信息增益率最大的属性作为当前分叉节点,递归直到形成决策树模型;

5）从构造的决策树中得到分类规则。

2.2 FCM-C4.5 检测算法

双过滤检测机制的两层犹如两个筛孔大小不一的筛子将正常数据逐层过滤掉,留下异常数据进行报警响应。其中第一层是大筛子,通过阈值 α 调节筛孔大小,让尽可能多的异常数据通过,附带着会有一定数量的正常数据跟随通过,使得这一层具有较高的检测率同时伴随着较高的误报率,这些数据都送入第二层进行进一步过滤。第二层是小筛子,首先利用 C4.5 算法对已知攻击检测率高的特点,让已知攻击通过这一层进入下步的报警响应,然后利用阈值 β 调节筛孔的大小和 C4.5 算法低误报率的特点,让正常数据尽可能多地过滤掉,附带着会过滤掉少量异常数据,最终留下大量异常数据和少量正常数据,得到一个高检测率低误报率的检测结果。具体检测算法如下:

算法 3　FCM-C4.5 检测算法

1）通过多次试验选定阈值 α 和 β;

2）给定测试集 $Test = \{x_1, x_2, \cdots, x_n\}$;

3）选择 $x_i \in Test$, $i = 1, 2, \cdots, n$;

4）计算 x_i 与训练阶段 FCM 算出的各正常类的聚类中心的距离,选取最小值,记为 d_{normal};

5）计算 x_i 与训练阶段 FCM 算出的各异常类的聚类中心的距离,选取最小值,记为 d_{unnormal};

6）如果 $\frac{d_{\text{unnormal}}}{d_{\text{normal}}} < \alpha$,将 x_i 暂时标记为异常,转7);否则 x_i 的最终检测结果为正常($result(x_i) = 0$),转 9);

7）利用决策树模型对 x_i 进行检测,如果决策树模型检测结果为异常,则 x_i 的最终检测结果为异常($result(x_i) = 1$),转9);

8）如果 $\frac{d_{\text{unnormal}}}{d_{\text{normal}}} > \beta$, 则 x_i 的最终检测结果为正常($result(x_i) = 0$),否则 x_i 的最终检测结果为异常($result(x_i) = 1$);

9）$Test = Test - \{x_i\}$;

10）如果 $Test = \Phi$,结束,否则转 3)。

3 实验及结果分析

3.1 实验数据描述

本文选用的数据集是 KDD Cup 99 数据中的"kddcup.data_10.percent"[17]。数据集中攻击数据共有四大类(DOS、U2R、R2L 和 Probe),每条数据有 41 个属性(3 个字符型属性,38 个数值型属性)。本文将攻击样本划分为两部分:一部分攻击样本放入训练集中作为已知攻击;另一部分攻击样本放入测试集中作为未知攻击。首先从训练集中随机抽取 30 400 条样本作为训练数据进行模型构建,其中正常样本 30 000 条,异常样本 400 条。再从训练集中和测试集中随机抽取六组数据样本(D1 – D6)作为测试数据,其中 D1 – D3 三组数据中包含已知攻击样本(训练集中的攻击样本)和未知攻击样本(测试集中的攻击样本,即训练集中没有的攻击样本),D4 – D6 三组只包含未知攻击样本,各组数据样本分布如表 1 所示。

表 1　测试数据样本分布

测试组	正常数据	已知攻击				未知攻击
		DOS	Probe	R2L	U2R	
D1	10 000	150	190	30	30	300
D2	10 000	200	140	40	20	200
D3	10 000	250	150	65	15	100
D4	10 000	0	0	0	0	700
D5	10 000	0	0	0	0	500
D6	10 000	0	0	0	0	300

3.2 数据预处理

本文使用文献[18]中利用贝叶斯分类器分出的 12 个关键数值属性和3 个字符型属性(protocol_type, service, flag)进行实验分析。

对于字符型属性,我们进行数值编码并归一化处理。对于数值型属性,由于采集的网络数据的属性值有不同的单位度量,各属性值之间的差别也可能很大,所以需要对其进行标准化处理[18]。设样本集 $X = \{x_1, x_2, \cdots, x_n\}$ 的容量为 n,维数为 d,x_{if} 表示第 i 个样本的第 f 个属性值,用 x'_{if} 表示标准化后的属性值,则:

$$x'_{if} = \frac{x_{if} - m_f}{s_f} \tag{1}$$

其中:

$$m_f = \frac{1}{n}\sum_{i=1}^{n} x_{if} \tag{2}$$

$$s_f = \frac{1}{n}\sum_{i=1}^{n} |x_{if} - m_f| \tag{3}$$

式中,m_f 表示样本集中第 f 个属性值的平均值,s_f 表示样本集中第 f 个属性值的平均绝对偏差。

3.3 实验结果

实验是在 CPU:2.53 GHz,内存:2048 MB, OS:Windows 8.1,开发环境:Matlab r2009a 下进行的。通过多次试验,选定迭代终止阈值 $\varepsilon = 1e-5$,$N = 100$。

本文使用 D1 组数据进行多次实验来选取合适的聚类个数 k、α 阈值和 β 阈值,其中检测率和误报率定义为:检测率 DR(Detection Rate) = 检测对的异常样本数/总的异常样本数,误报率 FAR(False Alarm Rate) = 正常样本检测为异常样本数/总的正常样本数。

图 3 表示聚类数目与检测率和误报率之间的关系。从图 3 可知,聚类数目 $k = 22$ 时,检测率和误报率都达到了较好的效果,所以本文选定 $k = 22$。

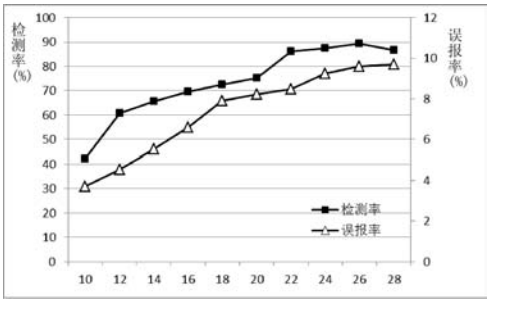

图 3　聚类数目与检测率和误报率之间的关系

图 4 表示 α 取不同值时第一层过滤的检测效果。从图 4 可知,第一层过滤的检测率和误报率大致随着 α 的增大而提高。这

和我们预想的情况是一致的,因为第一层过滤是通过阈值 α 来控制筛孔的大小,当 $\alpha < 0.6$ 时,此时筛孔太小,正常数据和异常数据都不能通过,形成图中检测率和误报率都接近零的情况。随着 α 增大时,筛孔变大,通过第一层的攻击数也就增多,附带着也会有较多的正常数据跟随通过,这样,检测率和误报率也就自然而然的提高了。当 $\alpha = 3.2$ 时,此时筛孔较大,几乎所有的异常数据和约 40% 左右的正常数据可以通过,也就是说有约 60% 的正常数据给过滤掉了。当 $\alpha > 4$ 时,此时筛孔过大,所有的正常数据和异常数据都能通过,也就导致图中的检测率和误报率都是 100%。

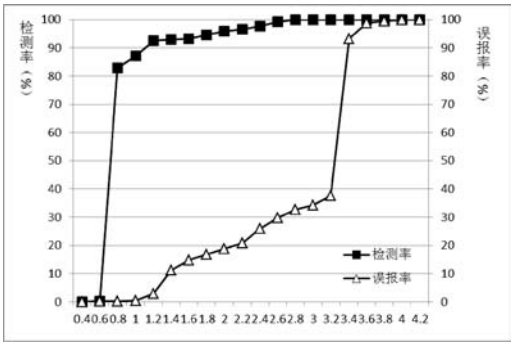

图 4 α 取不同值时第一层过滤的效果

由于本文检测模型的思想是希望通过第一层过滤掉明显的正常数据并尽可能多地筛选出异常数据,所以我们选择 $\alpha = 3.2$,此时检测率接近 100%,也能过滤掉约 60% 的正常数据。

图 5 表示 $\alpha = 3.2$ 时,β 取不同值时,第二层过滤的检测效果(即本文检测模型的最终检测效果)的变化情况。从图 5 可知,检测率和误报率大致随着阈值 β 的增大而升高,这和我们预想的情况是一致的。因为第二层过滤是首先利用 C4.5 算法对已知攻击检测率高的特点,让已知攻击通过这一层,然后利用阈值 β 调节筛孔的大小来对未知攻击和正常数据进行过滤。当 $\beta < 0.3$ 时,检测率维持在 57% 左右,这是因为筛孔太小,很多未知攻击和正常数据不能通过,但已知攻击是不受筛孔大小影响的,它是在调节筛孔之前就已通过 C4.5 算法挑拣出来了,所以此时的检测效果和单一 C4.5 算法的检测效果一致。随着 β 的慢慢增大,第二层的筛孔也慢慢增大,通过第二层的异常数据和正常数据也就慢慢多了起来,模型的检测率和误报率也就随之升高。当 $\beta = 3.2$ 时,第二层的筛孔基本接近于第一层筛孔大小,这样第二层的过滤基本不起作用,检测率和误报率与第一层也就基本相同了。显然 $\beta = 0.7$ 时,模型检测效果最好。

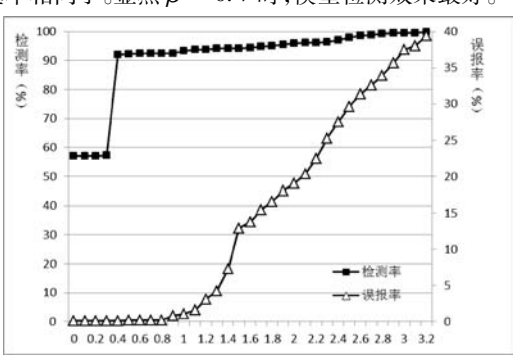

图 5 β 取不同值时模型的检测效果

表 2、表 3 为 $\alpha = 3.2$,$\beta = 0.7$,其他参数如上选取时单一算法和组合算法的性能比较,表 3 为各算法对未知攻击的检测情况。为降低实验过程中随机选取数据带来的实验差别,本文每组测试数据都进行了十次实验,取十次实验的平均值作为最终结果。

表 2 不同算法的检测性能比较试数据样本分布

测试组	FCM		C4.5		本文算法	
	检测率	误报率	检测率	误报率	检测率	误报率
D1	51.17	0.90	57.30	0.10	92.80	0.14
D2	42.57	0.94	66.20	0.12	93.98	0.23
D3	38.85	1.11	80.67	0.11	94.95	0.46

表 3 各算法对未知攻击的检测性能比较

测试组	FCM		C4.5		本文算法	
	检测率	误报率	检测率	误报率	检测率	误报率
D4	47.93	0.45	7.33	0.12	89.97	0.21
D5	47.43	0.71	7.72	0.12	89.56	0.23
D6	49.27	0.86	6.83	0.11	90.33	0.25

由表 2、表 3 可知,传统 FCM 算法对混合攻击和未知攻击的检测率大致在 50% 左右,检测率较低。这是因为 FCM 算法是通过找出网络中与正常网络行为模型偏离的行为来检测攻击的。但现实中有很多攻击行为和正常网络行为很相似,如 R2L 攻击是伪装合法用户的身份来实施攻击的,这种攻击数据与正常的数据包特征比较相似,传统 FCM 算法很难检测率出来。经调整参数 α 后,检测率和误报率都增大,从图 4 可知,在 FCM 检测率达到 100% 时,误报率达到 40% 左右,此时虽然检测率达到理想效果,但显然误报率太高。为此本文通过第二层过滤来降低误报率。

从表 2、表 3 可知,C4.5 算法对已知攻击和未知攻击的误报率都非常低,只有 0.1% 左右。这是因为 C4.5 算法是通过预先建立的入侵攻击行为模型来检测攻击行为的,显然对于已知攻击有较高的检测率,但由于未知攻击不存在于所建立的入侵攻击行为模型中,所以使得 C4.5 算法误报率低,但对未知攻击的检测能力弱。本文在第二层过滤中利用 C4.5 算法对已知攻击高检测率的特点,克服 FCM 对 R2L 类已知攻击检测难的问题,从而提高检测率。同时利用 C4.5 算法低误报率的特点,降低第一层过滤的误报率,从图 5 可知,经调整参数 β 后,检测率和误报率都增大,但在 $\beta = 0.7$ 时,本文算法能得到一个理想的效果。从表 2、表 3 可知,此时本文算法对混合攻击的检测率在 92% 以上,对未知攻击的检测率在 88% 以上,明显高于传统的 FCM 算法和 C4.5 算法。本文算法的误报率都在 0.5% 以下,贴近 C4.5 算法,明显低于传统的 FCM 算法。显然整个仿真实验结果和预计情况一致,FCM-C4.5 在入侵检测上显示出了比 FCM 和 C4.5 算法更佳的性能。

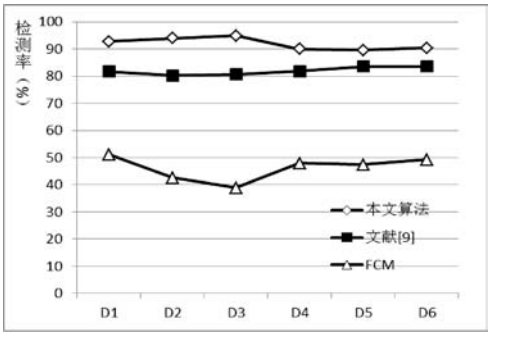

图 6 算法检测率对比

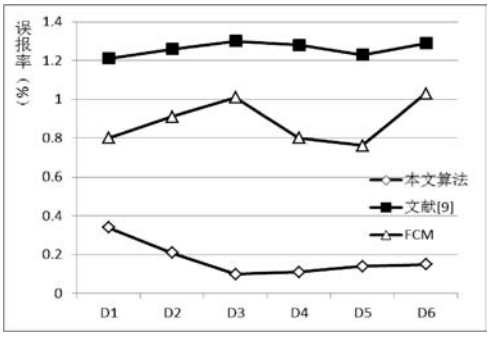

图 7　算法误报率对比

图 6、图 7 为不同算法对各测试集的检测率和误报率的对比,从对比中可以看出,本文提出的 FCM-C4.5 双过滤检测机制在 KDDCup99 数据集上的测试效果明显优于传统的 FCM 算法和文献[9]改进的聚类算法。

4　结　语

本文提出了一种基于 FCM-C4.5 的双过滤入侵检测机制,该机制首先利用 FCM 算法将明显的正常数据过滤掉,然后利用 C4.5 算法对剩下的正常数据和异常数据进行再次过滤。通过 KDD CUP 99 数据集实验表明,该机制充分发挥了 FCM 能检测到未知攻击的能力与 C4.5 低误报率和对已知攻击高检测率的优点,既克服了 FCM 检测率低的问题,又解决了 C4.5 对未知攻击检测能力差的问题;同时通过逐层过滤的方式减少了第二层需要过滤的数据量,达到效率与精度的共同提高。本文参数较多,实验中参数阈值是通过多次试验选取的,下一步研究是寻找一个自适应的方法来选取最优参数。

参　考　文　献

[1] Fisch D, Hofmann A, Sick B. On the versatility of radial basis function neural networks: A case study in the field of intrusion detection[J]. Information Sciences, 2010, 180(12): 2421 - 2439.

[2] Wang L, Jajodia S, Singhal A, et al. k-Zero day safety: A network security metric for measuring the risk of unknown vulnerabilities[J]. IEEE Transactions on Dependable and Secure Computing, 2014, 11(1): 30 - 44.

[3] Li Y, Xia J, Zhang S, et al. An efficient intrusion detection system based on support vector machines and gradually feature removal method [J]. Expert Systems with Applications, 2012, 39(1): 424 - 430.

[4] Wu S X, Banzhaf W. The use of computational intelligence in intrusion detection systems: A review[J]. Applied Soft Computing, 2010, 10(1): 1 - 35.

[5] Guo C, Zhou Y J, Ping Y, et al. Efficient intrusion detection using representative instances[J]. Computers & Security, 2013, 39(Part B): 255 - 267.

[6] Kim G, Lee S, Kim S. A novel hybrid intrusion detection method integrating anomaly detection with misuse detection[J]. Expert Systems with Applications, 2014, 41(4): 1690 - 1700.

[7] Park N H, Oh S H, Lee W S. Anomaly intrusion detection by clustering transactional audit streams in a host computer[J]. Information Sciences, 2010, 180(12): 2375 - 2389.

[8] Wang G, Hao J, Ma J, et al. A new approach to intrusion detection using Artificial Neural Networks and fuzzy clustering[J]. Expert Systems with Applications, 2010, 37(9): 6225 - 6232.

[9] 杜强,孙敏. 基于改进聚类分析算法的入侵检测系统研究[J]. 计算机工程与应用,2011,47(11):106 - 108.

[10] SivathaSindhu S S, Geetha S, Kannan A. Decision tree based light weight intrusion detection using a wrapper approach[J]. Expert Systems with applications, 2012, 39(1): 129 - 141.

[11] 蔡洪民,伍乃骐,滕少华. 分布式木马检测系统设计实现[J]. 计算机应用与软件,2012,29(5):278 - 280.

[12] Sengupta N, Sen J, Sil J, et al. Designing of on line intrusion detection system using rough set theory and Q-learning algorithm[J]. Neurocomputing, 2013, 111(2): 161 - 168.

[13] Lee S, Kim G, Kim S. Self-adaptive and dynamic clustering for online anomaly detection[J]. Expert Systems with Applications, 2011, 38(12): 14891 - 14898.

[14] Lei J Z, Ghorbani A A. Improved competitive learning neural networks for network intrusion and fraud detection[J]. Neurocomputing, 2012, 75(1): 135 - 145.

[15] 陈烨,刘渊. 基于参数优化 SVM 融合的网络异常检测[J]. 计算机应用与软件,2013,30(9):39 - 43.

[16] Teng S H, Du H L, Wu N Q, et al. A Cooperative Network Intrusion Detection Based on Fuzzy SVMs[J]. Journal of Networks, 2010, 5(4): 475 - 483.

[17] Stolfo S J, Fan W, Lee W, et al. Kdd CUP 1999 data[DB/OL]. [1999]. http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html.

[18] 肖立中,邵志清,马汉华,等. 网络入侵检测中的自动决定聚类数算法[J]. 软件学报,2008,19(8):2140 - 2148.

(上接第 253 页)

[3] Claes J, Poels G. Merging Event Logs for Process Mining: A Rule Based Merging Method and Rule Suggestion Algorithm[J]. Expert Systems with Applications, 2014, 41(16): 7291 - 7306.

[4] 王文娟,李炳龙. IDS 规则库构建与合并算法[J]. 计算机应用与软件,2010,27(11):259 - 261.

[5] 徐珊珊,董利达,朱丹,等. 一类活性 Petri 网控制器的冗余检测及结构简化[J]. 控制理论与应用,2013,30(6):673 - 682.

[6] 江先伟,王军祥. 规则不一致消解网格服务的设计与实现[J]. 电子科技,2013,26(12):138 - 140.

[7] 倪俊,陈晓苏,刘辉宇,等. 网络安全策略求精一致性检测盒冲突消解机制的研究[J]. 计算机科学,2011,38(2):32 - 37.

[8] 彭志平,夏战锋,周超. 多知识库整合技术在企业链中的应用[J]. 计算机工程,2012,38(2):82 - 84.

[9] Schmolze J G, Snyder W. Detecting Redundancy Among Production Rules Using Term Rewrite Semantics[J]. Knowledge-based System, 1999, 12(1 - 2): 3 - 11.

[10] 安莉,王建林. 用于专家系统规则库的冗余校验方法研究[J]. 计算机工程与应用,2008,44(34):191 - 193.

[11] 孙伟,郭莉,高天一,等. 一种基于有向超图的规则库冗余及环路检测方法[J]. 大连理工大学学报,2008,48(1):74 - 78.

[12] Valiente G. Verification of Knowledge based on Redundancy and Subsumption Using Transformations[J]. International Journal Expert Systems, 1993, 6(3): 341 - 355.

[13] Nazareth D L, Kennedy M H. Verification of Rule-Based Knowledge Using Directed Graphs[J]. Knowledge Acquisition, 1991, 3(4): 339 - 360.

[14] Ramaswamy M, Sarkar S, Chen Y S. Using Directed Hypergraphs to Verify Rule-Based Expert Systems[J]. Transactions on knowledge and Data Engineering, 1997, 9(2): 221 - 237.

[15] Nuffelen C V. On Adjacency Matrices for Hypergraphs[J]. Annals of Discrete Mathematics, 1980(9): 181.