基于改进支持向量机的计算机网络信息入侵检测研究

冷 静[1]
LENG Jing

摘 要 计算机网络信息数据具有高维度、复杂多变的特点，其中包含大量的噪声和无关信息，使得难以准确、高效地获取关键特征，进而导致入侵检测精度下降。为此，研究了一种基于改进支持向量机的计算机网络信息入侵检测方法。首先通过引入高效的映射函数将原始的高维、复杂的网络信息数据映射到一个更合适的特征空间。以自动提取和捕捉计算机网络信息入侵特征样本，有效过滤掉噪声和无关信息；然后利用粒子群优化算法对支持向量机的参数进行优化，以降低计算复杂度并提高模型的泛化能力。并采用优化后的支持量机对捕捉的计算机网络信息入侵特征样本分类；最后根据分类结果判定特征样本所属类别实现入侵检测。实验结果表明，该方法网络信息入侵检测精度最高时达到97%，当时间达到100 min时，本文方法的监测数据量达到了375 Mbit/s。

关键词 改进支持向量机；计算机网络信息；入侵检测；样本分类；粒子群优化

doi：10.3969/j.issn.1672-9528.2025.05.030

0 引言

网络入侵检测作为保障网络安全的重要手段，其研究具有重要的现实意义和应用价值。计算机网络信息入侵检测旨在通过对网络流量的实时监测和分析，及时发现并响应潜在的入侵行为，从而保护网络系统的安全稳定运行。然而，面对日益复杂多变的网络攻击手段，传统的入侵检测方法已难以满足当前网络安全防护的需求。

近年来，研究者们提出了一系列基于不同技术的网络入侵检测方法。例如，刘联海等人[1]研究的基于图像凸包特征的CBAM-CNN网络入侵检测方法，首先将数据集转化为RGB图像，通过提取图像中的凸包特征进行分类实现网络入侵检测。然而，将数据集转化为RGB图像的过程中，可能会丢失原始数据集中一些重要的特征信息，导致检测性能下降。王雪妍等人[2]研究的卷积神经网络结合特征融合的网络入侵检测方法将流量数据转化为灰度图像，提取纹理特征并与流量特征进行融合，利用优化后的卷积神经网络模型进行多分类实验提高入侵检测精度。然而，特征融合会导致数据维度和复杂度增加，在处理大规模网络数据时，会面临计算资源不足和训练时间过长的问题。张震等人[3]研究基于空间特征和生成对抗网络的网络入侵检测方法通过转换流量数据

1. 深圳市第二职业技术学校 广东深圳 518107

Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2020: 11534-11542.

[12] CHEN X, HUANG Y F, XU L. Multi-scale attentive residual dense network for single image rain removal[C]//Computer Vision–ACCV 2020. Berlin:Springer,2021:286-300.

[13] DING X H, ZHANG X Y, MA N N, et al. RepVGG: making VGG-style convnets great again[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2021: 13728-13737.

【作者简介】

吴章云（1999—），男，湖南郴州人，硕士研究生，研究方向：人工智能，email:741124643@qq.com。

吴传宇（1976—），通信作者（email:360626964@qq.com），男，福建建瓯人，博士，高级工程师，研究方向：农业机械化。

叶智国（1999—），男，江西上饶人，硕士研究生，研究方向：人工智能。

熊盛佳（2000—），女，湖南益阳人，硕士研究生，研究方向：人工智能。

王硕（2001—），男，黑龙江绥化人，硕士研究生，研究方向：农业工程与信息技术。

林凯旋（2002—），男，福建福州人，硕士研究生，研究方向：集成电路工程。

（收稿日期：2024-12-23 修回日期：2025-05-08）

转化为灰度图，并采用双向生成对抗网络模型进行异常检测。然而，生成对抗网络容易受到对抗样本的攻击。攻击者可能通过构造特定的对抗样本，使模型误判，从而绕过入侵检测系统。江荣等人[4]研究的基于集成学习的无监督网络入侵检测方法，使用基于3种不同异常检测理念的深度学习检测器，在3种不同集成逻辑下对各单检测器的检测结果进行检测判定。然而，在3种不同集成逻辑下对各单检测器的检测结果进行检测判定，需要选择合适的集成逻辑。在确定最佳的集成逻辑时，需要进行大量的实验和评估，导致算法复杂性提高。

针对上述问题，本文提出了一种基于改进支持向量机的计算机网络信息入侵检测方法。

1 计算机网络信息入侵检测设计

1.1 捕捉计算机网络信息入侵特征样本

通过引入高效的映射函数，将原始的高维、复杂的网络信息数据映射到一个更合适的特征空间。这个映射函数能够自动提取和捕捉计算机网络信息入侵特征样本，有效过滤掉噪声和无关信息，突出与入侵行为密切相关的关键特征，从而更准确地捕捉到入侵特征样本。为后续的入侵检测提供基础数据。首先获取计算机网络中的 n 个信息入侵数据样本，记作 $S=\{s_1, s_2, \cdots, s_n\}$，这些数据样本包含了丰富的网络活动信息，是检测入侵行为的基础。为了从这些复杂的数据中提取出有用的特征，引入一个高效的映射函数。这个函数将整体入侵特征数据映射到一个高维且可用的空间中，获得一系列映射值入侵特征样本集，记作 $\psi(s_1), \psi(s_2), \cdots, \psi(s_n)$。这里的映射值需要满足一定的条件，以确保数据的准确性和可用性[5]。具体条件用公式表示为：

$$\sum_{i=1}^{n}\psi(s_n)-\sum_{i=1}^{n}\psi(s_i)^2\neq 0 \tag{1}$$

式中：i 表示入侵特征的各个维度。

在确保这些条件得以满足的前提下，从不同维度中进一步筛选出入侵特征的关键部分。具体公式为：

$$\psi(s_k)=T\times\left[\phi(s_i)-\phi(s_n)\right] \tag{2}$$

式中：T 表示入侵特征的捕捉周期，代表收集和分析入侵特征数据的时间范围；ϕ 表示一个集合，其中包括了所有被监测到的入侵特征样本；s_i 表示这些入侵样本之间存在的特征关联与相互影响；s_n 表示在特征分析过程中被识别为不必要或重复的冗余特征样本。

通过深入分析特征提取周期内的全部入侵特征样本，揭示入侵样本之间的复杂特征关系，基于上述特征关系构建一个有效的入侵特征样本筛选矩阵，具体表示为：

$$\boldsymbol{A}=\psi(s_k)\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \tag{3}$$

式中：$\boldsymbol{A}$ 表示入侵特征样本筛选矩阵；a 表示入侵特征的分量。

根据上述矩阵挑选出最具代表性和有效性的特征样本，具体公式表示为：

$$y=\sum_{i=1}^{n}\varepsilon_i^k\psi(s_i)\boldsymbol{A} \tag{4}$$

式中：y 表示筛选出来的最优特征样本集；k 表示幂次调节参数，用于控制特征权重的非线性缩放。

通过上述计算过程排除了计算机网络信息入侵特征中那些不必要或重复的特征样本，准确捕捉计算机网络信息入侵特征样本。

1.2 基于改进支持向量机的特征样本分类

针对传统支持向量机（SVM）处理大规模样本时的计算复杂度和过拟合问题，采用粒子群优化（PSO）算法对其参数进行优化。利用参数优化后的支持向量机对捕捉的计算机网络信息入侵特征样本分类，为后续的入侵识别提供基础信息。

1.2.1 支持向量参数优化

针对传统支持向量机（SVM）在处理大规模样本时可能面临的计算复杂度和过拟合问题，对此采用粒子群优化（PSO）算法对支持向量机的参数进行优化[6-7]。粒子寻优公式表示为：

$$x_i^{'}=\lambda x_i y+\alpha_1(\text{pbest}_i-x_i)+\alpha_2(\text{gbest}_i-x_i)+\alpha_3(p_i^{'}-x_i) \tag{5}$$

式中：x_i 表示粒子位置；λ 表示初始权值；pbest_i 表示粒子的历史最优位置；gbest_i 表示群体中所有粒子的历史最优位置；α_1、α_2 和 α_3 表示学习因子；p_i' 表示适应度；x_i' 表示当前粒子位置。

在上述参数优化算法的基础上增加一个扰动项 Δx_i，以避免后期振荡问题。扰动项的具体公式表示为：

$$\Delta x_i=\alpha_1(\text{pbest}_i-x_i)+\alpha_2(\text{gbest}_i-x_i)+\alpha_3(p_i^{'}-x_i) \tag{6}$$

上述扰动项根据当前粒子位置与全局最优解之间的距离动态调整，以确保粒子在搜索过程中既能保持多样性又能逐步逼近最优解。根据上述改进策略，不断更新粒子的位置，并计算 SVM 在给定参数下的分类性能。具体计算公式为：

$$x_i^{''}=\lambda x_i+\Delta x_i^{'}+\sigma\Delta x^{'} \tag{7}$$

式中：σ 表示扰动量影响因子。

逐一比较每个粒子的当前适应度与其个体极值。如果当前适应度表现更佳，即对 pbest 进行更新。与此同时将每个粒子的 pbest 与全局最优值进行比较，如果 pbest 表现得更为优越，那么 gbest 也会被相应地更新。

通过这一系列的迭代过程找到支持向量机（SVM）的最优参数组合。通过上述操作完成支持向量机参数优化。

1.2.2 入侵特征样本分类

使用参数优化后的支持向量机对捕捉的计算机网络信息入侵特征样本分类。优化后的支持向量机具体公式表示为：

$$\min W=\sum_{i,j=1}^{n}\alpha_1\alpha_2K(y_1,y_2)-\sum_{i=1}^{n}\alpha_i+x_ib(\sum_{i=1}^{n}y_i) \tag{8}$$

式中：K 表示核函数；(y_1,y_2) 表示网络信息入侵特征样本；b 表示入侵特征样本分类属性。

支持向量机的训练样本集 X 具体表示为：

$$X=\{(x_1,x_1),(x_2,x_2),\ldots,(x_n,x_n)\}\min W \tag{9}$$

这些样本用于训练 SVM 模型，使其能够学习到不同类别之间的区分边界。在训练过程中，采用不同网络结构下的反转训练方法对网络信息入侵特征样本进行自适应分类，从而得到入侵特征样本分类的支持向量聚类中心。具体表示为：

$$G_i=\sum_j\alpha_iK(y_1,y_2)+\alpha_ib-2X \tag{10}$$

式中：G_i 表示入侵特征样本分类的支持向量聚类中心。这些聚类中心代表了不同类别在特征空间中的分布。

根据凸优化聚类约束条件推导出网络信息入侵特征样本分类的判决表达式，具体表示为：

$$\begin{cases}G_i\geq 0,\alpha_i=0 & S_R\\ G_i=0,0<\alpha_i<C & S_S\\ G_i\leq 0,\alpha_i=C & S_E\end{cases} \tag{11}$$

式中：C 表示惩罚函数；S_R、S_S 和 S_E 分别表示权值、均值和阈值 [8-9]。对样本集进行特征分解，特征分解公式具体表示为：

$$\boldsymbol{Q}^-=\begin{bmatrix}S_R\\S_S\\S_E\end{bmatrix}\begin{bmatrix}0 & y_1 & \cdots & y_n\\ y_1 & Q_{11} & \cdots & Q_{1n}\\ \vdots & \vdots & \ddots & \vdots\\ y_n & Q_{n1} & \cdots & Q_{nn}\end{bmatrix}=\begin{bmatrix}0 & y^T\\ y & \boldsymbol{Q}\end{bmatrix} \tag{12}$$

式中：$\boldsymbol{Q}$ 表示分类属性矩阵。利用正定的分类属性矩阵及其逆矩阵，提高分类的准确性和效率。通过上述操作完成计算机网络信息入侵特征样本分类，分类结果如图 1 所示。

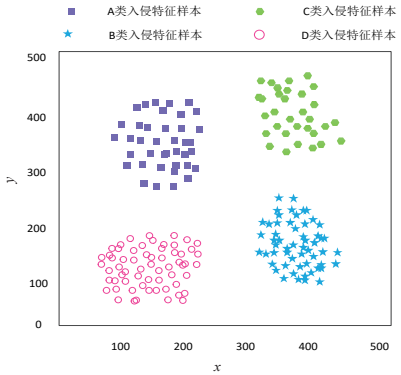

图 1 分类结果

根据图 1 可以看出，改进支持向量机能够有效实现入侵样本特征分类，为后续的入侵识别提供基础信息。

1.3 判定特征样本所属类别实现入侵检测

在得到分类结果后，根据这些结果来判定特征样本所属类别，从而实现入侵检测。首先计算待检测样本数据到支持向量机分类超平面(ω^*,c^*)之间的距离大小，具体计算公式 [10] 为：

$$r_i=y_i\left(\omega^*x_i+c^*\right)\boldsymbol{Q}^- \tag{13}$$

若该距离值小于预先设定的阈值，那么判定待检测的样本信息为正常的网络流量；相反，如果距离值大于预设的阈值，那么该待检测信息被归入不确定集合中。对于不确定集合内的入侵数据样本，进行更为深入地判断。将这些集合中的数据样本的属性特征和决策值整合成决策信息表，随后根据属性特征来划分数据的等价类。在等价类中计算出样本归属于某一类别的概率。对于边界信息集合中的数据样本，由于其判定结果不确定，采用归一化处理来进一步判断。具体计算公式为：

$$m=\frac{\omega r_i+(1-\omega)P([X(x)])}{4} \tag{14}$$

式中：$P([X(x)])$表示入侵特征样本类别归属概率。

将归一化处理后的结果与预设的域值进行对比，如果小于域值，则判定为规则信息；反之，则判定为攻击信息。通过上述步骤准确地判定样本所属类别，从而实现计算机网络信息的入侵检测。

2 实验测试

为了研究本文提出的基于改进支持向量机的计算机网络信息入侵检测方法的有效性。选用文献 [1] 基于图像凸包特征的 CBAM-CNN 网络入侵检测方法、文献 [2] 基于卷积神经网络结合特征融合的网络入侵检测方法和文献 [3] 基于空间特征和生成对抗网络的网络入侵检测方法作为对比方法，进行测试分析。

2.1 实验环境

实验环境方面，采用了高性能的服务器作为实验平台，配置了先进的处理器和大容量的内存，以确保实验的顺利进行。

服务器的具体配置包括但不限于多核高性能 CPU，以及足够的 RAM，以支持大数据量的处理和存储。此外，服务器还连接了高速的网络设备，包括千兆或万兆以太网接口，以确保数据传输的流畅性和实时性。这样的网络配置能够模拟高速数据流，从而模拟真实网络环境中的信息入侵检测，能够更准确地评估检测方法的性能和效率。

2.2 实验数据集

准备实验数据集，其中涵盖了多种类型的网络入侵行为，还包含了大量的正常网络流量数据，具体如表 1 所示。

表 1 各种计算机网络信息入侵样本数据

序号	入侵类型	样本数目
1	DOS	1 000
2	U2L	600
3	U2R	200
4	PROBE	200

2.3 检测精度结果分析

检测精度作为衡量入侵检测方法性能的核心指标之一，能直观地反映检测方法在识别正常流量与入侵流量方面的准确性。在上述实验环境中，使用本文方法和 3 种对比方法对上述实验数据集进行入侵检测，各方法的检测精度结果如图 2 所示。

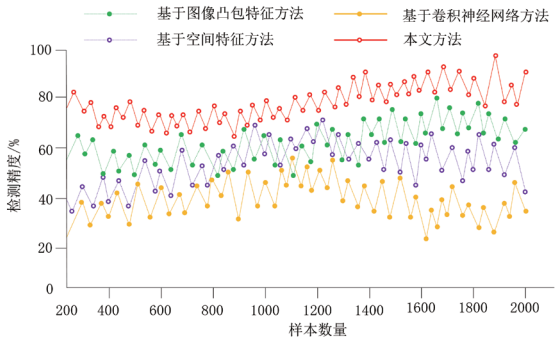

图 2 检测精度结果图

由图 2 可以清晰地看出，本文所提出的检测方法在入侵检测精度方面展现出了卓越且稳定的性能。在整个测试周期内，该方法的检测精度一直维持在较高水平，最高时达到 97%。其他 3 种方法的检测精度在不同测试阶段存在一定的波动，时而高时而低，而本文方法的检测精度曲线相对平稳，波动幅度较小，这表明该方法在面对不同类型的入侵行为和复杂网络环境时，都能够保持较高的识别准确性和稳定性。

2.4 检测效率结果分析

在评估网络入侵检测方法的整体性能时，检测效率是一个至关重要的考量因素。它直接关联到检测方法在实际应用中的响应速度、资源消耗以及用户体验等多个方面。一个高效的检测方法能够在短时间内完成大量数据的分析，从而及时响应潜在的安全威胁。4 种方法的检测效率如图 3 所示。

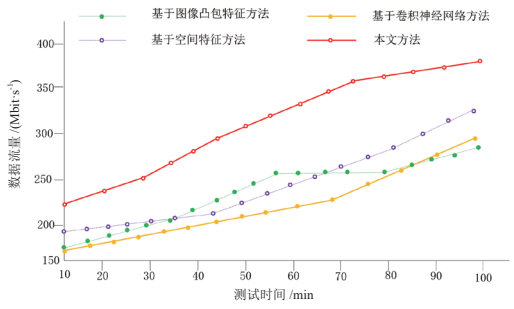

图 3 检测效率结果图

由图 3 可以看出，相较于其他 3 种检测方法，本文所提出的方法在每秒处理的数据流量上始终保持在一个较高的水平。当时间达到 100 min 时，本文方法的监测数据量达到了 375 Mbit/s。这一数据对比有力地证明了本文方法具有更为出色的检测效率。这是由于本文方法利用粒子群优化算法对支持向量机的参数进行了精细调整，从而有效降低了计算复杂度，并显著提升了模型的泛化能力，有效提高了检测精度和检测效率。

3 结语

本文提出的基于改进支持向量机的计算机网络信息入侵检测方法，通过引入高效的映射函数、粒子群优化算法以及反转训练方法和特征分解技术解决了传统入侵检测方法在计算复杂度和模型泛化能力方面的不足。实验结果表明，该方法在检测精度和检测效率上均表现出显著优势，为网络安全防护提供了更加有效和可靠的技术手段。未来，将继续深入研究，优化算法参数，拓展应用场景，为网络安全防护提供更加高效、可靠的技术支持，为构建安全稳定的网络环境贡献力量。

参考文献：

[1] 刘联海，黎汇业，毛冬晖．基于图像凸包特征的 CBAM-CNN 网络入侵检测方法 [J]. 信息网络安全，2024, 24(9): 1422-1431.

[2] 王雪妍，温蜜，李晋国，等．一种卷积神经网络结合特征融合的网络入侵检测方法 [J]. 计算机应用与软件，2024, 41(8): 359-366.

[3] 张震，周一成，田鸿朋．基于空间特征和生成对抗网络的网络入侵检测 [J]. 郑州大学学报 (工学版), 2024, 45(6): 40-47.

[4] 江荣，刘海天，刘聪．基于集成学习的无监督网络入侵检测方法 [J]. 信息网络安全，2024, 24(3): 411-426.

[5] 高倩．基于 IHGS-SVM 算法的网络入侵检测方法 [J]. 成都工业学院学报，2023, 26(5): 39-42.

[6] 周湘贞，李帅，隋栋．线性判别分析优化孪生支持向量机的网络入侵检测 [J]. 济南大学学报 (自然科学版), 2023, 37(4): 466-471.

[7] 席钰，侯致武．基于传统象群优化 - 粒子群优化 - 支持向量机的网络入侵检测 [J]. 天津理工大学学报，2022, 38(6): 58-64.

[8] 张海锋．基于无线传感网络的机房环境监控系统 [J]. 传感器世界，2024, 30(6):38-45.

[9] 杨晓文，张健，况立群，等．融合 CNN-BiGRU 和注意力机制的网络入侵检测模型 [J]. 信息安全研究，2024, 10(3): 202-208.

[10] 谭郁松，王伟，蹇松雷，等．基于异常保持的弱监督学习网络入侵检测模型 [J]. 计算机工程与科学，2024, 46(5): 801-809.

【作者简介】

冷静（1985—），女，吉林蛟河人，硕士，讲师、高级网络技术工程师，研究方向：计算机应用、计算机网络、程序设计、人工智能。

（收稿日期：2025-01-02　修回日期：2025-04-09）