文章编号：1674-0629(2025)01-0074-11 中图分类号：TM73 文献标志码：A
DOI：10. 13648/j. cnki. issn1674-0629. 2025. 01. 008

基于改进卷积神经网络的电网虚假数据注入攻击定位方法

席磊[1,2]，程琛[1]，田习龙[1]
（1. 三峡大学电气与新能源学院，湖北 宜昌 443002；2. 三峡大学梯级水电站运行与控制湖北省重点实验室，湖北 宜昌 443002）

摘要：虚假数据注入攻击通过篡改数据采集与监视控制系统采集的数据，进而破坏电力系统的稳定运行。传统虚假数据注入攻击检测方法无法对受攻击位置进行定位，亦或定位精度低。首先提出一种改进海鸥优化卷积神经网络的虚假数据注入攻击检测方法，所提方法利用具有共享权值和局部连接特性的卷积神经网络来对高维历史量测数据进行高效的特征提取及分类。然后引入具备平衡全局搜索和局部搜索能力的改进海鸥优化算法进行超参数寻优，以获得虚假数据检测的高度匹配网络结构，进而对不良数据进行检测和定位。最后通过对IEEE-14和IEEE-57节点系统进行大量攻击检测实验，验证了所提方法的有效性，并与其他多种检测方法对比，验证了所提方法的具有更优的分类性能、更高的准确率、精度、召回率和F1值。

关键词：虚假数据注入攻击；电力系统；卷积神经网络；海鸥优化；数据检测

Improved Convolutional Neural Network-Based Localization Method for False Data Injection Attacks on Power Grids

XI Lei[1,2], CHENG Chen[1], TIAN Xilong[1]
（1. College of Electrical Engineering and New Energy, China Three Gorges University, Yichang, Hubei 443002, China；2. Hubei Provincial Key Laboratory for Operation and Control of Cascaded Hydropower Station, China Three Gorges University, Yichang, Hubei 443002, China）

Abstract: False data injection attacks disrupt the stability of power systems by tampering with the data collected by data acquisition and monitoring control systems. Traditional methods for detecting false data injection attacks are unable to locate the attacked location or have low accuracy. Firstly, an improved method for detecting false data injection attacks using seagull optimized convolutional neural networks is proposed. The proposed method uses a convolutional neural network with shared weights and local connectivity to efficiently extract and classify features from high-dimensional historical measurement data. Secondly, an improved seagull optimization algorithm with balanced global and local search capabilities is introduced to perform hyperparametric optimization to obtain a highly matched network structure for false data detection. The network structure is then used to detect and locate bad data. Finally, the effectiveness of the proposed method is verified through extensive attack detection experiments on IEEE-14 and IEEE-57 node systems, and compared with various other detection methods to verify that the proposed method has better classification performance, higher accuracy, precision, recall, and F1 value.

Key words: false data injection attacks; power systems; convolutional neural networks; seagull optimization; data detection

基金项目：国家自然科学基金资助项目(52277108,S2477104)。
Foundation item: Supported by the National Natural Science Foundation of China (52277108,52477104).

0　引言

融合了传感设备、计算设备、通信系统的电力系统逐渐发展成为一个高度复杂的电力信息物理系统[1-3]。近年，针对信息完整性的虚假数据注入攻击(false data injection attack，FDIA)通过篡改数据采集与监视控制系统(supervisory control and data acquisition，SCADA)数据(如母线电压、母线有功无功功率注入、支路有功无功潮流等的量测值)，对量测值注入错误向量，并可成功避免不良数据检测机制[4]的识别，从而使得被篡改的量测数据产生偏离实际值的估计结果，导致状态估计[5]器向控制中心输出错误的结果，以至于严重威胁电力系统的安全稳定运行[6-8]。

目前，面向FDIA的检测方法可分为基于模型驱动的检测方法和基于数据驱动两大类检测方法。针对基于模型驱动的检测方法已经有众多学者进行深入研究，文献[9]提出一种分布式状态估计算法，即根据网络相邻节点估计值之差是否位于置信区间内，来计算节点状态，进而检测FDIA。当电力信息物理系统受到FDIA时，其拓扑连接相关性也会发生改变，所以文献[10]基于图论中节点度的概念，通过对状态估计结果进行离群点检测来识别FDIA的存在与否。还有学者考虑到FDIA会对电力系统的物理特性指标产生影响，从物理特性的角度寻求FDIA的检测方法[11-12]。虽然上述模型驱动的检测方法可以成功检测出FDIA，但是其检测性能强依赖于系统模型和参数，当攻击场景改变时，检测方法无法适用，可拓展性较低，最重要的是上述方法只能检测出FDIA的存在，而无法对受攻击位置进行定位。

而基于数据驱动的FDIA检测方法能够通过对大量量测数据进行特征提取及分类来实现FDIA检测，不需要系统的模型和参数，适用于不同场景，拓展性好。所谓分类，即提取具有电气特征的量测数据的显著特征进而可以区分母线节点为受攻击和不受攻击两种状态，例如文献[13]先使用PCA将历史量测数据投影到低维空间，然后训练分布式支持向量机(support vector machine，SVM)对数据进行二元分类。同样为解决维度问题，文献[14]设计了一种结合KPAC技术的极随机树检测方法，在减小计算量的同时成功实现FDIA定位检测。文献[15-16]利用K最近邻分类算法的多数投票机制对数据进行分类，确定测试样本的标签，进而实现FDIA检测。然而，无论是支持向量机还是KNN算法，都无法避免“维数诅咒”，即随着数据规模的指数型增长，会消耗大量训练时间和存储空间。因此，面对处理高维数据时，上述分类方法效果并不理想。

深度学习所具有的强特征提取能力在处理高维数据[17]的问题上具有显著优势，因此学者们采用神经网络对FDIA定位检测进行探索。文献[18]提出了多层感知机与SVM结合的混合算法，并在标准IEEE-6总线和IEEE-24总线系统进行仿真测试，验证了所提检测方法的有效性。文献[19]提出一种扩展的深度置信网络(deep believe network，DBN)架构，利用捕获的行为特征实时检测FDIA行为，并与其他机器学习算法作对比验证了所提检测方法性能更强。文献[20]利用灰狼优化算法的全局搜索能力对多隐层极限学习机(multiple hidden layers extreme learning machine，MELM)做出改进，不仅检测到FDIA的存在，而且实现了虚假数据的定位。

然而，由于上述经典神经网络的每个神经元都与其相邻两层的所有神经元两两相连，因此面临的一个潜在问题是参数数量的膨胀，同样会导致定位检测精度低、效率低。

不同于此，卷积神经网络(convolutional neural network，CNN)的神经元之间以“卷积核”作为介质，具有权值共享和局部连接的特点，使其不仅在处理高维数据时具有高效性，并且无需对输入数据进行降维预处理，在训练过程中即可有效自动提取数据特征，避免了复杂的人工特征选择的同时限制了模型参数，大大提高了训练速度。学者们已对基于CNN的FDIA定位检测进行探索。文献[21-22]利用CNN解决FDIA定位检测问题，并通过多标签二分类评价指标印证了CNN在处理关于高维度数据问题方面的优越性，所谓多标签二分类问题在FDIA检测中的体现即通过将每个系统状态变量分类为存在攻击或者不存在攻击两种类型识别具体受攻击的节点位置[19]。然而，随着学者的深入研究发现，在处理FDIA定位检测问题时，CNN的一些超参数对特征提取能力具有重要影响[23]，而人为设置超参数具有一定的盲目性和偶然性，从而影响检测精度。海鸥优化算法[24](seagull optimization algorithm，SOA)是一种具有超参数寻优特性的群

智能优化算法，广泛应用于超参数寻优问题[25-26]。然而，学者们通过长期探索发现海鸥优化算法容易陷入局部最优解。

文献[27]利用tent混沌映射策略改进海鸥优化算法的种群初始化，增加了种群多样性的同时使得种群分布更均匀。同时将t-分布扰动策略引入种群位置更新上，使其具备了平衡全局搜索和局部搜索的能力。进而提出具备平衡全局搜索和局部搜索能力的改进海鸥优化算法(improved SOA，ISOA)进行超参数寻优。

故本文针对电力信息物理系统的FDIA定位检测问题，从解决多标签二分类问题的角度[21]提出一种改进海鸥优化CNN的虚假数据注入攻击检测方法，即ISOA-CNN，来构建虚假数据检测的高度匹配网络结构，进而实现不良数据的精确定位检测。通过使用标准IEEE-14和IEEE-57系统对所提方法进行大量仿真，验证了所提方法的有效性。且分别与DBN、ELM、CNN、反向传播神经网络(backpropagation，BP)及海鸥优化卷积神经网络(SOA-CNN)检测方法进行对比分析，验证了所提方法在准确度、精度、召回率和F1值均具有最优性能。

1 FDIA模型构建

控制中心利用远程智能仪表采集量测数据，以此对电力系统进行状态估计进而做出实时决策，维持电力系统运行安全和稳定。在采集量测数据方面，同步相量测量单元(phasor measurement unit，PMU)相比于SCADA系统，由于其采样周期短和数据精度高的优点同样越来越多的应用于采集量测数据。因此，状态估计器已经朝着混合SCADA和PMU量测发展，构建FDIA模型也应考虑PMU采集的量测数据。

交流FDIA攻击目的是使得注入虚假数据前后的残差小于状态估计坏数据检测的检验阈值：

$$\left\| r_a \right\| = \min \left\| z^{\sum_A} - \boldsymbol{h}^{\sum_A}\left(x^a\right) \right\|_0 < \tau \tag{1}$$

式中：$z^{\sum A}$为在攻击区域$\sum_A$内的SCADA和PMU混合量测向量；x^a为攻击后的状态变量；$\boldsymbol{h}^{\sum_A}$为受攻击后的量测值和状态变量关系的雅可比矩阵；τ为最大标准化残差检验的阈值。

攻击者通过向量测数据注入攻击向量，针对SCADA量测数据的方程式如下。

$$P_i + \Delta P_i^{\mathrm{a}} = V_i^{\mathrm{a}} \sum_{j \in i} V_j^{\mathrm{a}} \left(G_{ij}\cos\theta_{ij}^{\mathrm{a}} + B_{ij}\sin\theta_{ij}^{\mathrm{a}} \right) \tag{2}$$

$$Q_i + \Delta Q_i^{\mathrm{a}} = V_i^{\mathrm{a}} \sum_{j \in i} V_j^{\mathrm{a}} \left(G_{ij}\sin\theta_{ij}^{\mathrm{a}} - B_{ij}\cos\theta_{ij}^{\mathrm{a}} \right) \tag{3}$$

$$P_{ij} + \Delta P_{ij}^{\mathrm{a}} = \left(g_i + g_{ij} \right)\left(V_i^{\mathrm{a}} \right)^2 - V_i^{\mathrm{a}} V_j^{\mathrm{a}} \left(g_{ij}\cos\theta_{ij}^{\mathrm{a}} + b_{ij}\sin\theta_j^{\mathrm{a}} \right) \tag{4}$$

$$Q_{ij} + \Delta Q_{ij}^{\mathrm{a}} = -\left(b_i + b_{ji} \right)\left(V_i^{\mathrm{a}} \right)^2 - V_i^{\mathrm{a}} V_j^{\mathrm{a}} \left(g_{ij}\sin\theta_{ij}^{\mathrm{a}} - b_{ij}\cos\theta_{ij}^{\mathrm{a}} \right) \tag{5}$$

针对PMU量测数据的方程式如下。

$$I_{ij}^{\mathrm{re}} + \Delta I_{ij}^{\mathrm{re,a}} = V_i^{\mathrm{a}} \left[\left(g_i + g_{ij} \right)\cos\theta_i^{\mathrm{a}} - \left(b_i + b_{ij} \right)\sin\theta_i^{\mathrm{a}} \right] - V_j^{\mathrm{a}} \left[g_{ij}\cos\theta_i^{\mathrm{a}} + b_{ij}\sin\theta_j^{\mathrm{a}} \right] \tag{6}$$

$$I_{ij}^{\mathrm{im}} + \Delta I_{ij}^{\mathrm{im,a}} = V_i^{\mathrm{a}} \left[\left(g_i + g_{ij} \right)\sin\theta_i^{\mathrm{a}} + \left(b_i + b_{ij} \right)\cos\theta_i^{\mathrm{a}} \right] - V_j^{\mathrm{a}} \left[g_{ij}\sin\theta_j^{\mathrm{a}} + b_{ij}\cos\theta_j^{\mathrm{a}} \right] \tag{7}$$

式中：下标$j \in i$表示节点j与节点i直接相连，并包括$j=i$的情况；上标a表示受到攻击后对应的量测数据；P_i和Q_i分别为注入有功和无功功率；P_{ij}和Q_{ij}分别为有功和无功功率流出；V_i^{a}和θ_i^{a}分别为索引节点处受攻击后的电压幅值和相角；V_j^{a}和θ_j^{a}分别为节点j受攻击后的电压幅值和相角；I_{ij}^{re}和I_{ij}^{im}分别为节点i与节点j相连支路电流的实部和虚部；ΔP_i^{a}、ΔQ_i^{a}、ΔP_{ij}^{a}、ΔQ_{ij}^{a}、$\Delta I_{ij}^{\mathrm{re,a}}$和$\Delta I_{ij}^{\mathrm{im,a}}$为受攻击后相应的增量变化；$g_i$和$b_i$分别为连接在节点$i$上支路的电导和电纳。

本文考虑了SCADA和PMU混合量测数据以及不完整的网络拓扑，在第4节仿真验证了从攻击者角度可以成功实现隐蔽的网络攻击。

2 ISOA-CNN检测方法

2.1 改进海鸥优化算法

种群的初始化是优化过程中重要的一步，可以保证算法的有效性，而SOA的初始个体位置是随机的，这会导致其收敛速度慢。所以本文结合文献[27]引入Tent混沌映射产生混沌粒子序列，使得初始化种群位置分布更均匀，Tent混沌映射过程如下。

$$x_{n+1,i} = \begin{cases} \dfrac{x_{n,i}}{a}, 0 \leqslant x_{n,i} \leqslant a \\ \dfrac{(1 - x_{n,i})}{1 - a}, a < x_{n,i} \leqslant 1 \end{cases} \tag{8}$$

式中：n>1表示种群规模；$x_{n,\ i}$为第n个时间步时的混沌序列；i≥1表示种群维度；a∈(0，1)，本文取a为0.5。

根据式(8)生成的混沌序列$x_{n,\ i}$，逆映射到种群空间即可得到分布相对均匀的初始化种群。

$$y_{n,i} = l_{\mathrm{b}i} + (u_{\mathrm{b}i} - l_{\mathrm{b}i})x_{n,i} \tag{9}$$

式中l_{bi}和u_{bi}分别为搜索的下界和上界。

该优化算法的全局搜索阶段是根据海鸥的迁徙行为构建，海鸥个体在满足3个移动条件后，到达新位置。

$$\boldsymbol{D}_{\mathrm{s}}(n) = \left|\boldsymbol{C}_{\mathrm{s}}(n) + \boldsymbol{M}_{\mathrm{s}}(n)\right| \tag{10}$$

式中：$\boldsymbol{D}_{\mathrm{s}}(n)$为当前个体接近最佳海鸥的位置；$\boldsymbol{C}_{\mathrm{s}}(n)$为不会与其他海鸥发生冲撞的新位置；$\boldsymbol{M}_{\mathrm{s}}(n)$为最佳位置所在方向。

局部搜索阶段是根据海鸥在空中调整飞行速度和角度逮捕猎物的攻击行为构建的，海鸥种群在三维坐标空间中更新位置的过程如下。

$$\begin{cases} x = r \times \cos(\beta) \\ y = r \times \sin(\beta) \\ z = r \times \beta \\ r = u \times \mathrm{e}^{\beta v} \end{cases} \tag{11}$$

式中：x、y、z分别为横坐标、纵坐标和竖坐标，用于描述海鸥种群在三维坐标空间中的位置；r为每个螺旋半径；β为[0，2π]范围内的随机角度；u和v为螺旋相关常数。

最终，位置更新可表示为：

$$\boldsymbol{P}'_{\mathrm{s}}(n) = \boldsymbol{x} \times \boldsymbol{y} \times \boldsymbol{z} \times \boldsymbol{D}_{\mathrm{s}}(n) + \boldsymbol{P}_{\mathrm{best}}(n) \tag{12}$$

式中：$\boldsymbol{x}$、$\boldsymbol{y}$、$\boldsymbol{z}$为海鸥种群的位置坐标向量；$\boldsymbol{P}_{\mathrm{best}}(n)$为最佳海鸥位置；$\boldsymbol{P}'_{\mathrm{s}}(n)$为迭代后海鸥的位置。

但是在优化迭代中，算法在全局搜索和局部搜索之间难以平衡。为使SOA在迭代范围内更均衡的控制全局搜索和局部搜索之间的关系，引入t-分布扰动策略，其概率密度函数$p_t(x)$如式(13)所示。

$$\begin{cases} p_t(x) = \dfrac{\Gamma\left(\dfrac{n+1}{2}\right)}{\sqrt{n\pi} \times \Gamma\left(\dfrac{n}{2}\right)} \times \left(1 + \dfrac{x^2}{n}\right)^{\frac{n+1}{2}} \\ -\infty < x < +\infty \end{cases} \tag{13}$$

本文将t-分布扰动概率设置为0.5。对每个海鸥个体都生成一个[0，1]的随机数，得到随机数小于变异概率海鸥个体的执行t-分布扰动，位置更新如式(14)所示。

$$Y^t = Y + Y \times t(iteration) \tag{14}$$

式中：Γ为伽马函数；Y^t为经过t-分布扰动更新后的海鸥个体位置；Y为海鸥个体前一阶段的位置；$t(iteration)$表示与迭代次数相关的t-分布扰动。

在算法迭代前期，t-分布扰动和Cauchy分布相似，此时对海鸥种群扰动较强，有助于开发新空间，提高全局搜索能力；在算法迭代后期，t-分布扰动和Gaussian分布相似，此时扰动能力减弱，有助于在周围空间进行搜索，提高局部搜索能力；在算法迭代中期，t-分布扰动结合Cauchy分布和Gaussian分布的优点，达到平衡全局搜索能力和局部搜索能力的目的。

2.2　ISOA-CNN

卷积神经网络是一种前馈式网络，与传统的神经网络不同的是，在全连接层前有若干卷积层和池化层，从而达到深度学习的目的，通过大量训练学习即可获得输入与输出的映射关系。

然而，学者们发现在处理FDIA定位检测问题时，CNN的关键超参数对特征提取能力具有重要影响，而人为设置超参数具有一定的盲目性和偶然性，从而影响检测精度。

所以构建ISOA-CNN检测模型解决上述问题，首先，建立用于量测数据的特征提取和降维的CNN网络，由输入层、三层卷积层、两层池化层和两层全连接层构成。

1）输入层。输入层具有n个神经元，连接遭受FDIA节点系统的n维量测数据。

2）卷积层。卷积层的功能就是对量测数据进行特征提取，本文所提检测方法中卷积层使用ReLU激活函数，卷积操作提取数据特征的过程如下所示。

$$c_{1,j} = \mathrm{ReLU}(\boldsymbol{z} * \boldsymbol{h}_{1,j} + b_{1,j}) \tag{15}$$

式中：$c_{1,\ j}$为提取的输入量测数据的隐藏特征；$\boldsymbol{z}$为输入量测数据；$\boldsymbol{h}_{1,\ j}$为第j个卷积核；$b_{1,\ j}$为相应的偏置；符号*表示卷积运算。

相似的，前一层卷积层生成的隐藏特征被用于下一层卷积层的输入，其过程如下所示。

$$c_{i,j} = \mathrm{ReLU}(\boldsymbol{c}_{i-1} * \boldsymbol{h}_{i,j} + b_{i,j}) \tag{16}$$

3）池化层。由于本文量测数据的高维性质，为提取最显著的特征信息，所以本文选取最大池化层，并且通过优化算法获得池化层的大小S_{p}。

4）全连接层。本文所设计检测模型包含两个

全连接层，其中具有ReLU激活函数的第一层全连接层处理过程如式(17)所示。

$$c_{f,j} = \mathrm{ReLU}(\boldsymbol{w}_f \times \boldsymbol{c}_{f-1} + b_f) \tag{17}$$

式中：$c_{f,j}$表示全连接层对输入特征的非线性映射；$\boldsymbol{c}_{f-1}$为由最后卷积层输出经过扁平化后的特征向量；$\boldsymbol{w}_f$和b_f分别为权重和偏置。

为实现对量测数据的多标签二分类，第二层全连接层即输出层采用Sigmoid函数，最终输出结果可以表示为：

$$\boldsymbol{y} = \mathrm{sigmoid}(\boldsymbol{w}_l \times \boldsymbol{c}_f + b_l) \tag{18}$$

式中$\boldsymbol{w}_l$和b_l分别为此全连接层的权重和偏置。第一层全连接层中的神经元个数C由优化算法确定，第二层全连接层用于分类，交流FDIA下的两个标准IEEE系统的量测数据节点状态变量即为标签数量，则所提检测模型进行仿真测试时设置输出层神经元个数分别为28和114。

总的来说，本文将对三层卷积层中卷积核大小S_f和数量N_i(i=1，2，3)，两层池化层的池大小S_p，以及全连接层神经元C 6个超参数寻优来构建最佳网络结构，以提高所提检测方法的分类精度和泛化性能。

所提方法的具体流程如图1所示。对海鸥种群进行编码时，设定ISOA中每个个体为一个六维向量X，种群数量为N，迭代次数M设置为20。以检测终端设备SCADA和PMU采集的混合量测向量作为ISOA-CNN的输入，并进行维度转化预处理，相应的预测输出结果是含“正常状态”(用0表示)和“攻击状态”(用1表示)两个决策边界[28]的标志向量。

3 模型设计

3.1 种群规模对模型的影响

在优化过程中，若种群过小，容易陷入局部最优；若种群过大，迭代的效率又会大大降低，因此选取合适大小的种群数至关重要。本文分别取种群数量M= {5，10，15，20}，通过实验分析了ISOA-CNN中种群大小对实验结果的影响。为了避免单次划分训练集和测试集导致的偶然性和随机性，本实验采用十折交叉验证，即将数据集随机打乱顺序划分为10个规格一致的互斥子集，轮流将1个子集作为测试集，剩余9个子集作为训练集，将10次实验测试集准确率均值作为选取种群数量大小的判断依据。

从图2中可以看出，在IEEE-14节点和IEEE-57节点系统测试中，选取种群大小分别为10和20时平均准确率最高，检测效果最好。

3.2 ISOA-CNN结构

在最佳种群大小下进行的十折交叉实验运行10

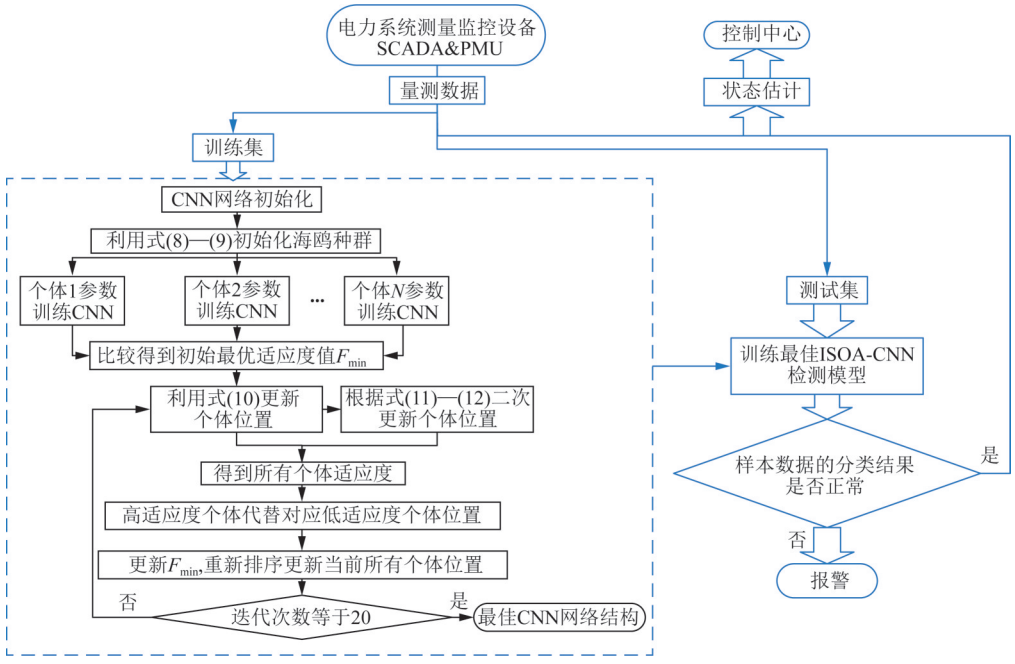

图1 基于ISOA-CNN的FDIA攻击检测流程图

Fig. 1 Flowchart of FDIA attack detection based on ISOA-CNN

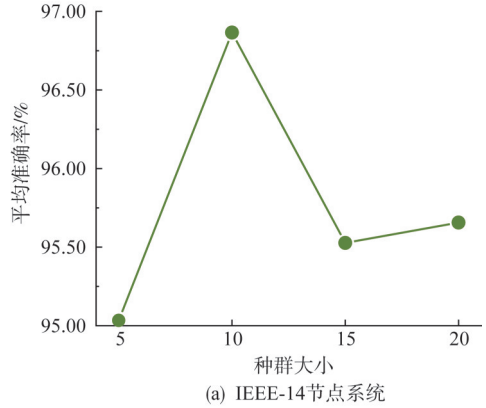

图2　不同种群大小下的检测精度

Fig. 2　Detection accuracy at different population sizes

次后，最终在两个节点系统中得到 10 组不同的 CNN 超参数最优解组合，具体如表 1 所示。

表 1　10 组超参数最优解组合

Tab. 1　Ten combinations of hyperparametric optimal solutions

运行次数	IEEE-14 节点	IEEE-57 节点
	(S_f,N_1,N_2,N_3,S_p,C)	
1	6,100,15,82,2,100	2,4,2,2,2,115
2	2,34,10,20,2,34	4,2,60,2,2,120
3	2,5,10,20,2,20	3,100,2,2,2,65
4	2,15,82,22,5,82	2,37,9,2,2,20
5	2,5,10,20,1,20	6,100,100,2,2,120
6	2,5,10,20,5,100	14,100,67,2,3,120
7	10,48,100,20,5,100	10,18,10,4,2,120
8	4,5,16,20,2,39	2,18,10,3,2,120
9	10,16,10,20,1,100	8,2,20,2,2,107
10	2,5,30,20,2,100	2,2,5,3,5,20

由于初始化权重和偏差的差异，即便使用相同的超参数对网络进行多次训练，也会导致训练结果的不同和分类性能的差异[23]。因此，为了获得性能最优的网络结构，对表 1 中的最优解组合进行 5 次重复训练，在此实验过程中，把量测数据集按照 8∶2 的比例分为训练集和测试集，结果如图 3—4 所示。

如图 3 和图 4 所示，在 IEEE-14 节点系统中组合 2 的准确率最高，在 IEEE-57 节点系统中，组合 9 的性能最好，根据超参数最优解构造的 CNN 结构如图 5 所示。

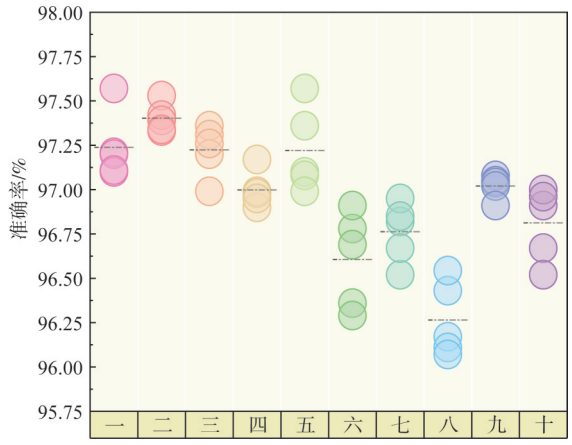

图3　IEEE-14 节点 5 次实验结果

Fig. 3　Results of five experiments on the IEEE-14 bus system

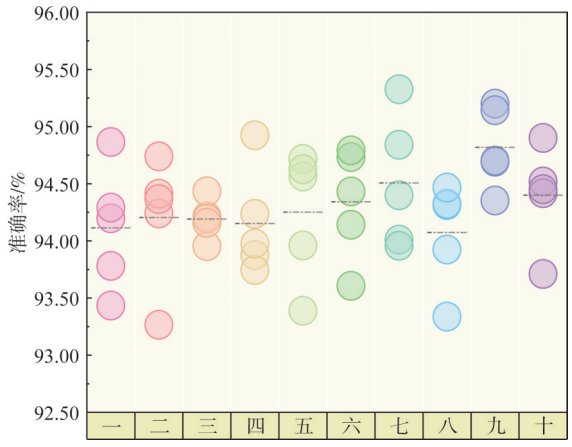

图4　IEEE-57 节点 5 次实验结果

Fig. 4　Results of five experiments on the IEEE-57 bus system

4　算例分析

4.1　性能评价指标

在多标签二分类问题中，仅有准确率并不能很好的度量检测方法性能，这里我们将量测数据标签值预测为 0 的称为正类，预测为 1 的为负类，根据混淆矩阵计算常用的二分类评价指标，如准确率 A、精确率 P、召回率 R、F_1 值来衡量模型的性能：

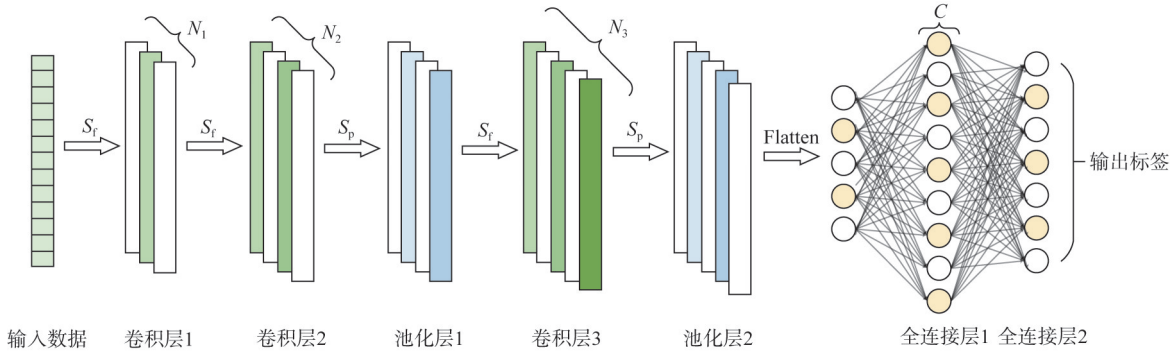

图5 构造的CNN结构

Fig. 5 Constructed CNN structure

$$A = \frac{T_P + T_N}{T_P + F_P + T_N + F_N} \tag{19}$$

$$P = \frac{T_P}{T_P + F_P} \tag{20}$$

$$R = \frac{T_P}{T_P + F_N} \tag{21}$$

$$F_1 = \frac{2 \times P \times R}{P + R} \tag{22}$$

式中：T_N为模型预测为正类，实际为正类的样本数量；F_N为模型预测为负类，实际为正类的样本数量；T_P为模型预测为负类，实际为负类的样本数量；F_P为模型预测为正类，实际为负类的样本数量。

此外本文采用接受者操作特征曲线(receiver operating characteristic，ROC)曲线来观测不同检测方法的分类性能。

4.2 通用攻击模型仿真

基于本文提出的攻击模型构建两种不同场景下的攻击算例，将具有电网完整信息和具有部分信息的攻击模型分别对IEEE-14节点和IEEE-57节点实施攻击，并对IEEE-57节点系统采取随机攻击的模式，其攻击方式更灵活多变、隐蔽性更强[29]。

为验证所提攻击模型的通用性，在某一工作日24 h内对节点系统实施攻击，每15 min采样一次，一共是96个采样点。在IEEE-14节点系统中的攻击策略是使6-13线路过载，而在IEEE-57总线系统中，每次都在攻击区域内随机选取攻击目标。为证明所提出的通用FDIA模型的可以成功躲避不良数据检测机制实施性，对两个测试系统仿真得出攻击前后的最大归一化残差如图6—7所示。

4.3 检测结果和分析

本文从解决多标签二分类问题的角度进行

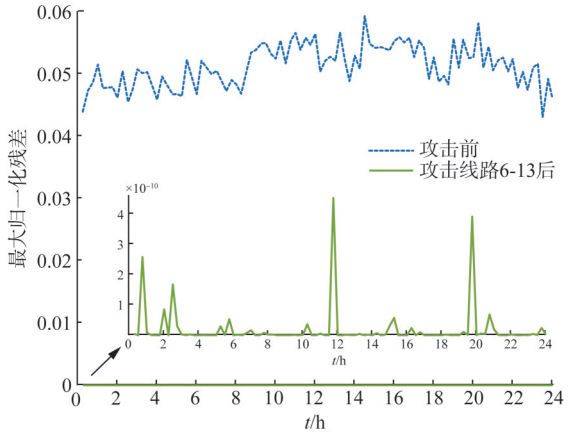

图6 IEEE-14节点系统日内攻击前后的残差

Fig. 6 Residuals before and after the attack for the IEEE-14 bus system under a workday

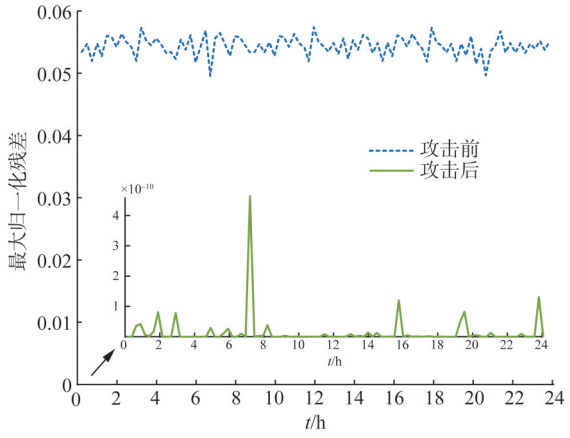

图7 IEEE-57节点系统日内攻击前后的残差

Fig. 7 Residuals before and after the attack for the IEEE-57 bus system under a workday

FDIA定位检测，当攻击某一线路或节点会导致节点系统中多个相连节点受到攻击污染，所提检测方法在攻击时段内对每个母线节点的幅值和相角状态进行分类预测，其中0表示状态正常，1表示遭受

攻击，因此通过输出的二值来定位到受到攻击的某个母线节点的幅值和相角的精确位置。成功实施攻击后获得检测终端设备的母线电压幅值，支路有功、无功潮流和母线有功、无功注入功率等量测数据，每个系统各生成2 480组量测数据，将作为检测模型的输入，且分别与基于BP神经网络[30]、SVM[31]、DBN[32]、DBN-ELM[33]、CNN[22]及SOA-CNN的检测方法进行对比分析。

本文所提检测方法最终输出的二值结果表明了每个母线状态情况，而准确率越高，表明所提检测方法的辨别能力越强，其定位到具体遭受污染的节点的结果可信度越高。从图8中可以看出，经典的二分类SVM检测方法准确率最低，且波动极大，最低准确率仅达到86. 875 %。此外，其他神经网络相关检测算法均有不同幅度的提升，例如BP、CNN以及DBN-ELM的最低准确率可达到90. 833 %、90. 416 %、89. 791 %，但在各节点状态量上的准确率仍然不稳定。而改进过的SOA-CNN检测方法在波动上有很大改善，最低准确率92. 292%，最高准确率为98. 75 %。ISOA-CNN进一步优化了检测性能，范围波动小且准确率远远高于对比算法，最低和最高准确率为95 %和98. 958 %，相对于CNN高出4. 584 %和1. 25 %，此结果验证了所提检测方法的优越性。

由图9可知，所有检测方法在IEEE-57节点系统下的准确率均有不同幅度的降低，这是由于IEEE-57节点的量测数据维度高达419，相对于IEEE-14节点数据维度的104，其数据结构更复杂，特征表达更抽象。其中，DBN检测方法的表现性能最差，最低准确率仅为39. 375%，这说明其在训练过程中无法充分提取重要特征。3种检测方法BP神经网络、SVM和DBN-ELM的分类效果接近，其最低和最高准确率分别为82. 708 %、80. 625 %、78. 125 %和98. 750 %、99. 582 %、99. 375 %，然而BP神经网络检测过程中耗时过长。所提检测方法在大部分节点状态量精度都在93 %以上，最高准确率高达99. 583 %。

ROC曲线直观展示所提算法和对比算法的鲁棒性和分辨能力，曲线越接近左上角，AUC值越趋近于1，其分类效果越好，更具有优越的检测性能。在IEEE-14节点和IEEE-57节点系统上，所提方法与对比方法的ROC曲线如图10—11所示。

从图10—11可以看出，所提方法的可以在较低的误报率情况下识别出虚假数据，在IEEE-14节点系统中，当误报率为10 %时，各种方法的正报率分别为98. 972 %、94. 427 %、93. 305 %、97. 630 %、97. 766 %、95. 982 %和99. 096 %，所提方法的AUC值为0. 99，相比于SOA-CNN和CNN分别提高了0. 005 5和0. 002。在IEEE-57节点系统中，当误报率为10%时，各个方法的正报率分别为91. 143 %、92. 501 %、30. 701 %、90. 884 %、83. 307 %、93. 627 %、96. 625 %，同样所提方法ROC曲线更接近坐标轴左上角，其AUC值最大为0. 970 9，DBN方法的AUC值最小为0. 693 2。从以上数据容易看出，在处理不同场景下不同节点数据时，本文所提方法在不良数据检测方面均优于传统神经网络CNN、BP及DBN的检测方法，在FDIA检测和定位问题上具有良好的泛化能力。

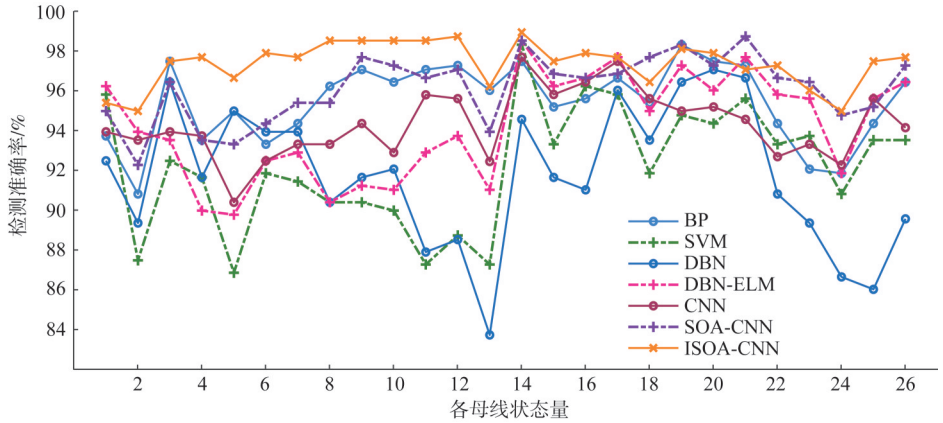

图8　IEEE-14节点系统受攻击状态准确率

Fig. 8　Attacked state detection accuracy with various methods for IEEE-14 bus system

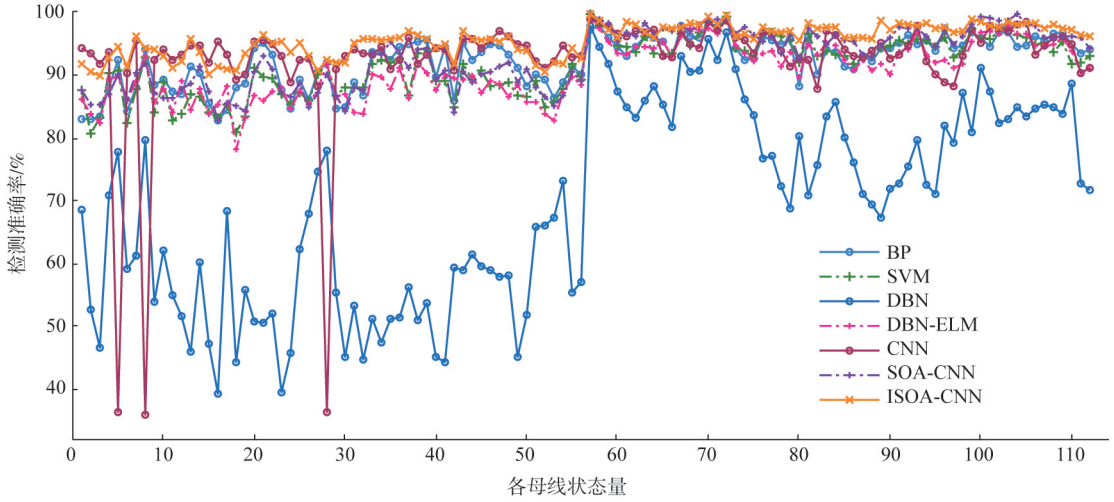

图 9 IEEE-57 节点系统受攻击状态准确率

Fig. 9 Attacked state detection accuracies with various methods for IEEE-57 bus system

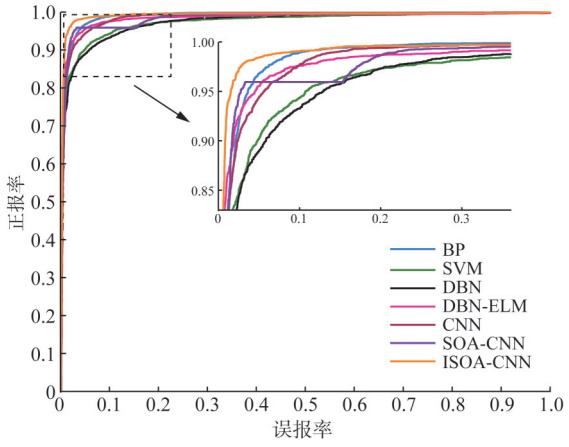

图 10 IEEE-14 节点系统 ROC 曲线

Fig. 10 ROC curves for IEEE-14 bus system

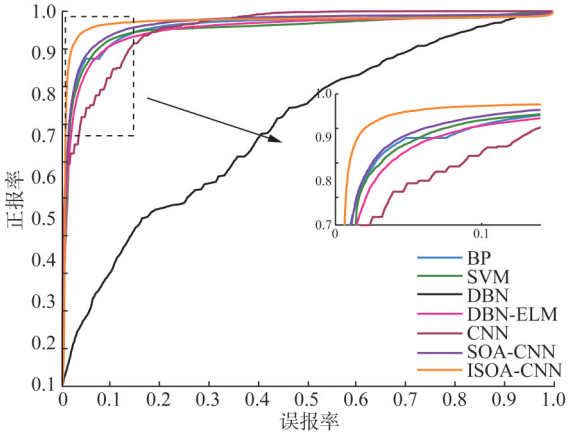

图 11 IEEE-57 节点系统 ROC 曲线

Fig. 11 ROC curves for IEEE-57 bus system

不同节点系统下，所提方法和对比方法的各项评价指标如图 12—13 所示，为避免单次实验结果偏差性，每个方法进行 5 次独立重复实验，各项结果取平均值。

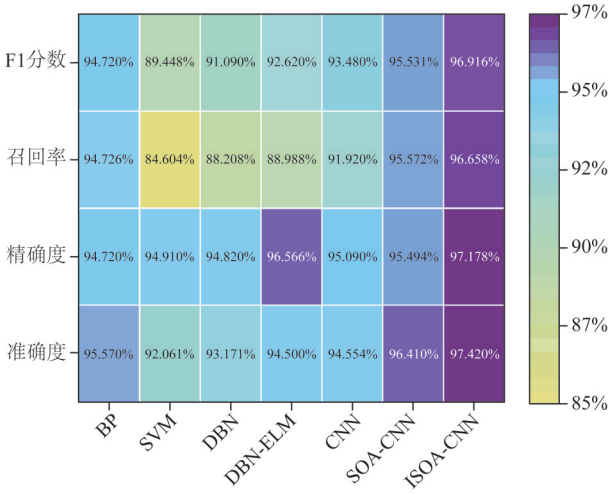

图 12 IEEE-14 节点系统性能评估指标结果

Fig. 12 Performance evaluation index results for IEEE-14 bus system

可以看到所提方法在准确率、精度、召回率和 F_1 值 4 种指标上均明显优于对比方法，甚至在 IEEE-57 节点系统中，所提方法的 F1 值相比于其他对比方法优越性更明显，分别高出 3.363 %、4.308 %、40.280 %、6.684 %、10.366 %、3.774 %。

5 结论

为解决电力信息物理系统中的 FDIA 的精确定位检测问题，提出了基于 ISOA-CNN 方法。根据 FDIA 检测问题数据复杂且规模庞大的特性，所提

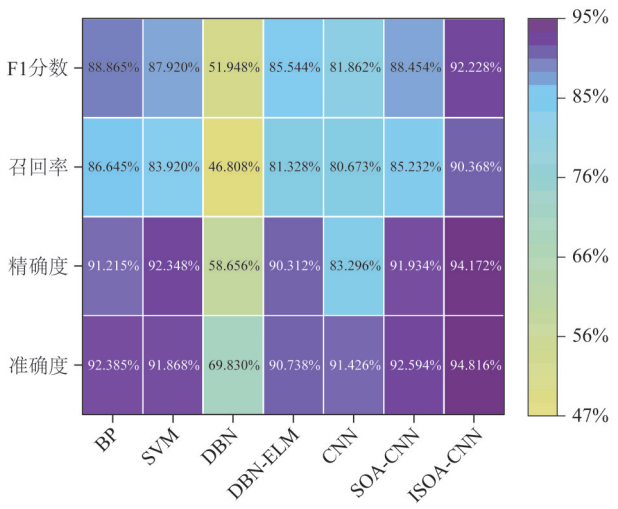

图 13　IEEE-57 节点系统性能评估指标结果

Fig. 13　Performance evaluation index results for IEEE-57 bus system

方法将检测不良数据转化为多标签二分类问题进行求解。

所提方法利用 CNN 的神经元之间以“卷积核”作为介质以及具有权值共享和局部连接的特点，使其不仅在处理高维数据时具有高效性，且无需对输入数据进行降维预处理，在训练过程中即可有效自动提取数据特征并进行分类，避免了复杂的人工特征选择的同时，限制了模型参数，大大提高了训练速度；引入了具备平衡全局搜索和局部搜索能力的 ISOA 进行超参数寻优，以获得虚假数据检测的高度匹配网络结构，进而对不良数据进行检测和定位。

针对 IEEE-14 和 IEEE-57 节点系统进行大量仿真实验，结果表明，所提方法适用于不同的攻击场景且不依赖系统拓扑结构。且与多种方法相比，所提方法在各项评价指标上的性能明显优于其他方法，具有更精确的定位检测性能。

后续工作主要针 FDIA 检测之后的状态重构，在定位到具体受攻击的位置后，应设法从量测数据中剔除掉虚假数据，进而减少对状态估计结果的影响，可以利用历史量测数据进行预测，采用预测值替换虚假数据。

参考文献

[1]　林峰，梅勇，朱益华，等. 网络攻击对电力系统典型场景全过程影响综述[J]. 南方电网技术，2023，17(11)：61－75.
LIN Feng，MEI Yong，ZHU Yihua，et al. Overview of the overall influence of cyber attack on typical scenarios of power systems [J]. Southern Power System Technology，2023，17(11)：61－75.

[2]　陈碧云，陆智，李滨. 考虑故障处理过程信息系统连通性和准确性的配电网可靠性评估[J]. 电网技术，2020，44(2)：742－750.
CHEN Biyun，LU Zhi，LI Bin. Reliability assessment of distribution network considering information system connectivity and accuracy of fault handling process [J]. Power System Technology，2020，44(2)：742－750.

[3]　田波，张越，蒙飞，等. 电网故障处置信息自适应理解框架及关键技术[J]. 中国电力，2024，57(7)：188－195.
TIAN Bo，ZHANG Yue ，MENG Fei ，et al. Adaptive understanding framework and key technology of power grid fault disposal information [J]. Electric Power，2024，57(7)：188－195.

[4]　汪际峰，吴小辰，林火华，等. 数字电网的概念、特征与架构[J]. 南方电网技术，2023，17(12)：36－41.
WANG Jifeng，WU Xiaochen ，LIN Huohua，et al. Concepts，features and architectures of digital grids [J]. Southern Power System Technology，2023，17(12)：36－41.

[5]　DENG Ruilong，XIAO Gaoxi，LU Rongxing，et al. False data injection on state estimation in power systems-attacks，impacts，and defense：a survey [J]. IEEE Transactions on Industrial Informatics，2017，13(2)：411－423.

[6]　LIU Zhaoxi，WANG Lingfeng. Leveraging network topology optimization to strengthen power grid resilience against cyber-physical attacks [J]. IEEE Transactions on Smart Grid，2021，12(2)：1552－1564.

[7]　李刚，张博，赵文清，等. 电力设备状态评估中的数据科学问题：挑战与展望[J]. 电力系统自动化，2018，42(21)：10－20，177.
LI Gang，ZHANG Bo，ZHAO Wenqing，et al. Data science issues in power equipment condition assessment：challenges and perspectives [J]. Automation of Electric Power Systems ，2018，42(21)：10－20，177.

[8]　单瑞卿，盛阳，苏盛，等. 考虑攻击方身份的电力监控系统网络安全风险分析[J]. 电力科学与技术学报，2022，37(5)：3－16.
Shan Ruiqing，SHENG Yang，SU Sheng，et al. Network security risk analysis of power monitoring system considering the identity of attacker [J]. Journal of Electric Power Science and Technology，2022，37(5)：3－16.

[9]　XIE Bin，PENG Chen，YANG Minjing，et al. A novel trust-based false data detection method for power systems under false data injection attacks [J]. Journal of the Franklin Institute，2021，358(1)：56－73.

[10]　JORJANI M，SEIF H，VARJANI A Y. A graph theory-based approach to detect false data injection attacks in power system AC state estimation [J]. IEEE Transactions on Industrial Informatics，2021，17(4)：2465－2475.

[11]　PAL S，SIKDAR B，CHOW J H. Classification and detection of PMU data manipulation attacks using transmission line parameters [J]. IEEE Transactions on Smart Grid，2018，9(5)：5057－5066.

[12]　AMELI A，HOOSHYAR A，EL-SAADANY E F. Development of a cyber-resilient line current differential relay [J]. IEEE Transactions on Industrial Informatics，2019，15 (1)：305－318.

[13] ESMALIFALAK M, LIU Lanchao, NGUYEN N, et al. Detecting stealthy false data injection using machine learning in smart grid[J]. IEEE Systems Journal, 2017, 11(3):1644 - 1652.

[14] CAMANA M R, AHMED S, GARCIA C E, et al. Extremely randomized trees-based scheme for stealthy cyber-attack detection in smart grid networks[J]. IEEE Access, 2020, 8(3): 19921 - 19933.

[15] MOHAMMADPOURFARD M, WENG Y, PECHENIZKIY M, et al. Ensuring cybersecurity of smart grid against data integrity attacks under concept drift[J]. International Journal of Electrical Power & Energy Systems, 2020(119):105947.

[16] OZAY M, ESNAOLA I, YARMAN VURAL F T, et al. Machine learning methods for attack detection in the smart grid[J]. IEEE Transactions on Neural Networks and Learning Systems, 2016, 27(8):1773 - 1786.

[17] 苏超, 杨强. 基于多视图稀疏特征选择的架空输电线路故障原因判别[J]. 智慧电力, 2023, 51(3): 96 - 103.
SU Chao, YANG Qiang. Fault cause discrimination of overhead transmission lines based on multi-view sparse feature selection[J]. Smart Power, 2023, 51(3):96 - 103.

[18] ALIMI O A, OUAHADA K, ABU-MAHFOUZ A M. Real time security assessment of the power system using a hybrid support vector machine and multilayer perceptron neural network algorithms[J]. Sustainability, 2019, 11(3):19921 - 19933.

[19] HE Youbiao, MENDIS G J, WEI Jin. Real-time detection of false data injection attacks in smart grid: a deep learning-based intelligent mechanism[J]. IEEE Transactions on Smart Grid, 2017, 8(5):2505 - 2516.

[20] 席磊, 何苗, 周博奇, 等. 基于改进多隐层极限学习机的电网虚假数据注入攻击检测[J]. 自动化学报, 2023, 49(4): 881 - 890.
XI Lei, HE Miao, ZHOU Boqi, et al. False data injection attack detection for power grids based on improved multi-hidden limit learning machine[J]. Journal of Automation, 2023, 49(4):881 - 890.

[21] WANG Shuoyao, BI Suzhi, ZHANG Yingjun Angela. Locational detection of the false data injection attack in a smart grid: a multilabel classification approach[J]. IEEE Internet of Things Journal, 2020, 7(9):8218 - 8227.

[22] 李元诚, 曾婧. 基于改进卷积神经网络的电网假数据注入攻击检测方法[J]. 电力系统自动化, 2019, 43(20): 97 - 104.
LI Yuancheng, ZENG Jing. An improved convolutional neural network-based method for detecting fake data injection attacks in power grids[J]. Automation of Electric Power Systems, 2019, 43(20):97 - 104.

[23] TIJSKENS A, JANSSEN H, ROLES S. Optimising convolutional neural networks to predict the hygrothermal performance of building components[J]. Energies, 2019, 12(20):3966 - 3984.

[24] DHIMAN G, KUMAR V. Seagull optimization algorithm: theory and its applications for large-scale industrial engineering problems[J]. Knowledge-Based Systems, 2018, 165(2019): 169 - 196.

[25] 盛四清, 关皓闻, 雷业涛, 等. 基于混沌海鸥优化算法的含光伏发电系统负荷模型参数辨识[J]. 太阳能学报, 2022, 43(7): 64 - 72.
SHENG Siqing, GUAN Haowen, LEI Yetao, et al. Parameter identification of load models for PV-containing power systems based on chaotic seagull optimization algorithm[J]. Journal of Solar Energy, 2022, 43(7):64 - 72.

[26] 许乐, 莫愿斌, 卢彦越. 基于改进海鸥优化算法的PID控制器参数优化[J]. 机床与液压, 2021, 49(16): 17 - 23.
XU Le, MO Yuanbin, LU Yanyue. Parameter optimization of PID controller based on improved seagull optimization algorithm[J]. Machine Tools and Hydraulics, 2021, 49(16): 17 - 23.

[27] 王娟, 秦江涛. 混沌映射与t-分布变异策略改进的海鸥优化算法[J]. 计算机应用研究, 2022, 39(1): 170 - 176, 182.
WANG Juan, QIN Jiangtao. Chaotic mapping and t-distribution variational strategy improvement of seagull optimization algorithm[J]. Computer Application Research, 2022, 39(1):170 - 176, 182.

[28] XUE Wenli, WU Ting. Active learning-based XGBoost for cyber physical system against generic AC false data injection attacks[J]. IEEE Access, 2020(8):144575 - 144584.

[29] 王琦, 邰伟, 汤奕, 等. 面向电力信息物理系统的虚假数据注入攻击研究综述[J]. 自动化学报, 2019, 45(1): 72 - 83.
WANG Qi, TAI Wei, TANG Yi, et al. A review of research on false data injection attacks for power information physical systems[J]. Journal of Automation, 2019, 45(1):72 - 83.

[30] 郑心仪,唐波,张嵩阳,等. 交直流并行输电线路无线电干扰的预测算法[J]. 南方电网技术, 2021, 15(10):72 - 79.
ZHENG Xinyi, TANG Bo, ZHANG Songyang, et al. Prediction algorithm for radio interference of AC/DC parallel power transmission lines[J]. Southern Power System Technology, 2021, 15(10):72 - 79.

[31] 李刚, 张博, 赵文清, 等. 电力设备状态评估中的数据科学问题:挑战与展望[J]. 电力系统自动化, 2018, 42(21): 10 - 20, 177.
LI Gang, ZHANG Bo, ZHAO Wenqing, et al. Data science issues in power equipment condition assessment: challenges and perspectives[J]. Automation of Electric Power Systems, 2018, 42(21):10 - 20, 177.

[32] 李刚, 于长海, 范辉, 等. 基于多级决策融合模型的电力变压器故障深度诊断方法[J]. 电力自动化设备, 2017, 37(11): 138 - 144.
LI Gang, YU Changhai, FAN Hui, et al. A deep diagnosis method for power transformer faults based on multi-level decision fusion model[J]. Power Automation Equipment, 2017, 37(11):138 - 144.

[33] 胡聪, 洪德华, 张翠翠, 等. 一种基于特征映射与深度学习的虚假数据注入检测方法[J]. 现代电力, 2023, 40(1): 125 - 132.
HU Cong, HONG Dehua, ZHANG Cuicui, et al. A false data injection detection method based on feature mapping and deep learning[J]. Modern Electric Power, 2023, 40(1):125 - 132.

收稿日期: 2023-06-27; **网络首发日期**:2024-04-22

作者简介:

席磊(1982), 男, 通信作者, 教授, 博士, 从事电力系统运行与控制、自动发电控制、信息物理系统网络攻击与防御、智能控制方法的研究, xilei2014@163.com;

程琛(2000), 女, 硕士研究生, 研究方向为信息物理系统网络攻击与防御, chengchen20212024@163.com;

田习龙(1997), 男, 硕士研究生, 研究方向为信息物理系统网络攻击与防御, tianxilong1024@163.com。