DOI:10. 3969/j. issn. 2097-0706. 2025. 11. 005

基于多元检测模型的信息物理系统网络攻击防御机制

Multivariate detection model-based defense mechanism against cyber attacks on cyber-physical power systems

薛雯丽[1],洪晓燕[1],杨文杰[1],吴婷[2]
XUE Wenli[1],HONG Xiaoyan[1],YANG Wenjie[1],WU Ting[2]

(1.广东省技师学院 机电工程学院,广东 惠州 516100; 2.哈尔滨工业大学(深圳) 机器人与先进制造学院,广东 深圳 518055)
(1.College of Mechanical and Electrical Engineering,Guangdong Technician College,Huizhou 516100,China;2.School of Robotics and Advanced Manufacture, Harbin Institute of Technology (Shenzhen),Shenzhen 518055, China)

摘 要: 在新型信息物理系统的发展过程中,虚假数据注入攻击构成了严重威胁,它可以通过篡改电网数据信息形成虚假的电网状态,诱导运行人员做出错误的运行决策,进而干扰电力系统的稳定运行。现有的防御手段无法解决复杂数据类型的攻击,也无法精确定位异常的状态,因此,提出多情形交流虚假数据注入攻击策略,构建更符合实际电网环境且具有强隐蔽性的攻击模型。在此基础上,设计基于多元检测模型的防御机制,有效结合极限学习机、极端梯度提升树、轻量级梯度提升器3种检测器的优点,以多情形的攻击情况为训练数据,形成高效、精准定位异常状态的攻击检测模型。攻击和防御模型都在IEEE 14和IEEE 57节点系统中仿真。试验结果验证了攻击的有效性、隐蔽性、多样性以及检测机制的实时性和准确性。

关键词: 信息物理系统;电力信息安全;虚假数据注入攻击;多元检测模型;极限学习机;极端梯度提升树;轻量级梯度提升器

中图分类号: TK 01:TM 73 **文献标志码:** A **文章编号:** 2097-0706(2025)11 - 0052 - 10

Abstract: False data injection attacks pose a severe threat that cannot be overlooked during the development of new cyber-physical power systems. These attacks can tamper with power grid data to create false grid states, mislead operators into making incorrect operational decisions, and consequently disrupt the stable operation of the power system. Moreover, existing defense methods are incapable of addressing attacks involving complex data types or pinpointing abnormal states. Therefore, a multi-scenario AC false data injection attack strategy was proposed, and an attack model better aligning with actual power grid environments and exhibiting strong stealthiness was constructed. On this basis, a defense mechanism based on a multivariate detection model was designed, effectively integrating the advantages of three detectors: extreme learning machine, extreme gradient boosting, and light gradient boosting machine. Using multi-scenario attack cases as training data, an efficient attack detection model capable of pinpointing abnormal states was formed. Both the attack and defense models were simulated in IEEE 14-bus and IEEE 57-bus systems. The experimental results verified the effectiveness, stealthiness, and diversity of the attacks, as well as the real-time performance and accuracy of the detection mechanism.

Keywords: cyber-physical power system; electric power system information security; false data injection attack; multivariate detection model; extreme learning machine; extreme gradient boosting; light gradient boosting machine

0 引言

近年来,国家大力推进“双碳”目标以及各类关于电力系统调节能力优化的专项行动,旨在加快构建新型电力系统,进一步提升电力系统调节能力和调用水平,这使得传统能源系统被进一步优化为信息物理系统(Cyber-Physical Power System, CPPS),以提升系统效率和可靠性。构建稳定高效的CPPS需要在传统能源系统的基础上建立复杂的网络基础设施,包括全方位通信网络和传感设备等。网络

基金项目: 广东省基础与应用基础研究基金项目(2024A1515011012)
Basic and Applied Basic Research Foundation of Guangdong Province (2024A1515011012)

系统部分存在暴露的风险，更容易出现网络安全威胁，因此，针对CPPS可能遭受的网络攻击进行分析，从而提出相应的防御手段，对研究CPPS的稳定运行具有重要意义。

目前，对CPPS网络攻击的研究主要分为2类。第1类研究旨在针对电力系统的实时测量和网络拓扑结构，设计低成本、高隐蔽性的攻击。例如，文献[1]中的拒绝服务攻击和虚假数据注入攻击（False Data Injection Attack，FDIA），通过破坏可再生能源发电波动下的微电网频率调节能力，破坏系统的整体稳定性。文献[2]通过实施数据完整性攻击，误导系统运营商在电力市场中非法获利。文献[3]提出了一种共振攻击，根据频率变化率来调节发电厂的输入，从而导致篡改攻击，即线路添加攻击、线路移除攻击和线路切换攻击，以干扰电力系统的运行，威胁系统安全和经济运行。文献[4]介绍了一种分布式联合攻击策略以影响和改变能源系统的运行状态。此外，引入传输线额定值攻击[5]、延迟攻击[6]等，都可实现相应的非法攻击目的。需要特别指出的是，构建电网攻击模型的核心目的在于真实模拟潜在威胁场景，通过对抗性测试验证和提升电网防御系统的鲁棒性，因此，针对CPPS的高逼真度FDIA仿真模型具有重要的研究价值。

第2类研究则侧重于为CPPS制定安全防御措施，例如保护、检测和缓解措施，以抵御网络攻击。保护措施集中在通过安装冗余监测设备和构建加密通信信道来保护关键区域测量。例如，文献[7]首先对网络攻击的不利影响进行了理论分析和定量评估，然后采用相量测量单元（Phasor Measurement Unit，PMU）配置方法，以最大限度地提高网络对此类攻击的安全性。文献[8]开发了一个双层混合整数线性规划模型，以找到需要保护的最小单元数量，并在计算时间和求解质量方面取得了令人满意的性能。文献[9]对攻击者发起的攻击向量的自由度进行研究，提出一种识别关键测量值的方法，从而最大限度地减少了攻击向量中的自由度。文献[10]提出了一种基于数据复制框架的分散式移动目标防御，通过引入2层不确定性来增强CPPS的安全性，有效限制了网络攻击。此外，一些检测措施采用传统的统计方法和人工智能方法来识别CPPS中的网络攻击，例如，文献[11]开发了一种基于无迹卡尔曼滤波器的区间状态估计方法，用于检测电力系统中的动态攻击。传统检测方法通常仅针对特定攻击类型的数据特征进行分析，缺乏泛化能力。相比之下，基于机器学习的检测方法通过自动学习攻击特征，无需人工分析数据模式，因而能够有效应对多样化的网络攻击[12-13]。例如，文献[14]采用孤立森林算法和局部线性嵌入算法提取异常分值特征和属性特征，将2种特征作为样本数据输入基于梯度提升决策树的攻击检测模型，实现虚假数据注入攻击检测，文献[15]提出一种基于图像编码与多头自注意力卷积神经网络的攻击定位检测方法，文献[16]提出联邦学习方法可能通过其高泛化能力实现对不同类型的电网异常数据进行检测。

然而，现有攻击检测方法普遍存在2个关键局限：（1）多采用单一检测模型，难以应对复杂多变的攻击场景[17]；（2）检测目标往往局限于特定攻击类型，缺乏综合防御能力。这种单一化的检测范式严重制约了电网安全防护的全面性和实时性，因此，亟须构建基于多模态融合的协同检测机制[18]，通过集成多种检测算法的优势，形成互补的防御体系，从而实现对复杂攻击的全面识别与精准定位。

针对以上问题，本文提出了一种基于多元检测方式的FDIA定位检测模型。首先，提出一种多情形的动态FDIA模型，能够更好地模拟实际运行状态下CPPS受到攻击的情形以及更多类型的攻击情况；其次，针对该攻击模型提出多元检测模型（Multivariate Detection Model，MDM），通过组合不同类型的攻击检测器，尽可能减少检测模型的偶然性，同时可以应对更多类型的攻击数据，实现全方位攻击检测。通过IEEE系统仿真数据，验证了所构建FDIA的攻击效果及该检测模型的高效性。

1　多情形FDIA模型

FDIA下的CPPS整体结构如图1所示。图中SCADA为数据采集与监视控制系统（Supervisory Control and Data Acquisition）。一般情况攻击者的目标可以概括为3个类型[19]：（1）在系统拓扑资源受限的情况下，以较低的成本发起攻击；（2）以尽可能大的冲击力对电力系统发动攻击；（3）以尽可能少的痕迹发起攻击。

1.1　攻击的隐蔽性

攻击者可能针对不同类型的目标构建相应的目标函数实施攻击，而隐蔽性是发起FDIA的关键前提[20]。隐蔽性FDIA使得攻击者可以潜伏在CPPS中，并在不被察觉的情况下持续对系统发生影响，甚至可以在关键时刻发动攻击，造成更大的破坏或获取更多的非法利益，因此，通过优化不同的攻击模型，可以实现不同程度的隐蔽性攻击，其隐蔽性程度如图2所示。

1.2 交流 FDIA 模型

一般情况下，构建通用直流潮流模型是电力系统中常用的估计方法，忽略了无功功率和电压幅值的变化[21]，主要关注有功功率和电流，模型相对简单，也更容易实现和验证，属于成本较低的攻击。然而，直流模型可能无法完全反映电力系统的实际运行状态，特别是在涉及无功功率和电压稳定性方面，甚至可能导致在某些情况下攻击效果不准确[22]。因此，针对 CPPS 构建交流攻击模型能够更准确地描述电力系统的动态行为和稳定性问题，从而模拟系统的实际运行状态。

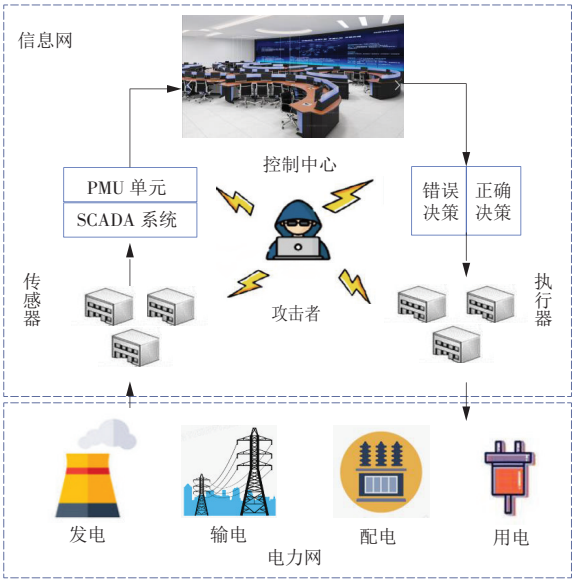

图 1 CPPS 整体结构

Fig. 1 Overall structure of a CPPS

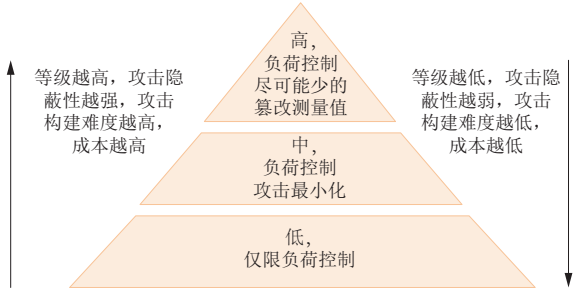

图 2 攻击隐蔽程度

Fig. 2 Levels of attack stealthiness

图 1 中 SCADA 能够实时采集电力系统的各种测量数据，并用于状态估计，为操作人员提供操作指引[23]。但是，其测量数据可能会由于传感器故障、通信延迟或中断等，导致数据不准确或丢失，甚至因为其较低的采样频率而无法准确捕获电力系统的瞬时动态变化，这也给攻击者提供了漏洞。而 PMU 则以其高精度的同步相量测量和高数据的采样率成为 SCADA 系统的有力补充[24]。因此，构建交流 FDIA 模型需要考虑 SCADA 和 PMU 的混合电力系统，以提高其隐蔽性和攻击强度。

在综合考虑系统拓扑、负载数据及线路参数的基础上，该交流 FDIA 模型以综合考虑最小化攻击成本、最大化冲击力和隐蔽性为目标，其目标函数可分别相应表述。

(1)假设攻击者可以访问全局网络拓扑结构信息，在考虑攻击隐蔽性程度的基础上，尽可能少地篡改测量值，即最小化虚假数据与正常数据之间的距离

$$\min\left\| \boldsymbol{z}_{\mathrm{H}} - \boldsymbol{h}_{\mathrm{H}}\left(\boldsymbol{x}^{\mathrm{c}}\right)\right\|_{0}, \tag{1}$$

式中：$\boldsymbol{z}_{\mathrm{H}}$ 为攻击后的混合测量值，包括 SCADA 和 PMU 数据；$\boldsymbol{h}_{\mathrm{H}}(\cdot)$ 为与攻击状态变量 $\boldsymbol{x}^{\mathrm{c}}$ 相关的非线性关系；上标 c 表示攻击后的状态。

(2)假设攻击者想要篡改总线 i—j(bus i—j)的功率，并选择 V_i 作为攻击状态量，则约束方程描述为

$$P_{ij}^{\mathrm{c}} = \left(G_{\mathrm{sh},i} + G_{ij}^{\mathrm{b}}\right)\left(V_i^{\mathrm{c}}\right)^2 - V_i^{\mathrm{c}}V_j\left(G_{ij}^{\mathrm{b}}\cos\theta_{ij} + B_{ij}^{\mathrm{b}}\sin\theta_{ij}\right), \tag{2}$$

式中：$P_{\mathrm{ij}}^{\mathrm{c}}$ 为攻击者的目标功率；G_{ij}^{b} 为串联支路的电导；B_{ij}^{b} 为串联支路的电纳；$G_{ij}^{\mathrm{b}}+\mathrm{j}B_{ij}^{\mathrm{b}}$ 为连接 bus i—j 的串联支路的导纳；$G_{\mathrm{sh},i}$ 为连接在 bus i 上的分流支路的电导率；V_i 和 θ_i 分别为 bus i 处电压幅值和相位的状态变量；V_i^{c} 为 SCADA 或 PMU 所采集的电压数据，即需要考虑对 PMU 所提供的实时数据进行攻击。

(3)攻击向量需遵守基尔霍夫定律等，即

$$Q_{ij}^{\mathrm{c}} = -\left(B_{\mathrm{sh},i} + B_{ij}^{\mathrm{b}}\right)\left(V_i^{\mathrm{c}}\right)^2 - V_i^{\mathrm{c}}V_j\left(G_{ij}^{\mathrm{b}}\sin\theta_{ij} - B_{ij}^{\mathrm{b}}\cos\theta_{ij}\right), \tag{3}$$

$$P_i^{\mathrm{c}} = V_i^{\mathrm{c}}\sum_{j\in\aleph_i}V_j\left(G_{ij}\cos\theta_{ij} + B_{ij}\sin\theta_{ij}\right), \tag{4}$$

$$Q_i^{\mathrm{c}} = V_i^{\mathrm{c}}\sum_{j\in\aleph_i}V_j\left(G_{ij}\sin\theta_{ij} - B_{ij}\cos\theta_{ij}\right), \tag{5}$$

$$I_{ij}^{\mathrm{re,c}} = V_i^{\mathrm{c}}\left[\left(G_{\mathrm{sh},i} + G_{ij}^{\mathrm{b}}\right)\cos\theta_i - \left(B_{\mathrm{sh},i} + B_{ij}^{\mathrm{b}}\right)\sin\theta_i\right] - V_j\left[G_{ij}^{\mathrm{b}}\cos\theta_j - B_{ij}^{\mathrm{b}}\sin\theta_j\right], \tag{6}$$

$$I_{ij}^{\mathrm{im,c}} = V_i^{\mathrm{c}}\left[\left(G_{\mathrm{sh},i} + G_{ij}^{\mathrm{b}}\right)\sin\theta_i + \left(G_{\mathrm{sh},i} + B_{ij}^{\mathrm{b}}\right)\cos\theta_i\right] - V_j\left[G_{ij}^{\mathrm{b}}\sin\theta_j + B_{ij}^{\mathrm{b}}\cos\theta_j\right], \tag{7}$$

式中：$B_{\mathrm{sh},i}$ 为连接在 bus i 上的分流支路的电纳；Q_{ij}^{c}，P_i^{c} 和 Q_i^{c} 分别为攻击后 SCADA 收集的无功功率、注入有功和无功功率；$I_{ij}^{\mathrm{re,c}}$ 和 $I_{ij}^{\mathrm{im,c}}$ 分别为攻击后 PMU 收集的电流相量的实部和虚部；G_{ij} 为母线导纳矩阵元素的实部(互电导)；B_{ij} 为母线导纳矩阵元素的虚部(互电纳)；$G_{ij}+\mathrm{j}B_{ij}$ 为复数母线导纳矩阵的第 i 个元素；$\aleph_i$ 为直接连接到 bus i 的总线集合。

(4)攻击模型需满足电网负载约束条件(8)，通过操纵状态估计结果诱导线路潮流偏移，进而引发连锁过载效应[25]。其设计目标是通过最大化攻击冲击力，最终导致发电机等关键物理设备损毁。不仅如此，负载约束还可以减少系统频率波动以提高

FDIA的隐蔽性，即

$$\sqrt{\left(P_l+\Delta P_l^{\mathrm{c}}\right)^2+\left(Q_l+\Delta Q_l^{\mathrm{c}}\right)^2}\geqslant P_{\mathrm{tr}\max_l},\tag{8}$$

式中：P_l和Q_l分别为攻击前支路l上的有功和无功功率流；ΔP_l^{c}和ΔQ_l^{c}分别为攻击后其相应增量；$P_{\mathrm{tr}\max_l}$为该支路l上所能承受的负荷最大值。

(5)为确保发电机l支路的视在功率有效超过其规定式(8)中的限值，还必须遵守母线电压运行限值和发电机容量限值，即

$$V_i^{\min}\leqslant V_i^{\mathrm{c}}\leqslant V_i^{\max},\tag{9}$$

$$P_{Gl}^{\min}\leqslant P_l+\Delta P_l^{\mathrm{c}}\leqslant P_{Gl}^{\max},\tag{10}$$

$$Q_{Gl}^{\min}\leqslant Q_l+\Delta Q_l^{\mathrm{c}}\leqslant Q_{Gl}^{\max},\tag{11}$$

式中：$V_i^{\min}$和$V_i^{\max}$分别为电压幅值的最小值和最大值；$P_{Gk}^{\min}$，$P_{Gk}^{\max}$，$Q_{Gk}^{\min}$和$Q_{Gk}^{\max}$分别为发电机的有功和无功功率容量的最小值和最大值。

在满足式(1)—(11)约束条件下，攻击者可构造虚假测量值$\boldsymbol{z}_{\mathrm{H}}$，并通过状态估计求得攻击向量和相应的攻击状态值$\boldsymbol{x}^{\mathrm{c}}$。由于该攻击严格遵循基尔霍夫定律等物理规律，其计算残差仍符合正态分布特性，即可以避免坏数据检测机制(Bad Data Detector，BDD)的监测。

2　基于多元检测方式的FDIA定位检测模型

基于多元的MDM检测机制，通过系统结合极端梯度提升树(Extreme Gradient Boosting，XGBoost)、轻量级梯度提升器(Light Gradient Boosting Machine，LightGBM)和极限学习机(Extreme Learning Machine，ELM)3种技术实现异常状态值的定位，其训练过程如图3所示。其中，历史CPPS数据集$\{(\boldsymbol{\nu}_i,\boldsymbol{o}_i)\}_N^{i=1}$由第1节中构建的攻击模型生成的多情形攻击数据和攻击前的原始数据构成：输入向量$\boldsymbol{\nu}_i=[v_{i1},v_{i2},\cdots,v_{iN_1}]$为CPPS所采集的混合测量数据，包括电压幅值、功率流、注入功率等；输出向量$\boldsymbol{o}_i=[o_{i1},o_{i2},\cdots,o_{iN_2}]$为一个标志向量，包括$N_2$个“正常状态”(用0表示)和“异常状态”(用1表示)。因此，该数据集组成较为复杂且具有显著特点：(1)输入数据维度较大；(2)输入数据和输出数据非一一映射的关系，即$N_1\neq N_2$；(3)电力系统拓扑结构关系导致输入数据对于不同的输出状态的重要程度不同。此外，由于动态FDIA模型生成的虚假状态向量不仅有高度结构化特征，还能严格满足基尔霍夫电路定律的物理约束，即基于残差分析的传统防御策略难以有效识别。为此，引入基于MDM的检测机制，通过集成多种机器学习算法构建协同检测体系，为电力系统防御架构提供新的技术路径。

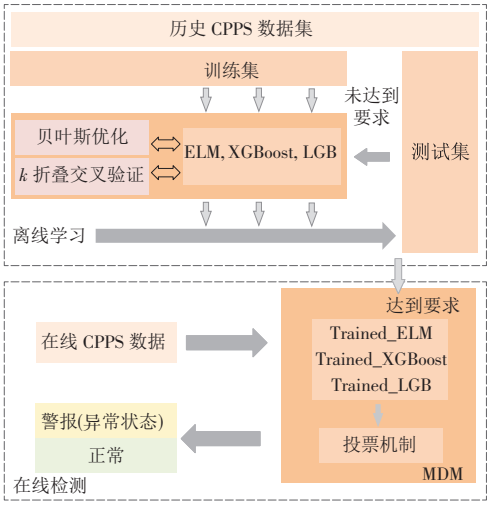

图3　MDM检测模型框架

Fig. 3　Framework of MDM

2.1　基于XGBoost算法的攻击检测模型

XGBoost算法是一种基于梯度提升决策树(Gradient Boosting Decision Tree，GBDT)的机器学习算法，即在训练过程中，低等级的决策树模型会通过迭代合成高等级的预测模型[26]，并通过梯度下降的方法减少预测误差，形成强预测模型。其中，寻找最优特征分裂点通常需要对特征值进行排序，这一步骤计算开销较大，而XGBoost采用预排序优化策略，即在训练前对所有特征值进行排序，并存储为块结构，同时利用压缩稀疏列格式进行高效存储[27]。这种设计使得针对数据集$\{(\boldsymbol{\nu}_i,o_i)\}_N^{i=1}$的训练过程可以复用排序结果，大幅减少计算量，从而提升训练效率。此外，特征排序机制还能有效衡量$\{(\boldsymbol{\nu}_i,\boldsymbol{o}_i)\}_N^{i=1}$中不同特征$\boldsymbol{\nu}_{iN}^{i=1}$对输出$\boldsymbol{o}_{iN}^{i=1}$的贡献度，有助于提升模型的解释性和泛化能力。该目标函数表示为

$$O_{\mathrm{bj}}^{(t)}=\sum_{i=1}^{M}L\left(\boldsymbol{o}_i,\hat{\boldsymbol{o}}_i^{(t)}\right)+\sum_{i=1}^{t}\Omega\left(f_i\right)=\sum_{i=1}^{M}L\left(\boldsymbol{o}_i,\hat{\boldsymbol{o}}_i^{(t-1)}+f_t\left(\boldsymbol{\nu}_i\right)\right)+\Omega\left(f_t\right),\tag{12}$$

$$\Omega\left(f_u\right)=\gamma T+0.5\lambda\|\omega\|^2,\tag{13}$$

式中：M为给定训练数据集总实例数；L为损失函数，用于衡量t次迭代中目标值$\boldsymbol{o}_i$和预测值$\hat{\boldsymbol{o}}_i^{(t)}$之间的差；$\Omega$为正则化项，用于控制模型的复杂度以防止过拟合；$f_i$为决策树；$T$和$\omega$分别为叶节点的数量和分数；$\gamma$和$\lambda$分别为用于控制叶节点的数量和分数的参数。为了加快收敛速度，对式(12)进行二阶泰勒展开，即

$$O_{\mathrm{bj}}^{(t)}=\sum_{i=1}^{M}\left[L\left(\boldsymbol{o}_i,\hat{\boldsymbol{o}}_i^{(t-1)}\right)+g_if_i\left(\boldsymbol{\nu}_i\right)+\frac{1}{2}h_if_i^2\left(\boldsymbol{\nu}_i\right)\right]+\Omega\left(f_t\right),\tag{14}$$

式中：g_i和h_i分别为损失函数的一阶和二阶梯度统计量，$g_i = \partial_{\hat{o}_i^{(t-1)}} L\left(o_i, \hat{o}_i^{(t-1)}\right)$，$h_i = \partial^2_{\hat{o}_i^{(t-1)}} L\left(o_i, \hat{o}_i^{(t-1)}\right)$。根据卡特树理论，即通过递归二分将特征空间划分为多个子区域，并在每个子区域内进行预测，可得

$$f_t(v) = \omega_j。 \tag{15}$$

式(15)可以优化为

$$O_{\mathrm{bj}}^{(t)} = \sum_{j=1}^{T}\left[\left(\sum_{i\in I_j} g_i\right)\omega_j + \frac{1}{2}\left(\sum_{i\in I_j} h_i + \lambda\right)\omega_j^2\right] + \gamma T = \sum_{j=1}^{T}\left[G_j\omega_j + \frac{1}{2}\left(H_j + \lambda\right)\omega_j^2\right] + \gamma T, \tag{16}$$

式中：$G_j = \sum_{i\in I_j} g_i$；$H_j = \sum_{i\in I_j} h_i$；$I_j$为叶节点$j$的样本集。当$\omega_j = -G/(H_j + \lambda)$时，最小目标函数表示为

$$O_{\mathrm{bj}}^{(t)} = -\frac{1}{2}\sum_{j=1}^{T}\frac{G_j^2}{H_j + \lambda} + \gamma T。 \tag{17}$$

通过贪婪算法执行决策树分割，可以求得模型的最大增益，当分割增益小于固定值或分割时间达到指定的最大深度时，可求得最终的二分类模型。

2.2 基于LightGBM算法的攻击检测模型

检测模型中输入为系统采集的混合测量数据，输出为状态变量异常与否的判断结果。输入数据维度随着系统变大指数倍增加及实时检测的需求，都对训练速度和精度提出更高要求。因此，选择异于按行分裂树(XGBoost模型)进行存储和计算的LightGBM模型，可以更好地处理大规模数据和稀疏数据[28]。

作为GBDT的改进实现，LightGBM采用了4种策略，即基于梯度的单侧采样(Gradient-based One-Side Sampling，GOSS)、逐叶生长、直方图算法和排他性特征捆绑(Exclusive Feature Bundling，EFB)，以降低传统GBDT实现的计算复杂度。(1)由于梯度较大的实例在信息增益计算中更重要，所以GOSS通过保持梯度较大的数据实例和随机采样梯度较小的样本来提高效率。(2)采用叶片生长策略的LightGBM会分裂出分裂增益最大的叶片。与水平增长策略相比，在相同的分割时间下，该策略可以获得更好的准确性。此外，该策略中增加了最大深度限制以确保高效率，同时防止过拟合。(3)FDIA检测直方图算法的核心思想是将系统测量值划分为s个区间，并在s个值中选择最佳分割点。与GBDT中基于预排序的决策树方法相比，直方图算法的内存消耗可以大大减少。(4)EFB将互斥特征绑定在一起，形成稀疏特征空间，从而降低特征维度。

总之，LightGBM通过求解近似函数$f(x)$的导数，使特定损失函数$L(y,f(x))$的期望值最小，即

$$f^{\Lambda}(x) = \arg\min E_{y,x} L(y, f(x)), \tag{18}$$

式中：$E_{y,x}$为输入数据x和目标输出y联合分布下对损失函数L的期望值。同理，通过优化决策树的目标函数得以更精确地拟合损失函数的梯度，即

$$O_{\mathrm{bj}}^{(t)} = \sum_{j=1}^{T}\left[g_i f(x_i) + \frac{1}{2}h_i f^2(x_i)\right] + \gamma(f), \tag{19}$$

式中：$\gamma(f)$为所有样本的正则化项。

2.3 基于ELM算法的攻击检测模型

鉴于数据集的维度较大，梯度提升类算法的训练速度较低，且内存占用较大，因此，采用单隐层前馈神经网络的ELM算法作为补充具有显著优势。该算法仅需设定隐层节点数，无学习率、树深度、正则化项等复杂参数，降低调优成本[29]。对于具有M个总实例的给定训练数据集，具有K个隐藏神经元节点和激活函数ϑ的单个隐藏层前馈网络的ELM模型可以表示为

$$\sum_{i=1}^{K}\boldsymbol{\beta}_i\vartheta\left(\boldsymbol{w}_i\boldsymbol{v}_j + \eta_i\right) = \boldsymbol{o}_j \quad j = 1, \cdots, M, \tag{20}$$

式中：$\boldsymbol{w}_i$和$\boldsymbol{\beta}_i$分别为将第i个隐层节点连接到输入和输出节点的权重向量；η_i为该节点的偏差；$\boldsymbol{o}_j$为第j个目标值。式(20)也可表示为

$$\boldsymbol{A}\boldsymbol{\beta} = \boldsymbol{o}, \tag{21}$$

式中：$\boldsymbol{A}$为ELM的隐层输出矩阵；$\boldsymbol{\beta} = [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_K]^{\mathrm{T}}$，为输出权重矩阵；$\boldsymbol{o} = [o_1, o_2, \cdots, o_N]^{\mathrm{T}}$，为目标向量，即

$$\boldsymbol{A} = \begin{bmatrix} \vartheta\left(\boldsymbol{w}_1\boldsymbol{v}_1 + \eta_1\right) & \cdots & \vartheta\left(\boldsymbol{w}_K\boldsymbol{v}_1 + \eta_K\right) \\ \vdots & \ddots & \vdots \\ \vartheta\left(\boldsymbol{w}_1\boldsymbol{v}_N + \eta_1\right) & \cdots & \vartheta\left(\boldsymbol{w}_K\boldsymbol{v}_N + \eta_K\right) \end{bmatrix}。 \tag{22}$$

此外，ELM学习过程需注意：(1)输入权重$\boldsymbol{w}_i$和偏差η_i为随机生成；(2)$\boldsymbol{A}$需要计算获得；(3)输出权重矩阵导出为$\boldsymbol{\beta} = \boldsymbol{A}^{-1}\boldsymbol{o}$，且通过使用奇异值分解方法获得$\boldsymbol{A}$的Moor-Penrose广义逆矩阵$\boldsymbol{A}^{-1}$。

2.4 多元检测机制

借助所提出的MDM方法整合ELM模型、Xgboost模型和LightGBM模型后，对于输出$\boldsymbol{o}_i$中的第N个状态都可以得到k_1个“1”和k_2个“0”(k_1+k_2=3)，则输出状态o_{iN}的最终表达式为

$$o_{iN} = \begin{cases} 1 & k_1 > k_2 \\ 0 & k_1 < k_2 \end{cases}。 \tag{23}$$

最终，通过上述机制，可以实现异常状态的警告从而实现防御目的。

3 仿真实例

3.1 试验说明

(1)攻击情形。为了验证本文所提出的攻击模

型，试验采用IEEE 14，IEEE 57节点系统以及2016年中国东莞电网采集的四季典型日负荷数据集，日负荷曲线如图4所示。数据集采样间隔为15 min/次。试验系统的拓扑结构和网络参数来自文献[30]，其测量设置见表1。此外，交流FDIA的攻击情形分别考虑全局攻击、局部攻击、单线路攻击、多线路攻击、逐步过载攻击等的随机组合情况。其中，全局攻击即电网全线路范围内可以发起攻击[31]；局部攻击即根据拓扑结构划分区域，可以对特定区域内的线路进行攻击；单线路攻击即针对区域中的某条线路进行特定过载攻击[32]；多线路攻击即针对区域中的多条线路进行过载攻击[15]；逐步过载攻击即在某时间区间内缓慢攻击直至线路发生过载[15]。

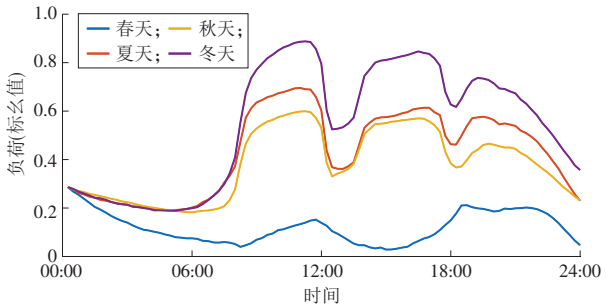

图4　四季典型日负荷曲线

Fig. 4　Typical daily load curves of four seasons

表1　测试系统的相关信息

Table 1　Relevant information of test systems

项目	IEEE 14	IEEE 57
电压幅值	14	57
功率流	40	156
注入功率	28	114
支路	14	78

（2）检测情形。本次训练检测模型所采用的数据集为IEEE 14节点和57节点系统正常运行状态下的数据（无攻击情形）和多种攻击策略下的数据（有攻击情形），样本数据集$\{(\boldsymbol{v}_i, \boldsymbol{o}_i)\}_N^{i=1}$的输入$\boldsymbol{v}_i$为所采集的混合测量值，输出$\boldsymbol{o}_i$为对应该测量值的状态情况（无攻击情形时，$\boldsymbol{o}_i$为零向量）。

3.2　仿真试验

（1）攻击情形。为验证攻击效果，采用定量指标偏差指数V_1计算攻击前后的数据。对IEEE 14节点系统在春季14:00时发起全局单线路和多线路攻击（分别攻击线路4—7，2—5，6—13，12—13），其攻击效果如图5所示。任取2种情况求其在某时间段内的最大标准化残差，结果如图6所示。对IEEE 57节点系统在冬季08:00时发起局部多线路逐步过载攻击（即08:00 — 08:15对线路36—40，37—39造成逐步过载），其动态攻击效果如图7所示。V_1计算式为

$$V_1 = \left| \frac{P_{ij} - P_{ij}^{c}}{P_{ij}} \right| = \left| \frac{\Delta P_{ij}^{c}}{P_{ij}} \right|, \tag{24}$$

式中：P_{ij}为线路i—j的功率。

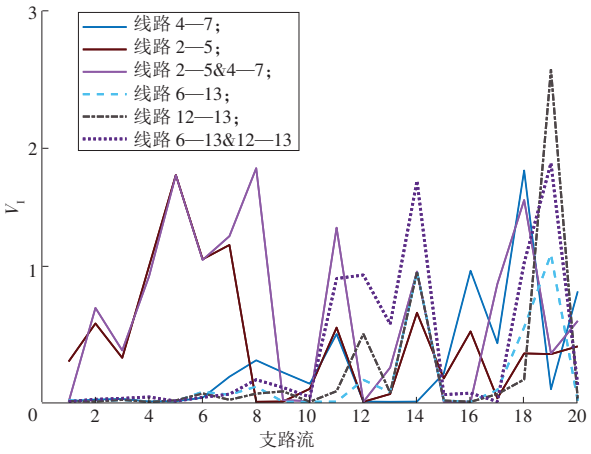

图5　IEEE 14节点系统的V_1

Fig. 5　V_1 of IEEE 14-bus system

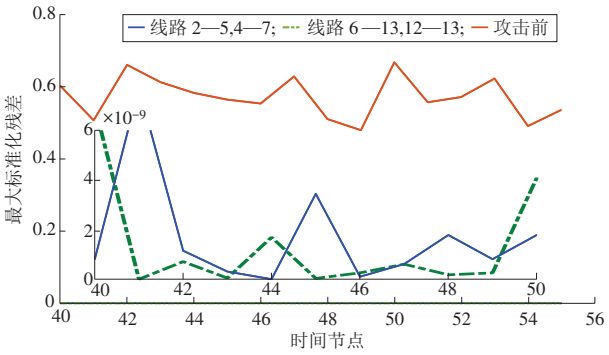

图6　IEEE 14节点系统最大标准化残差

Fig. 6　Largest normalized residuals of IEEE 14-bus system

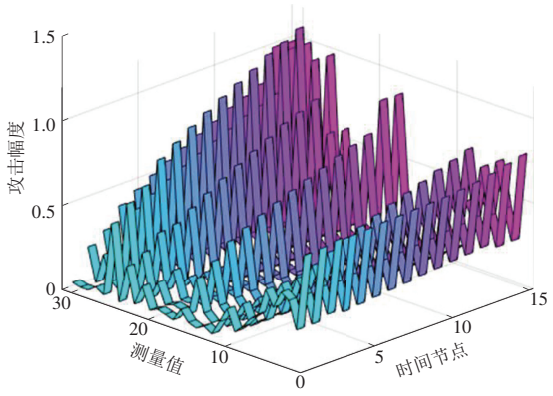

图7　IEEE 57节点系统攻击幅度

Fig. 7　Attack magnitude of IEEE 57-bus system

从图5中14节点20条支路各支路流的V_1幅度可以发现，所发起的攻击可以成功造成测量值偏差，且攻击线路不同时对线路造成的影响不同；其次，可以发现多线路攻击对测量值造成的偏差较单线路的大，影响范围也更广。由图6可得，任意选择的2种攻击情形的最大标准化残差远小于报警阈

值，即能够有效躲过BDD检测。由图7可知，动态攻击是在一段时间内对特定线路造成过载攻击，且只对局部线路造成影响；攻击幅度也随着时间的增长而逐渐增加，因而更加隐蔽；此外，整个动态攻击过程的残差也都远小于10^{-5}，即能完美避开BDD。综上所述，该动态FDIA模型可以通过不同攻击情形篡改混合测量数据，并避开BDD检测，最终使得电网运行人员做出不利于电网实际运行的操作，使得CPPS受到严重伤害。

（2）检测情形。为验证攻击效果，将单一检测模型、其他类型的攻击检测器和该多元检测器进行对比试验，对2个系统的实时检测效果和训练效果如图8—9所示。

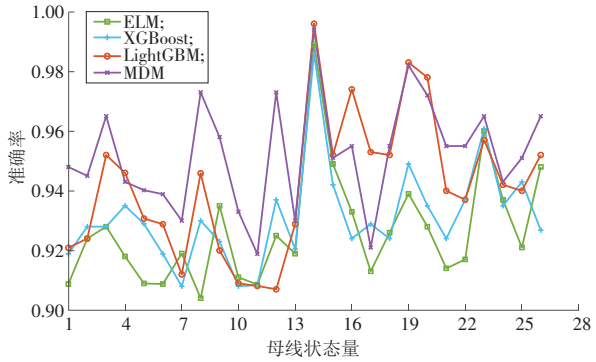

图8 IEEE 14节点系统攻击检测准确率

Fig. 8 Attack detection accuracy rate of IEEE 14-bus system

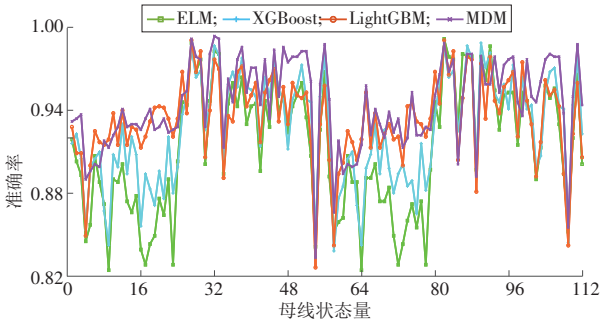

图9 IEEE 57节点系统攻击检测准确率

Fig. 9 Attack detection accuracy rate of IEEE 57-bus system

图8中，针对14节点系统的26个状态变量，对比了4种检测模型的分类准确率。以第12号状态为例，MDM方法的检测准确率达到97.8%，显著优于其他3种单一检测方法（分别为93.8%，92.3%和91.5%）。这一结果表明，基于MDM的检测策略在异常状态定位方面具有明显优势，这是由于攻击情形的多样性使得数据具有一定的复杂性，而单一检测模型一般能够高效处理某种特定类型的攻击数据。此外，可以看出多元检测模型可以显著提高识别易错状态值的准确率，例如，单一检测第8个状态值时，3个检测模型的准确率分别为94.6%，93.0%和90.4%，而多元检测模型的准确率为97.3%。从图9的IEEE 57节点系统实时检测效果可以看出，单一检测模型对于该系统数据的检测率较IEEE 14节点系统明显降低，这是由于大节点系统收集的测量数据量明显增加，线路增多导致的攻击情形更复杂等，因而对模型训练的准确率、效率都提出了更高的要求。而采用XGBoost和LightGBM检测模型后，能够利用特征重要性排序功能将数据中的重要特征进行筛选，从而在训练时增加权重而使得准确率得到进一步提升。

检测模型的检测率见表2。由表2可以看出该MDM模型对于异常状态定位的能力，例如，对于IEEE 14节点系统，能够100%定位异常状态的概率为83.1%；对于IEEE 57节点系统，能够100%定位异常状态的概率为78.6%，其中，“100%准确定位”指模型能够完整识别所有异常状态的能力。除此之外，单一检测模型容易出现检测率较低的情况（准确率<80%），而MDM方式能够有效避免这一情况。此外，由表2可知单一检测模型的平均检测率也普遍低于该MDM方法，对IEEE 14和IEEE 57节点系统的平均检测准确率达到98.4%和96.5%以上，即可以在IEEE 14节点系统中的26个状态值可以正确识别25个以上；IEEE 57节点系统中的112个状态值可以正确识别108个以上。这验证了MDM方法通过多元组合的方式能够应对单一检测模型无法处理的某些数据，从而大幅提高检测率。

表2 模型的检测率

Table 2 Detection rate of detection models %

系统	检测模型	检测率				平均检测率
		100	[90,100)	[80,90)	<80	
IEEE 14	ELM		20.8	4.3	4.3	96.8
	XGBoost		14.2	3.1	2.1	97.2
	LGB	81.5	13.2	2.8	2.5	97.8
	MDM	83.1	13.7	2.6	0.6	98.4
IEEE 57	ELM	68.6	10.8	11.5	9.1	90.5
	XGBoost	74.6	11.5	5.6	8.3	94.5
	LGB	72.5	9.8	6.2	11.5	95.5
	MDM	78.6	12.5	4.3	4.6	96.5

最终，考虑到系统数据存在一定的冗余度，而操作员也是根据全局状态信息做出相应的运行措施，因此，在MDM方法的准确率下，能够为运行人员提供有效数据参考，从而避免做出错误操作，验证了该MDM方式的有效性和高效性。

4 结论

本文针对CPPS，提出了一种能够抵御高隐蔽性、强攻击性多情形FDIA的MDM，并通过试验验证

了其有效性。主要结论如下。

(1)验证了 CPPS 的脆弱性与防御的必要性。通过构建可绕过传统 BDD 的高隐蔽性攻击模型,实证了 CPPS 面临严重安全威胁的脆弱性,凸显部署高级防御机制的必要性。

(2)检测有效性高。MDM 能有效融合 3 种单一检测模型的优势,成功解决了因数据集维度高、数据重要性不均及复杂度大导致的检测精度差的问题,对复杂攻击实现了高效识别。

(3)满足实时性需求。相较于单一检测模型,MDM 在保持高精度的同时,显著提升了检测效率,能够满足 CPPS 对安全防御的实时性要求。

(4)为系统稳定运行提供保障。该检测机制能为运行人员提供准确预警,避免其因受到攻击误导而做出错误决策,从而有效保障 CPPS 的安全稳定运行。

未来的研究将侧重于进一步提升攻击方案的逼真度与破坏性,以更全面地测试系统韧性,并据此开发更完善的检测机制以实现更加综合的防御措施。

参考文献:

[1]李瑶虹, 黄伟, 孙贝贝, 等. 拒绝服务攻击下基于分布式事件触发一致性预测补偿的微电网能量优化管理[J]. 现代电力, 2021, 38(2): 178-186.
LI Yaohong, HUANG Wei, SUN Beibei, et al. Optimal energy management in micro-grid based on distributed event-triggered consensus predictive compensation under DoS attack [J]. Modern Electric Power, 2021, 38(2): 178-186.

[2]王胜锋, 丁洲, 吴劲松, 等. 基于拓扑篡改的电力市场虚假数据注入攻击方案[J]. 电力自动化设备, 2021, 41(11): 147-152.
WANG Shengfeng, DING Zhou, WU Jinsong, et al. False data injection attack scheme of electricity market based on topology tampering [J]. Electric Power Automation Equipment, 2021, 41(11): 147-152.

[3]WU Y D, WEI Z, WENG J, et al. Resonance attacks on load frequency control of smart grids[J]. IEEE Transactions on Smart Grid, 2018, 9(5): 4490-4502.

[4]RUAN J Q, YANG C, WANG Q H, et al. Assessment of spatiotemporally coordinated cyberattacks on renewable energy forecasting in smart energy system [J]. Applied Energy, 2023, 347: 121470.

[5]YE H X, GE Y Y, LIU X, et al. Transmission line rating attack in two-settlement electricity markets [J]. IEEE Transactions on Smart Grid, 2016, 7(3): 1346-1355.

[6]YAO W T, WANG Y, XU Y, et al.Cyber-resilient control of an islanded microgrid under latency attacks and random DoS attacks [J]. IEEE Transactions on Industrial Informatics, 2023, 19(4): 5858-5869.

[7]黄崇鑫, 洪明磊, 伏帅, 等. 考虑虚假数据注入攻击的有源配电网分布式状态估计[J]. 电力工程技术, 2022, 41(3): 22-31.
HUANG Chongxin, HONG Minglei, FU Shuai, et al. Distributed state estimation of active distribution network considering false data injection attack [J]. Electric Power Engineering Technology, 2022, 41(3): 22-31.

[8]黄鸿志. 恶意攻击威胁下电力系统脆弱性分析模型与方法[D]. 北京: 华北电力大学, 2012.
HUANG Hongzhi. Analysis of power system vulnerable under deliberate attack threat [D]. Beijing: North China Electric Power University, 2012.

[9]伍虹, 杨超, 鲁杰, 等. 基于自适应 UKF 算法的虚假数据注入攻击检测研究[J]. 智能计算机与应用, 2023, 13(6): 168-173.
WU Hong, YANG Chao, LU Jie, et al. Research on false data injection attack detection based on adaptive unscented Kalman filter algorithm [J]. Intelligent Computer and Applications, 2023, 13(6): 168-173.

[10]GIRALDO J, HARIRI M E, PARVANIA M. Decentralized moving target defense for microgrid protection against false-data injection attacks [J]. IEEE Transactions on Smart Grid, 2022, 13(5): 3700-3710.

[11]WANG H Z, MENG A J, LIU Y T, et al. Unscented Kalman Filter based interval state estimation of cyber physical energy system for detection of dynamic attack [J]. Energy, 2019, 188: 116036.

[12]郑凯源. 基于机器学习的网络安全威胁检测系统设计[J]. 电脑编程技巧与维护, 2025(2): 158-160.

[13]汪锦, 张啸宇. 基于 LightGBM 的家庭负荷虚假数据注入攻击检测模型[J].综合智慧能源,2024, 46(11):1-9.
WANG Jin, ZHANG Xiaoyu. False data injection attacks detection model based on Light GBM for household load data [J]. Integrated Intelligent Energy, 2024, 46(11): 1-9.

[14]郭敬东, 刘文亮, 罗富财, 等. 基于梯度提升决策树的网络虚假数据注入攻击检测方法[J]. 自动化技术与应用, 2025, 44(5): 85-89.
GUO Jingdong, LIU Wenliang, LUO Fucai, et al. Detection method of network false data injection attacks based on gradient lifting decision tree [J]. Techniques of Automation and Applications, 2025, 44(5): 85-89.

[15]席磊, 李宗泽, 刘治洪, 等. 基于图像编码与多头自注意力卷积神经网络的电网虚假数据注入攻击检测[J/OL]. 中国电机工程学报, 2025: 1-13(2025-01-06)

[2025-03-14]. https://doi. org/10.13334/j. 0258-8013. pcsee.242225.
XI Lei, LI Zongze, LIU Zhihong, et al. False data injection attack detection of power grid based on image coding and multi-head self-attention convolutional neural network [J/OL]. Proceedings of the CSEE, 2025: 1-13 (2025-01-06) [2025-03-14]. https://doi. org/10.13334/j.0258-8013.pcsee.242225.

[16]吕永升, 张啸宇, 王榕夕, 等. 联邦学习在新型电力系统中的应用与展望[J]. 综合智慧能源, 2024, 46(11): 54-64.
YongshengLYU, ZHANG Xiaoyu, WANG Rongxi , et al. Application and prospect of federated learning in new power systems[J]. Integrated Intelligent Energy, 2024, 46(11): 54-64.

[17]曹磊, 温蜜, 何蔚. 基于深度学习的车联网的路网监测系统的DoS和DDoS攻击的入侵检测方法[J]. 计算机应用与软件, 2025, 42(1): 303-311.
CAO Lei, WEN Mi, HE Wei. Deep learning based DoS and DDoS attack detection method in the highway monitoring system of IOV [J]. Computer Applications and Software, 2025, 42(1): 303-311.

[18]李华, 陆明璇, 佟永吉, 等. 态势感知技术在新型电力系统运行中的应用[J]. 综合智慧能源, 2023, 45(3): 24-33.
LI Hua, LU Mingxuan, TONG Yongji, et al. Application of situational awareness technology in the safe and stable operation of new power systems [J]. Integrated Intelligent Energy, 2023, 45(3): 24-33.

[19]唐云泽,苏晓茜.电力系统网络攻击方法研究综述[J]. 中国信息化,2020(12):57-60.
TANG Yunze, SU Xiaoqian. A Review of research on cyberattack methods in power systems [J]. China Informatization, 2020(12): 57-60.

[20]王琦, 郜伟, 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 on false data injection attack toward cyber-physical power system [J]. Acta Automatica Sinica, 2019, 45(1): 72-83.

[21]刘石川, 慕腾, 郭裕, 等. 基于直流潮流模型的电网结构优化方法研究与应用[J]. 内蒙古电力技术, 2022, 40(2): 45-49.
LIU Shichuan, MU Teng, GUO Yu, et al. Research and application of network structure optimization method based on DC power flow model [J]. Inner Mongolia Electric Power, 2022, 40(2): 45-49.

[22]赵红嘎. 电力系统状态估计中相量量测应用及直流模型处理问题[D]. 济南: 山东大学, 2005.
ZHAO Hongga. Study on phasor measurements and DC model in power system state estimation [D]. Jinan: Shandong University, 2005.

[23]王文杰, 房海腾, 朱成杰, 等. 基于KNN分类算法的电力系统网络虚假数据注入攻击防御方法[J]. 微型电脑应用, 2024, 40(10): 130-134.
WANG Wenjie, FANG Haiteng, ZHU Chengjie, et al. A defense method for false data injection attacks in power system networks based on KNN classification algorithm[J]. Microcomputer Applications, 2024, 40(10): 130-134.

[24]朱斌, 胡威. 基于同步相量测量的电力系统稳定性分析与控制策略分析[J]. 电工技术, 2024(S2): 484-486.
ZHU Bin, HU Wei. Stability analysis and control strategy analysis of power system based on synchronous phasor measurement [J]. Electric Engineering, 2024 (S2): 484-486.

[25]伏坚, 胡博, 谢开贵, 等. 应对协同攻击的电力系统发输电拓展随机规划[J]. 电力系统自动化, 2021, 45(2): 21-29.
FU Jian, HU Bo, XIE Kaigui, et al. Stochastic planning of generation and transmission expansion for power system against coordinated attacks [J]. Automation of Electric Power Systems, 2021, 45(2): 21-29.

[26]黄冬梅, 杨旭, 胡安铎, 等. 基于CNN-BiGRU-XGBoost的新型电力系统虚假数据注入攻击检测[J]. 电网技术, 2025, 49(5): 2119-2127.
HUANG Dongmei, YANG Xu, HU Anduo, et al. Detection of false data injection attack in new power system based on CNN-BiGRU-XGBoost [J]. Power System Technology, 2025, 49(5): 2119-2127.

[27]翟千惠, 朱萌, 俞阳, 等. 基于XGBoost算法的电力虚假数据注入攻击残差检测[J]. 电子设计工程, 2025, 33(1): 109-112, 117.
ZHAI Qianhui, ZHU Meng, YU Yang, et al. Residual detection of power false data injection attacks based on XGBoost algorithm [J]. Electronic Design Engineering, 2025, 33(1): 109-112, 117.

[28]李俊颉, 高莲, 李鹏, 等. 基于DDPM-LightGBM的电力CPS多标签不平衡虚假数据注入攻击的检测[J]. 昆明理工大学学报(自然科学版), 2025, 50(3): 49-57.
LI Junjie, GAO Lian, LI Peng, et al. Detection of imbalanced multi-class false data injection attacks in cyber-physical systems based on DDPM-Light GBM [J]. Journal of Kunming University of Science and Technology (Natural Science), 2025, 50(3): 49-57.

[29]常英贤, 郭阳, 王越越, 等. 基于混合集成学习的电网安全评估模型研究[J]. 自动化仪表, 2025, 46(3): 43-48.
CHANG Yingxian, GUO Yang, WANG Yueyue, et al. Research on grid security assessment model based on hybrid integrated learning [J]. Process Automation

Instrumentation, 2025, 46(3): 43-48.
[30]LI Y F, XUE W L, WU T, et al. Intrusion detection of cyber physical energy system based on multivariate ensemble classification[J]. Energy, 2021, 218: 119505.
[31]阮兆文, 孟干, 周冬青, 等. 智能电网中的虚假数据注入攻击检测方法研究[J]. 自动化与仪器仪表, 2019(3): 49-52.
RUAN Zhaowen, MENG Gan, ZHOU Dongqing, et al. Research on false data injection attack detection method in smart grid[J]. Automation & Instrumentation, 2019(3): 49-52.
[32]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. Risk anlysis of power system cyber securityc onsidering identity of malicious adversaries[J]. Journal of Electric Power Science and Technology, 37(5): 3-16.

(本文责编:王庆霞)

收稿日期:2025-03-15;**修回日期**:2025-10-16
上网日期:2025-11-25;**附录网址**:www.iienergy.cn

作者简介:

薛雯丽(1996),女,硕士,从事机电一体化、智能电网安全等方面的研究,283793143@qq.com;

洪晓燕(1976),女,高级讲师,从事机电一体化、自动控制、模具设计等方面的研究,2352054433@qq.com;

杨文杰(1986),男,高级讲师,从事机电一体化、智能电网数据分析等方面的研究,234595761@qq.com;

吴婷(1987),女,副教授,博士,从事信息物理融合系统攻防博弈、交通电力系统优化规划与运行控制等方面的研究,twu920@hotmail.com。

****

广 告 索 引