人工智能技术在电网数据安全防护中的应用

常雅楠[1] 张云鹏[2] 房宏基[3]
CHANG Yanan ZHANG Yunpeng FANG Hongji

摘 要 随着智能电网的快速发展，其数据规模与复杂性急剧增加，传统安全防护手段难以有效应对动态化、多样化的安全威胁，数据安全防护面临严峻挑战。为提升智能电网数据安全防护的智能化水平与防御效能，文章聚焦人工智能（AI）技术在智能电网数据安全防护中的应用，探讨其核心技术在数据分级、异常检测、加密算法等场景下的实践价值。实验结果表明，AI 方案在异常流量检测准确率、响应时间和数据恢复效率等方面显著优于传统方法，为智能电网的安全运行提供了高效、智能化的解决方案。

关键词 人工智能（AI）；电网数据；安全防护

doi：10.3969/j.issn.1672-9528.2025.12.032

0 引言

在能源变革与科技进步的双重驱动下，可靠、安全、高效、环保和用户友好的新型电力网络逐渐成为应对能源需求增长和能源结构变化的必然选择[1]。智能电网（Smart Grid），是先进传感、通信、控制技术与传统电网深度融合的产物，被普遍认为是新一代电网技术解决方案。随着“双碳”战略的推进和新型电力系统建设的深入，智能电网正逐步从概念走向规模化部署，其运行模式也从单向供电向“源 - 网 - 荷 - 储”协同互动演进。智能电网将数据、通信及控制技术深度应用于电力系统发、输、变、配、用各个环节。其演进大致可分为三个阶段，第一阶段是通过信息化，实现信息流与电力流在电网内的一体化双向流动，在满足用户侧电力需求的基础上，优化资源配置及调度，实现节能减排，提高经济效益；第二阶段则是在业务线上化、信息化的基础之上，发挥数据驱动赋能作用，实现业务的数字化，从而推动业务赋能增效；第三阶段是融合人工智能（AI）大模型的推演及预测仿真能力，实现业务的智能化升级，从而辅助甚至替代人工进行作业、决策。智能电网作为现代能源系统的核心，承载着海量敏感数据，包括用户用电信息、电网运行参数及设备状态日志等。然而，其复杂的网络架构和开放的业务场景使其成为网络攻击的主要目标[2]。近年来，随着电网的数字化转型、智能化升级的进程加速，其数据规模呈现爆炸式增长。据行业统计，一个省级电网年增数据量可达数十 PB，甚至更多，涵盖用户用电行为、设备状态监测、能源交易记录、调度控制指令等多类数据，这些数据不仅体量大、类型多，还具有强时效性和高价值密度，一旦遭到泄露、篡改或破坏，将直接威胁电网稳定运行和用户隐私安全。因此，数据安全问题也随之增长，智能电网安全和隐私问题成为日益关注的焦点。近年来，针对电力系统的网络攻击事件频发，攻击手段日趋复杂和隐蔽。除传统的病毒、木马、DDoS 攻击外，还出现了针对工控系统的高级持续性威胁（APT）、虚假数据注入（FDI）攻击、勒索软件攻击等新型威胁。2023 年国家电网安全报告显示，电力行业已成为国家级网络攻击的重要目标，其中数据窃取与系统瘫痪是两大主要风险。传统安全防护技术（如静态加密、规则型防火墙、入侵检测系统）虽在一定程度上保障了基础安全，但其响应滞后、规则僵化，难以适应动态多变的攻击模式，尤其在面对零日漏洞和内部威胁时显得力不从心。人工智能（AI）技术凭借其在数据处理、模式识别、预测决策等方面的独特优势，为构建动态、自适应、智能化的电网数据安全防护体系提供了全新路径。AI 不仅能够实现对海量数据的实时分析与异常感知，还能通过机器学习、深度学习、强化学习等方法[3]，从历史攻击中学习并不断优化防护策略，形成“感知 - 研判 - 响应 - 进化”的闭环安全能力。本文结合 AI 算法与电网业务特性，设计了一种动态、自适应的安全防护框架，旨在提升电网数据的保密性、完整性和可用性。

1 人工智能技术

人工智能作为新一轮科技革命与产业变革的核心驱动力，正深刻改变着各行各业的发展模式与技术架构。其技术体系涵盖机器学习、深度学习、自然语言处理、计算机视觉、

1. 山东省工程咨询院（山东省政府投资项目评审中心）山东济南 250013
2. 国网山东省电力公司 山东济南 250013
3. 山东高速信息集团有限公司 山东济南 250013

强化学习等多个分支，共同构建起“感知 - 认知 - 决策 - 行动”的智能闭环。在数据安全领域，AI 的应用已从早期的辅助分析工具，逐步演进为支撑系统性安全运维的核心技术。近年来，国网、南网先后发布了电力专用大模型，为 AI 在电力行业各个领域的实用化、规模化应用奠定了坚实的技术基础。机器学习（Machine Learning）是 AI 在安全领域中应用最广泛的技术之一。监督学习方法（如支持向量机、随机森林、梯度提升树等）通过对历史攻击样本进行训练，建立起特征与标签之间的映射关系，可用于恶意流量分类、入侵行为识别等任务。例如，基于随机森林的 DDoS 检测模型，可通过分析网络流量的源 IP 分布、数据包大小、请求频率等特征，快速判别是否遭受分布式拒绝服务攻击。无监督学习（如 K-means 聚类、主成分分析、孤立森林等）则适用于未标注数据的异常发现，能够从海量日志中自动识别出偏离正常模式的行为，适用于内部威胁检测与未知攻击发现。

深度学习（Deep Learning）作为机器学习的重要分支，通过构建多层神经网络模型，实现对复杂数据的高层次特征提取与表达，可以进一步突破特征工程的限制，卷积神经网络（CNN）可有效解析电网设备监控图像中的异常发热点，而长短期记忆网络（LSTM）擅长捕捉电力负荷数据的时序关联特性，提前预警潜在过载风险。

值得关注的是，强化学习（Reinforcement Learning）技术在动态防御与主动响应方面的潜力日益凸显。其通过构建智能体（Agent）与环境之间的交互机制，使系统能够在不断试错中学习最优安全策略。例如，在面对持续变化的网络威胁时，基于强化学习的密钥管理系统可自主调整加密算法强度与密钥更换频率，以平衡安全性与性能开销；在多智能体对抗场景中，AI 还可以模拟攻击者与防御者之间的博弈过程，生成更具鲁棒性的防护方案。技术发展脉络显示，AI 正从单点工具向系统性解决方案进化，其与边缘计算、区块链的融合将重塑电网安全生态。

2 电网数据安全现状

从隐私数据类型看，智能电网中的数据呈现出多源性、高价值性和强关联性特征，其安全与隐私保护需求也因此存在显著差异。智能电表采集的用户实时数据包括有功 / 无功功率、电压电流值、功率因数等，能为控制中心制定发电计划和调度策略提供实时负荷状态数据；电动汽车则是电气化交通网的重要负荷，用户将行驶中的车辆位置和电池电量信息通过无线信号发射器发送给道路侧的数据采集单元进行数据交互，数据采集单元将收集的数据传输至控制中心，从而发挥电动汽车的储能特性达到削峰填谷、平衡负荷的效果，车辆轨迹与充电行为的结合，足以揭示用户的活动范围、通勤规律乃至社交关系，对数据脱敏和访问控制提出了极高要求；电力通信过程中的隐私数据主要指智能电网将部分数据信息存储于本地网关操作，网关收集地区用户用电量，传输给控制中心进行分析与处理，根据用户用电量信息，为用户制定个性化阶梯电价策略；能源交易的隐私数据主要关注交易过程中的交易信息安全，能源互联网支持多中心、分布式管理及能源的就地消纳，在能源互联网中，用户既是能源消费者，同时也可将自己能提供的可再生能源交易量发布，与有需求的用户建立协议，以去中心化方式完成电力交易。

根据智能电网的业务特点与数据类型，其数据安全防护必须系统性地兼顾保密性、完整性与可用性（CIA Triad）三重核心目标，三者缺一不可，共同构成电网安全稳定运行的“铁三角”。保密性是智能电网数据安全的首要屏障，意味着智能电网中的数据应只对拥有权限的用户开放，任何未授权用户不能获取数据，缺少保密性会导致用户及电网隐私数据泄露，侵害利益；完整性关乎电网决策与控制的正确性，意味着应保证电网信息真实性，避免在数据存储和传输中被非法篡改和损坏或及时发现篡改数据，此类数据包括调配信息、传感器数据、控制指令、智能电表数据等，缺少完整性会造成控制中心误判进而发出错误指令；可用性则强调了安全服务与业务连续性的平衡，意味着控制系统、安全系统、执行系统等各类通信元件应正常访问，不被攻击者阻断，缺少可用性会导致失去对相关设备的及时监控和有效控制，破坏电网安全运行。

从智能电网的网络架构来看，智能电网的设备多部署于复杂威胁环境中，易受到攻击者攻击。在用户侧，用户用电行为数据，如高频采样的负荷曲线能精确反映家庭的生活作息，可能暴露家庭生活习惯，构成隐私泄露风险；设备状态数据（如变压器油温、线路电流相位）的篡改将直接引发继电保护误动作；而调度控制指令若遭劫持，可能导致区域性停电事故。当前，威胁来源呈现出多元化与专业化的特征。外部攻击者常利用工控系统固有的协议漏洞（如 Modbus TCP/IP 缺乏加密与强认证机制）实施中间人攻击，窃听或注入控制指令；内部人员则可能因权限管理疏漏或恶意行为，越权访问敏感数据库，造成“堡垒从内部攻破”的局面。更严峻的挑战源于电网边缘的扩张。数以亿计的智能电表、智能家居控制器、分布式能源控制器等终端设备的大规模部署，使得攻击面急剧扩大，这些设备通常资源受限，难以运行复杂的安全软件，且物理防护相对薄弱，极易成为攻击者渗透内网的跳板。传统基于边界防护的“城堡式”安全模型已无法适应去中心化架构。这迫切要求安全体系向数据本体防护转型，将安全能力嵌入到数据生成、传输、处理的每一个环节。而 AI 技术，凭借其强大的实时分析、异常感知和自适应决

策能力，正成为实现这一深刻转型的核心驱动力。

3 人工智能技术在电网数据安全防护中的应用

3.1 数据分级与安全防护

电网数据的价值密度差异决定了分级防护的必要性。这一问题的核心在于，智能电网中产生的数据具有显著的非均质性：从用户身份信息、实时控制指令到设备运行日志，其敏感性、实时性和价值密度差异巨大[4]。传统的数据分级方法依赖人工标注和静态规则，难以应对动态变化的业务场景。

基于 AI 的数据分类体系首先通过自然语言处理（NLP）解析数据元语义，例如识别“用户身份证号”“断路器分合状态”等高敏感字段；继而采用改进的 DBSCAN 聚类算法，结合数据访问频率、存储位置等上下文信息动态划分安全等级，其核心公式为：

$$\text{Cluster}(x_i)=\begin{cases}\text{Core} & |N_\epsilon(x_i)|\geq\text{MinPts}\\ \text{Border} & 0<|N_\epsilon(x_i)|<\text{MinPts}\\ \text{Noise} & |N_\epsilon(x_i)|=0\end{cases} \quad (1)$$

式中：x_i 表示待分类数据点；ϵ 表示邻域半径；MinPts 表示最小样本数；$N_\epsilon(x_i)$表示以 x_i 为中心、ϵ 为半径的邻域内数据集合。

通过动态调整 ϵ 和 MinPts，结合数据访问频率（f）和存储位置权重（ω），实现安全等级划分。高敏感数据采用轻量级同态加密算法（如 CKKS 方案），允许在密文状态下进行负荷预测计算；中敏感数据引入差分隐私机制，在馈线拓扑数据发布时添加拉普拉斯噪声以模糊关键节点信息；低敏感数据则通过生成对抗网络（GAN）合成脱敏样本供科研分析。

3.2 异常检测与实时监控

电网数据流的异常检测是构建其主动安全防御体系的关键环节，其目标是从海量、高速、多源的运行数据中，精准、实时地识别出潜在的攻击行为和系统故障。然而，这一任务面临着三大核心难题：

（1）高维度与强噪声：是数据复杂性带来的检测困境。电网传感器每秒生成数万维数据这些数据不仅维度极高，涵盖了电压、电流、相位、频率、功率因数、谐波分量等多种物理量，而且其数据质量深受现场复杂电磁环境的影响，不可避免地掺杂着大量的随机噪声与瞬时干扰，传统阈值检测方法误报率高达 30% 以上。

（2）概念漂移：是动态环境中正常行为的时变性，电网的运行状态并非一成不变，而是随着季节更替、天气变化、节假日安排以及社会活动呈现出强烈的周期性和趋势性波动，导致正常行为基线不断偏移，静态模型易将合法波动误判为异常。

（3）隐蔽性攻击：是智能化威胁对检测算法的挑战。随着攻击技术的演进，针对电网的高级持续性威胁（APT）和虚假数据注入（FDI）攻击变得愈发隐蔽和狡猾。不同于传统 DDoS 攻击的“狂轰滥炸”，这些高级攻击为了规避检测，会有意模仿合法操作的通信模式、数据包格式和交互逻辑，传统基于流量统计的方法难以识别。

基于混合式 AI 的检测框架，首先利用孤立森林算法过滤设备传感器噪声数据，降低误报率[5]。随后时序卷积网络（TCN）TCN 通过因果卷积和扩张卷积提取电压、电流波形的时空特征：

$$\boldsymbol{y}_t=\varnothing\sum_{k=0}^{K-1}\boldsymbol{W}_k\cdot\boldsymbol{x}_{t-d\cdot k} \quad (2)$$

式中：y_t 为时间步 t 的输出；$\boldsymbol{W}_k$ 为卷积核权重；$\boldsymbol{x}_{t-d\cdot k}$为输入序列；$d$ 为扩张系数；K 为卷积核大小；$\varnothing$为激活函数。通过堆叠多层 TCN 构建正常行为基线。

最后采用基于注意力机制的 LSTM 模型动态评估数据流偏离度，用公式表示为：

$$\boldsymbol{a}_t=\text{softmax}(\boldsymbol{W}_a\cdot\tanh(\boldsymbol{W}_h h_t+\boldsymbol{W}_c c_t)) \quad (3)$$

式中：h_t 为 LSTM 隐藏状态；c_t 为单元状态；$\boldsymbol{W}_a$、$\boldsymbol{W}_h$、$\boldsymbol{W}_c$ 为权重矩阵；$\boldsymbol{a}_t$ 为注意力权重。

最终异常分数为：

$$S=\sum_{t=1}^{T}\boldsymbol{a}_t\cdot\left\|h_t-\hat{h}_t\right\|_2 \quad (4)$$

式中：$\hat{h}_t$为基线预测值；T 为时间窗口长度。

当 S 超过阈值时触发告警，系统自动触发溯源机制——图神经网络（GNN）分析网络流量日志，重构攻击路径并定位 compromised 节点。

3.3 人工智能算法加密

传统加密算法（如 RSA、AES）依赖固定密钥策略与确定性数学难题（如大数分解），面临两大威胁：

（1）量子计算威胁：Shor 算法可在多项式时间内破解 RSA 加密，Grover 算法使 AES-256 的有效密钥空间减半。

（2）模式固定性：密钥生成与更换周期由人工预设，无法动态适应攻击强度变化。

基于 AI 的动态加密技术通过两方面实现突破：一是利用强化学习框架训练密钥生成器，根据当前网络威胁指数（如攻击频率、类型）自适应调整 AES 密钥长度与更换周期，使暴力破解成本提升 3 个数量级；二是设计神经网络驱动的加密算法，将明文数据映射到高维混沌空间，受限玻兹曼机（RBM）生成非线性变换矩阵：

$$P(\boldsymbol{v},h)=\frac{1}{Z}\exp(-E(\boldsymbol{v},h)) \quad (5)$$

$$E(\boldsymbol{v},h)=-\boldsymbol{v}^T\boldsymbol{W}h-b^T\boldsymbol{v}-c^Th \quad (6)$$

式中：$\boldsymbol{v}$ 为明文向量；h 为隐藏层状态；$\boldsymbol{W}$ 为权重矩阵；b、c 为偏置项；Z 为配分函数；$E(\boldsymbol{v},h)$ 为能量函数。

加密过程通过吉布斯采样生成混沌密文，即使密文被截获也无法逆向推导密钥。

4 实验结果分析

4.1 实验数据

实验数据由三部分构成，涵盖仿真运行数据、模拟攻击流量与对抗样本。模拟真实的运行场景生成 2 000 条仿真数据，包括 SCADA 数据、PMU 数据和设备日志数据。CyberRange 平台构建电网攻击沙盒，模拟 DDoS 攻击、中间人攻击和隐蔽数据注入攻击场景。基于（projected gradient descent, PGD）算法，对正常 PMU 数据添加微小扰动（ε=0.03），构造对抗性样本以测试模型鲁棒性。实验中数据集将被划分为三部分：训练集为 80% 正常数据 +20% 已知攻击样本；测试集为 15% 正常数据 +5% 新型攻击样本；验证集为 5% 数据用于超参数调优。

4.2 性能指标对比分析

为全面评估 AI 模型的综合性能，设计以下五类指标：

（1）准确率：正确识别正常与攻击样本的比例，反映整体分类能力。

（2）精确率：针对攻击样本的判定可靠性，避免过度防御干扰正常业务。

（3）召回率（Recall）：攻击样本的检出完整性，衡量漏报风险。

（4）F_1 分数：精确率与召回率的调和均值，平衡误报与漏报的权重。

（5）安全增益系数（SGC）：

量化 AI 模型对抗未知威胁的能力，定义为：

$$\mathrm{SGC}=\frac{\text{新型攻击检出率}}{\text{传统方法检出率}}\times\log(\theta) \tag{7}$$

式中：θ 表示模型的参数；SGC ＞ 1 表示 AI 方案具备正向安全增益。

测试数据如表 1 所示。

表 1 测试数据表

方案	准确率 /%	精确率 /%	召回率 /%	F_1 分数 /%	SGC
传统方案	86.5	78.2	72.4	75.1	0.8
AI 方案	97.8	95.6	93.7	94.6	4.7

实验结果表明，AI 模型在攻击检测准确率、精确率、召回率、F_1 分数及 SGC 上全面超越传统方案。AI 方案的准确率高达 97.8%，较传统方案的 86.5% 提升了超过 11 个百分点，证明 AI 模型在区分正常与恶意行为方面优秀的综合能力。AI 模型的精确率达到 95.6%，远高于传统方案的 78.2%，意味着系统在发出告警时具有极高的可信度。同时，其召回率也达到了 93.7%，相较于传统方案的 72.4% 有巨大提升，这表明模型对真实攻击的检出更为全面，极大降低了因漏报而带来的潜在安全风险。作为精确率与召回率的调和均值，F_1 分数最为直观地体现了 AI 模型在误报与漏报之间取得的优异平衡。这表明其 TCN 模块有效捕捉了负荷曲线的长期依赖关系，而 LSTM 的注意力机制强化了对短期突变的敏感性。人工智能技术通过多层级检测框架与动态防护策略，显著提升了电网数据安全防护的精准性与实时性。

5 结语

人工智能与电网数据安全的深度融合，标志着电力系统防护从“被动响应”向“主动免疫”的范式转变。本研究验证了 AI 在数据分类、异常检测、动态加密等环节的技术可行性，但同时也揭示出待解决的深层矛盾：一方面，深度学习模型的黑箱特性可能掩盖决策漏洞，需发展可解释 AI（XAI）增强安全审计透明度；另一方面，联邦学习等隐私计算技术虽然能保护本地数据，但其通信开销可能影响实时控制指令下发。未来研究应聚焦于构建“AI+ 区块链”的混合架构，通过智能合约实现防护策略的分布式验证，同时探索神经形态计算芯片在边缘侧的应用，最终形成覆盖“云 - 边 - 端”的全场景安全防护体系。

参考文献：

[1] 杨泽渠 . 智能电网中用户数据安全及隐私保护研究 [D]. 杭州 : 浙江大学 , 2017.

[2] 吴盼 , 杨洋 . 智能电网数据安全防护策略研究 [J]. 信息与电脑 , 2024, 36(24): 63-65.

[3] 张翔 . 人工智能技术在医院数据安全防护中的应用 [J]. 智能物联技术 , 2024, 56(6): 101-104.

[4] 李杨 , 刘晓冬 . 人工智能的计算机数据网络安全防护系统设计 [J]. 信息与电脑 (理论版), 2024, 36(10): 115-117.

[5]LECUN Y, BENGIO Y, HINTON G. Deep learning-nature[J]. Nature, 2015, 521: 436-444.

【作者简介】

常雅楠（1992—），女，山东济南人，硕士研究生，中级职称，研究方向：人工智能、建筑工程、电子信息，email: 547111579@qq.com。

张云鹏（1993—），男，山东济南人，硕士研究生，中级职称，研究方向：人工智能、电力数字化，email: zhyp_1993 @hotmail.com

房宏基（1987—），男，山东济南人，本科，中级职称，研究方向：智慧交通、交安物联、北斗产业化应用，email: fuki61@163.com。

（收稿日期：2025-09-11 修回日期：2025-12-08）