基于深度学习算法的电力运行数据隐私保护方法

李 兴[1] 吴天宇[1] 马光明[1]
LI Xing WU Tianyu MA Guangming

摘 要 目前电力系统的数据往往包含大量的隐私信息，一旦泄露或被滥用，将会对个人隐私和企业利益造成严重威胁，因此提出基于深度学习算法的电力运行数据隐私保护方法。首先，基于深度学习算法提取电力运行数据特征，有效地保护电力运行数据的隐私；其次，构建电力运行数据加密模型，在处理电力运行数据时，防止未经授权的用户获取敏感数据；最后，保护电力运行的隐私数据，实现平衡数据利用和隐私保护的需求。实验结果表明，传统加密方法在 40 s 内，对电力数据进行加密后的成功数量达到了 4700 条，而基于深度学习算法的电力运行数据隐私保护方法达到了 5800 条，可见其数据隐私保护的效率相较于传统加密方法的效率更佳。

关键词 深度学习算法；电力运行；数据保护；隐私管理

doi：10.3969/j.issn.1672-9528.2024.03.046

0 引言

随着智能电网的发展和电力信息化的推进，电力运行数据已经成为电力系统运行管理、调度控制、市场交易和客户服务等领域的重要信息来源。然而，这些数据往往包含大量的隐私信息，如用户的个人信息、用电信息等，一旦泄露或被滥用，将会对个人隐私和企业利益造成严重威胁。因此，如何保护电力运行数据的隐私信息，成为一个亟待解决的问题。同时，深度学习算法作为一种强大的机器学习技术，已经在图像识别、语音识别、自然语言处理等领域取得了巨大的成功。近年来，越来越多的学者开始尝试将深度学习算法应用于电力运行数据的分析和处理中，以实现更加准确和高效的电力系统运行和调度控制。然而，深度学习算法的应用也带来了一些隐私保护方面的问题，如何确保深度学习算法在电力运行数据中的应用不会泄露用户的隐私信息，也是一个需要解决的问题 [1]。

1 电力运行数据隐私保护方法的目的意义

电力运行数据隐私保护方法的目的意义在于确保电力运行数据的安全性和可靠性，防止数据泄露和非法获取。电力运行数据是电力系统中非常重要的信息，涉及到电力系统的运行状态、设备状况、能源供需等方面的内容，一旦泄露可能会对电力系统造成严重的安全威胁和稳定风险。

同时，电力运行数据隐私保护还有助于保护个人隐私和商业机密。在电力系统中，个人的用电数据和企业的用电数据都是非常重要的隐私信息，泄露出去可能会对个人和企业造成不良影响。因此，电力运行数据隐私保护方法的实施是非常必要的。

电力运行数据隐私保护方法的必要性在于其对于整个电力系统的安全稳定运行具有至关重要的作用。这些数据不仅反映了电力系统的实时状态，还涉及到众多用户的个人信息和商业机密。因此，有效地保护电力运行数据的隐私，不仅能够维护电力系统的安全，还能够保护用户的合法权益。

首先，电力运行数据隐私保护是保障国家安全的重要组成部分。电力系统是国家基础设施的核心，其安全稳定运行直接关系到国家的经济、社会发展和人民生活的各个方面。一旦电力运行数据泄露或被非法获取，可能会对国家安全造成严重影响。因此，通过实施电力运行数据隐私保护方法，可以有效防止这种风险，保障国家安全。

其次，电力运行数据隐私保护对于维护用户权益具有重要意义。在电力系统中，用户的用电数据等个人信息是极为敏感的，一旦泄露可能会对用户造成财产损失或隐私侵犯。因此，电力企业有责任采取措施保护用户的隐私信息，维护用户的合法权益。

此外，电力运行数据隐私保护还有助于提升电力企业的形象和信誉。通过采取严格的隐私保护措施，电力企业可以向用户展示其对用户隐私的重视和保护，从而赢得用户的信任和好感。这不仅可以提高电力企业的市场竞争力，还有助于提升企业的社会形象和信誉度。

综上所述，电力运行数据隐私保护方法的实施是保障电力系统安全稳定运行、维护用户权益以及提升电力企业形象和信誉的重要举措。因此，我们应该高度重视电力运行数据隐私保护工作，采取切实有效的措施确保数据的安全性和可

1. 甘肃同兴智能科技发展有限责任公司 甘肃兰州 730050

靠性。

2 基于深度学习算法的电力运行数据隐私保护方法的设计

2.1 基于深度学习算法提取电力运行数据特征

深度学习算法通常采用多层神经网络结构，这种结构使其能够从数据中提取复杂的特征，同时将高层的抽象特征表示为底层特征的非线性函数。一种典型的神经网络结构包括输入层、隐藏层和输出层。网络中的每个节点都是一个神经元，而每一个神经元都会从上面的一个神经元中获得一个权值。将一个非线性活化函数加到一个权重相加的输入值上，得到神经元的输出。图 1 为深度学习算法中的神经网络结构示意图[2]。

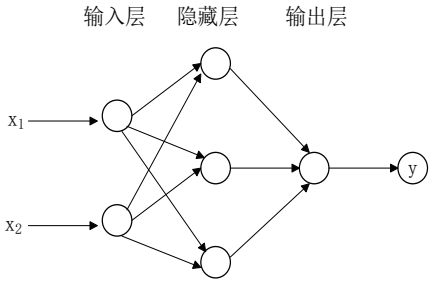

图 1 神经网络结构

在电力系统运行的过程中，隐私数据保护是一个非常重要的环节，定期对电力系统中的重要数据进行备份，以防止数据丢失或损坏。同时，应确保备份数据的安全性，例如使用加密备份或远程备份数据[3]。

对电力数据进行特征抽取，可以从多个角度深入挖掘时序电力运行数据中隐含的信息。深度学习分解适合处理非平稳、非线性的电力数据，能够将这部分数据转为平稳线性数据。其主要思想是将数据序列分解为各个尺度的局部特征信息和残差两个部分，其计算公式为：

$$x=\sum a+R \tag{1}$$

式中：x 表示电力运行序列数据，a 表示原始电力数据中各个尺度的局部特征信息，R 表示残差分量，该数值表示原始电力数据序列的趋势。

当原始电力数据序列 x 满足以下条件时，x 的极值点与零点个数相同或不大于 1，在此基础上，利用最大值和极小值所构成的上下包络线的平均值为 0，然后对原始电力数据进行过滤处理，即可对 x 进行分解，其表达公式为：

$$m=\frac{h+l}{2} \tag{2}$$

式中：m 表示均值包络线，h 代表 x 的局部极大值对应的上包络线，l 表示 x 的局部极小值对应的下包络线。

设定在电力运行数据中设置 n 个训练样本，各个样本中含有 z 个属性特征，电力数据集合为 $T_{n\times z}$，电力序列标签为 $y\in\{y_1,y_2,\cdots,y_n\}$，子序列的长度为 p，在 $T_{n\times z}$ 中的滑动长度 p 的窗口中，得出子序列 s_i^p，i 代表了电力数据序列中滑动窗口的原始位置，在 T 中提取长度 p 的子序列集合表示为 $S=\{S_1^p,S_2^p,\cdots,S_{n-p+1}^p\}$。

基于以上分析，长度为 ∂ 的两个电力数据序列 T_1、T_2 之间距离的计算公式为：

$$Dist(T_1,T_2)=\sqrt{\frac{1}{\partial}\sum_{i=1}^{\partial}(T_{1i}-T_{2i})^2} \tag{3}$$

设定一个电力序列数据集 D 由 α、β 两个类型组成，则该数据集的信息熵的表达公式为：

$$I(D)=q(\alpha)\log(q(\alpha))-q(\beta)\log(q(\beta)) \tag{4}$$

式中：$q(\alpha)$ 表示 α 类的比例，$q(\beta)$ 代表 β 类的比例。

电力运行数据的分析主要通过关注最值、峰极值、均值等统计特征进行，电力运行数据的复杂性和高维特征增加了数据挖掘的难度[4]。

2.2 构建电力运行数据加密模型

在电力系统中，使用对称加密或非对称加密算法对电力运行数据进行加密，以防止敏感数据被未经授权的用户访问和窃听。电力数据的隐私推断是一种利用历史数据对电力用户的敏感信息进行预测与推断的过程，该问题可被形式化地描述为：给定电力时序数据 $\boldsymbol{x}$，攻击者试图从 x 中挖掘出其中所包含的用户行为模式，而这些电力数据序列往往含有一定的规则性，攻击者能够推断出在将来某个时间段内的电力用户敏感数据，不利于电力运行数据的隐私保护[5]。然而，最新研究指出，这些基于深度学习网络的模型容易受到对抗攻击。这些攻击样本难以察觉，但会给模型带来严重的安全隐患。因此，为保护电力运行过程中的隐私数据，需要对攻击进行对抗预测。对抗的样本 $\hat{x}$ 由电力时序数据 x 与扰动 δ 生成，其表达公式为：

$$\hat{x}=x+\delta \tag{5}$$

式中：δ 表示扰动过程，其表达公式为：

$$\delta=\varepsilon\cdot sign(J(x,y)) \tag{6}$$

式中：ε 代表扰动程度，J 表示损失函数，x 代表原始电力数据，$sign$ 代表获取梯度的方向，y 表示原始电力数据对应的标签[6]。

x 与 $\hat{x}$ 之间的差距应该尽量缩小，使电力数据隐私推理模型基于 $\hat{x}$ 进行推理时，模型的性能明显下降。经过训练后的深度学习模型，原始电力时序数据 x 以及生成的对抗样本 $\hat{x}$，若满足下列公式，则表示该模型对攻击实行有效对抗：

$$f(\hat{x})=\hat{Y},f(x)=Y',Y\neq Y' \tag{7}$$

式中：$\hat{Y}$ 代表深度学习模型在对抗样本 $\hat{x}$ 中输出的错误结果，Y' 表示深度学习模型在原始电力时序数据 x 中输出的

结果[7]。

基于以上分析，深度学习算法在解决隐私泄露问题中表现出良好的性能，同时很好地适用于电力数据预测问题，其中包括卷积层、池化层、全连接层。卷积层通过卷积核提取电力数据特征后，对其特征进行浓缩采样，并保留特征的关键信息，能够准确保护隐私特征，全连接层进行卷积层、池化层处理后的数据加密，最后进行输出。其电力数据加密模型的输出结果的计算公式为：

$$s(x)=R\left(wx-\delta+\frac{e}{\hat{x}}\right) \tag{8}$$

式中：R代表激活函数，x表示输入电力数据，w代表权重向量，e表示偏置单元。

基于此，同时进行定期备份敏感数据，确保数据在遭受攻击或意外丢失后能够及时恢复。备份数据也需要进行适当的加密和存储，确保备份数据的安全性[8]。

2.3 保护电力运行的隐私数据

隐私保护的实施方式是向查询函数中加入适量的噪声，以实现保护数据的目的。而噪声的大小是由敏感度决定的，它是指数据集中任意一条记录的删除都会对查询结果产生最大的改变。当电力系统运行过程中抽取两个数据集 U、U'，两个数据集拥有相同的特征结构，U、U'最多相差一条数据记录时，则 U、U'成为相邻数据集[9]。将上述章节中的电力数据加密模型代入相邻数据集中，则此时电力数据加密模型需要满足以下公式：

$$P\left[s(U_x)\in\mu\right]\le\exp(\sigma+g)\left[s(U'_x)\in\mu\right] \tag{9}$$

式中：P代表电力数据加密模型中所有可能输出构成的集合，μ表示P中的子集，g代表噪声，σ参数代表隐私保护的预算，该参数可用于表现出隐私保护的程度。

此时 $s(U_x)$加入噪声后，g提供电力系统数据的隐私保护，隐私度随着预算的减小而增加，可用性随之减小[10]。加入噪声是完成数据隐私保护的基础步骤，则此时隐私度的密度函数的表达公式为：

$$p(x)=\frac{b\cdot g}{2\exp\left(-\frac{|x|}{b}\right)} \tag{10}$$

式中：b代表拉普拉斯噪声机制的尺度。

全局隐私度取决于函数本身，不同函数对应不同的全局隐私度。隐私度直接决定实现差分隐私保护所需的噪声，隐私度越小，所需噪声越少[11]。经过隐私度计算后的 $s(U'_x)$数据集结果，即实现整个数据集的隐私保护结果。

最后，还可以建立完善的数据安全管理机制，包括制定合理的安全策略、培训员工提高安全意识、定期进行安全审计等，有助于提高整个电力系统的数据隐私保护意识和能力[12]。总之，保护电力运行隐私数据需要从制度、技术、人员等多个方面入手，建立全方位的数据保护体系，确保电力运行数据的安全性和可靠性。

3 实验测试与分析

在实验测试之前，需要一些准备工作，保证本次实验的准确性。

3.1 实验准备

为验证本文提出的基于深度学习算法的电力运行数据隐私保护方法的有效性，使用 Matlab 软件构建实验环境，本次测试环境、参数的设置如表 1 所示。

表 1 实验环境配置

项目	参数
CPU	i5 7200U
编程环境	PyCharm2020.3.3
硬件配置	3.40 GHz，8 GB
操作系统	Windows10
内存	8 G
开发语言	Python

基于此次测试需要，将从 A 市的 500 家企业的用电量中抽取样本数据。本次实验所用用电量数据集如表 2 所示。

表 2 测试数据集

数据集	企业用电数据集 1	企业用电数据集 2	企业用电数据集 3
属性数量	12	6	10
类型	Real	Real	Real
记录数量	600	800	1200
别名	D1	D2	D3

少量数据集可能造成实验精准度较低，因此，本次测试选择数据集中数据数量较多的企业用电数据集 3。首先，收集该企业用电数据集 3 的传统数据加密方法的效果情况，其变化情况如图 2 所示。

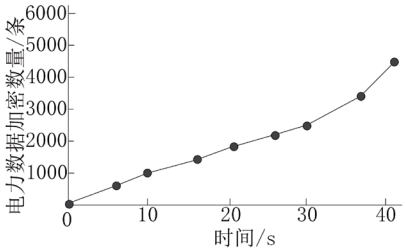

图 2 传统数据加密方法效果

3.2 实验结果与分析

将本文提出的基于深度学习算法的电力运行数据隐私保护方法代入实验测试中，使用本文方法进行隐私保护，其加密效果如图 3 所示。

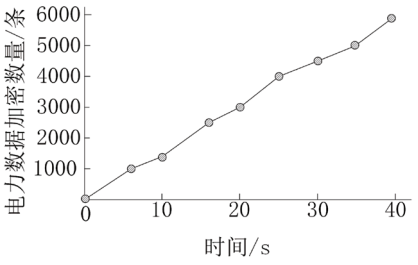

图 3 实验结果

从上述实验结果中可以看出，传统加密方法在 40 s 内，对电力数据进行加密后的成功数量达到了 4700 条，而本文方法达到了 5800 条，因此，本文方法的数据隐私保护的效率相较于传统加密方法的效率更佳。

综上所述，通过深度学习算法，可以在不泄露敏感数据的情况下，更高效地对电力运行数据进行处理和分析。这证明该方法可以有效地保护电力运行数据的隐私，同时保证数据的安全性和可用性。在数据传输、保存或使用的过程中，该方法可以检测并抵制不合法的篡改行为，保障数据的完整性、可靠性和安全性。

4 结语

在数字化浪潮之下，电力运行数据的隐私保护问题愈发受到关注。近年来，深度学习算法在电力运行数据隐私保护领域取得了显著成果，但这仅仅是探索的起点。实际应用中，深度学习算法在电力运行数据隐私保护方面仍面临多重挑战，亟待持续的科研攻关和技术革新。

深度学习算法凭借其强大的特征提取与分类能力，为电力运行数据隐私保护提供了新的路径。通过构建精密的神经网络模型，能够实现对电力运行数据的自动化分析与处理，有效识别和过滤敏感信息。然而，现有的隐私保护方法仍存不足，如数据泄露风险、计算复杂度高、实时性能差等问题。因此，我们需进一步优化和改进深度学习算法，提升其在电力运行数据隐私保护领域的性能与效率。

为实现上述目标，我们需要制定更为具体可行的实施细则与评估标准。首要之务是明确电力运行数据隐私保护的具体需求与目标，从而有针对性地设计与优化深度学习算法。其次，建立完善的评估体系，对隐私保护方法的性能进行全面、客观的评价。这不仅能够确保数据隐私保护措施的有效实施，而且为未来的研究提供宝贵的参考与借鉴。

值得强调的是，电力运行数据隐私保护不仅涉及技术层面，更是一个涉及法律法规、伦理道德、社会安全等多方面的社会问题。在实际应用中，我们应充分考虑这些因素。因此，加强跨学科合作与研究，吸引计算机科学、法学、社会学等领域的专家与学者共同参与，显得尤为重要。通过集思广益、汇聚智慧，我们有望实现更为全面和有效的数据隐私保护。

综上所述，基于深度学习算法的电力运行数据隐私保护方法的研究，是一个充满挑战与机遇的领域。随着技术的不断进步和社会认识的日益提高，我们有信心在未来实现更加高效和可靠的电力运行数据隐私保护方法。这不仅有助于保障电力系统的安全稳定运行，还可为其他领域的数据隐私保护提供有益的参考与启示。让我们携手共进，共同推动电力运行数据隐私保护技术的创新与发展。

参考文献：

[1] 贾峰 . 基于深度学习算法的电力调度数据网络异常检测方法 [J]. 信息与电脑 (理论版), 2023,35(12):79-81.

[2] 罗涛 , 孙阔 , 张章 , 等 . 多能源数据驱动的电力信息物理系统综合态势感知模型 [J]. 可再生能源 ,2021,39(3):395-400.

[3] 吴桂龙 , 杨志敏 , 黄昱 . 电力通信运行管理中典型业务数据的智能关联方法 [J]. 电信科学 ,2021,37(2):164-172.

[4] 费雯丽 , 邓显波 , 王格 , 等 . 基于运行数据分析的高压电缆绝缘老化状态评估平台 [J]. 湖北电力 ,2020,44(5):59-64.

[5] 穆钢 , 陈奇 , 刘洪波 , 等 . 揭示电力系统运行数据中因果关系的逆信息熵因果推理方法 [J]. 中国电机工程学报 , 2022, 42(15): 5406-5417.

[6] 陈凡 , 张继聪 , 赵美莲 , 等 . 考虑负荷重分配攻击的电力系统运行可靠性评估 [J]. 电力系统保护与控制 , 2023, 51(8): 160-168.

[7] 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.

[8] 张旭东 , 谢民 , 黄建平 , 等 . 基于数据挖掘的电力自动化系统运行数据中台资源检索技术研究 [J]. 安徽师范大学学报 (自然科学版), 2023,46(2):119-125.

[9] 贺娇 , 夏帅 , 刘辉 , 等 . 加强电力通信电源系统安全稳定运行的方法探究 [J]. 科技与创新 ,2023(4):126-128.

[10] 姜杰 , 付申杰 , 杨君艺 , 等 . 负荷预测精准度对电力系统运行的影响 [J]. 集成电路应用 ,2023,40(2):188-189.

[11] 王文歆 , 刘璐 . 基于粒子群算法的电力电缆运行状态智能监测技术 [J]. 自动化应用 ,2022(9):103-105+108.

[12] 卢冠华 , 陈俊斌 , 丁茂生 , 等 . 知识图谱在电力系统调度运行中的应用与展望 [J]. 电力信息与通信技术 , 2023, 21(7): 27-35.

【作者简介】

李兴（1984—），男，甘肃甘谷人，硕士，副高级工程师，研究方向：信息化项目建设。

吴天宇（1981—），男，甘肃永登人，本科，中级工程师，研究方向：信息化项目运维。

马光明（1978—），男，回族，甘肃兰州人，本科，中级工程师，研究方向：信息安全运维。

（收稿日期：2024-01-05）