ICS 国际标准分类号
CCS 中国标准文献分类号

团　　体　　标　　准

T/CES XXX-XXXX

虚拟电厂云边协同数据安全与隐私保护技术规范

Technical Specification for Data Security and Privacy Protection in Cloud–Edge Collaboration for Virtual Power Plants

XXXX–XX–XX 发布　　XXXX–XX–XX 实

中国电工技术学会　发布

目　　次

前　言

本文件按照 GB/T 1.1—2020《标准化工作导则　第 1 部分：标准化文件的结构和起草规则》给出的规则起草。

请注意本文件的某些内容可能涉及专利，本文件的发布机构不承担识别专利的责任。

本文件由中国电工技术学会提出。

本文件由中国电工技术学会标准工作委员会工作组归口。

本文件起草单位：国网江苏省电力有限公司宿迁供电分公司、 东南大学、 双创中心、 国网信息通信产业集团有限公司、 南京邮电大学、国网江苏省电力有限公司

本文件主要起草人：王虎、薛风华、王秀茹、庞吉年

本文件为首次发布。

引　　言

随着虚拟电厂的快速发展，电力数据的处理和分析变得日益关键。在虚拟电厂环境中，大量多源异构电力数据不断产生，这些数据对于提升电力系统的效率和稳定性具有重要价值。然而，目前在电力数据处理中，隐私保护不足和通信开销大等问题日益突出，亟待有效解决方案。

传统预测方法存在多源异构数据整合难、隐私保护不足等问题，难以适应新能源接入的复杂场景。集中式数据处理模式在新能源场景下也暴露出诸多不足，如通信拥堵、传输延迟等，无法满足实时性要求。尽管基于云边端协同的联邦学习方法在一定程度上缓解了隐私和通信问题，但在新能源电力数据预测中仍存在缺陷，如数据分布差异大、模型偏差、隐私预算分配不合理、噪声添加过度等问题，难以兼顾隐私保护与模型性能。

为解决上述问题，迫切需要一种能够平衡用户隐私保护与电力数据预测性能，并适应智能配电网与新能源融合复杂环境的创新方法。本标准正是基于这样的背景和需求而制定，旨在提出一种虚拟电厂云边协同数据安全与隐私保护技术规范，为虚拟电厂中的电力数据处理提供创新的技术支持。

虚拟电厂云边协同数据安全与隐私保护技术规范

1 范围

本文件规定了虚拟电厂云边协同场景下多源异构电力数据在采集、传输、存储、处理及共享过程中的数据安全与隐私保护技术要求。

本文件适用于虚拟电厂中云边协同架构的构建、运维和评估，也可为分布式能源、储能及可控负荷接入系统提供参考。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中,注日期的引用文件,仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本(包括所有的修改单)适用于本文件

GB/T 20986-2023	信息安全技术网络安全事件分类分级指南
GB/T 22239-2019	信息安全技术网络安全等级保护基本要求
GB/T 33603-2017	电力系统模型数据动态消息编码规范
GB/T 33607-2017	虚拟电厂调度控制系统总体框架
GB/T 35273-2020	信息安全技术个人信息安全规范
GB/T 36572-2018	电力监控系统网络安全防护导则
GB/T 41373-2022	信息安全技术数据分类分级指南
DL/T 2473.2-2022	可调节负荷并网运行与控制技术规范
IEC 62351-2020	电力系统管理及其信息交换数据和通信安全（Power systems management and associated information exchange - Data and communications security）
ISO/IEC 27001-2022	信息安全、网络安全和隐私保护信息安全管理系统要求（Information security, cybersecurity and privacy protection — Information security management systems —Requirements）
ISO/IEC 27002-2022	信息安全管理体系实践规范（Code of Practice for Information Security Controls）
ISO/IEC 27018-2019	公有云中个人信息保护（Protection of Personally Identifiable Information in Public Clouds）

3 术语和定义

下列术语和定义适用于本文件。

3.1 多源异构电力数据 multi-source heterogeneous power data

来源于多种不同渠道、具有不同结构和特性的电力相关数据，包括但不限于家庭用电数据、工业用电数据、分布式能源发电数据等。

3.2 个性化隐私保护联邦学习 personalized privacy-preserving federated learning

一种结合联邦学习与个性化隐私保护技术的机器学习方法，旨在保护数据隐私的前提下，实现多个参与方的协同训练，提升模型性能。

3.3 联邦学习 federated learning

一种分布式机器学习方法，允许多个客户端在不共享原始数据的情况下，协作训练一个全局模型。

3.4 差分隐私 differential privacy

一种数学隐私保护框架，通过在数据发布或分析过程中添加噪声，确保单个数据记录的隐私不会被泄露。

3.5 隐私预算 **privacy budget**

在差分隐私保护中，用于衡量和控制隐私泄露程度的参数，通常用ε表示隐私预算。

3.6 自适应梯度裁剪 **adaptive gradient clipping**

一种优化技术，通过动态调整梯度的裁剪阈值，防止梯度爆炸和过度噪声添加，提高模型训练的效率和性能。

3.7 云边协同架构 **cloud-edge collaborative architecture**

一种结合云计算和边缘计算的架构，旨在实现数据的分布式处理与模型训练，降低通信开销，提高系统的实时性和效率。

3.8 数据敏感度评估 **data sensitivity assessment**

对数据的敏感程度进行量化评估的过程，以确定数据在隐私保护中的优先级和保护强度。

3.9 虚拟电厂 **virtual power plant**

一种将分布式能源、储能设备和可控负荷集成在一起，通过智能控制和优化调度，实现能源高效利用和灵活管理的虚拟实体。

4 缩略语

下列符号、代号和缩略语适用于本文件。

FL：联邦学习（Federated learning）

DP：差分隐私（Differential privacy）

VPP-CE-DSPP：虚拟电厂云边协同数据安全与隐私保护系统（Virtual Power Plant Cloud-Edge Collaborative Data Security and Privacy Protection System）

OPSD：开放电力系统数据（Open power system data）

TCN：时间卷积网络（Temporal convolutional network）

Non-IID：非独立同分布（Non-Independent and identically distributed）

MEC：多接入边缘计算（Multi-Access edge computing）

API：应用程序接口（Application programming interface）

AI：人工智能（Artificial intelligence）

SCADA：监控与数据采集系统（Supervisory control and data acquisition）

SSL/TLS：安全套接层协议/传输层安全协议（Secure sockets layer / transport layer security）

AES：高级加密标准（Advanced encryption standard）

5 总体框架

5.1 概述

本标准规定的虚拟电厂云边协同数据安全与隐私保护系统（VPP-CE-DSPP），包含服务器端、虚拟电厂、加密通信通道和半可信计算平台（可选）四个核心组件，构成云边协同架构。该架构通过数据敏感度评估与个性化隐私预算分配，结合自适应梯度裁剪技术，实现多源异构开放电力系统数据（OPSD）的安全、高效处理与模型训练，同时确保用户隐私得到充分保护。VPP-CE-DSPP 系统架构见图 1。

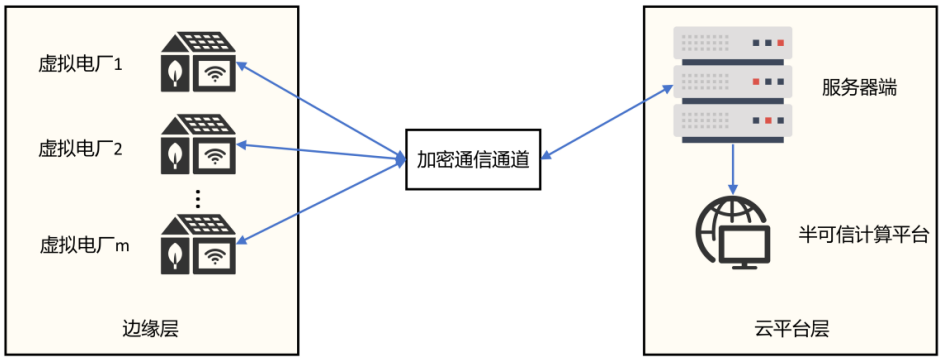

图1　VPP-CE-DSPP 系统架构图

5.2　系统框架

5.2.1　服务器端

服务器端作为 OPSD 管理与模型聚合中心，负责全局模型的管理和聚合。它接收来自虚拟电厂的加密数据或本地训练模型，并进行加权聚合以更新全局模型。系统初始化阶段，服务器端应将预训练全局模型或随机初始化模型分发至各虚拟电厂，用于本地首次训练。服务器端还负责管理隐私预算分配，监控数据流动和模型更新过程，确保整个系统的隐私保护和数据安全性。服务器端功能结构见图 2。

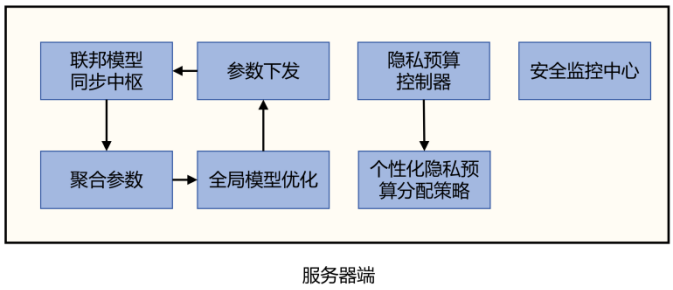

图 2　服务器端功能结构

5.2.2　虚拟电厂

虚拟电厂是由分布式能源、储能、可控负荷等聚合而成的虚拟实体，其边缘智能终端负责本地数据采集、加密与模型训练。它们负责本地非独立同分布（Non-IID）数据的采集和初步处理，包括数据清洗、格式转换等操作。虚拟电厂根据数据的敏感度和分配的隐私预算，进行本地模型训练或数据加密处理。训练完成后，虚拟电厂通过加密通信通道将本地模型或加密数据上传至服务器端。虚拟电厂通常具有一定的计算和存储能力，以支持本地数据处理和模型训练任务。虚拟电厂功能结构见图 3。

虚拟电厂系统应符合 GB/T 22239—2019、GB/T 35273—2020 及 IEC 62351-2020 的通用安全要求，并在以下方面补充虚拟电厂场景的特殊要求：

a）多源异构：数据来自分布式光伏、储能、可控负荷等多种异构设备，格式涵盖 JSON、XML、DL/T 634.5104 等。

b）高时效与多频度：一次调频业务采集频率≥1 次/s，调峰/需求响应≥1 次/min，功率曲线按 15 min 间隔上送。

c）双向互动：不仅上送运行数据（有功/无功、电压、电流、电量等），还需接收调度指令、价格曲线、日内控制计划等下行数据；

d）数据质量要求：日数据完整度≥99%，丢失点数≤0.5%，支持 2 点插值拟合，功率采集准确度不低于 1 级；

e）隐私敏感：用户发电、用电曲线可直接推断行为模式，须采用差分隐私、加密终端、签名与完整性校验等手段保护。

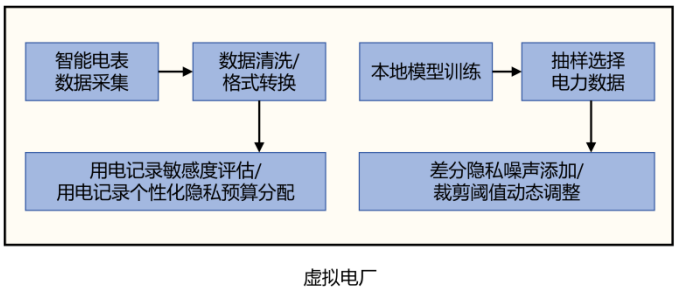

图 3　虚拟电厂功能结构

5.2.3　加密通信通道

加密通信通道是连接虚拟电厂和服务器端的加密安全传输信道。它确保数据在传输过程中的安全性，防止数据被窃取或篡改。通信通道采用加密技术，如 安全套接层协议/传输层安全协议（SSL/TLS） 或其他安全协议和应用程序接口（API）对齐，对数据进行加密传输，保障数据的保密性和完整性。加密通信通道的设计还需考虑低延迟和高带宽，以满足实时性要求和大量数据传输的需求。

5.2.4　半可信计算平台（可选）

半可信计算平台是一个可选组件，用于执行同态加密与安全多方计算任务。它可以增强系统的安全性和可信度，特别是在涉及多个参与方的数据协作场景中。半可信计算平台可以协助进行加密数据的计算，无需解密数据，从而保护数据隐私。它还可以作为独立的计算节点，参与安全多方计算过程，确保计算结果的正确性和隐私保护。

5.3　数据流动路径

5.3.1　数据采集与预处理

数据流动从新能源主体侧的智能终端开始。这些智能终端采集多源异构电力数据，包括发电量、用电量、储能状态等。采集到的数据首先在虚拟电厂进行本地预处理，预处理步骤包括数据清洗、格式转换、特征提取等操作，以确保数据的质量和一致性。预处理后的数据根据其敏感度进行分级，为后续的隐私保护措施提供依据。

5.3.2　本地训练与加密上传

预处理后的数据在虚拟电厂进行本地模型训练或加密处理。根据数据的敏感度和分配的隐私预算，虚拟电厂采用不同的隐私保护策略。对于高敏感数据，采用更强的加密算法和差分隐私（DP）技术进行保护。本地训练过程中，虚拟电厂使用自适应梯度裁剪技术，降低通信开销，同时提升模型训练效率和性能。训练完成后，虚拟电厂将本地模型或加密数据通过加密通信通道安全上传至服务器端。服务器端接收到上传的数据后，进行聚合处理，生成全局模型或综合分析结果。在确保隐私的前提下，服务器端对聚合后的模型或结果进行授权解密，使其可用于后续的电力系统分析和决策。授权解密过程严格遵循隐私保护要求，确保只有经过授权的用户或系统能够访问敏感信息。

6　数据交互与处理流程

6.1　系统初始化

6.1.1　数据收集与整理

从虚拟电厂的各个节点，包括家庭光伏储能系统、工业用电设备、分布式能源发电装置等，使用监控与数据采集系统（SCADA）收集多源异构电力数据。这些数据以时间序列的形式被记录下来，并存储在分布式的数据采集点。

6.1.2 参数初始设置

建议联邦学习系统初始学习率为 0.01，每 10 轮通信衰减 10%；批大小设置为 64；规划通信轮次为 25；客户端采样率为 0.2。这些参数可根据具体应用场景和数据特性进行调整优化。

6.2 数据采集与预处理

6.2.1 数据采集

数据来源于智能电能表、SCADA 系统、分布式能源管理系统等，采集频率依据数据类型和应用场景确定，如智能电能表数据每 15 分钟采集一次，SCADA 系统数据实时采集。数据质量指标应满足 5.2.2 d）的规定。

6.2.2 数据预处理步骤

对收集到的原始数据进行清洗，去除噪声和异常值；进行格式统一，将不同来源的数据转换为标准化格式；实施特征工程，提取关键特征以提高模型训练效率。步骤如下：

a） 数据清洗：识别并处理缺失值、错误值和重复值，采用插值、删除或修正等方法。
b） 数据归一化：将数据线性变换至[0,1]或[-1,1]区间，消除量纲影响。
c） 特征提取：选取与电力负荷预测相关的特征，如历史负荷、气象数据、节假日标识等。

6.2.3 数据敏感度评估

运用标准化的评估模型，考量数据的波动性、规律性及与其他数据的相关性，将电力数据划分为多个敏感度等级，为后续的隐私保护措施提供依据。

6.3 本地训练与加密

6.3.1 本地模型训练

本地模型训练应按下列步骤进行：

a） 数据抽样：按照预设的采样率和采样策略，从本地数据集中抽取用于训练的样本子集。
b） 梯度计算：基于抽取的样本子集，根据时间卷积网络（TCN）等模型计算参数的梯度更新值。
c） 噪声添加：依据分配的隐私预算，向梯度中添加适量高斯噪声以保护数据隐私。

6.3.2 数据加密处理

采用对称加密算法（如 AES）对训练数据进行加密，确保数据在本地存储和传输过程中的安全性；运用同态加密技术对模型参数进行加密，使得云端服务器能够在不解密的情况下对加密数据进行聚合操作。

6.4 数据安全上传

6.4.1 上传流程

本地训练完成后，多接入边缘计算（MEC）客户端将加密后的模型参数或数据通过加密通信通道上传至云端服务器。

6.4.2 完整性校验

在数据上传过程中，采用哈希算法对数据块生成摘要，云端服务器接收到数据后重新计算摘要并进行比对，确保数据的完整性。

6.4.3 保密性保障

通信通道采用 SSL/TLS 协议进行加密，防止数据在传输过程中被窃听或篡改。

6.5 云端服务器聚合与模型更新

服务器接收来自各个虚拟电厂的参数，并进行聚合，步骤如下：

a） 模型参数收集：云端服务器接收来自各个虚拟电厂上传的加密模型参数。

b） 加权聚合：根据各虚拟电厂的数据量、模型性能等权重因素，对收集到的模型参数进行加权平均，更新全局模型参数。

c） 隐私预算监控：实时跟踪每个虚拟电厂的隐私预算消耗情况，确保聚合过程中的隐私泄露风险控制在可接受范围内。

6.6 模型授权解密与应用

6.6.1 授权解密

云端服务器对聚合后的全局模型进行加密存储，只有获得合法授权的用户或系统，通过身份验证和权限验证后，才能申请对模型的解密操作。解密后的模型可在电力系统运行监控中心、调度决策支持系统等场景中投入使用，为电网的稳定运行提供决策依据。

6.6.2 模型应用流程

模型授权解密后的应用流程如下：

a） 模型部署：将解密后的模型部署到电力系统的核心业务系统中，如电网调度自动化系统、配电自动化系统等。

b） 推理计算：利用部署好的模型对实时电力数据进行分析和预测，为电网的调度运行提供决策支持。

c） 结果反馈：将模型的分析结果以可视化的方式呈现给电力系统运行人员，辅助其进行决策；同时，将结果反馈给相关业务部门，为电力市场的运营、电网规划等提供参考依据。

7 技术要求

虚拟电厂云边协同隐私保护系统应满足数据敏感度评估、隐私预算分配、自适应梯度裁剪及云边协同架构等技术要求。

7.1 数据敏感度评估

7.1.1 短期波动敏感度计算

短期波动敏感度计算应按下列步骤进行：

a） 将电力数据按时间顺序划分为多个短时间窗口，窗口长度可根据具体应用场景确定，如每小时一个窗口。

b） 计算每个窗口内用电量的变化率，具体公式为：

$$r_i = \frac{P_i - P_{i-1}}{P_{i-1}} \tag{1}$$

式中：P_i--第i个窗口的用电量，P_{i-1}--第$i-1$个窗口的用电量。

c） 计算所有n个窗口变化率的标准差，标准差越大，--短期波动越剧烈，敏感度越高。标准差的计算公式为：

$$\sigma_s = \sqrt{\frac{1}{n-1}\sum_{i=1}(r_i - \bar{r})^2 n} \tag{2}$$

式中：r_i--第i个窗口的变化率，$\bar{r}$--所有窗口变化率的平均值，n--窗口数量。

7.1.2 长周期规律敏感度计算

长周期规律敏感度计算应按下列步骤进行：

a） 按长周期（如周、月）聚合数据，分析其规律性，如工作日与周末差异、季节性趋势等。

b） 计算每个长周期段的用电量与典型模式的偏差度，具体公式为：

$$d_j = \frac{Q_j - Q_j^{typical}}{Q_j^{\text{typical}}} \tag{3}$$

式中：Q_j--第j个长周期段的实际用电量，Q_j^{typical}--第j个长周期段的典型模式用电量。

c） 根据偏差度评估长周期规律敏感度，偏差度越小，规律性越强，敏感度越高。

7.1.3 综合敏感度评分计算

综合敏感度评分计算应满足下列要求：

a） 结合短期和长周期指标，通过加权公式计算综合敏感度评分，具体公式为：

$$S_{\text{comp}} = w_s \cdot S_s + w_l \cdot S_l \tag{4}$$

式中：w_s和w_l分别为短期和长周期敏感度的权重，满足$w_s + w_l = 1$；S_s为短期敏感度评分，S_l为长周期敏感度评分。

b） 根据综合敏感度评分划分用电数据敏感等级。

规范中常用符号含义见符号说明表。

表 1：符号说明表

符号	含义	单位/范围	属性解释
D	电力数据集	大于 0 的整数	包含来自不同虚拟电厂或终端的多源异构数据
N	窗口数量/数据集数量	个	用于敏感度计算或隐私预算分配
x_t	第 t 个时间窗口的用电量	kWh	短期波动敏感度计算输入
Δ_t	第 t 个窗口用电量变化率	大于等于 0 的浮点数	$\Delta_t = (x_t - x_{t-1})/x_{t-1}$
$\bar{\Delta}$	所有窗口变化率的平均值	大于等于 0 的浮点数	用于标准差计算
σ_Δ	短期波动敏感度指标	大于 0 的整数	$\sigma_\Delta = \sqrt{(1/\mathrm{N})\Sigma_{t=1}^{N}(\Delta_t - \bar{\Delta})^2}$
y_i	第 i 个长周期段的实际用电量	kWh	例如周、月聚合值
y_i^{ref}	第 i 个长周期段的典型模式用电量	kWh	由历史规律或基准模型确定
δ_i	长周期规律敏感度偏差度	大于等于 0 的浮点数	$\delta_i = \lvert y_i - y_i^{ref} \rvert / y_i^{ref}$
S_s	短期波动敏感度评分	大于 0 的整数	来自σ_Δ
S_l	长周期规律敏感度评分	大于 0 的整数	来自 δ_i
S	综合敏感度评分	大于 0 的整数	$S = \alpha S_s + \beta S_l$，其中 $\alpha + \beta = 1$
α, β	综合评分权重系数	[0,1]	由应用场景设定
ε	总隐私预算	大于等于 0 的浮点数	DP 控制参数，数值越小保护越强
ε_i	分配给第 i 个数据集的隐私预算	大于等于 0 的浮点数	按敏感度与数据量分配
w_i	第 i 个数据集的权重	[0,1]	与数据敏感度或规模相关
T_i	第 i 个数据集的数据时效性	时间（天/月）	用于修正预算分配公式
$f(T_i)$	时效性修正函数	[0,1]	例如$f(T) = e^{-\lambda T}$
m	梯度裁剪阈值	[1-20]	控制每轮训练中梯度范数的最大值
η	学习率	(0-1]	模型训练参数

ξ	添加的噪声项	服从高斯分布 $N(0,\sigma^2)$	DP 保护中注入的随机噪声

7.2 隐私预算分配

7.2.1 分配原则

隐私预算分配应遵循以下原则：

a）敏感度越高，分配的隐私预算越低，以加强对高敏感数据的保护。

b）数据量越大，分配的隐私预算越高，以平衡数据效用和隐私保护。

7.2.2 计算公式

隐私预算分配的具体计算公式为：

$$\varepsilon_i = \frac{1/S_{\mathrm{i}} \cdot D_i \cdot f(T_i)}{\sum_{j=1}^{m}(1/S_{\mathrm{j}} \cdot D_j \cdot f(T_j))} \cdot \varepsilon_{\mathrm{total}} \tag{5}$$

式中：ε_i--第i个数据集的隐私预算，S_i--第i个数据集的敏感度评分（1-5），D_i--第i个数据集的数据量，$\varepsilon_{\mathrm{total}}$--总隐私预算，$m$--数据集数量，$T_i$--第$i$个数据集的数据时效性，$f(T_i)$--时效性修正函数。

7.2.3 分配流程

隐私预算分配流程应按下列步骤执行：

a） 计算每个数据集的敏感度评分S_i。

b） 统计每个数据集的数据量D_i。

c） 计算每个数据集的权重$1/S_i \cdot D_i$。

d） 根据上述公式计算每个数据集的隐私预算ε_i。

e） 将隐私预算分配给相应的数据集。

7.2.4 隐私预算量化指标

隐私预算分配应满足下列定量要求：

a）单节点总预算$\varepsilon_{total} \leq 5.0$。

b）中低敏感数据集$\varepsilon_i \leq 1.0$。

c）高敏感数据集$\varepsilon_i \leq 0.5$。

7.3 自适应梯度裁剪

7.3.1 技术原理

自适应梯度裁剪技术原理应符合下列说明：

a）自适应梯度裁剪通过动态调整梯度裁剪阈值，防止梯度爆炸和过度噪声添加，提高模型训练效率和性能。

b）在训练过程中，根据梯度的分布情况自动调整裁剪阈值，确保梯度的大小在合理范围内。

7.3.2 实现方法

自适应梯度裁剪实现方法如下：

a）初始化裁剪阈值，推荐初始取值范围$\theta \in [0.1-1.0]$和移动平均系数α。

b）在每次迭代中，计算当前批次梯度的范数$|\nabla w|$。

1. 更新裁剪阈值：

$$\theta_{\mathrm{new}} = \alpha \cdot \theta_{\mathrm{old}} + (1-\alpha) \cdot |\nabla w| \tag{6}$$

2. 根据新的裁剪阈值对梯度进行裁剪：

$$\nabla w_{\mathrm{clipped}} = \frac{\nabla w}{\max\left(1, \frac{|\nabla w|}{\theta_{\mathrm{new}}}\right)} \tag{7}$$

3.使用裁剪后的梯度更新模型参数。

7.3.3 动态调整策略

动态调整策略应符合下列规定：

a) 在训练初期，梯度变化较大，裁剪阈值应较快适应梯度分布。

b) 随着训练进行，梯度逐渐稳定，裁剪阈值的调整幅度应逐渐减小。

c) 可根据模型的收敛情况和隐私保护需求，动态调整裁剪阈值的更新频率和幅度。

7.4 云边协同架构

7.4.1 整体框架

云边协同整体框架应满足下列要求：

a) 云边协同架构包括云端服务器和多个边缘设备（如虚拟电厂），通过加密通信通道进行数据交互和模型更新。

b) 云端服务器负责全局模型的管理和聚合，边缘设备负责本地数据的采集、预处理和模型训练。

7.4.2 数据传输流程

数据传输流程应按下列步骤进行：

a) 边缘设备将采集到的数据进行预处理和加密。

b) 通过加密通信通道将加密后的数据或本地训练模型上传至云端服务器。

c) 云端服务器对上传的数据或模型进行聚合处理，生成全局模型。

d) 云端服务器将全局模型分发给各边缘设备，用于下一轮本地训练。

7.4.3 数据处理和存储流程

数据处理和存储流程应满足下列要求：

a) 边缘设备存储本地预处理后的数据，并进行本地模型训练。

b) 云端服务器存储全局模型和聚合后的数据，确保数据的安全性和可用性。

c) 数据存储应符合相关安全标准，采用加密技术保护数据隐私。

7.4.4 云边协同架构性能量化指标

云边协同架构应满足下列定量要求：

a) 端到端通信延迟≤200ms。

b) 单节点上行带宽≥10Mbps。

c) 联邦学习通信轮次≤50。

d) 模型聚合完成时间≤30s（节点数≤100）。

e) 边缘节点 CPU 占用率≤80%，内存占用率≤70%。

8 安全与隐私要求

8.1 数据传输安全

数据传输安全应符合下列规定：

a) 加密技术：数据在传输过程中必须采用加密技术，如 SSL/TLS 协议，禁止使用 SHA-1、RSA-1024、MD5 等弱算法，必须符合 GB/T 39786-2021、GB/T 20986-2023、GB/T 22239-2019 和 GM/T 0054-2018 要求，确保数据在云端服务器与虚拟电厂之间的传输过程中的保密性和完整性。

b) 通信通道安全：通信通道应采用安全的网络协议，防止数据在传输过程中被窃取或篡改。

8.2 数据存储安全

数据存储安全应满足下列要求：

a) 加密存储：在云端服务器和虚拟电厂，数据应以加密形式存储，采用高级加密标准（AES）等加密算法，参考 GB/T 36572-2018 的要求执行保护数据的机密性。

b） 访问控制：实施严格的访问控制机制，确保只有授权用户和系统能够访问敏感数据。

8.3 数据处理安全

数据处理安全应符合下列规定：

a） 隐私保护算法：在数据处理过程中，采用差分隐私等隐私保护算法，防止用户隐私数据泄露。

b） 模型安全：确保本地和全局模型的安全性，防止模型被篡改或用于恶意目的。

8.4 用户隐私保护

用户隐私保护应满足下列要求：

a） 差分隐私机制：采用差分隐私技术，在数据发布和共享过程中保护用户的隐私，确保单个用户的隐私信息不会被泄露。

b） 隐私预算管理：合理分配和管理隐私预算，确保在满足隐私保护要求的前提下，数据的效用最大化。

8.5 合规性要求

合规性要求应符合下列规定：

a） 法律法规遵守：本标准遵循《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》等相关法律法规，确保数据处理活动的合法性。

b） 行业规范遵循：同时应遵循 GB/T 35273-2020《信息安全技术 个人信息安全规范》和 GB/T 41373-2022《信息安全技术 数据分类分级指南》的条款等国家标准和行业规范，确保数据隐私和安全措施的合规性。

9 测试与验证

9.1 测试方法

9.1.1 功能测试

测试数据集应来源于真实电网运行数据和公开新能源数据集（如 Pecan Street、UCI 电力负荷数据集等），测试环境应明确包括云端服务器性能参数（CPU/GPU 型号、内存容量）与边缘节点配置（计算能力、存储容量、网络带宽），具体测试步骤如下：

a） 数据采集与预处理测试：验证系统能否准确采集多源异构电力数据，并进行有效的预处理，包括数据清洗、归一化、特征提取等操作。检查数据的质量和一致性，确保其满足后续模型训练的要求。

b） 本地训练与加密测试：验证虚拟电厂是否能够根据分配的隐私预算正确地进行本地模型训练，并对数据进行加密处理。检查加密算法的正确性和加密数据的完整性。

c） 数据上传与聚合测试：验证加密后的数据或本地训练模型能否通过加密通信通道安全上传至云端服务器，并在云端服务器进行正确的聚合处理。检查数据传输过程中的安全性和聚合结果的准确性。

d） 模型应用与反馈测试：验证聚合后的模型能否在电力系统分析和决策中有效应用，检查模型的输出结果是否准确，并评估其对电力系统运行的指导意义。

9.1.2 性能测试

性能测试应包括下列内容：

a） 模型预测准确性测试：使用测试集对训练得到的模型进行评估，计算测试准确率（$test_{acc}$）和测试损失（$test_{loss}$）等指标，验证模型对电力数据的预测性能。

b） 训练效率测试：记录模型训练过程中的时间消耗，包括本地训练时间和云端服务器聚合时间，评估系统的训练效率。

c） 通信效率测试：测量数据在虚拟电厂与云端服务器之间的传输时间、带宽占用等指标，评估系统的通信效率。

9.1.3 隐私保护测试

隐私保护测试应符合下列规定：

a） 隐私泄露风险评估：通过模拟攻击等方式，评估系统在数据采集、传输、存储和处理过程中是否存在隐私泄露风险。检查加密算法、差分隐私机制等隐私保护措施的有效性。

b） 隐私预算消耗测试：监测系统在运行过程中的隐私预算消耗情况，验证隐私预算分配策略的合理性和有效性。

c） 敏感数据保护测试：针对高敏感度数据，进行专项隐私保护测试，确保其在各个环节得到充分保护，不会被非法获取或泄露。

9.1.4 场景化测试

场景化测试应包括下列情形：

a）极端数据场景测试：模拟新能源出力剧烈波动、负荷突发异常等情况，验证系统在数据分布偏离时的稳定性与鲁棒性。

b）网络延迟与丢包场景测试：模拟高延迟、弱网环境下的通信，检验系统在传输效率和模型收敛上的表现。

c）数据缺失场景测试：构造部分时序数据缺失或错误的情况，评估系统在数据完整性不足时的预测精度和隐私保护效果。

9.2 验证指标

9.2.1 模型预测准确性

模型预测准确性验证应按下列指标计算：

a） 测试准确率（$test_{acc}$）：反映模型在测试集上的整体预测正确率，计算公式为：

$$test_{acc} = \frac{\sum_{i=1}^{N} \delta\left(\hat{y}_i, y_i\right)}{N} \tag{8}$$

式中：N--测试集的总样本数，$\hat{y}_i$--模型对第i个样本的预测结果，y_i--第i个样本的真实标签，$\delta(\hat{y}_i, y_i)$是指示函数，当$\hat{y}_i = y_i$时值为 1，否则为 0。

b） 测试损失（$\mathrm{test}_{\mathrm{loss}}$）：衡量模型在测试集上的误差程度，通常采用均方误差（MSE）等作为损失函数，计算公式为：

$$test_{loss} = \frac{1}{N}\sum_{i=1}(y_i - \hat{y}_i)^2 \tag{9}$$

式中：N--测试集的总样本数，$\hat{y}_i$--模型对第i个样本的预测结果，y_i--第i个样本的真实标签。

9.2.2 定量指标验证测试

定量指标验证测试应与第 7 章技术要求一一对应，按下列步骤执行：

a）数据敏感度评估指标

1. 短期波动敏感度σ_s——抽样窗口≥1000 个，$\sigma_s \geq 0.15$的检出率≥95%。
2. 长周期规律偏差度δ_i——取连续 4 周数据，$\delta_i \geq 10\%$的样本标识准确率≥95%。
3. 综合敏感度S——人工标注 200 条样本，$S \geq 75$分的分类准确率≥90%。

b）隐私预算分配指标

1. 单节点总预算$\varepsilon_{\mathrm{total}}$——读取系统配置，结果≤5.0 视为通过。
2. 极高/高敏感数据集ε_i——抽取 10 个数据集，$\varepsilon_i \leq 0.5$（极高）、$\varepsilon_i \leq 0.5$（高）的比例=100%。
3. 预算消耗率——模拟 25 轮训练，消耗≥90%时系统应给出告警，告警延迟≤1s。

c）自适应梯度裁剪指标

1. 初始阈值θ_0——检查配置文件，$0.3 \leq \theta_0 \leq 1.0$即为通过。
2. 平滑系数α——读取代码缺省值，$\alpha = 0.9 \pm 0.01$。
3. 梯度范数上限——注入梯度范数=25 的测试样本，裁剪后范数≤20。
4. 阈值变化量——记录 100 次迭代，$|\theta_{\mathrm{new}} - \theta_{\mathrm{old}}| \leq 0.1$的次数占比≥98%。

d）云边协同架构指标

1. 端到端延迟——发送 1000 条 32KB 模型参数，99 百分位延迟≤200ms。
2. 上行带宽——iPerf3 打流，单节点持续 10s 平均带宽≥10Mbps。
3. 通信轮次——完成一次全局收敛，实际轮次≤50。
4. 聚合时间——100 个节点同时上传，从最后一个包到达至聚合完成≤30s。
5. 资源占用——30min 监测，CPU≤80%、内存≤70%的持续时长≥95%。

9.2.3 隐私保护水平

隐私保护水平验证应包括下列指标：

a） 隐私泄露评估指标：采用隐私泄露概率、信息熵等指标评估系统的隐私保护水平，确保用户隐私信息在数据处理过程中不被泄露。

b） 隐私预算消耗指标：监测隐私预算的使用情况，确保其在合理范围内，以平衡隐私保护与数据效用。

9.2.4 通信效率

通信效率验证应包括下列指标：

a） 通信开销：测量数据传输过程中的通信量，包括上传和下载的数据量，评估系统的通信效率。

b） 传输时间：记录数据在虚拟电厂与云端服务器之间的传输时间，评估系统的实时性和响应速度。

10 管理维护要求

10.1 本地参数配置要求

本地参数配置应满足下列要求：

a） 终端内生认证模型配置界面：对于终端内生认证模型，应能通过统一的配置界面，对物理层特征参数和认证协议参数进行配置。物理层特征参数包括信号强度、信道状态信息等，认证协议参数包括认证超时时间（建议范围 10s-30s）、重试次数（建议不超过 5 次）等。

b） 网关代理认证模型参数分离配置：对于网关代理认证模型，应能分别配置物理层特征提取参数和认证协议参数。物理层特征提取参数如采样频率（建议不低于 10Hz）、滤波参数等，认证协议参数如握手消息间隔（建议 1s-5s）、会话密钥更新周期（建议 1min-10min）等。

10.2 远程参数配置要求

远程参数配置应符合下列规定：

a） 安全通道建立与协议使用：应支持通过安全通道对认证系统进行远程配置。应采用 TLS 或 DTLS 协议保护远程配置通道，确保配置指令传输的机密性和完整性。

b） 配置优先级设定：当本地参数配置与远程参数配置不一致时，应优先使用远程参数配置，以保证远程管理的有效性和一致性。

10.3 软件升级要求

软件升级应满足下列要求：

a） 安全远程升级能力：应具备认证系统组件的安全远程升级能力，支持自动或手动触发升级流程。

b） 标准协议支持：应支持基于 HTTPS 或 DTLS 保护的 CoAP 等标准协议进行软件升级，保障升级过程的安全性。

c） 版本兼容性检查与错误提示：软件升级前应进行版本兼容性检查，对不适用于当前设备的软件禁止进行升级操作，并返回明确的错误提示信息，如“软件版本与设备不兼容，升级终止”。

d） 完整性验证与恢复机制：升级过程中，应通过密码学方法（如哈希值校验）对升级文件进行完整性验证。当完整性验证失败时，应终止升级操作并恢复至升级前状态，确保系统稳定性。在

任何软件升级失败情况下，认证系统应保持基本功能，防止设备被锁定，如允许设备以降级模式运行或启用备份认证机制。

10.4 参数备份要求

参数备份应符合下列规定：

a) 备份功能与权限控制：认证系统应具备安全参数的本地或远程备份功能，应对参数备份操作实施权限控制，仅允许管理员通过身份验证后执行备份操作。

b) 加密存储与安全导入：应以加密格式存储导出的参数文件，加密算法建议使用 AES-256 或以上强度算法。应支持参数文件的安全导入功能，并进行适当的完整性验证，如通过数字签名验证参数文件的来源和完整性。

10.5 恢复默认配置要求

应提供本地和远程恢复认证系统默认参数配置的操作方法，本地操作可通过设备上的实体按键或操作菜单实现，远程操作应通过安全的管理界面进行。

10.6 系统重启要求

应提供本地和远程重启认证系统组件的操作方法，确保在系统异常或升级后能顺利重启。本地重启可通过设备管理界面或命令行工具实现，远程重启应支持通过网络管理平台发送重启指令，并返回重启状态信息。

10.7 系统日志要求

应具备日志功能，记录影响认证系统的本地和远程操作的时间、类型、操作员等信息，操作日志应保留不少于 180 天。日志应能本地保存，存储容量建议不少于 1GB，宜具有定期备份功能，备份周期可根据设备存储能力和应用场景设置为每周或每月。

10.8 认证服务器连接参数配置要求

认证服务器连接参数配置应符合下列规定：

a) 预置与本地配置：认证系统宜在出厂时，根据应用场景预置认证服务器连接参数，包括服务器地址、端口号、认证协议版本等。应支持通过本地配置修改认证服务器连接参数，如支持通过设备的配置界面或命令行工具进行修改，包括服务器 URL 和认证凭证的更新。

b) 远程配置与安全通信：对于终端内生认证模型，宜支持认证服务器参数的安全远程配置，确保参数更新的及时性和安全性。系统应支持 MQTT-TLS 或 CoAP-DTLS 等标准物联网安全协议与认证服务器进行安全通信，遵循标准化流程与认证服务器建立安全连接，如按照 TLS 握手协议完成身份验证和密钥协商。对于使用 TLS/DTLS 进行安全通信的系统，应实施标准的安全证书管理程序，包括证书的颁发、更新、撤销等操作，确保通信双方身份的可靠性和数据传输的保密性。

参 考 文 献

[1] Pei J, Liu W, Li J, et al. A review of federated learning methods in heterogeneous scenarios[J]. IEEE Transactions on Consumer Electronics, 2024, 70(3): 5983-5999.
[2] Chen J, Yan H, Liu Z, et al. When federated learning meets privacy-preserving computation[J]. ACM Computing Surveys, 2024, 56(12): 1-36.
[3] Yazdinejad A, Dehghantanha A, Karimipour H, et al. A robust privacy-preserving federated learning model against model poisoning attacks[J]. IEEE Transactions on Information Forensics and Security, 2024, 19: 6693-6708.
[4] Hallaji E, Razavi-Far R, Saif M, et al. Decentralized federated learning: A survey on security and privacy[J]. IEEE Transactions on Big Data, 2024, 10(2): 194-213.
[5] ElRobrini F, Bukhari S M S, Zafar M H, et al. Federated learning and non-federated learning based power forecasting of photovoltaic/wind power energy systems: A systematic review[J]. Energy and AI, 2024, 18: 100438.
[6] Lin W T, Chen G, Zhou X. Privacy-preserving federated learning for detecting false data injection attacks on power system[J]. Electric Power Systems Research, 2024, 229: 110150.
[7] Zheng R, Sumper A, Aragüés-Peñalba M, et al. Advancing power system services with privacy-preserving federated learning techniques: A review[J]. IEEE access, 2024, 12: 76753-76780.
[8] Li Y, Chen S, Hu X, et al. Power Data Analysis and Privacy Protection Based on Federated Learning[J]. Scalable Computing: Practice and Experience, 2025, 26(2): 630 - 640-630 - 640.
[9] Alshardan A, Tariq S, Bashir R N, et al. Federated learning (FL) model of wind power prediction[J]. IEEE Access, 2024, 12: 129575-129586.
[10]Zhao Y, Pan S, Zhao Y, et al. Ultra-short-term wind power forecasting based on personalized robust federated learning with spatial collaboration[J]. Energy, 2024, 288: 129847.
[11]Rajesh M, Ramachandran S, Vengatesan K, et al. Federated learning for personalized recommendation in securing power traces in smart grid systems[J]. IEEE Transactions on Consumer Electronics, 2024, 70(1): 88-95.