

智能电网中隐私保护的多维多函数数据聚合方案

王圆圆¹, 韩益亮^{1,2}, 吴立强¹, 李鱼¹, 姚武军¹, 吴旭光¹, 朱率率¹

(1. 武警工程大学密码工程学院, 陕西 西安 710086; 2. 武警工程大学反恐指挥信息工程教育部重点实验室, 陕西 西安 710086)

摘要: 智能电网中精细的用户数据可以有效支撑电网运行。目前, 如何在不影响用户隐私的前提下, 实现数据的功能性分析, 尤其是数据的多维化和多函数操作, 是一个亟待解决的问题。现有数据聚合方案大多基于 Paillier 和 BGN 加密系统, 无法同时兼顾数据的多维化和多函数操作。针对这个问题, 设计了一种基于半同态加密的多维多函数数据聚合方案, 旨在保护用户隐私的同时, 实现电表数据的多维化和多函数操作。该方案通过引入部分同态加密的机制, 利用盲化因子和哈希函数技术, 增强了数据的机密性和利用率。通过安全性分析证明该方案能够抵抗选择明文攻击, 并且在计算效率方面优于基于 Paillier 加密的方案。

关键词: 隐私保护; 数据聚合; 同态加密; 智能电网; 多函数运算

中图分类号: TP393

文献标志码: A

DOI: 10.11959/j.issn.2096-109x.2025037

Privacy-preserving multi-dimensional multi-function data aggregation scheme in smart grid

WANG Yuanyuan¹, HAN Yiliang^{1,2}, WU Liqiang¹, LI Yu¹, YAO Wujun¹,
WU Xuguang¹, ZHU Shuaishuai¹

1. College of Cryptography Engineering, Armed Police Engineering University, Xi'an 710086, China
2. Key Laboratory of Counter-Terrorism Command Information Engineering of the Ministry of Education,
Armed Police Engineering University, Xi'an 710086, China

Abstract: In smart grid, fine-grained user data can be effectively leveraged to support grid operation. However, achieving functional analysis of such data, especially its multi-dimensionalization and multi-functional operation, without compromising user privacy has been identified as an urgent problem. Most existing data aggregation schemes were based on the Paillier and BGN encryption systems, which failed to simultaneously address the multi-dimensionalization and multi-functional operation of data. To tackle this issue, a multi-dimensional and multi-functional data aggregation scheme was designed based on semi-homomorphic encryption. This scheme aimed to enable the multi-dimensionalization and multi-functional operation of meter data while protecting user privacy. The mechanism of partial homomorphic encryption was introduced, and blinding factor and hash function techniques were utilized. These enhancements significantly improved data confidentiality and utility. Security analysis indicates that the scheme is robust against chosen ciphertext attacks. In terms of computational efficiency, it outper-

收稿日期: 2025-02-10; 修回日期: 2025-03-30

通信作者: 韩益亮, hanyil@163.com

基金项目: 国防科技创新计划

Foundation Item: The National Defense Technology Innovation Program

引用格式: 王圆圆, 韩益亮, 吴立强, 等. 智能电网中隐私保护的多维多函数数据聚合方案[J]. 网络与信息安全学报, 2025, 11(3): 163-174.

Citation Format: WANG Y Y, HAN Y L, WU L Q, et al. Privacy-preserving multi-dimensional multi-function data aggregation scheme in smart grid[J]. Chinese Journal of Network and Information Security, 2025, 11(3): 163-174.

forms existing schemes based on the Paillier encryption system.

Keywords: privacy preservation, data aggregation, homomorphic encryption, smart grid, multi-function arithmetic

0 引言

智能电网 (smart grid, SG) 作为电力系统的创新, 通过先进的信息通信技术实现了电力资源的高效管理和分配。智能电表 (smart meter, SM) 作为智能电网, 尤其是智能配电网中数据采集的核心设备, 承担着原始电能数据的采集、计量与传输等关键任务。其中, 电表数据的采集及其多函数操作对于电网的运行维护、电力市场定价、用户行为分析等方面具有重要意义^[1]。在电力市场定价方面, 通过电表数据的内积计算, 可以评估不同用户间的用电关联性, 为市场双边交易提供科学的决策支持^[2]。此外, 方差分析揭示了电网负荷的波动性, 为电力市场的定价策略提供了坚实的数据基础, 并有助于识别不同的用户用电模式, 构建精细化的用户画像。用户行为分析同样能够从电表数据的多维统计分析中获益^[3]。深入挖掘电表数据的潜力, 不仅可以提高电网的运行效率和可靠性, 还能为电力市场的稳定发展和用户的精细化管理提供坚实的数据支撑。

然而, 智能电表所收集的精细数据同样引发了隐私保护方面的担忧。这些数据的泄露可能侵犯用户的隐私权, 例如, 通过分析用户的用电模式, 不法分子或未经授权的第三方可能推断出用户的日常行为习惯^[4]。因此, 能否对用户的隐私信息保密同时不影响系统功能, 成为了智能电网在现实生活中推广的决定性因素之一。

本文设计了一种基于环学习误差 (ring learning with errors, RLWE) 的多维多函数数据聚合方案, 旨在智能电网领域中实现数据隐私保护与高效聚合运算的双重目标。本方案结合多维数据处理与多函数运算的能力, 借助半同态加密 (semi-homomorphic encryption, SHE) 机制, 并引入盲化因子和容错机制, 支持对加密数据进行加法运算和有限范围内乘法操作, 为智能电网中的数据隐私保护和高效聚合运算问题提供了一种新的解决方案。

1 相关工作

在智能电网领域, 隐私保护和数据聚合的高效

性是两个核心的研究方向, 其大致分为以下几类。

1) 基于 Paillier 加密系统的方案: 这类方案能够实现数据的加法同态操作, 但受限于只能执行加法或乘法中的一种, 无法同时进行。例如, FPDA^[4]方案通过 Shamir 秘密共享机制增强了系统的容错性; Shang 等^[1]提出了一种结合 Paillier 同态加密和 ECDSA (elliptic curve digital signature algorithm) 签名的隐私保护数据聚合方案, 用于边缘计算支持的 IIoT (industrial internet of thing), 有效解决了数据隐私和完整性问题; Chen 等^[2]结合了 ElGamal 同态加密系统和 Paillier 同态加密技术, 提出了一种弹性的基于分组的子集数据聚合方案, 解决了智能电网中实时数据可能泄露用户隐私信息的问题。

2) 基于 BGN 加密系统的方案: 这类方案支持无限次加法操作和有限次乘法操作, 为数据隐私性提供了新的解决方案^[5-8]。Khan 等^[6]结合了 BGN (Boneh-Goh-Nissam) 同态加密技术和椭圆曲线数字签名算法, 提出了一种雾计算支持的智能电网中的隐私保护数据聚合方案, 解决了在存在故障智能电表情况下的数据聚合问题, 确保了数据的隐私性、完整性和源认证。

3) 基于双线性配对的隐私保护方案: 这类方案利用双线性配对和同态加密实现高效的数据处理, 但通常只允许一阶同态乘法, 限制了其在智能电网中的应用^[9-12]。Abdallah 和 Shen^[9]结合了轻量级基于格的同态加密技术和智能家电之间的直接通信, 提出了一种保护智能电网中居民用户电耗数据隐私和完整性的聚合方案, 实现了在不增加智能电表计算负担的情况下保护用户隐私和数据安全的目标; Wang 等^[10]结合了同态加密技术和双线性配对技术, 提出了一种在雾计算环境中保护数据隐私和用户身份的匿名安全聚合方案, 成功实现了在公共云服务器和终端设备之间安全、高效地上传和处理加密数据; PDA-SP^[11]是基于 SHE 加密系统的智能电网数据聚合方案, 它使用控制中心的公钥对数据进行加密, 这种方式不能保证单个用户数据的隐私。

然而,上述方案大多根据智能电网的现实需要,对智能电表上传的数据进行更为精细的处理,使得用户数据可以在加密的情况下进行更多的处理分析。主要包括两个主要方向。

1) 多维数据统计:通过对用户耗电量的细粒度划分,使得控制中心能够根据多维电力分析优化电力分配方案^[13-16]。Shen等^[13]结合Paillier密码体系和Horner法则,提出了一种高效的隐私保护数据聚合方案,解决了智能电网中多维电能消耗数据的安全聚合问题,实现了对大规模用户群体的灵活和高效管理,但该方案无法抵抗差分攻击;Liu等^[15]结合Paillier同态加密技术和HMQV密钥协商协议,提出了一种无须可信第三方的多子集数据聚合方案,解决了智能电网中用户隐私保护和动态用户管理的问题,实现了用户数据隐私保护和高效的多子集聚合功能。

2) 多函数数据操作:允许在加密状态下对用户耗电量进行求和、方差、平均等操作,为用户耗电量分析提供了更多的维度。Zhao等^[17]结合了部分同态加密技术和打包方法,提出了一种智能且实用的隐私保护数据聚合方案,解决了雾计算基础智能电网中的多样化计费、多函数统计和效率问题,实现了在保护用户隐私的同时进行高效的数据聚合和安全计算,但在实际部署中可能需要进一步优化以应对更复杂的网络环境 and 安全威胁^[18]。

然而,现有方案难以同时满足多维数据统计和多函数数据操作的需求^[19]。多维数据聚合方案主要基于指数上的编码,虽然能够处理多维数据,但无法对加密后的数据执行多函数操作,如数据平方、数据方差等运算。同时,多函数数据聚合方案虽然能够执行加密数据的复杂运算,但无法实现对用户数据的多维细粒度处理。

2 模型描述

2.1 系统模型

本文方案基于智能电网框架,主要涉及4类实体:可信中心(authority, TA)、控制中心(control center, CC)、雾节点(fog node, FN)和智能电表。系统模型如图1所示。

可信中心是系统中最可靠的实体,负责以下任务:在初始化阶段,生成系统参数;在注册阶段,通过安全信道向用户、控制中心发送密钥;

在聚合过程中,为FN提供在线用户的掩码信息。

控制中心负责解密过程,得到正确的聚合数据结果,并对聚合的数值进行处理。

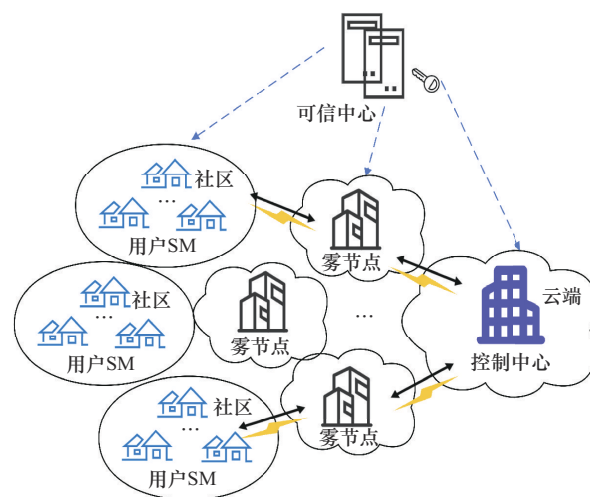


图1 系统模型

Figure 1 System model

雾节点具有较强的计算能力,主要负责对用户发送的密文进行聚合并将结果转发给CC。

智能电表在规定的时间内,对用户使用的电量进行收集。此外,执行用于隐私保护的加密操作,并将加密信息发给相连的雾节点。

2.2 攻击模型

在本文的攻击者模型中,假设智能电表、雾节点、控制中心以及可信中心在某种程度上是“诚实但好奇”的。这些实体在正确遵循协议的情况下,可能会试图获取用户的隐私信息。本文的攻击模型基于以下假设。

1) 前端节点和控制中心均为“诚实但好奇”类型。这意味着,它们会遵循协议,但同时也渴望获取智能电表读数的值。

2) 尽管每个用户希望电费尽可能少,但本文假定所有智能电表用户都是诚实的。

3) 外部敌手 $\mathcal{A}$ 可能会攻破控制中心,但只能得到最终的聚合结果而无法获取到单个用户的隐私信息。

2.3 设计目标

基于上述的系统模型和攻击模型,本文达到了以下设计目标。

1) 机密性:敌手 $\mathcal{A}$ 不能获得任何单个用户的实时用电信息,即在用户与FN之间和FN与CC

之间的信道窃听，或入侵 FN、CC 的数据库，敌手 $\mathcal{A}$ 均无法获得任何单个用户的实时用电信息。

2) 多维多函数运算：智能电网对多维数据进行多函数的运算，如求和、内积、方差等，这符合智能化发展的需求，并在功能上丰富电网的应用。

3) 高效性：轻量级且抗量攻击的数据聚合。基于 RLWE 的半同态加密方案，聚合过程的时间开销比基于 Paillier 的方案更快，同时具有抗量子攻击的特性。

3 方案描述

本文设计了一个多维多函数数据聚合方案，在一个聚合的周期内，智能电表测量用户的耗电量，并用私钥进行加密，随后智能电表通过与雾节点的通信，将加密后的耗电量发送给雾节点。雾节点进行聚合，并将聚合后的数值发送给控制中心，最后，控制中心对聚合结果进行解密。智能电表根据电器的种类将总的耗电量形成一个多维的数据。该方案中使用的符号及其描述见表 1。

表 1 符号及其描述

Table 1 Symbols and descriptions

参数符号	描述
SM_i	用户
FN	雾节点
CC	控制中心
TA	可信中心
$\mathcal{A}$	安全参数
$d_{i,0}, d_{i,1}, \dots, d_{i,l}$	多项式系数
M_{sum}	聚合明文
C_{sum}	聚合后的密文
cc_{variance}	平方差密文
$c_i, cc_\eta, cc_\xi, \tilde{c}, \hat{c}$	密文形式
$\overline{cc}$	平均密文
$\oplus$	同态加法
$\otimes$	同态乘法
PP	公共参数
$\hat{m}, \tilde{m}$	二次解密密文
m_i	耗电量
p_i	用户私钥
Ts	时间戳
H	哈希函数
n	用户总数
n_1	在线用户数量
U_1	全体用户集合
U_2	在线用户集合

3.1 系统初始化阶段

所有 SM、FN 和 CC 需要在 TA 处进行注册。TA 随机选择私钥 $SM_i p_i \leftarrow \mathbb{Z}_N^*$ ($i = 1, \dots, n$)，并将私钥通过安全信道分别发送。在每一个 Ts 下， SM_i 更新自己的掩码为 $H(Ts)^{p_i}$ 。TA 为 CC 生成私钥和公钥。TA 在 $s \leftarrow \chi$ 中选取一个元素，其中 χ 表示离散高斯误差分布，定义为 $D_{\mathbb{Z}^N, \sigma}$ ，其中 $\mathbb{Z}^N$ 为 N 维整数空间， σ 为标准差。选择 $a_1 \leftarrow R_q$ ， $R_q = R/qR = \mathbb{Z}_q[x]/(f(x))$ ，其中 q 为素数。作为随机元素， $e \leftarrow \chi$ 作为错误。令 $sk = s$ 是私钥，计算 $a_0 = -(a_1 \cdot s + t \cdot e)$ ，其中 t 是一个整数，满足 $t < q$ 的条件，用于确定明文空间 $R_t = \mathbb{Z}_t[x]/(f(x))$ ， $pk = (a_0, a_1)$ 是公钥。将哈希函数 SHA-256 记作 H 。

SM_i 将其 $l+1$ 位数据类型按最大位数编码为 $d_{i0}, d_{i2}, \dots, d_{il}$ 。按如下方式构造： $m_i(x) = d_{i0} + d_{i1}x + d_{i2}x^2 + \dots + d_{il}x^l$ ，编码后， SM_i 所含有的多维数据表示为 $m_i(x)$ ，简称为 m_i 。图 2 给出多维数据编码和解码的过程。

3.2 用户报告阶段

CC 需要实时地收集用户的耗电信息，SM 每隔 Ts（如 Ts=15 min）向 FN 报告当前的耗电信息，具体步骤如下。

1) SM_i 随机选择 $R \leftarrow u, f \leftarrow \chi, g \leftarrow \chi$ ，并计算耗电信息 m_i 的密文

$$c_i = \text{Enc}(m_i, pk) = (c_{0,i}, c_{1,i}) = (a_0 u + t g + m_i + H(Ts)^{p_i}, a_1 u + t f) \quad (1)$$

2) 将密文 (c_i, Ts, ID_i) 发送给 FN。

3.3 数据聚合阶段

FN 将在 Ts 内收到消息的用户集合记为 U_2 ，注意 $U_2 \in U_1$ ，并通知 TA 在线用户集合 U_2 。TA 将在线用户的噪声 $H(Ts)^{p_i}$ ($i = 1, 2, 3, \dots, n_1$) 发送给 FN。根据 CC 的要求，在保护客户隐私的同时，FN 可以聚合客户的用电数据 c_i 。

FN 通过同态加法 “ $\oplus$ ” 对所有的密文 c_i ($i = 1, 2, 3, \dots, n_1$) 进行聚合。同态加法 “ $\oplus$ ”，即将两个任意长度的密文 $(R_q)^u \ni (v_0, \dots, v_{u-1}) \mapsto$ 定义为 $cc_\eta = (c_0, \dots, c_\eta)$ 和 $cc_\xi = (c_0, \dots, c_\xi)$ 。而后，计算

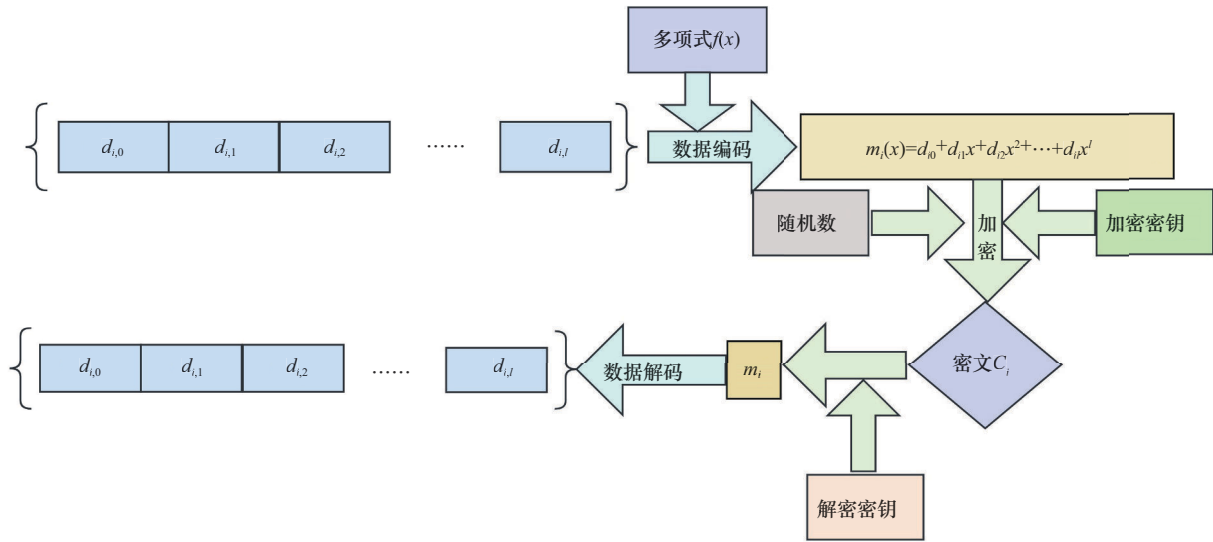


图2 多维数据编码和解码过程

Figure 2 The process of multi-dimensional data encoding and decoding

同态加法“ $\oplus$ ”，即：

$$\text{cc}_\eta \oplus \text{cc}_\xi = \left(c_0 + c'_0, \dots, c_{\max(\eta, \xi)} + c'_{\max(\eta, \xi)} \right) \in R_q^{\max(\eta, \xi) + 1} \quad (2)$$

具体步骤如下：

1) 计算 C_{sum} 为

$$C_{\text{sum}} = \oplus_{i=0}^{n_1} c_i = c_0 \oplus \dots \oplus c_{n_1} - \sum_{i=0}^{n_1} H(\text{Ts})^{p_i} \quad (3)$$

2) 将聚合结果 $(C_{\text{sum}}, \text{Ts})$ 发送给 CC。

3.4 多函数运算阶段

为了多维度分析用户电量，下面进行多函数运算。根据 CC 对数据运算的要求，TA 将在线用户的噪声 $H(\text{Ts})^{p_i} (i = 1, 2, 3, \dots, n_1)$ 发送给 FN，然后进行多函数操作。

3.4.1 内积运算

FN 可通过同态乘法“ $\otimes$ ”进行运算。同态乘法“ $\otimes$ ”，即 $\text{ct} \otimes \text{ct}' = (\hat{c}_0, \dots, \hat{c}_{\eta+\xi})$ ，设置 x 为一个符号变量和 ct 为其中的元素， $\text{ct}' \leftarrow R_q[x]$ ，然后通过 $\sum_{i=0}^{u-1} v_i x^i \in R_q[x]$ 计算任何 $u \geq 1$ 的嵌入映射。因此，所有元素 $\hat{c}_i$ 的计算式为：

$$\sum_{i=0}^{\eta+\xi} \hat{c}_i x^i = \left(\sum_{i=0}^{\eta} c_i x^i \right) \cdot \left(\sum_{i=0}^{\xi} c'_i x^i \right) \in R_q[x] \quad (4)$$

$$\text{FN 计算 } \sum_{i=0}^{n_1} \tilde{c}_i x^i = \sum_{i=0}^{n_1} (\hat{c}_i x^i) \cdot (\hat{c}_i x^i) \in R_q[x],$$

得到密文内积的结果。

具体步骤如下：

1) FN 先消除掩码 $H(\text{Ts})^{p_i}$ ， $\hat{c}_{i,0} = c_{i,0} - H(\text{Ts})^{p_i}$ 。

2) 由 $\sum_{i=0}^{n_1} \tilde{c}_i x^i = \sum_{i=0}^{n_1} (\hat{c}_i x^i) \cdot (\hat{c}_i x^i) \in R_q[x]$ 得到

密文内积的结果。

3.4.2 方差运算

FN 可通过同态方差运算“ $\otimes_{\text{var}}$ ”来计算密文的方差。

首先，定义同态平均运算，对于一系列密文 $\text{cc}_1, \text{cc}_2, \dots, \text{cc}_n$ ，计算平均密文 $\overline{\text{cc}}$ ：

$$\overline{\text{cc}} = \left(\frac{1}{n} \sum_{i=1}^n c_{i,\eta_0}, \dots, \frac{1}{n} \sum_{i=1}^n c_{i,\eta_l} \right) \in R_q^{\max(\eta_l) + 1} \quad (5)$$

接着，定义同态方差运算 $\otimes_{\text{var}}$ ，对于每个密文 cc_i 与其对应的平均密文 $\overline{\text{cc}}$ 的差 $\Delta \text{cc}_i = \text{cc}_i \ominus \overline{\text{cc}}$ ，然后计算平方差 $\Delta \text{cc}_i^2 = \Delta \text{cc}_i \otimes \Delta \text{cc}_i$ ，并对所有平方差进行同态加法，最后除以 $n - 1$ ：

$$\text{cc}_{\text{variance}} = \left(\frac{1}{n-1} \sum_{i=1}^n \Delta \text{cc}_i^2 \right) \in R_q^{\max(\eta_l) + 1}$$

其中， n 是密文的数量， η_l 是每个密文 cc_i 的维度， $\ominus$ 表示同态减法， $\otimes$ 表示同态乘法。具体步骤如下。

1) FN 根据 $\hat{c}_1, \hat{c}_2, \dots, \hat{c}_{n_1}$, 计算平均密文 $\bar{c}$ 。

$$\bar{c} = \left(\frac{1}{n} \sum_{i=1}^{n_1} \hat{c}_{i, \eta_i}, \dots, \frac{1}{n} \sum_{i=1}^{n_1} \hat{c}_{i, \eta_i} \right) \in R_q^{\eta_{n_1} + 1} \quad (6)$$

2) 对于每个密文 c_i 与其对应的平均密文 $\bar{c}$ 的差 $\Delta c_i = c_i \ominus \bar{c}$ 。

3) 对所有平方差进行同态加法, 最后除以 $n - 1$:

$$cc_{\text{variance}} = \frac{1}{n - 1} \sum_{i=1}^{n_1} \Delta c_i^2 \in R_q^{\eta_{n_1} + 1} \quad (7)$$

其中, n 是密文的数量, η_i 是每个密文 c_i 的维度。

3.5 数据恢复阶段

在收到 FN 的聚合消息后, CC 通过式(8)计算总用电量, 而不会暴露单个客户的隐私。

$$M_{\text{sum}} = \text{Dec}(C_{\text{sum}}, \text{sk}) =$$

$$\begin{aligned} \sum_{i=0}^{n_1} (a_{i,0}u + \text{tg} + m_i + (a_1u + \text{tf})s) \bmod q &= \\ \sum_{i=0}^{n_1} (a_{i,0}ut + \text{tg} + m_i + a_{i,1}us + \text{tfs}) \bmod t &= \\ \sum_{i=0}^{n_1} (a_{i,0}u + m_i + a_{i,1}us) \bmod t = \sum_{i=0}^{n_1} m_i \bmod t \end{aligned} \quad (8)$$

其中, s 为 CC 私钥, CC 对得到的内积密文进行如下解密:

$$\begin{aligned} M = \sum_{i=0}^{n_1} \hat{C}_i \bmod q &= \sum_{i=0}^{n_1} ((a_{i,0}u + \text{tg} + m_i)^2 + \\ (a_1u + \text{tf})^2 s^2) \bmod q &= \sum_{i=0}^{n_1} (t\bar{e} + m_i^2) \bmod t = \\ \sum_{i=0}^{n_1} m_i^2 \bmod t \end{aligned}$$

其中, $\bar{e}$ 表示其余项。

对方差密文进行相同的解密步骤。

3.6 正确性分析

CC 接收到聚合密文后进行解密操作:

$$\begin{aligned} M &= \sum_{i=0}^{n_1} C_i \bmod q = \\ \sum_{i=0}^{n_1} (a_{i,0}u + \text{tg} + m_i + (a_1u + \text{tf})s) \bmod q &= \\ \sum_{i=0}^{n_1} (a_{i,0}ut + \text{tg} + m_i + a_{i,1}us + \text{tfs}) \bmod t &= \\ \sum_{i=0}^{n_1} (a_{i,0}u + m_i + a_{i,1}us) \bmod t = \sum_{i=0}^{n_1} m_i \bmod t \end{aligned} \quad (9)$$

得到聚合的正确数据, 解密成功。

4 安全性分析

4.1 方案的 IND-CPA 安全

4.1.1 IND-CPA 安全模型

下面通过挑战者和攻击者 $\mathcal{A}$ 之间的安全游戏来定义本文方案的 IND-CPA (indistinguishability under chosen plaintext attack) 安全性, 游戏共分为以下 3 个阶段。

1) 初始阶段: 挑战者将安全参数 λ 输入

$\text{KG}()$, 密钥生成算法 $\text{KG}(\lambda)$ 根据输入的安全参数 λ 输出公共参数 pp 。并将 pp 发送给攻击者 $\mathcal{A}$, 根据攻击者 $\mathcal{A}$ 是否知道用户的私钥信息, 将用户分为诚实用户和腐化用户。

2) 学习阶段: 攻击者 $\mathcal{A}$ 可以向挑战者发起任意多项式次以下查询。

私钥查询: 攻击者 $\mathcal{A}$ 向挑战者发起私钥查询, 若查询的用户为腐化用户, 挑战者返回给攻击者 $\mathcal{A}$ 相应的用户私钥 sk_i , 否则返回终止符 $\perp$ 。

加密查询: 攻击者 $\mathcal{A}$ 向挑战者发起明文 m_i 的加密查询, 挑战者利用 CC 公钥运行加密算法 $\text{Enc}()$ 得到加密密文 C_i , 并将其返回给攻击者 $\mathcal{A}$, 其中加密算法 $\text{Enc}(\text{pp}, m)$ 利用公共参数 pp 对明文 m 进行加密, 生成密文 C 。如果询问的明文是测试明文, 则挑战者忽略攻击者 $\mathcal{A}$ 的请求, 返回一个随机值。

3) 挑战阶段: 当攻击者 $\mathcal{A}$ 认为学习阶段可以结束时, 选择两个明文 (m_0, m_1) 和攻击目标用户 i^* (i^* 不能是腐化用户) 并提交给挑战者, 挑战者随机选择一个随机数 $d \in \{0, 1\}$, 然后返回用户 i^* 的加密密文 $C_{i^*} = \text{Enc}(\text{pk}_c, m_d)$ 给攻击者 $\mathcal{A}$ 作为挑战密文。

攻击者 $\mathcal{A}$ 在收到挑战密文之后可以继续发起学习阶段中的私钥查询、加密查询, 但不能针对用户 i^* , 挑战者的应答保持一致。

猜测: 攻击者 $\mathcal{A}$ 停止查询并输出一个猜测值 $d' \in \{0, 1\}$, 如果 $d' = d$ 则宣布 $\mathcal{A}$ 获胜, 其优势为:

$$\text{Adv}_{\Lambda}^{\text{IND-CPA}} = \left| \Pr[d' = d] - \frac{1}{2} \right| \quad (10)$$

如果对于任意多项式时间攻击者, 攻击者的优势是可忽略的, 则称基于 RLWE 的多维多函数方案是 IND-CPA 安全的。

4.1.2 IND-CPA 安全性证明

根据 Kerckhoffs 原则, 攻击者即使监听所有参与方之间的通信, 也无法通过这些信息来破解用户个人的耗电数据或者篡改传输的数据信息。所有窃听可获取的数据都是已经加密的数据, 攻击者无法从中获取实质性的数据。只有持有正确的密钥, 才能解密数据的真实内容。

定义 1 RLWE 假设：对于给定的参数 $(\kappa, N, q, t, \sigma)$ ，RLWE 假设 $E_{\kappa, N, q, t, \chi}$ 表明以下两个分布是不可区分的。

- 1) 从 $(R_q)^2$ 中均匀抽取的 (a_i, b_i) 。
- 2) 首先从 $D_{\mathbb{Z}^N, \sigma}$ 中均匀抽取一个秘密随机向量 $\mathbf{s}$ ，然后对于每个 i ，从 R_q 中均匀抽取 b_i ，从 $\mathbb{Z}_q^*$ 中选择 t (t 和 q 互质)，从 χ 中抽取 e_i ，并计算 $a_i = -(b_i \cdot \mathbf{s} + t \cdot e_i)$ 。

定理 1 假设存在一个 IND-CPA 攻击者 $\mathcal{A}$ 能够以 ε 的优势攻破本文方案，那么一定存在一个多项式时间算法 β 能够以 $\frac{\varepsilon}{2}$ 的优势解决 RLWE 问题。

证明 1) 初始化阶段：根据攻击者 $\mathcal{A}$ 是否知道某个用户的私钥信息，将用户分为诚实用户和腐化用户，算法 β 通过访问预言机 O_{sample} 随机均匀采样攻击者 $\mathcal{A}$ 查询的信息， O_{sample} 以 $\frac{1}{2}$ 的概率选中 $R_q \times R_q$ 上的随机均匀分布和上 RLWE 分布 $A_{s, \chi}$ 进行采样。

2) 学习阶段：攻击者 $\mathcal{A}$ 可向算法 β 进行任意多项式次以下查询。

私钥查询：攻击者 $\mathcal{A}$ 向算法 β 发起私钥查询，若查询的用户不是攻击目标用户 i^* ，算法 β 运行 KG 算法返回给攻击者相应的用户私钥 p_i ，否则返回终止符 $\perp$ 。

加密查询：攻击者 $\mathcal{A}$ 向算法 β 发起明文 m_i 到密文 C_i 的加密查询，算法 β 利用加密密钥 sk_e 运行 Enc 算法得到加密密文 C_i ，并将其返回给攻击者 $\mathcal{A}$ 。如果询问的明文是测试明文，则挑战者忽略 $\mathcal{A}$ 的请求，返回一个随机值。

3) 挑战阶段：当攻击者 $\mathcal{A}$ 认为学习阶段可以结束时，选择两个明文 (m_0, m_1) 和攻击目标用户 i^* (i^* 不能是腐化用户)，并提交给挑战者，挑战者随机选择一个随机数 $d \in \{0, 1\}$ ，访问预言机 O_{sample} 得到采样值 (a_0, a_1, C_{i^*}') ，计算

$$C_{i^*}' = (a_0 u + \text{tg} + m_i + H(\text{Ts})^{p_i}) \quad (11)$$

作为挑战密文发送给攻击者 $\mathcal{A}$ 。攻击者 $\mathcal{A}$ 在收到挑战密文后可以继续进行上述查询，但

不能针对用户 i^* ，算法 β 应答的方式与上阶段一致。

猜测：攻击者 $\mathcal{A}$ 在经过多次查询后停止查询，并输出一个猜测值 $d' \in \{0, 1\}$ ，如果 $d' = d$ ，则宣布攻击者 $\mathcal{A}$ 获胜。算法 β 输出 1 并判定 O_{sample} 采样的分布为 RLWE 分布 $A_{s, \chi}$ ，否则判定 O_{sample} 输出采样的分布为 $R_q \times R_q$ 上的随机均匀分布。RLWE 问题的概率分析如下。

1) 预言机 O_{sample} 采样 (a_0, a_1, C_{i^*}') 的分布为 RLWE 分布 $A_{s, \chi}$ ，采样值的形式为 (a_0, a_1, C_{i^*}') ，挑战密文 $C_{i^*}' = (a_0 u + \text{tg} + m_i + H(\text{Ts})^{p_i})$ ，此时攻击者 $\mathcal{A}$ 猜测正确概率为 $\frac{1}{2} + \varepsilon$ ，算法 β 判定 O_{sample} 采样的分布为 RLWE 分布 $A_{s, \chi}$ 的概率为 $\frac{1}{2} + \varepsilon$ 。因此，挑战密文在 R_q 上随机均匀分布且独立于随机值 d ，此时攻击者 $\mathcal{A}$ 猜测正确概率为 $\frac{1}{2}$ ，算法 β 判定 O_{sample} 采样的分布为 RLWE 分布 $A_{s, \chi}$ 的概率为 $\frac{1}{2}$ 。

综上，因为 O_{sample} 选中上述公布的概率均为 $\frac{1}{2}$ ，当攻击者 $\mathcal{A}$ 能够以不可忽略的优势 ε 攻破方案时，算法 β 能够对 RLWE 问题进行破解的概率为 $\Pr[\beta = \text{success}] = \frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \cdot \left(\frac{1}{2} + \varepsilon\right) = \frac{1}{2} + \frac{\varepsilon}{2}$ ，优势为 $\frac{\varepsilon}{2}$ 。

上述结果与 RLWE 问题的困难性假设相冲突，因为 ε 是非忽略的。因此，本文设计的密码系统在 RLWE 假设下对于 CPA 是语义安全的。

4.1.3 对比其他加密技术

在智能电网的数据安全方面，同态加密技术是确保用户隐私和数据完整性的关键。然而，随着量子计算的发展，传统加密技术（如 Paillier 和 BGN）面临着潜在的安全威胁。这些方案在经典计算环境下虽能提供一定程度的安全保障，但它们不具备抵抗量子攻击的能力。

而基于 RLWE 的加密方案，由于其算法的数学基础，被认为具有抗量子攻击的特性。

RLWE 方案的安全性建立在当前认为即使量子计算机也无法有效解决的数学问题之上,这使得它在后量子时代仍能保持数据的安全性和隐私性。

4.2 隐私性

本文方案保证了用户隐私数据能够得到有效保护。由于 SM 的私钥 p_i 是通过安全信道分发的,且仅用于加密用户数据,因此可以保证即使在 SM 被攻击者 $\mathcal{A}$ 攻破的情况下,私钥也不会泄露。加密函数 Enc 确保了数据在传输过程中的机密性,因为只有拥有正确私钥的实体才能解密信息。而且通过引入随机掩码因子 p_i ,每个用户的用电数据 m_i 在加密前都会与掩码 p_i 相结合,生成一个新的掩码值 c_i 。 p_i 是随机选取的,且每次通信都可能不同,增加了方案的安全性。

FN 在聚合数据时,仅对加密数据进行操作,不涉及解密,因此无法获取用户的私有信息。TA 在计算噪声和时间戳时,仅对时间戳进行哈希处理,不涉及用户数据的明文,因此无法从这一过程中获取隐私信息。FN 在聚合过程中,保持数据的加密状态,确保了数据在传输过程中不被篡改。

CC 通过其私钥对聚合数据进行解密,由于私钥的保密性,只有 CC 能够正确解密并获取总的统计量,这保证了消息的机密性和真实性。

本文方案的内积和方差运算均为一种同态加密操作,同态加密允许在加密数据上直接进行算术运算,运算结果解密后与在明文上直接进行相同运算的结果相同。因此,在内积计算中,即使数据 C_1 和 C_2 都是加密的,它们的同态乘法 $C_{\text{prod}} = C_1 \cdot C_2$ 仍然能够正确反映明文数据的内积。由于掩码的随机性和同态加密的数学性质,通过挑战-应答游戏,攻击者无法区分不同用户的加密数据,也无法从加密的内积结果中提取任何有关个别用户的信息。

5 性能评估

本节通过测量加密开销、解密开销、内积开销、聚合开销来评估方案的性能,同时也给

出通信开销。在进行实验分析时,代码基于 Helib 进行开发,仿真系统为 Intel(R)Core i5-8265U, 1.60 GHz CPU, 16 GB 内存的便携式计算机。SHE 系统使用密文长度为 2 048 bit;达到了可证明安全的安全参数。Paillier 同态密码系统使用 1 024 位素数 p 和 q ,以满足可证明安全参数,密文长度为 4 096 bit。使用 SHA-256 实现伪随机函数,选择的 Hmac 为 Javax.Crypto.Mac 库中的 HmacSHA1。在聚合周期内,每个电表耗电量为 1 000 之内的一个随机值,每隔 15 min 聚合一次。

5.1 计算开销比较

本文方案在于效率优于基于 Paillier 的加密方案,满足智能电网对数据传输的高效性的要求。在该方案中,SM 对数据进行加密以及完成内积操作,FN 对收到的数据进行聚合,而 CC 则完成解密任务。具体来说,SM 在加密过程中面临 1 次指数运算,主要发生在掩码生成中。FN 在聚合过程中主要进行加法运算,其执行次数与数据量成正比。而对于 FPDA^[4]、SEMDA^[1] 加密方案,在加密过程需要进行多次乘法和指数运算,其计算开销高于本文方案。

图 3 为加密时间对比,可以看出,本文方案在加密时间上的优势,随着用户数量的增加,本文方案的加密时间增长平缓,这得益于其优化的算法和高效的数据处理能力,加密过程中本文方案只涉及加法这一种操作,因此计算效率高。相比之下,FPDA^[3] 和 SEMDA^[4] 方案涉及复杂的乘法和指数运算,在用户数量增多时,加密时间开销显著上升,尤其在大规模数据处理时效率较低。本文方案在加密效率上优于 FPDA 和 SEMDA 方案,特别是在用户数量较大时,这种优势更为明显。这表明本文方案更适合于大规模的用户环境。

在 CC 端,解密操作相对简单,仅涉及一次取模和一次乘法运算,这对于具备较强计算能力的 CC 来说不是难题。智能电表作为资源受限的设备,减轻其计算压力是至关重要的,而 FN 则可以利用其计算能力来支持更复杂的操作。

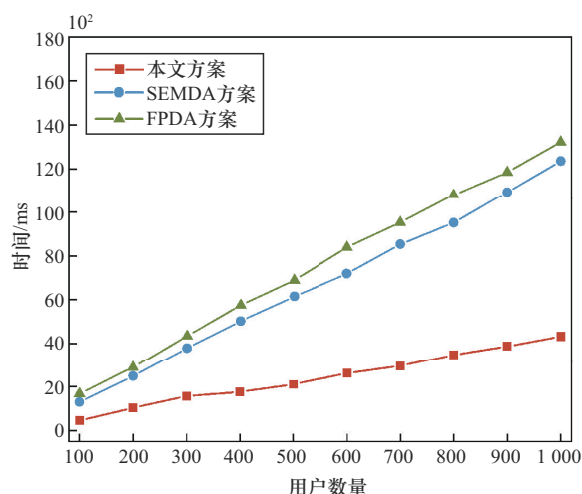


图3 加密时间对比

Figure 3 Comparison of encryption time

图4的解密时间对比中可以看出本文方案在不同用户数量上均展现出了较低的解密时间，主要是两个方案解密过程仅涉及一次取模和一次乘法运算，从而减少了计算资源的需求，适合资源受限的环境，如智能电表。相比之下，SEMDA方案解密时间稍长，尤其是在用户数量较大时，这一差异更加明显。这是由于SEMDA方案在解密过程中需要执行更复杂的计算。总的来说，本文方案在解密效率上明显优于SEMDA方案。这一结果强调了在设计智能电网数据聚合方案时考虑计算效率的重要性，尤其是在资源受限的智能电表上。

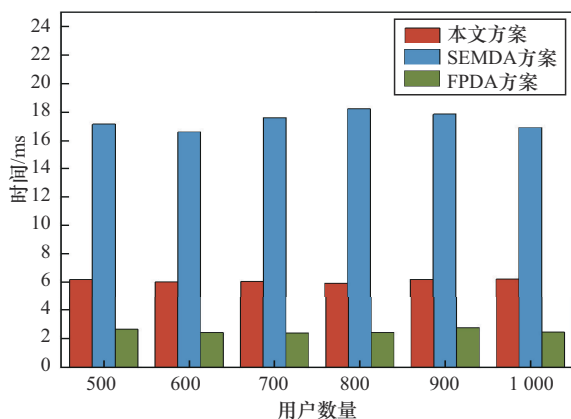


图4 解密时间对比

Figure 4 Comparison of decryption time

综上所述，本文方案实现了在保证数据安全的前提下，对计算资源的合理分配和利用，使得

整个系统更适用于资源受限的智能电网环境。

聚合时间比较见图5。在不同用户数量下，本文方案与基于Paillier加密的FPDA和SEMDA方案在数据聚合时间上的性能差异较大。随着用户数量的增加，本文方案始终保持较低的聚合时间，这凸显了SHE在同态加法运算中的高效率，尤其是在大规模数据处理时的优势。相比之下，FPDA和SEMDA方案的聚合时间随着用户数量的增多而增长，整体高于本文方案，这表明Paillier加密在聚合效率上低于SHE方案。因此，这一结果强调了在智能电网等需要处理大量用户数据的环境中，本文方案的适用性和效率优势。

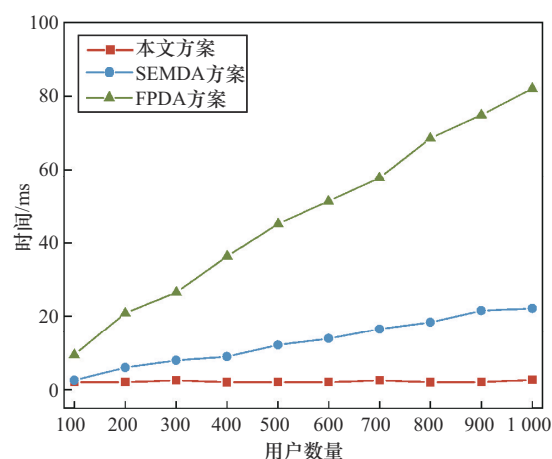


图5 聚合时间对比

Figure 5 Comparison of aggregation time

如图6所示，对本文方案、FPDA和SEMDA方案在不同用户数量下的总时间开销进行对比分析。在用户数量增加的过程中，本文方案的运行时间缓慢增加，增长幅度显著低于SEMDA与FPDA方案。进一步分析表明，本文方案的总运行时间约为SEMDA方案的39.47%和FPDA方案的27.85%。这些数据表明，无论用户数量如何变化，本文方案始终保持更低的运行时间与更稳定的性能表现，相较对比方案展现出更优的时间效率优势。

图7为本文方案各类时间开销对比，可以看到，在系统实现过程中，主要的时间开销在用户加密过程，而用时最少的是解密过程，这与边缘计算特性相一致。因此需要通过减轻控制中心的负担，提高整个系统的效率。

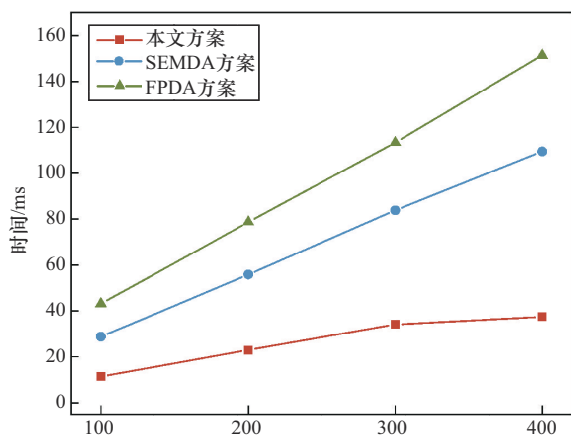


图6 方案总时间开销对比

Figure 6 Comparison of schemes all time overhead

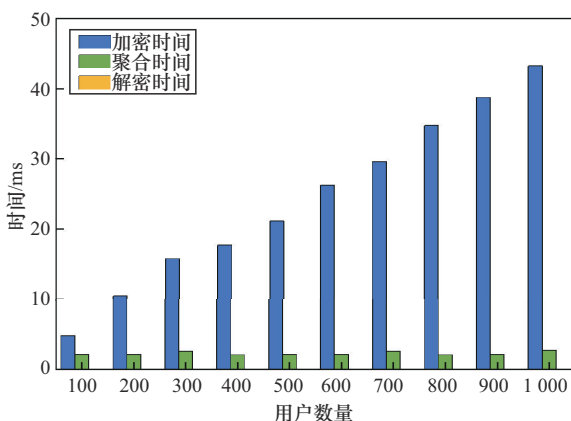


图7 方案各类时间开销对比

Figure 7 Comparison of scheme time overhead

5.2 通信开销和存储开销

本节对智能电表、雾节点以及控制中心之间的通信和存储负荷与现有方案进行了详细评估。设定用户数量为 n ，掉线率为 m 。

本文方案每个SM生成的数据包大小为2 048 bit，包括序列号、时间戳、属性对、标签和MAC值，每个用户每天发送96次，因此每天的通信量为 $96 \times 2\,048$ bit，FPDA方案在生成SM数据包后，还需要生成5份掩码份额，因此每天的通信量为 $(1-m) \times 5 \times 96 \times 2\,048$ bit。对于本文方案，FN聚合后发送给CC的数据量为 $96 \times 2\,048 \times n$ bit，FPDA方案需要将掉线SM的掩码重构，因此数据量为 $5 \times 96 \times (1-m) \times 2\,048 \times n$ bit。可以看出FPDA方案的通信开销远高于本文方案。

在存储需求方面，本文方案中SM需要保存

自己的私钥2 048 bit和每次生成掩码2 048 bit，共4 B，而FPDA方案还需要对SM掩码生成份额，共12 B。本文方案中FN则需存储密钥 $(n+1) \times 2\,048$ bit， n 为SM数量，总计 $2(n+1)$ B，对于FPDA方案，FN还需要存储重构的掩码份额 $2\,048 \times mn$ bit，总计 $2(n+1+mn)$ B。本文方案CC仅需存储私钥2 048 bit，而FPDA方案需要存储所有SM的掩码信息，总计 $2(n+1)$ B。通信开销和存储开销见表2。

表2 通信开销和存储开销

Table 2 Communication overhead and storage overhead

开销	方案	SM/B	FN/B	CC/B
通信开销	FPDA	$960(1-m)$	$960(1-m)n$	—
	本文方案	192	$192n$	—
存储开销	FPDA	12	$2(n+1+mn)$	$2(n+1)$
	本文方案	4	$2(n+1)$	2

分析结果表明，在聚合过程中主要的通信开销在SM和FN之间，然而，由于SM和FN只距离1跳，因此，产生的通信开销是可以接受的。此外，FN承担了较多的存储任务，这符合最初的设计方案，即通过FN的存储和计算资源来缓解CC的压力，有效利用雾计算的优势，体现了雾计算分散计算和存储的特点。与现有方案对比，本文方案具有较低的通信开销与存储开销，使得数据处理更加高效。

6 结束语

为了在智能电网领域中实现数据隐私保护与高效聚合运算的双重目标，本文提出了一种基于RLWE的多维多函数数据聚合方案。通过对智能电表数据的多维化和多函数操作，本文方案不仅提高了电网数据分析效率，还强化了数据的机密性和利用率。方案基于SHE加密系统，可提供数据的多函数操作，如聚合、内积和方差，进一步提高了数据分析的利用率。在安全性方面，本文方案证明了其能够抵抗选择明文攻击，并通过实验验证了在计算效率方面优于基于Paillier加密的方案。本文未来的工作

将集中在通过算法的并行化处理和硬件加速技术来进一步提升方案的计算速度和处理能力，同时在不同规模的智能电网中进行广泛测试，以验证其可扩展性和实用性。此外，还将探索本文方案在医疗健康数据分析、金融交易记录汇总等其他数据密集型应用场景中的潜在应用，以期为更多领域提供数据安全和隐私保护的技术支持。

参考文献：

- [1] SHANG S, LI X, GU K, et al. A robust privacy-preserving data aggregation scheme for edge-supported IIoT[J]. IEEE Transactions on Industrial Informatics, 2024, 20(3): 4305-4316.
- [2] CHEN Y, YANG S, MARTÍNEZ-ORTEGA J F, et al. A resilient group-based multisubset data aggregation scheme for smart grid[J]. IEEE Internet of Things Journal, 2023, 10(15): 13649-13661.
- [3] SONG Z, ZHONG W, ZHOU T, et al. SEMDA: Secure and efficient multidimensional data aggregation in smart grid without a trusted third party[J]. Security and Communication Networks, 2023, 2023: 1-13.
- [4] WU L, XU M, FU S, et al. FPDA: Fault-tolerant and privacy-enhanced data aggregation scheme in fog-assisted smart grid[J]. IEEE Internet of Things Journal, 2022, 9(7): 5254-5265.
- [5] WANG X, LIU Y, CHOO K K R. Fault-tolerant multisubset aggregation scheme for smart grid[J]. IEEE Transactions on Industrial Informatics, 2021, 17(6): 4065-4072.
- [6] KHAN H M, KHAN A, JABEEN F, et al. Privacy preserving data aggregation with fault tolerance in fog-enabled smart grids[J]. Sustainable Cities and Society, 2021, 64: 102522.
- [7] XUE K, ZHU B, YANG Q, et al. An efficient and robust data aggregation scheme without a trusted authority for smart grid[J]. IEEE Internet of Things Journal, 2020, 7(3): 1949-1959.
- [8] LYU L, NANDAKUMAR K, RUBINSTEIN B, et al. PPFA: Privacy preserving fog-enabled aggregation in smart grid[J]. IEEE Transactions on Industrial Informatics, 2018, 14(8): 3733-3744.
- [9] ABDALLAH A, SHEN X S. A lightweight lattice-based homomorphic privacy-preserving data aggregation scheme for smart grid[J]. IEEE Transactions on Smart Grid, 2018, 9(1): 396-405.
- [10] WANG H, WANG Z, DOMINGO-FERRER J. Anonymous and secure aggregation scheme in fog-based public cloud computing[J]. Future Generation Computer Systems, 2018, 78: 712-719.
- [11] LIU J N, WENG J, YANG A, et al. Enabling efficient and privacy-preserving aggregation communication and function query for fog computing-based smart grid[J]. IEEE Transactions on Smart Grid, 2020, 11(1): 247-257.
- [12] CHEN D, ZHOU T, LIU W, et al. MDA-FLH: Multidimensional data aggregation scheme with fine-grained linear homomorphism for smart grid[J]. IEEE Internet of Things Journal, 2024, 11(2): 3524-3538.
- [13] SHEN H, ZHANG M, SHEN J. Efficient privacy-preserving cube-data aggregation scheme for smart grids[J]. IEEE Transactions on Information Forensics and Security, 2017, 12(6): 1369-1381.
- [14] ZENG Z, WANG X, LIU Y, et al. MSDA: multi-subset data aggregation scheme without trusted third party[J]. Frontiers of Computer Science, 2022, 16(1): 161808.
- [15] LIU Y, GUO W, FAN C I, et al. A practical privacy-preserving data aggregation (3PDA) scheme for smart grid[J]. IEEE Transactions on Industrial Informatics, 2019, 15(3): 1767-1774.
- [16] MERAD-BOUDIA O R, SENOUCI S M. An efficient and secure multidimensional data aggregation for fog-computing-based smart grid[J]. IEEE Internet of Things Journal, 2021, 8(8): 6143-6153.
- [17] ZHAO S, LI F, LI H, et al. Smart and practical privacy-preserving data aggregation for fog-based smart grids[J]. IEEE Transactions on Information Forensics and Security, 2021, 16: 521-536.
- [18] SALEEM A, KHAN A, MALIK S U R, et al. FESDA: Fog-enabled secure data aggregation in smart grid IoT network[J]. IEEE Internet of Things Journal, 2020, 7(7): 6132-6142.
- [19] CHEN Y, MARTINEZ-ORTEGA J F, LOPEZ L, et al. A dynamic membership group-based multiple-data aggregation scheme for smart grid[J]. IEEE Internet of Things Journal, 2021, 8(15): 12360-12374.

作者简介



王圆圆（2001-），女，江西上饶人，武警工程大学硕士生，主要研究方向为密码学和机器学习。



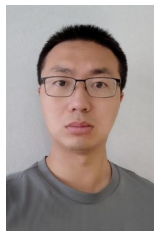
韩益亮（1977-），男，甘肃兰州人，武警工程大学教授、博士生导师，主要研究方向为密码学和机密计算。



吴立强（1986-），男，陕西蓝田人，武警工程大学讲师，主要研究方向密码学和信息安全。



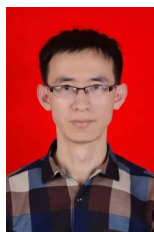
李鱼 (1995-), 男, 重庆丰都人, 武警工程大学博士生, 主要研究方向为密码学和机器学习。



吴旭光 (1986-), 男, 河南商丘人, 武警工程大学博士生、讲师, 主要研究方向为密码学和信息安全。



姚武军 (1986-), 男, 湖北随州人, 武警工程大学博士生、讲师, 主要研究方向为密码学和信息安全。



朱率率 (1985-), 男, 山东淄博人, 武警工程大学教授、硕士生导师, 主要研究方向为后量子密码和密码分析。