doi:10.11835/j.issn.1000-582X.2023.219
引用格式:庞博,张凌浩,梁晖辉,等.面向智能电网的数据聚合隐私保护方案[J].重庆大学学报,2025,48(3):38-49.

面向智能电网的数据聚合隐私保护方案

庞 博[1],张凌浩[1],梁晖辉[1],常政威[1],刘泽伟[2],胡春强[2]
(1.国网四川省电力公司电力科学研究院,成都 401331;2.重庆大学 大数据与软件学院,重庆 400030)

摘要:数据聚合是智能电网通信中的一项关键技术,能够以高效节能的方式收集用户用电数据。随着智能电表的大规模部署,这引发了诸多用户隐私方面的担忧,例如对个人生活习惯的监测。提出了一种高效且保护隐私的数据聚合方案(efficient and privacy-preserving data aggregation,EPPDA)。首先,提出基于区块链的智能电网4层架构支持电力数据聚合。在架构的采集层中,改进了基础的Boneh-Goh-Nissim加密系统,使其更适合于电网隐私保护场景。在架构的平台层,利用区块链的防篡改特性对聚合数据进行有效的存储及查询。性能分析结果表明:提出的EPPDA可以满足智能电网系统的几种隐私特性。实验数据表明:EPPDA在保证数据隐私和安全的条件下降低了计算与通信成本,提高了整个方案的效率。

关键词:智能电网;隐私保护;数据聚合;Boneh-Goh-Nissim加密;区块链
中图分类号:TP309 **文献标志码:**A **文章编号:**1000-582X(2025)03-038-12

A novel privacy-preserving data aggregation scheme for smart grids

PANG Bo[1], *ZHANG Linghao*[1], *LIANG Huihui*[1], *CHANG Zhengwei*[1], *LIU Zewei*[2], *HU Chunqiang*[2]
(1. Electric Power Research Institute of State Grid Sichuan Electric Power Company, Chengdu 401331, P. R. China; 2. School of Big Data & Software Engineering, Chongqing University, Chongqing 400030, P. R. China)

Abstract: Data aggregation is a key technology in smart grid communication, enabling efficient collection of essential data while optimizing energy usage. However, the large-scale deployment of smart meters raises significant privacy concerns, as it may expose users' lifestyle habits. To address this issue, this paper proposes an efficient and privacy-preserving data aggregation(EPPDA) scheme for IoT-enabled smart grid, leveraging smart contracts. First, a four-layer blockchain-based architecture is introduced to facilitate secure data aggregation. At the collection layer, the Boneh-Goh-Nissim system is improved to better suit privacy protection scenarios in smart grids. At the platform layer, blockchain's tamper-proof features are utilized for secure storage and efficient querying of aggregated data. Performance analysis indicates that the proposed EPPDA satisfies key privacy requirements of smart grid systems. Finally, experimental results show that the proposed EPPDA reduces computational and communication costs while improving overall system efficiency.

Keywords: smart grid; privacy-preserving; data aggregation; Boneh-Goh-Nissim encryption; blockchain

收稿日期:2023-06-07 **网络出版日期:**2023-12-05
基金项目:国网四川省电力公司科技项目(SGSCDK00LYJS2200130)。
Supported by State Grid Sichuan Power Company Science and Technology Project(SGSCDK00LYJS2200130).
作者简介:庞博(1994—),男,硕士,主要从事隐私计算应用研究,(E-mail)pang-bo@outlook.com。
通信作者:胡春强,男,教授,博士生导师,(E-mail) chu@cqu.edu.cn。

传统电力系统一般由发电、输电、配电和电力用户组成[1]。发电厂提供的电力通常是超高电压,通过低压配电网将高压转换为低压,才能供应给最终用户。然而,由于传统电力系统各个环节间信息不流通,电力设备部署成本很高。一旦问题出现,则需大量的故障排除工作,修复成本极高。为此,近十年来,众多国家持续致力于发展下一代电网——智能电网[2]。与传统电网相比,智能电网将数据通信分析、精确控制、传感等多种技术集成到传统电网系统中,大大增强了电网企业感知用户端用电量情况的能力,为企业提高用电数据应用效率和提升服务质量提供支持[3]。

智能电表(smart meter,SM)是智能电网(smart grid,SG)的重要组成部分,用于测量、收集和传输分布式用户的能耗信息[4]。SM的工作是收集家庭区域网络(home area network,HAN)中所有电子设备的用电数据。同一地理区域中的用电消费者集合成一个邻域网络(neighbor area network,NAN),每个NAN都有一个本地聚合器。智能电表通过无线通信技术定期向NAN中的聚合器发送家庭用户的耗电量信息[5]。聚合器的作用是聚合来自NAN的测量数据,对其进行简单处理后转发到控制中心服务器进行进一步的分析和处理。例如,中心服务器做出实时定价决策和检测电力欺诈等。以上描述的整个过程即数据聚合的完整步骤[6]。

电力数据的聚合使访问方(如电网控制中心)能够及时了解客户端的用电量,并采取下一步措施(调度或定价等)。通过分析智能电表数据(例如煤气、水、电的消耗)可以很容易地感知居民的行为[7]。例如,一个家庭的日用电量异常低可能表明用户可能不在家。智能电表测量已构成了严重的隐私问题,因此需保护此类隐私敏感信息,防止未经授权用户访问。此外,在智能电网中,网络安全同样至关重要。在网络上传输的所有数据都必须经过身份验证,防止恶意修改。文中考虑了智能电表用户端的计算和通信能力受限,探讨一种新颖的智能电网隐私保护数据聚合方案,在保证用电数据安全和隐私的同时,具有较小的计算开销和通信成本。

1　相关工作

隐私保护数据聚合(privacy-preserving data aggregation,PPDA)是一种能够在不泄露任何个人身份或敏感信息的情况下对数据进行聚合和分析的数据隐私保护技术。该技术能在对用户数据进行有效保护的同时,能够降低数据采集和传输过程中的通信成本,因此吸引了研究者的广泛关注,并应用于智能电网中端到端的数据传输安全与隐私保护的研究领域。根据方案技术特点的不同,现有的智能电网PPDA方案可分为3大类:基于密码学的数据聚合方案、基于掩码屏蔽的数据聚合方案和基于差分隐私的数据聚合方案。

1.1　基于密码学的数据聚合方案

Yu等[8]提出了基于环签名的方案对用户的身份进行隐藏,使得攻击者无法将用电数据与住宅用户进行关联,然而该算法的计算成本随着智能电表数量的增长呈线性增长,无法适应大规模的智能电表部署。Diao等[9]使用匿名技术在通信过程中为用户生成假名实现用户隐私保护,同时该方案还支持控制中心对匿名用户的消息进行合法性验证。此外,随着同态加密算法的提出与发展,其具备密文可计算的特点被研究者所发现并成功应用于智能电网隐私保护数据聚合方案的设计中。其中,应用最为广泛的有Paillier同态加密算法、Boneh-Goh-Nissim同态加密算法(BGN)。Wen等[10]将属性决策树与Pailler同态加密算法相结合,提出了一种基于属性决策树的数据聚合方案,能够实现聚合过程的细粒度的访问控制。Lu等[11]结合Paillier同态加密和单向哈希链技术设计了一种3层架构的智能电网数据汇聚方案,并引入区块链对中间结果进行存证。Mohammadali等[12]提出了一种具有同态性质的隐私保护协议,该方案支持多类别数据聚合,并为边缘服务器和控制中心提供了批量验证的能力。虽然基于Pailler同态加密在一定程度上较好地保护了用户隐私和数据的安全,但是执行该算法消耗的计算开销是智能电表设备所无法承担的,并且由于电力数据收集频度高、时效性强,因此基于Pailler同态加密的隐私保护数据聚合方案的实行性较低。为解决该问题,He等[13]基于BGN公钥加密系统提出了一种隐私保护聚合方案,然而该方案需要高成本双线性配对操作,严重限制了该方案的实用性。Zhang等[14]基于改进的BGN密码,提出了一种在边缘辅助智能电网中具有轻量级验证的密钥泄漏弹性加密数据聚合方案。该方案能有效地检查聚合数据的完整性,获取特定区域的统计数据。随着

计算成本的降低，使得PPDA的功能得到了进一步丰富。为满足既能抵御恶意数据挖掘攻击，又能获得准确的融合结果的需求，Shen等[15]提出了一种有效抵抗恶意数据挖掘攻击的智能电网聚合方案。该方案给出了一种判断恶意数据挖掘攻击是否发生的方法，如果在某个时刻发现此攻击，相应时间内的所有计量数据就会被丢弃。此外，该方案利用Paillier加密和BLS (Boneh-Lynn-Shacham)签名来保证接收到的数据是有效的，且来自合法实体。

1.2 基于掩码屏蔽的数据聚合方案

Bohli等[16]提出了简易的加法掩码数据聚合方案，通过在用电数据中加入从特定数据分布中选择的随机数实现对用户数据的保护。为增强其安全性，Kursawe等[17]提出基于双线性映射的密钥协商方案，该方案允许用户以密钥协商的方式生成用于保护数据隐私的安全掩码，却忽略了智能电表有限的计算能力无法承担DDH密钥协商和双线性映射所产生的计算开销。为了解决该问题，在不需要进行大量计算的前提下Danezis等[18]和Knirsch等[19]结合秘密共享、多方安全计算等方式安全快速地生成掩码。此外，生成掩码的方式还包括基于公钥的掩码生成和基于聚合数的掩码生成。

在大多数掩蔽方案中[20-23]，屏蔽值通常由可信任的第三方生成和分发，另外，第三方也会协助控制中心执行用户注册和撤销步骤。第三方的存在为系统的安全性带来了额外的威胁，使得协议的执行和实用化推广带来不便。为了解决第三方存在所引发的安全问题，Xue等[24]提出了一些没有任何可信任机构的数据聚合的屏蔽值方法，该方案虽然减少了第三方的干扰，但缺少对数据真实性和完整性验证。Zhou等[25]提出基于任意单向陷门置换（one-way trapdoor permutation，OWTP）的高效多方数据聚合方案，该方案使用OWTP对掩码的安全进行保护。

1.3 基于差分隐私的数据聚合方案

差分隐私作为主流的隐私保护技术，通过向数据中添加服从特定分布的噪声的方法对用户数据隐私进行保护，该技术可以视为一种特殊的掩码方案。Shi等[26]通过在用户的用电数据中加入给定分布（如拉普拉斯分布）的噪声，抵御攻击者对用户数据实施的差分攻击。Baloglu等[27]则将高斯噪声加入原始数据并对数据进行同态加密，对数据进行双重保护，该方案能够抵抗的攻击也比其他方案更完备。Zheng等[28]提出了保护电力数据的用电行为模式，即用户用电的时间行为。在智能电表数据中，通过Fisher-Yates随机置乱算法对原始数据的时序进行扰动，将电量数据的测量时刻与发布时刻分离，能够有效地破坏原始数据中的负荷印记等与负载设备运行状况有关的属性信息。经过时序扰动后，智能电表发布与真实电量消耗情况有差别的电能量数据序列，从而达到对用户用电行为模式进行隐私保护目的。研究者们根据差分隐私对数据处理的方式不同，进一步提出了基于中心化差分隐私和基于本地化差分隐私的隐私保护数据方案。

尽管如此，在上述的聚合方案中，例如基于Paillier和基于掩码值的聚合方案，智能电表由于需要承担较为复杂的计算过程和频繁的通信，给电表造成了沉重的负担，聚合效率低下。鉴于此，笔者以BGN同态加密系统为基础，提出4层架构。通过在BGN密码系统中引入秘密共享技术，使它更加适合智能电网的聚合场景。

2 方案模型及目标

2.1 方案模型

智能电网通信的EPPDA模型包括初始层、采集层、聚合层、平台层4层架构，如图1所示。

可信授权机构（trust authority，TA）：它由信任的政府机构组成。它负责在初始化阶段引导整个系统，为通信中涉及的每个实体生成和发布必要的公共和私有参数。此外，初始化阶段结束后，TA将离线，不直接参与用户的数据上传。

采集层：安装在每个居民住宅的SM构成了采集层。根据地理位置的不同，每个住宅中的所有智能电表组成了一个NAN。智能电表定期收集用户的实时耗电量数据（如每15 min），并通过NAN中的无线通信技术将数据发送到上层聚合层进行处理。考虑智能电表的计算能力有限，因此本方案在设计时应最小化电表端的运算操作。

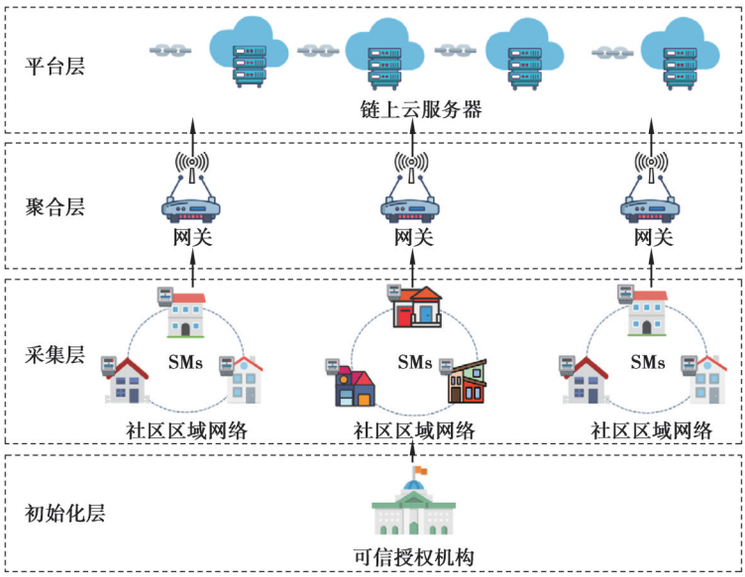

图 1　EPPDA 模型图

Fig. 1　The model of EPPDA

聚合层:该层的目的是将特定区域的聚合结果安全地上传到平台层,同时保护用户个人数据的私密性。每个NAN都有一个网关(gateway,GW),它是一个诚实但令人好奇的实体。它负责聚合NAN中的所有用电数据,并在聚合后将其报告给平台层。不过,它也对个别用户的用电数据感到好奇。

平台层:由链上的云服务器 $C_S=\{C_{S1},C_{S2},\cdots,C_{Sk}\}$ 组成。为了实现工作负载共享和容错,需要多个云服务器并结合秘密共享技术实现电力数据的聚合。作为一个具备强大计算能力的实体,云服务器被普遍认为是诚实且具有探索性的。云服务器主要承担着汇总来自数据聚合器的所有聚合数据,并进行总耗电量数据的计算任务。此外,它对个人用户的用电数据也表现出浓厚的兴趣。然而,不可忽视的是,潜在的攻击者有可能对某些云服务器造成破坏或使其瘫痪。不过,由于 C_S 的每个成员都是一个强大的实体,所以攻击者即使是破坏单个云服务器的代价也很高。因此,假设强大的攻击者只能破坏少数的云服务器,即不超过 $d=\lceil k/2\rceil-1$ 的云服务器。聚合完成后,云服务器将聚合结果和一些必要的信息存储在区块链上,可供查询[29]。

2.2　攻击模型

监听攻击[30]:在传输过程中,存在攻击者通过通信渠道窃听用户数据的可能,从而侵犯用户隐私。

半诚实攻击:除授权机构和智能电表外,所有其他参与者都是半诚实地遵守。也就是说,他们会遵循协议去做事情,但他们会尝试各种方法去寻找和推断用户的私有数据,从而造成隐私侵犯。

退出攻击:攻击者很可能通过破坏小于 $d=\lceil k/2\rceil-1$ 的云服务器来摧毁整个数据聚合系统。

2.3　设计目标

结合上述系统模型和攻击模型,文中的设计目标为区块链辅助下基于4层架构的智能电网隐私保护数据聚合方案,实现以下设计目标。

隐私保护:保护用户的数据隐私是本方案的主要特性之一。外部攻击者可以窃听用户的通信通道,但它不能透露用户的私有使用数据。

认证和完整性:为了确保接收到的报表是由合法用户生成的,保证在传输过程中不被恶意篡改,提出的方案应该提供认证和完整性保护。

容错能力:即使存在 d 个云服务器出现故障的情况或者被攻击者攻击,系统仍然可以有效和高效地聚合用电数据。

效率:考虑到智能电表计算资源有限,数据采集频繁,所提出的方案必须满足低计算和低通信开销的要求。

3 EPPDA方案描述

智能电网通信的EPPDA方案：在初始化层，TA需要经历“系统初始化”过程来初始化整个系统；在采集层，智能电表进行“用户报告生成”过程，生成加密的电力数据并转发给网关；在聚合层，网关进行“隐私保护的报告聚合”过程，负责聚合接收到的仪表数据并将其转发到云服务器；在平台层，云服务器进行“安全的数据解读”过程，解密并存储从网关接收到的聚合数据。

3.1 系统初始化

初始化阶段，由TA来引导整个聚合系统。首先，给定安全参数τ，TA运行Gen(τ)来获取三元组(p,q,G)，其中 G是生成元为 g的乘法循环群。随后，TA利用Boneh-Goh-Nissim密码系统生成元组(n,G,g,h)，其中$n = pq$,$h = g^q$是p阶群G子群的随机生成元，$g \in G$是群的随机生成元。其次，选取一个单向哈希函数H:$\{0,1\}^* \rightarrow Z_n^*$。之后，TA发布($n$,$G$,$g$,$h$,$H$)作为公钥，并将$K_S = p$作为私钥安全地存储的TA中。最后，TA需要按照以下步骤为用户$U = \{U_1,U_2,\cdots,U_n\}$，网关$W = \{W_1,W_2,\cdots,W_m\}$和云服务器$C_S = \{C_{S1},C_{S2},\cdots,C_{Sk}\}$分配公私密钥。

步骤1：对于HAN中的每一个用户$U_i \in U$，TA选取一个随机数$k_{u_i} \in Z_n^*$，计算$Y_{u_i} = g^{k_{u_i}}$，并将k_{u_i}和Y_{u_i}作为公私密钥发布给指定用户$U_i \in U$。

步骤2：对于每一个网关$W_i \in W$，TA选取一个随机数$k_{w_i} \in Z_n^*$，计算$Y_{w_i} = g^{k_{w_i}}$，并将k_{w_i}和Y_{w_i}作为公私密钥发布给指定网关$W_i \in W$。

步骤3：对于平台层的每一个云服务器$C_{Si} \in C_S$，TA运行Shamir秘密共享协议中的分片算法SS.share，将系统的私钥$K_S = p$分离成k份。TA首先随机生成一个d阶多项式函数$G(x) = p + a_1x + \cdots + a_dx_d$，其中$a_i \in Z_n^*(i = 1,2,\cdots,d)$。计算$G(i)$与$Y_{s_i} = g^{G(i)}$的值，并将$G(i)$与$Y_{s_i}$作为公私钥发布给指定云服务器$C_{Si} \in C_S$。

3.2 用户报告生成

假设用户的报告时间点定义为$T = \{t_1,t_2,\cdots,t_{\max}\}$。为了在每个固定时间点报告住宅用户的用电数据，每个用户$U_i \in U$在时间点$t_\gamma \in T$收集其使用数据$m_{i,\gamma} \in \{0,1,\cdots,i\}$，同时执行以下步骤。

步骤1：U_i计算当前时间点t_γ上的哈希值$\theta_{i,\gamma} = H(t_\gamma)$。

步骤2：U_i产生一个随机数$r_{i,\gamma}$作为盲因子，进行计算$C_{i,\gamma} = g^{m_{i,\gamma}} \cdot h^{r_{i,\gamma} \cdot \theta_{i,\gamma}}$。

步骤3：U_i选择一个随机数$\beta_{u_i} \in Z_n^*$，计算$R_{u_i} = g^{\beta_{u_i}}$和$h_{u_i} = H(C_{i,\gamma}||ID_{u_i}||t_\gamma)$。基于此，$U_i$对加密数据$C_{i,\gamma}$进行签名

$$S_{U_i} = \beta_{U_i} + h_{U_i} \cdot k_{U_i}。$$

步骤4：U_i通过无线通信（如WiFi）向GW报告($C_{i\gamma}||ID_{U_i}||t_\gamma||R_{U_i}||S_{U_i}$)。其中，$ID_{u_i}$是一个实体的唯一标识符。

3.3 隐私保护的报表聚合

在接收到w个加密用电加密数据($C_{i,\gamma}||ID_{U_i}||t_\gamma||R_{U_i}||s_{U_i}$)后，GW主要负责验证数据的有效性，并聚合同一邻居区域网络NAN中的所有数据，最后将其转发到平台层服务器。实施步骤如下。

步骤1：GW$_i$判断收到的加密数据是否满足$|t_{\gamma'} - t_\gamma| \leqslant \Delta t$，其中$t_{\gamma'}$是当前时间戳，$\Delta t$是预定义的阈值。如果满足，则GW$_i$进行判断签名$S_{U_i}$是否满足

$$g^{S_{U_i}} \stackrel{?}{=} R_{U_i} \cdot Y_{U_i}^{H(C_{i,\gamma}||ID_{U_i}||t_\gamma)}。 \tag{1}$$

如果它满足式(1)，则接受签名S_{U_i}；否则拒绝接受该用电加密数据。

步骤2：GW$_i$利用用户的加密数据$C_{i,\gamma}$计算聚合后的总用电加密数据A_γ为

$$A_\gamma = \prod_{i=1}^{w} C_{i,\gamma} = \prod_{i=1}^{w} (g^{m_{i,\gamma}} \cdot h^{\theta_{i,\gamma} * r_{i,\gamma}}) = g^{\prod_{i=1}^{w} m_{i,\gamma}} \cdot h^{Z_\gamma}。$$

式中，$Z_\gamma = \sum_{i=1}^{w} (\theta_{i,\gamma} \cdot r_{i,\gamma})$。

步骤 3:U_i选择一个随机数$\beta_{W_i} \in Z_n^*$,计算$R_{W_i} = g^{\beta_{w_i}}$和$h_{W_i} = H(C_{i,\gamma} \| ID_{w_i} \| t_\gamma)$。基于此,$\mathrm{GW}_i$对聚合后的加密数据$A_\gamma$进行签名

$$S_{W_i} = \beta_{W_i} + h_{W_i} \cdot k_{W_i}。$$

步骤 4:GW_i通过无线通信向 CS 报告$(C_{i,\gamma} \| ID_{W_i} \| t_\gamma \| R_{W_i} \| S_{W_i})$。其中$ID_{W_i}$是一个实体的唯一标识符。

3.4　安全的数据解读

当接收到相应的聚合数据A_γ时,平台层的云服务器$C_S = \{C_{S1}, C_{S2}, \cdots, C_{Sk}\}$主要负责在不暴露用户的隐私的前提下高效地计算所需的统计信息。最后,云服务器发起链上交易将计算出的信息保存在区块链上,供实体查询。

步骤 1:随机选择$(d+1)$个云服务器$S \subset C_S$来检查接受到数据是否满足$|t_{\gamma'} - t_\gamma| \leqslant \Delta t$,其中$t_{\gamma'}$是当前时间戳,$\Delta t$是预定义的阈值。接着,进一步判断签名$S_{W_i}$是否满足

$$g^{S_{W_i}} \stackrel{?}{=} R_{W_i} \cdot Y_{W_i}^{H(C_{i,\gamma} \| ID_{W_i} \| t_\gamma)}。$$

如果它满足上述等式,则接受签名S_{W_i};否则,拒绝接受该用电加密数据。

步骤 2:$(d+1)$个云服务器解密聚合的数据。根据 Shamir 秘密共享协议中的重构算法 SS.recon 来恢复秘密值$SK = p$。每个云服务器$S_j \in S$,首先计算

$$\beta_j = \prod_{i \subset S, i \neq j} \frac{i}{i-j},$$

然后产生

$$D_{j,\gamma} = A_\gamma^{\beta_j G(j)}。$$

步骤 3:其中一个$(d+1)$工作云服务器收集每个$S_j \in S$中的$D_{j,\gamma}$,并进行计算

$$P_\gamma = \prod_{S_j \in S} D_{j,\gamma} = \prod_{S_j \in S} A_\gamma^{\beta_j G(j)} = A_\gamma^{\sum_{S_j \in S} \beta_j G(j)} = A_\gamma^p =$$

$$(g^{\sum_{i=1}^{w} m_{i,\gamma}} \cdot h^{Z_\gamma})^p = (g^p)^{\sum_{i=1}^{w} m_{i,\gamma}} \cdot (g^{Z_\gamma})^{pq} = (g^p)^{\sum_{i=1}^{w} m_{i,\gamma}} \cdot 1 = \hat{g}^{\sum_{i=1}^{w} m_{i,\gamma}},$$

式中,$\hat{g} = g^p$。注意到多项式函数$G(x)$是$p + a_1 x + \cdots + a_d x^d$,其中$a_i \in Z_n^*$$(i = 1,2,\cdots,d)$。根据拉格朗日插值多项式,则

$$G(x) = \sum_{j=0}^{d} \left(\prod_{i=0, i \neq j} \frac{x_i - x}{x_i - x_j} \right) G(x_j),$$

因此,

$$\sum_{S_j \in S} \beta_j G(j) = \sum_{j=0}^{d} \left(\prod_{i=0, i \neq j} \frac{i-0}{i-j} \right) G(j) = G(0) = p。$$

步骤 4:通过以$\hat{g}$为底计算P_γ的离散对数,C_S在预期时间$O\left(\sqrt{(w+1)I}\right)$内采用 pollard's labmda 方法可以得到用户聚合的用电数据$M_{\mathrm{sum}} = \sum_{i=1}^{w} M_{i,\gamma}$。

步骤 5:云服务器发起链上交易将$\{t_\gamma, ID_{\mathrm{CS}_j}, Y_{\mathrm{CS}_j}, M_{\mathrm{sum}}\}$发布出去。经过其他链上节点的验证后,生成一个新的块,聚合后的信息数据在区块链上记录成功,以备查询。

4　理论分析

如 2.3 小节所述,为了抵御攻击者$\mathcal{A}$的攻击,需要满足安全需求,特别是保护用户的私有数据免受强大对手$\mathcal{A}$的攻击。在本节中,主要将分析提议的方案中涉及的一些安全与隐私问题。

4.1　抵抗监听攻击

挑战:在层与层之间的通信通道中可能会出现攻击者$\mathcal{A}$。一方面,攻击者通过窃听渠道获取数据,从而暴露了用户的隐私。另一方面,攻击者可以篡改数据,从而损害聚合数据的真实性。此外,攻击者还可以对通信信道发起重放攻击。

证明：假设攻击者$\mathcal{A}$在t_γ时间点监听了用户U_i的密文$C_{i,\gamma}=g^M\cdot h^Z$。由于用电数据在小范围时间内通常很小，因此电表每隔一定时间报告用电数据值通常在一定范围内。基于此，攻击者$\mathcal{A}$可能试图发起暴力攻击，以测试用电数据M的每一个可能值。由于文中方案基于Boneh-Goh-Nissim[31]密码，因此在语义上对所选密文攻击是安全的，攻击者$\mathcal{A}$在不知道用户U_i的私有参数的情况下无法恢复U_i的用电数据。

文中方案引入了签名方法和哈希函数来防止数据被篡改。无论是从收集层到聚合层，还是从聚合层到平台层，数据接收方都要验证数据发送方身份的合法性和数据的真实性。一旦数据被篡改，验证就无法成功。在此基础上，可以有效避免数据在传输过程中被篡改的问题。当然，对于防止重放攻击，数据的接收者首先检查时间戳$t_{\gamma'}$，计算是否满足$|t_{\gamma'}-t_\gamma|\leqslant\Delta t$。其中，$t_{\gamma'}$是当前时间戳，$\Delta t$是预定义的阈值。由于只有最新收集时间$t_{\gamma'}$的新报告才能通过验证，因此该方案能够抵抗重放攻击。

4.2 抵抗半诚实攻击

挑战：攻击者$\mathcal{A}$可能是整个系统中的一个参与者，例如网关GW或云服务器CS，他们试图从获得的数据中主动查找或推断用户U_i的私有用电数据。

证明：由于用户U_i的盲因子是私有的，网关或其他云服务器无法获取，因此网关或其他云服务器不可能从密文$C_{i,\gamma}$获取用户U_i的数据。另一方面，因为需要私钥$K_S=p$来获取用户用电数据的值。不过由于私钥已被TA碎片化，因此任何一个云服务器无法获取到完整私钥，进而获取到每个用户的用电数据。基于此，任何诚实而好奇的攻击者$\mathcal{A}$都无法推导出有助于揭示秘文的有价值的信息。

4.3 抵抗退出攻击

挑战：平台层的云服务器CS因损坏退出，可能导致系统崩溃或用户数据隐私泄露。

证明：如果系统中的服务器数量$k=1$，那么服务器故障或被攻击者破坏，整个系统将遭受单点故障；如果系统的服务器数量$k=2$，那么当强大的攻击者遇到其中一个服务器并攻破它时，就可以获得系统的所有机密信息；因此，选择服务器个数$k\geqslant 3$，并分配不同的私钥$G(j)$。

存在不超过$d=\lceil k/2\rceil-1$个CS失败或被破坏，系统也会保护用户的用电数据不受攻击。特别地，假设攻击者$\mathcal{A}$并获得它们的私钥$G(j)$，($j=1,2,\cdots,d$)。而$\mathcal{A}$仍然不能获得私有秘密p，因为根据秘密共享的“全或无”属性，极少需要$(d+1)$个CS才能恢复私钥p。类似地，为了解密用户的聚合数据，$(d+1)$个CS需要计算$(d+1)$解密共享$D_{\gamma,j}=A_\gamma^{\beta_j G(j)}$。攻击者$\mathcal{A}$只能获得$d$份解密份额，这不足以获得$P_\gamma=A_\gamma^p$，此$\mathcal{A}$无法获得用户的聚合数据。此外，本文系统可以支持CS容错，只要被破坏的CS数量小于d，仍然有$k-d\geqslant d+1$工作的CS可以保持系统正常工作。根据上面的讨论，同时d个CS失败，强大的攻击者$\mathcal{A}$仍然不能泄露用户的私有和聚合数据。

5 方案性能分析

在本节中，将通过实验方法对文中所提出的EPPDA性能进行评估，以证明其实际应用价值。。

5.1 实验设置

方案的设计目标之一是减少数据聚合中各个实体的计算开销和通信成本，从而提高整个聚合过程的效率。同时，在聚合过程中，用电数据需要时刻得到有效的保护。因此，首先进行了计算开销和通信成本的对比分析实验。在比较方案的选择上，由于文中采用的Boneh-Goh-Nissim加密算法属于同态加密类，因此选择同样采用同态加密算法来进行数据聚合的方案——LVPDA（lightweight and verifiable privacypreserving data aggregation）[32]；由于文中将聚合的信息存储在区块链上以供查询，选择了具有相同查询功能的聚合方案——ACFQ（aggregation communication and function query）[33]进行比较。

使用的笔记本电脑配置为：Windows系统(Win 11，64位)，16.0 GB RAM和2.3 GHz的Intel(R) Core(TM) i7-10510U CPU。程序代码用Java编程语言编写，基于JPBC（基于Java配对的加密）库。JPBC库是一个包容和高效的加密操作和协议库。

5.2 计算开销

在评估该方案的计算复杂度之前，考虑了聚合过程中涉及的加密操作。为了简单起见，使用了一些符号

来表示加密操作,如表1所示。由于指数运算和乘法运算相比,哈希运算的计算代价可以忽略不计,所以在计算复杂度评估中没有考虑哈希运算。假设聚合系统中有w个用户,并且所有方案都处于相同的安全级别。

EPPDA方案中,在对于用户报告生成阶段,每个用户U_i需要运行3个E_G和1个M_G用于将计量数据加密为C_i,1个E_G用于生成签名S_{U_i}。在安全报告聚合阶段,GW需要运行$2w$个E_G和w个M_G来认证每个数据发送方的用户身份U_i和数据完整性,使用w个M_G将用户的报告聚合为A_γ。接下来,GW只需要1个E_G来生成签名S_{W_i}。在安全的数据解读阶段,CS需要执行2个E_G和1个M_G来认证每个数据发送方的网关身份GW和数据完整性,接着花销$(d+1)$个M_G,将用户的碎片数据重构为P_γ。最后,CS采用pollard's labmda方法得到功率数据M_{data}的聚合并上链保存。

LVPDA方案中,每个用户使用2个E_{N^2}、1个M_{N^2}、6个E_G和2个M_G来生成报告。GW用$(w-1)$个M_{N^2}和4个E_G,$(w+1)$个M_G,$(w-1)$个M_{G_T}和$(w+1)$个$\widehat{Bp}$来检查数据报告完整性和聚合。最后,CS使用2个$\widehat{Bp}$、1个E_{N^2}和1个M_{N^2}来验证GW的报告和解密。

ACFQ方案中,每个用户使用2个E_{N^2}将使用数据加密为$C_{i,t}=(C_{i,t,1},C_{i,t,2})$,1个$M_G$生成签名$\sigma_{i,t}$。对于聚合报表,GW首先运行$w$个$M_G$、$(w+1)$个$\widehat{Bp}$和$w$个$E_{G_T}$,检查每个用户的报表来源和数据完整性,然后使用$w$个$E_G$完成聚合。最后,GW需要1个$E_G$来生成签名$\sigma_{i,t}$。在聚合检索和反馈阶段,CS运行2个$\widehat{Bp}$检查报告来源和数据完整性,然后取1个$E_G$和1个$E_{N^2}$恢复聚合使用情况。表2显示了3类聚合方案的计算复杂度。

表1　密码运算符号及描述

Table 1　Symbols and descriptions of cryptographic operations

符号	描述	符号	描述
E_{N^2}	E_{N^2}上的模幂运算	M_N	Z_N上的模乘法运算
$\widehat{Bp}$	双线性配对运算	M_{N^2}	Z_{N^2}上的模乘法运算
E_G	群G上的指数运算	M_G	群G上的模乘法运算
E_{G_T}	群G_T上的指数运算	M_{G_T}	群G_T上的模乘法运算

表2　计算复杂度对比

Table 2　Computational complexity comparison

方案	类别	理论耗时
LVPDA	用户端U_i	$2E_{N^2}+M_{N^2}+6E_G+2M_G$
	网关GW	$(w-1)M_{N^2}+4E_G+(w-1)M_G+(w-1)M_{G_T}+(w+1)\widehat{Bp}$
	云服务器CS	$E_{N^2}+2M_{N^2}+2\widehat{Bp}$
ACFQ	用户端U_i	$2E_{N^2}+M_G$
	网关GW	$2wM_G+(w+1)\widehat{Bp}+wE_{G_T}+E_G$
	云服务器CS	$2\widehat{Bp}+E_G+E_{N^2}$
EPPDA	用户端U_i	$4E_G+M_G$
	网关GW	$(2w+1)E_G+2wM_G$
	云服务器CS	$2E_G+(d+2)M_G$

表3为安全强度τ为80时相应耗时操作的运行时间。例如,执行一个双线性成对映射操作(即$\widehat{Bp}$)大约需

要 7.937 ms，而执行一个模幂运算（即 E_{N^2}）大约需要 3.542 ms。值得注意的是，表 3 所示的实验结果是每个操作运行 1 000 次后的平均运行时间。

表 3　安全强度 τ = 80 下密码运算耗时

Table 3　Security strength password operation in τ = 80

描述	耗时/ms	描述	耗时/ms
E_{N^2}	3.542	M_N	0.008 1
$\widehat{Bp}$	7.937	M_{N^2}	0.016 4
E_G	0.810 6	M_G	0.008 4
E_{G_T}	4.458 8	M_{G_T}	0.040 5

图 2 为 EPPDA 方案与 LVPDA、ACFQ 方案在用户端的计算复杂度对比。可知，EPPDA 方案在用户端耗时为 3.250 8 ms，分别比 LVPDA、ACFQ 的花销降低 72.87%、54.17%。为了便于比较云服务器的计算成本，文中设 d=1 000。图 3 为 3 种方案在云服务器端计算成本的比较。可知，EPPDA 花费了 10.038 ms，分别比 LVPDA、ACFQ 的花销降低 48.39% 和 50.37%。由于对比方案中在云服务端有双线性配对运算的存在，因此对比方案计算复杂度较高。EPPDA 方案在云服务端未采用该运算的同时也达到了相同的聚合目的，故 EPPDA 方案有效地降低了云服务端的计算复杂度。

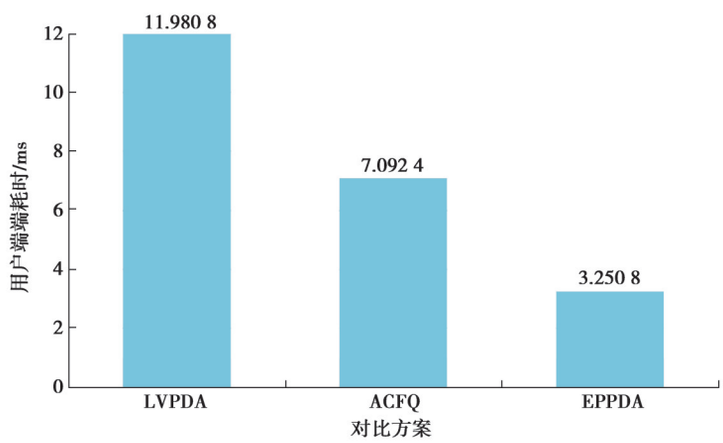

图 2　用户端的计算复杂度对比

Fig. 2　Comparison of computational complexity on the client side

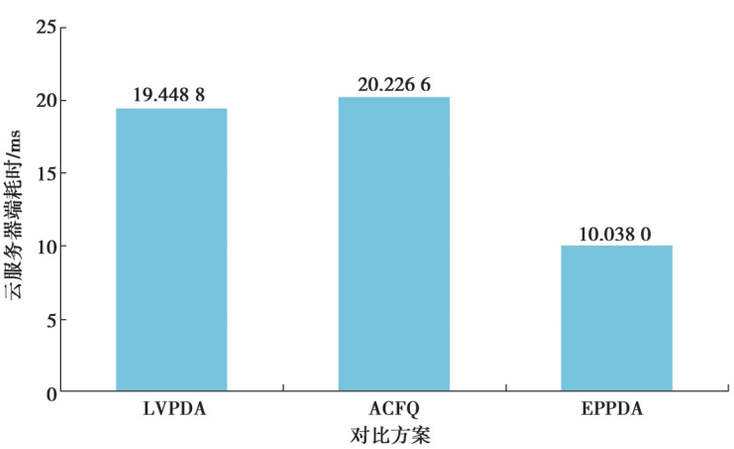

图 3　云服务器端计算复杂度对比

Fig. 3　Comparison of computational complexity on the cloud server

5.3　通信成本

通信开销包括从用户端到网关和网关到云服务器端通信。为了简单起见,假设TA基于椭圆曲线密码(elliptic curve cryptography,ECC)机制完成初始化 Gen(κ)。这样做的优点是,与RSA(rivest-shamir-adleman)加密等其他方法相比,使用ECC需要更小的密钥来提供更高级别的安全性。因此,与相关方案相比,EPPDA方案可以实现更低的带宽和通信过程的公平性。根据参考文献[1],为了达到τ=80的安全级别,将用于加密的RSA模N的长度设为1 024位,即$|N|$=1 024。实验中椭圆曲线采用Barreto-Naehring(BN)曲线且在F_{p_1}-160上进行,使$|G|$=320位。此外,由于SHA-1已经被证明为数字签名提供了低于80位的安全级别,EPPDA中应用了SHA-256哈希函数。

在EPPDA中,每个用户只需要在每一轮中传输消息($C_{i,\gamma}||ID_{u_i}||t_\gamma||R_{u_i}||s_{u_i}$),这将需要($|N|+3|G|+|ID|+|T|$)位。然后GW将花费($|N|+3|G|+|ID|+|T|$)来发送消息|($A_\gamma||ID_{W_i}||t_\gamma||R_{w_i}||S_{W_i}$)到云服务器CS。

对于LVPDA,用户需要执行2个阶段(即离线签名和在线签名生成)来生成认证信息和加密数据,其中可能需要($2|N|+5|G|+|T|+|ID|$)位,而GW传输聚合报告可能只需要($|2|N|+|G|+|T|+|ID|$)位。在ACFQ情况下,用户需要|($2|N|+|G|+|ID|+|T|$)位向GW节点(雾节点)发送消息($T,U_i,C_{i,T},\sigma_{i,T}$)。处理完成后,节点将($\delta,t,F_k,\widehat{c_{k,t}},\overline{\sigma_{k,t}}$)转发给云服务器CS,其通信成本为($2|N|+|G|+2|ID|+|t|$)位。

表4概述了方案之间的通信复杂性。与计算复杂度类似,EPPDA的通信复杂度低于其他方案。注意表格中的ID和T在文中被设置为32位。接下来,分别计算用户端到网关通信和网关到云服务器通信的成本,并进行成本对比。用户端到网关通信成本对比如图4所示。从图中可以看出,EPPDA方案从用户到网关的通信为704位,分别比LVPDA、ACFQ的通信成本降低了81.03%、71.05%;从图5可以看出,EPPDA方案从网关到云服务器的通信为704位,分别比LVPDA、ACFQ的通信成本降低了71.05%、50%。综上可以分析,EPPDA方案在保证用户数据隐私和安全的前提下,降低了各参与方的时间成本和通信开销,因此,文中提出的EPPDA聚合方案在整体性能上有较大提升优势。

表4　通信复杂度对比

Table 4　Communication complexity comparison

方案	用户端到网关	网关到云服务器端
LVPDA	$2\|N\|+5\|G\|+\|T\|+\|\mathrm{ID}\|$	$2\|N\|+\|G\|+\|T\|+\|\mathrm{ID}\|$
ACFQ	$2\|N\|+\|G\|+\|T\|+\|\mathrm{ID}\|$	$2\|N\|+\|G\|+2\|T\|+\|\mathrm{ID}\|$
EPPDA	$\|N\|+3\|G\|+\|T\|+\|\mathrm{ID}\|$	$\|N\|+3\|G\|+\|T\|+\|\mathrm{ID}\|$

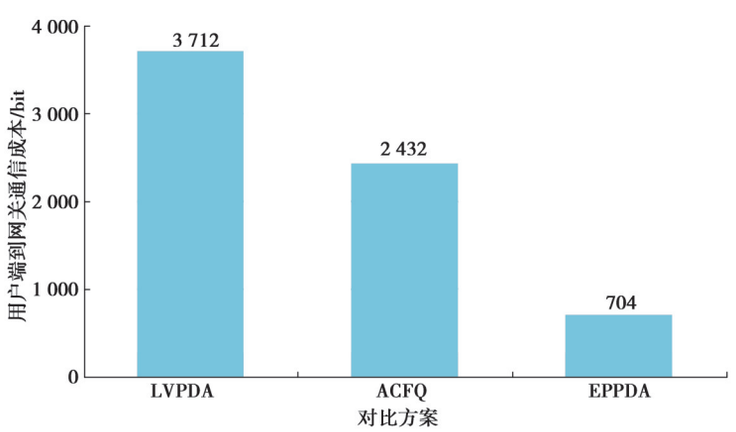

图4　用户端到网关通信成本对比

Fig. 4　Client-to-gateway communication cost comparison

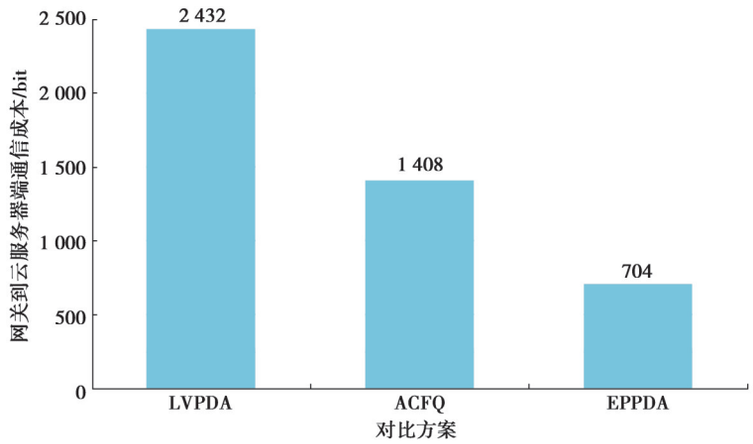

图5 网关到云服务器端通信成本对比

Fig. 5 Comparison of communication cost between gateway and cloud server

6 结束语

基于Boneh-Goh-Nissim密码系统和签名机制,提出了一种新的基于4层架构的智能电网安全数据聚合方案EPPDA,解决了智能电网数据安全聚合和隐私泄露问题。为了证明EPPDA的有效性,文中进行了详细的理论分析,并与LVPDA、ACFQ方案进行比较。结果表明,EPPDA方案在用户端计算和通信能力受限的前提下分别降低了72.87%与54.17%的计算耗时,减少了81.03%与71.05%的通信成本。在未来,将通过探索更高效的加密工具来进一步提高EPPDA的效率。

参考文献

[1] Wang J, Wu L B, Zeadally S, et al. Privacy-preserving data aggregation against malicious data mining attack for IoT-enabled smart grid[J]. ACM Transactions on Sensor Networks, 2021, 17(3): 1-25.

[2] Oakes G. Enabling renewable energy through smarter grids[J]. Communications of the ACM, 2021, 64(3): 11.

[3] 白雪杰, 郭雷岗, 姜丽鸽. 物联网技术在智能电网中的应用研究[J]. 物联网技术, 2022 (3): 12.
Bai X J, Guo L G, Jiang L G. Application of Internet of things in smart grid[J].Internet of things technologies, 2022 (3): 12. (in Chinese)

[4] Al-Waisi Z, Agyeman M O, Al-Waisi Z, et al. On the challenges and opportunities of smart meters in smart homes and smart grids[C]//Proceedings of the 2nd International Symposium on Computer Science and Intelligent Control. September 21-23, 2018, Stockholm, Sweden. ACM, 2018: 1-6.

[5] 丁勇, 王冰尧, 袁方, 等. 支持第三方仲裁的智能电网数据安全聚合方案[J]. 电子学报, 2020, 48(2): 350-358.
Ding Y, Wang B Y, Yuan F, et al. Secure data aggregation scheme in smart grid with third-party arbitration[J]. Acta Electronica Sinica, 2020, 48(2): 350-358. (in Chinese)

[6] 夏喆, 罗宾, 徐桂彬, 等. 智能电网中支持细粒度访问控制的隐私保护数据聚合方案[J]. 信息网络安全, 2021, 21(11): 28-39.
Xia Z, Luo B, Xu G B, et al. Privacy-preserving data aggregation with fine grained access control for smart grid[J]. Netinfo Security, 2021, 21(11): 28-39. (in Chinese)

[7] 朱嵩, 王化群. 基于Paillier算法的智能电网数据聚合与激励方案[J]. 计算机工程, 2021, 47(11): 166-174.
Zhu S, Wang H Q. Paillier-based data aggregation and stimulation scheme in the smart grid[J]. Computer Engineering, 2021, 47 (11): 166-174. (in Chinese)

[8] Yu C M, Chen C Y, Kuo S Y, et al. Privacy-preserving power request in smart grid networks[J]. IEEE Systems Journal, 2014, 8 (2): 441-449.

[9] Diao F, Zhang F G, Cheng X G. A privacy-preserving smart metering scheme using linkable anonymous credential[J]. IEEE Transactions on Smart Grid, 2015, 6(1): 461-467.

[10] Wen M, Zhang X, Li H W, et al. A data aggregation scheme with fine-grained access control for the smart grid[C]//2017 IEEE 86th Vehicular Technology Conference (VTC-Fall). September 24-27, 2017, Toronto, ON, Canada. IEEE, 2017: 1-5.

[11] Lu W F, Ren Z H, Xu J, et al. Edge blockchain assisted lightweight privacy-preserving data aggregation for smart grid[J]. IEEE

Transactions on Network and Service Management, 2021, 18(2): 1246-1259.

[12] Mohammadali A, Haghighi M S. A privacy-preserving homomorphic scheme with multiple dimensions and fault tolerance for metering data aggregation in smart grid[J]. IEEE Transactions on Smart Grid, 2021, 12(6): 5212-5220.

[13] He D B, Kumar N, Zeadally S, et al. Efficient and privacy-preserving data aggregation scheme for smart grid against internal adversaries[J]. IEEE Transactions on Smart Grid, 2017, 8(5): 2411-2419.

[14] Zhang X J, Huang C, Xu C X, et al. Key-leakage resilient encrypted data aggregation with lightweight verification in fog-assisted smart grids[J]. IEEE Internet of Things Journal, 2021, 8(10): 8234-8245.

[15] Shen H, Liu Y J, Xia Z, et al. An efficient aggregation scheme resisting on malicious data mining attacks for smart grid[J]. Information Sciences, 2020, 526: 289-300.

[16] Bohli J M, Sorge C, Ugus O. A privacy model for smart metering[C]//2010 IEEE International Conference on Communications Workshops. May 23-27, 2010, Cape Town, South Africa. IEEE, 2010: 1-5.

[17] Kursawe K, Danezis G, Kohlweiss M. Privacy-friendly aggregation for the smart-grid[M]//Privacy Enhancing Technologies. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011: 175-191.

[18] Danezis G, Fournet C, Kohlweiss M, et al. Smart meter aggregation *via* secret-sharing[C]//Proceedings of the First ACM Workshop on Smart Energy Grid Security. 8 November 2013, Berlin, Germany. ACM, 2013: 75-80.

[19] Knirsch F, Eibl G, Engel D. Error-resilient masking approaches for privacy preserving data aggregation[J]. IEEE Transactions on Smart Grid, 2018, 9(4): 3351-3361.

[20] Lyu L J, 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.

[21] Gope P, Sikdar B. Lightweight and privacy-friendly spatial data aggregation for secure power supply and demand management in smart grids[J]. IEEE Transactions on Information Forensics and Security, 2019, 14(6): 1554-1566.

[22] Mohammed H, Tonyali S, Rabieh K, et al. Efficient privacy-preserving data collection scheme for smart grid AMI networks [C]//2016 IEEE Global Communications Conference (GLOBECOM). December 4-8, 2016, Washington, DC, USA. IEEE, 2016: 1-6.

[23] Tonyali S, Cakmak O, Akkaya K, et al. Secure data obfuscation scheme to enable privacy-preserving state estimation in smart grid AMI networks[J]. IEEE Internet of Things Journal, 2016, 3(5): 709-719.

[24] Xue K P, Zhu B, Yang Q Y, 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.

[25] Zhou J, Zhang Y F, Cao Z F, et al. PPSAS: lightweight privacy-preserving spectrum aggregation and auction in cognitive radio networks[C]//2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS). July 7-10, 2019, Dallas, TX, USA. IEEE, 2019: 1127-1137.

[26] Shi Z G, Sun R X, Lu R X, et al. Diverse grouping-based aggregation protocol with error detection for smart grid communications[J]. IEEE Transactions on Smart Grid, 2015, 6(6): 2856-2868.

[27] Baloglu U B, Demir Y. Lightweight privacy-preserving data aggregation scheme for smart grid metering infrastructure protection[J]. International Journal of Critical Infrastructure Protection, 2018, 22: 16-24.

[28] Zheng Z G, Wang T, Bashir A K, et al. A decentralized mechanism based on differential privacy for privacy-preserving computation in smart grid[J]. IEEE Transactions on Computers, 2022, 71(11): 2915-2926.

[29] 应臣浩, 夏福源, 李颉, 等 . 区块链群智感知中基于隐私数据真值估计的激励机制[J]. 计算机研究与发展, 2022, 59(10): 2212-2232.
Ying C H, Xia F Y, Li J, et al. Incentive mechanism based on truth estimation of private data for blockchain-based mobile crowdsensing[J]. Journal of Computer Research and Development, 2022, 59(10): 2212-2232. (in Chinese)

[30] Jo H J, Kim I S, Lee D H. Efficient and privacy-preserving metering protocols for smart grid systems[J]. IEEE Transactions on Smart Grid, 2016, 7(3): 1732-1742.

[31] Bao H Y, Lu R X. A new differentially private data aggregation with fault tolerance for smart grid communications[J]. IEEE Internet of Things Journal, 2015, 2(3): 248-258.

[32] Zhang J L, Zhao Y C, Wu J, et al. LVPDA: a lightweight and verifiable privacy-preserving data aggregation scheme for edge-enabled IoT[J]. IEEE Internet of Things Journal, 2020, 7(5): 4016-4027.

[33] Liu J N, Weng J, Yang A J, 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.

（编辑　詹燕平）