Lightweight Privacy-Preserving Data Aggregation Scheme Based on Elliptic Curve Cryptography for Smart Grid Communications

Thokozani Felix Vallent[1(✉)], Damien Hanyurwimfura[1], Jayavel Kayalvizhi[2], Hyunsung Kim[3,4], and Chomora Mikeka[5]

[1] College of Science and Technology, African Center of Excellence in Internet of Things (ACEIoT), University of Rwanda, KN Street Nyarugenge, P.O. Box 3900, Kigali, Rwanda
tfvallent@gmail.com, dhanyurwimfura@ur.ac.rw

[2] School of Computing, Department of Information Technology, SRM Institute of Science and Technology, Potheri, Kattankulathur Campus, Chennai, India
kayalvij@srmist.edu.in

[3] School of Computer Science, Kyungil University, Gyeongsan-si, South Korea
kim@kiu.kr

[4] Chancellor College, Mathematical Science Department, University of Malawi, P.O. Box 280, Zomba, Malawi

[5] Chancellor College, Department of Physics, University of Malawi, P.O. Box 280, Zomba, Malawi
cmikeka@cc.ac.mw

Abstract. Smart grid (SG) is a modern electricity grid based on bi-directional flow of electricity and information for efficient energy management. Due to dependence on information communication, the system is prone to potential cyber-security attacks such as, user identity theft and data privacy breach. Addressing these cyber-security issues with optimal efficiency in smart grid is an open research problem. From this perspective this paper proposes a lightweight scheme for robust information security and privacy-preservation in data aggregation in SG. The proposed scheme utilizes elliptic curve variant of El Gamal encryption cryptosystem and signcryption techniques to achieve user anonymity with greater efficiency. The scheme satisfies the standard security requirements proven in the random oracle model and does not need a trusted third-party or certificate issuance during scheme run. Performance evaluation analysis shows that the proposed scheme has a better overall performance to most relevant comparable schemes, since it does not use heavy computation operations such as bilinear pairings, map-to-point hash operations, exponentiation among others. Furthermore, the proposed scheme does not depend on trusted authority (TA) neither suffers from coalition attack nor insider attacks.

Supported by University of Rwanda.

Y.-B. Lin et al. (Eds.): SGIoT 2021, LNICST 447, pp. 71–97, 2022.
https://doi.org/10.1007/978-3-031-20398-5_7

Keywords: Smart grid · Data aggregation · Bi-directional communication · Elliptic curve cryptography · Internet of things · Privacy-preserving

1 Introduction

As the future power grid, smart grid (SG) surpasses the legacy grid with a greater advantage due to incorporation of advanced communication technology in the electricity system [5,13,14,42,50,59]. In accordance to the model of National Institute of Standards and technology (NIST), smart meter (SM) based at the residential area collects and sends real-time electricity usage data to the operation center (OC), and in a similar manner a SM receives command messages related to control and management of electricity from OC [26]. In SG the electricity data bears economic value as it can be used for electricity marketing, management and regulation in applications such as demand response [20,24,30]. An existing challenge with the electricity consumer's consumption data is that it contains privacy information such that it can reveal the user's identification, lifestyle pattern and habits of electricity usage of the consumers, which is not a desirable thing [6,10,15,54,60].

In SG, recording of power consumption data is done not on month-to-month basis as is the case with the legacy grid, but rather is based on times of the day and gives details of the power consumption of individual appliances on the customers side in near real-time intervals [48,65]. On the negative note this detailed information has a potential of being used maliciously by an attacker to invade privacy and security. For instance, by knowing the times house owners' sleep or are absent from home based on analysis of SM information flow, thieves can plan to break into the homes when nobody is at home. Hence there is great need to protect the consumer's side against privacy and security breaches associated with smart meter information, with respect to the basic security requirements. That is to ensure SM communications should be supported with user authentication and identification measures when transacting on the open network between the SM and the monitoring device. Encryption with efficient algorithm for the customer consumption information transmitted by the SM to the utility or third party service providers will also be needed. The communication scheme should also protect individual energy consumer information from third parties for commercial purposes not related to services provided by the utility that is the scheme should be resilient to all sort of insider attacks [67].

However, the current challenge is ensuring security and privacy of smart grid AMI communications and balancing between lightweight cryptographic measures and ideal computational complexity for resource constrained devices. Once lightweight privacy-preserving and security mechanism are ascertained, consumers and utility are likely to have a widely varying preferences on how they wish to control and monitor third party to access their information. Thus, the challenge to protect user's privacy information effectively and efficiently has attracted researchers attention to find a lasting solution [8,34,39,67,70]. Such solutions would ensure the adoption of internet of things(IoT) applications over smart grid, since they consist of resource-constrained devices. Privacy-preserving data aggregation approach is one of viable mechanisms feasible for achieving data privacy [9,56,71]. Privacy-protection safeguards the fine-grained user data from disclosure to unwanted parties like the gateway or the system service provider during transmission and in some cases even protected from insider attack within the control center itself. Data aggregation is an ideal solution for securing data from eavesdroppers and has the advantage of improving network performance by virtue of reducing communication traffic. This approach goes with the following requirements: 1) the data aggregation operator can obtain sum of usage data in a region; 2) the data aggregation operator should know nothing about individual usage data in the data collection region [36,40,45].

In this regard, homomorphic encryption (HE) is a prospective mechanism for ensuring data aggregation privacy because it allows ciphertext manipulation without divulging plain-text [51]. The HE technique allows performing of addition operation or multiplication operation on encrypted cipher without requiring decryption, which is the desirable property for the aggregation operator to do, since it is a semi-trusted entity. There are several other data aggregation techniques such as using random numbers secrete sharing, Boneh-Goh-Nissim homomorphic encryption, Paillier homomorphic encryption, data slicing, differential privacy among other techniques [4]. In random number secret sharing design a series of random numbers with underlying properties are initially distributed to all the entities in the network that is to, all users and the data aggregation operators in advance, which are later used to obfuscate the usage data transmitted in the network [32]. The drawback of such a mechanism is reliance on Trusted Third Party (TTP) responsible for generating and distributing such random numbers. Differential privacy is also a technique for achieving privacy which adds random noise of Laplacian distribution or other distributions to mask the original value [17]. However, in many studies, this technique of differential privacy is deemed less accurate. Secret sharing which was proposed by Shamir in [55], is also another mechanism used to achieve privacy data aggregation. In this method group secret is split into shares and distributed amongst the participants and each one is kept highly confidential. The secret can only be reconstructed with a sufficient number of participants colluding and combining their secret values together, as such it is good for storing highly sensitive information [53]. However, most of these approaches are coupled with unbearable computation burden, which is an issue of major concern.

In this paper, a Lightweight Privacy Preserving Data Aggregation Scheme Based on Elliptic Curve for Smart Grid Communications is proposed, with the following main contribution are:

- The proposed scheme design precludes certificate and TTA or TA dependency, hence ensures system overall management requirements significantly.
- The scheme uses lightweight mathematical building blocks that are bilinear pairing-free, thus resulting in reduction of transmission delay from node to node communications.

- The proposed scheme has optimal computation and communication efficiency based on the security analysis and performance evaluation, and so the scheme has comparative advantage over other variant works, by its merits.
- Most importantly the proposed scheme ascertain user anonymity when communicating over a public channel beside providing content privacy.

In this vain, our work endeavors to design and incorporate these main contributions, and performance comparison with some state-of-the-art privacy-preserving data aggregation techniques provided. Thus, the proposed work achieves better overall communication and computation efficiency to the best comparable scheme besides better satisfaction of main security requirements.

The rest of the paper is organized as follows: Sect. 2, reviews related works in data aggregation over smart grid environment and pin-points the research gap and limitations in other researches. Then we present the generic system design model, basic mathematical and cryptographic preliminaries required for understanding the proposed scheme in Sect. 3. The proposed lightweight privacy-preserving data aggregation scheme is presented in Sect. 3. Further Sect. 4 gives the analysis and evaluation of results respectively. Finally, the conclusion and suggestions for further research are presented in Sect. 5.

2 Related Works and Limitations

A variety of techniques and approaches for privacy-preserving data aggregation (PPDA) have been proposed in literature in a quest to address the challenge of secure data aggregation to be practical in SG network environment. In [43] Liu Y., et al., proposed a 3PDA data aggregation scheme for SG, however the scheme uses computationally expensive bilinear pairings in the data collection units (DCUs) and the OC which is not ideal for a system of systems comprising of myriads of communicating devices. While in [32] He D., et al., designed an efficient PPAD scheme but is based on TTP which is not also a desirable feature in a highly populated network since it bears a bottleneck for the TTP to manage the messages with efficiency. Similarly in [37] Jo H.J., et al., asserts that their construction is efficient at the expense of using bilinear computations which contradicts the claim for a practical scheme. Additionally, in [31] is it disclosed that their scheme does not provide user privacy as the identity is transmitted in plaintext hence prone to human-factor-aware data aggregation (HDA) attacks and all sorts of privacy breaches. Other researches based their construction on computationally expensive Paillier's homomorphic cryptosystem [12,46,59], which still need some improvement to be deemed practical for smart grid because the Pailler's primitives is comparative heavier although is it widely used technique [16]. Achieving acceptable level of computational as well as communication efficiency is a default requirement for any real-time based communication technology like IoT-enhanced smart grid. In the schemes [12,46,59], theirs security is anchored in the gateway as it plays a trusted role and it provides the aggregated data to the OC. The proposed scheme in [37], does not rely on trusted gateway but rather uses a group approach to sent encrypted messages. The private key is secretly held by a group of smart meters and they collude to send an encrypted message to the AMI. In order for the AMI or OC to decrypt data, it requires the selected smart meters to help in the decryption which adds on overhead and clearly this is impractical design. Although the scheme removes trust-ship, it incurs unnecessary communication overhead in the collusion process. Also the design is prone to differential attack hence can suffer from privacy breach. To achieve this the AMI collude with participating smart meter to decrypt two different messages representing the sets which differ only by a single user. At the end the AMI can deduce the user's data from the difference of the two decrypted messages. Although the scheme [65] claims to be robust in security and efficiency, we analyzed it and found it lacking to provide user anonymity as the identity is transmitted in plain-text, hence is prone to eavesdropping, and also in [60] it was found to have its properties unattributed. In 2011, Wu and Zhou [68] proposed a key exchange scheme for smart grid based on elliptic curve cryptography (ECC) which required a trusted authority (TA) and a public key infrastructure (PKI) for key management. The authors of [68] claim their scheme is resistant to replay attacks and man-in-the-middle attacks. However in [69] Xia and Wang discovered that the scheme in [68] cannot provide replay attack and man-in-the-middle attack resilience and has even vulnerable session key. In 2018, Mahmood et al., [47] proposed an authentication scheme between two communicating entities in smart grid, however in [1] it is pointed out that the scheme in [47] has weakness of lack of perfect forward secrecy. Although in [63], it was assessed that the proposed data obfuscation approach provides good data throughput, good packet delay and delivery ratio under a variety of conditions, in [15] they found it bears high communication latency due to large bandwidth requirement and also in [60] they point out that the scheme in [63] fails to explicitly state its focusing field, whether is it for customer billing applications or grid operations. Another ECC based multi-dimensional data aggregation scheme was proposed by Boudia et al. in [10], which does not require bilinear pairings hence having efficient computation overhead. However, the scheme is flawed by being TA dependent. In [6], Badra and Zeadally, proposed an ECC based privacy-preserving data aggregation scheme by utilizing homomormphic encryption and Diffie-Hellman techniques, however it bears heavy communication overhead.

Different schemes in literature have been proposed based on different approached and methodologies. Their short-falls varies widely ranging from; having high computation overhead [2,7,43,46,59,61,63], prone to common attacks [11,47,57,68] and high demand of system computation resources due to need of additional requirements for TTP, PKI and TA management in their designs [28,30,32,33].

2.1 Generic System Model

In this section, we briefly describe the proposed data aggregation scheme's network architecture, the preliminaries building blocks for the scheme and subsequently we define the security requirements for the model. The system design of our model is depicted in Fig. 1, which consists of three main entities: SMs, data aggregation point (DAP) and the utility OC in a hierarchical structure, which follows the standard smart grid architecture [23,25]. An OC has a number of neighborhood area networks (NANs), in turns each NAN has numerous home area networks (HANs) [21,27,29,30,35,38]. The DAP is the gateway for a particular NAN which does data aggregation and relaying of information between OC and SM. Whereas for each HAN there is a SM enabling bi-directional communication with the DAP. In this model SMs collect real-time consumption data from the HAN and sends it to OC via DAP at 15 min regular intervals [47,49,60]. The communication between SM and DAP is through wireless technology such as Wi-Fi, RF mesh, 6LowPAN, ZigBee, Z-wave among other wireless communication protocols [2,12,65]. Smart meter transmitted data includes the bill, real-time electricity report, consumer identification and regular instruction dispatches. So, the DAP acts as a gateway for the neighborhood area network and bridges the SMs and the OC by utilizing long range and high bandwidth communication technology with low latency like WiMAX or, 3G, 4G communication technologies and wired links among others [3,18,19,52,62,64,65]. As a gateway the DAP aggregates the collected individual SM data to leverage the computational overhead of OC in decrypting each SM's information. Similarly, communication efficiency is still a challenging issue that requires permanent solution, since there will be hundreds or thousands of smart meters in regions reporting their electricity related information almost at the same time to the OC through the DAP.

The OC being in the utility side, is deemed honest-but-curious entity and can know customer's electricity usage during a billing period of pricing and power management. This means the OC executes operations according to the scheme without launching any active attack. On the other hand, the DAP is assumed not fully trusted since it can be easily controlled by an adversary. The user or SM is usually assumed as dishonest with some extent of trust or common interest to win incentives from the OC orders. A good property needed in this architecture is to conceal individual's electricity usage to neighboring SMs and the DAPs even in an attempt of smart meters launching a collusion attack. Thus, is necessary to secure the system from all sorts of attacks such as: data privacy attacks, relationship attacks, false data injection attacks and distortion attacks. Data privacy attack occurs when some malicious smart meter node collude with the gateway or another smart meter to obtain the real-time or total power consumption data of another uncompromised smart meter during a billing session. In relationship attack the gateway attempts to deduce user behavior and habits from the relationship between electricity consumption of different reporting intervals for one specific user. While false data injection refers to attack on the integrity of the data where an adversary introduces a code that compromises the correctness of the data. In distortion attack the gateway tries to forge a user's real-time power usage in order to disturb the billing system in smart grid.

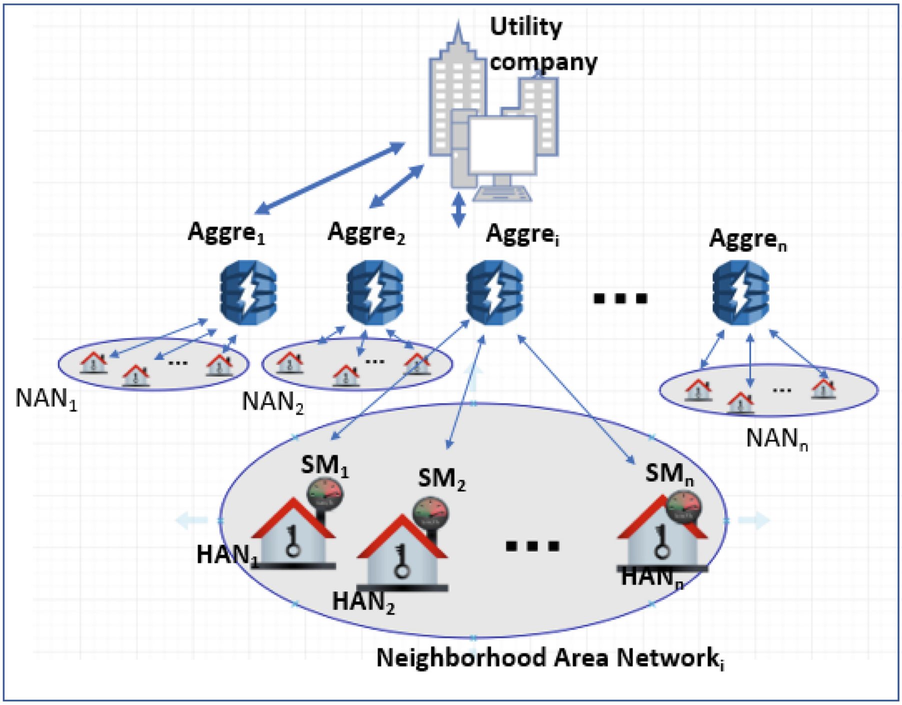

Fig. 1. Generic system model.

2.2 Preliminaries

Now, we will formalize the background knowledge of the building blocks for the proposed scheme. ECC is a public key cryptosystem based on elliptic curve theory and has an advantage for being a structure for faster and more efficient cryptosystems with robust security. ECC cryptosystems have low computational requirement hence making it as viable mechanism for securing IoT based systems with numerous resource constrained devices and real-time operations like in SG system [57,66].

- *Elliptic curve*: Given a prime number q, the equation $y^3 = x^2 + ax + bmodp$ defines an elliptic curve over a prime field $E(F_p)$, where $p > 3, a, b \in F_q$ and satisfies $\triangle = 4a^3 + 27b^2 \neq 0 \ modp$. The points on F_p together with the point at infinity $\mathcal{O}$ form an additive cyclic group G. Let P be the generator point of order n, the scalar multiple operation is defined as, $nP = P + P + \cdots + P$, for n times addition, where $n \in Z_q^*$, is a positive integer. So, there are a number of intractable problems in an elliptic curve group, G, of order n, suitable for cryptographic purposes as there is no polynomial time algorithm to solve them efficiently by brute-force within probabilistic polynomial time.
- *Elliptic Discrete Logarithm (ECDL) Problem*: Given an element $Q \in G$, the ECDL problem is to extract an element $x \in Z_q^*$, such that $Q = xP$.

- *Elliptic Curve Computational Diffie-Hellman (ECCDH) Problem*: Given two elements $xP, yP \in G$, with unknown elements $x, y \in Z_q^*$, the ECCDH problem is to compute $Q = xyP$.
- *Elliptic Curve Decisional Diffie-Hellman (ECDDH) Problem*: No any probabilistic polynomial time algorithm can distinguish between the tuples (P_1, xP_1, yP_1, T) and (P_1, xP_1, yP_1, xyP_1) where $P_1, T \in G$, with unknown elements $x, y \in Z_q^*$.

2.3 Adversary Model

The public channel is open for an adversary, $\mathcal{A}$, which can forge, replay, modify and intercept plying message between communicating parties. However, $\mathcal{A}$, has no full access to private information from observing the public channel. Thus, the SM, DAP and OC are regarded as secure by themselves. The focus in this work is to deal with a strong enough adversary which has access to the communication media able to perform the following malicious actions:

- Eavesdropping the communication channel between SM and DAP communications to get an idea of the smart devices usage data and other associated details.
- Impersonation attack on a particular user's smart meter and probably send falsified data on behalf of the targeted user.
- Replay attack of legitimate transmitted messages to resend them after being intercepted. This attack can overload the authentication process and can result into transmission delay, denial of service (DoS) attack and communication bandwidth.
- Man-in-the-middle attack where by an adversary can actively eavesdrop on legitimate user's communications and relay modified messages between legitimate users making them believe they are communicating with intended counterpart.

2.4 Security Requirements

There is need to satisfy basic security requirements for the proposed scheme to ascertain the claimed security of the transmitted electricity reports from the consumer to the data aggregation points over the public channel as well as regulation and control commands originating from operation center to the consumer. Obviously, the data aggregation scheme faces all kinds of attacks from an adversary in between SM and DAP. The proposed scheme, achieves these attributes without relying on TTP or certificate issuance to facilitate the authentication process. Thus, our scheme design ensures conformity of the following prescribed requirements.

- *Authentication*: The DAP has to validate the true identity of the smart meter to make sure the data it receives really comes from the smart meter. Thus, the communicating entities should ascertain that they are communicating with the intended counterpart. This measure checks possible malicious acts such as message forgery, impersonation and masquerading attacks.

- *Data Confidentiality*: The mechanism must avoid any leakage of an individual's electricity usage data that could pose privacy breach. So, it is necessary to ascertain that no-one either internal or external attacker, extracts any individual's electricity usage data. So, confidentiality of user's electricity consumption data is very necessary since its leakage can reveal user's habit or identity which eventually exposes the user to attacks.
- *Data Integrity*: This mechanism ensures consistency and trustworthiness of the transmitted data, so that data is not altered nor modified by unauthorized party while in transit or elsewhere. This feature is of significantly importance in smart grid so that the data aggregation process should uphold integrity and message modification or forgery, and should there be any attempt for such attacks, should be detected.
- *Consumer Privacy and Anonymity*: The actual identity of a consumer in a community should not be known by any malicious party eavesdropping on the communications between consumer and OC. Thus, even if two instances of consumption data reports are eavesdropped the adversary should not distinguish if the two consumption reporting data are from the same user or not. Thus, a scheme zshould uphold user anonymity during the message flow of the concerned parties.
- *Attack Resilience*: Due to communication over a public channel, the data aggregation scheme must ensure security to withstand common attacks such as: impersonation attack, replay attack, modification attack and man-in-the-middle-attack [22,41].

In regard to the discussed system model and security requirements, our main goal is to design a secure and efficient scheme for privacy-preserving data aggregation for smart grid AMI communications.

3 Proposed Scheme

We will present the framework of the proposed Lightweight Privacy-Preserving Data Aggregation (LPPDA) Scheme Based on Elliptic Curve Cryptography for Smart Grid Communications. The notation description of the symbols used in the proposed scheme is outlined in Table 1. Our construction suggests the variant construction of an El Gamal encryptosystem and an additive homomorphic encryption algorithm applied over elliptic curve field whose properties are:

(a) In order to encrypt a message m into a ciphertext C using an El Gamal approach, the sender, uses the public key of a receiver $pk = xP$, generates a random number $r \in Z_q^*$, with it computes $C_a = rP$ and $C_b = (pk)r + mP$. Thus an encryption process proceeds as:

$$E_{pk}(m) = (C_a, C_b) = (rP, r(pk) + mP) = C$$

Table 1. Notations used in our scheme

Symbols	Meanings of symbols in the scheme
$p,\ q$	Two large primes
E	Is the chosen elliptic curve, $y^3 = x^2 + ax + b mod p$ where $a, b \in Z_q^*$
$E(F_p)$	Is the prime field of an elliptic curve E order p
P	Is the generator of $E(F_p)$ with large prime order q
G	A cyclic group generated by a point P on a non-singular elliptic curve E
ID_s	Identity of smart meter user, SM_i
$t_s,\ t_d$	Timestamps for SM and DAP
e_i	Electricity consumption generated by a smart meter at t_s
δ	The time lapse for an interval of time from t_s to t_d
$pk,\ sk$	Public key and private key respectively
$E_{pk}(.)$	Encryption algorithm by using a public key pk
$D_{sk}(.)$	Decryption algorithm by using a private key sk
$X_s,\ x_s$	Public key and private key of smart meter user respectively
$X_d,\ x_d$	Public key and private key of gateway respectively
$X_\alpha,\ x_\alpha$	Public key and private key of utility company
$H_1,\ H_2$	Hash function: $H_1,\ H_2 : \{0,1\}^* \rightarrow\ Z_q^*$
$\oplus$	Exclusive-OR operation (XOR)
$\|\|$	Concatenation

(b) So to decrypt a ciphertext C, the receiver uses its private key $sk = x$, and carry out the following operation.

$$\begin{aligned} D_{sk}(C) = D_{sk}(E_{pk}(m)) &= (r(pk) + mP) - (sk)(rP) \\ &= (r(xP) + mP) - x(rP) \\ &= mP \end{aligned}$$

After which the message m can be retrieved by using the Pollards lambda method.

(c) For additive homomorphic property, we suppose there are two messages encryption instances m_1 and m_2 transformed into two respective ciphertext instances C_1 and C_2 as follows.

$$\begin{aligned} E_{pk}(m_1) + E_{pk}(m_2) &= (r_1P, r_1(pk) + m_1P) + (r_2P, r_2(pk) + m_2P) \\ &= ((r_1P + r_2P), ((r_1pk + m_1P) + (r_2pk + m_2P)) \\ &= ((r_1 + r_2)P, (r_1pk + r_2pk) + (m_1P + m_2P)) \\ &= ((r_1 + r_2)P, (r_1 + r_2)pk + (m_1 + m_2)P) \\ &= E_{pk}(m_1 + m_2) \\ &= (C_1 + C_2) \end{aligned}$$

(d) Scalar multiplication property.

$$E_{pk}(cx) = \underbrace{E_{pk}(x) + E_{pk}(x) + \cdots + E_{pk}(x)}_{\text{c-times}}$$
$$= \sum_{i=1}^{c} E_{pk}(x)$$

where $+$ is the additive homomorphic encryption operation and c is a constant for scalar multiplication.

The construction of the scheme consists of the following five algorithms, namely System Initialization, Key Generation, Smart Meter Data Reporting, Data Aggregation and Data Recovery as explained below.

(a) **System Initialization**: OC runs a system initialization algorithm with 1^k as a security parameter and outputs a cyclic group G of prime order q, with $P \in G$ as its generator, an elliptic curve $E : y^2 = x^3 + ax + b\ modp$, where $a, b \in F_p$, for F_p a prime field of order p. The OC then chooses P from the elliptic curve E, by using P generates a group G of order q. Later OC chooses $x_\alpha \in Z_q^*$ as its master secret key (private key) and $X_\alpha = x_\alpha P$ as its public key. OC selects secure hash functions: $H_1 : \{0,1\}^* \rightarrow Z_q^*$, $H_2 : \{0,1\}^* \rightarrow Z_q^*$. After initialization the following public parameters for system management are published, $params = \{P, p, q, E, F_p, G, X_\alpha, H_1, H_2\}$.

(b) **Key Generation**: After system parameter initialization each smart meter is issued with a private key x_s whose corresponding public key is, $X_s = x_s P$. Similarly the neighborhood or residential area gateway, DAP is securely issued with a pair of public and private key as, X_d and x_d respectively, where $X_d = x_d P$ and $x_d \in Z_q^*$.

(c) **Smart Meter Data Reporting**: The smart meter uses the public keys of the DAP and OC to send a cipher-text for smart metering data report for a particular period of time. For management purposes, the electricity reports from SMs to OC are transmitted in 10 - 15 min intervals. Let e_i represents an individual user's electricity consumption report in the specified time interval for $i = 1, 2, \cdots, n$. The SM deployed at household is in charge of all home appliances, and it collects the usage data, and then encrypts it before transmission to the DAP. Before SM sends the data, e_i to the receiver, it is necessary to map the message e_i to a point on the elliptic curve using a homomorphic mapping as $m_i = e_i G$. Then later, the SM carries out data encryption procedure by using the public keys of DAP and OC which are, x_d, X_α respectively as follows:

- Step 1: Each smart meter generates random numbers, $r_d, r_\alpha \in Z_q^*$, and then computes $C_1 = r_d P$, $C_2 = r_\alpha P$, $C_3 = r_\alpha X_\alpha \oplus ID_s$ then encrypts the data as, $C_i = m_i + ID_s + r_\alpha X_\alpha + r_d X_d$. Furthermore, the SM computes the authentication and integrity check component $\sigma_1 = H_1(C_i||C_1||X_s||t_s)x_s$, by using its own private key as a signature and public with its timestamp t_s

- Step 2: Then the smart meter sends $M_1 = \{C_1, C_2, C_3, C_i, X_s, t_s, \sigma_1\}$ to the DAP of its residential area network. Thus C_i is the actual cipher-text of the reporting data.

Furthermore, its worth to take note that, C_1, C_2 and C_3 can be pre-computed to expedite the smart meter data reporting process.

(d) **DAP Data Aggregation**: Upon receipt of the message M_1 from SM, the gateway (DAP) has to perform aggregation on the encrypted data and partially decrypt the message to leverage the OC of some computational overhead by using its private key. DAP proceeds as follows:

- step 1: Selects the timestamps t_d, and checks the validity by computing $|t_d - t_s| < \delta$ otherwise it aborts the session, where δ is the acceptable threshold of time lapse.
- step 2: Checks both the source and message authenticity and validity of the received message by verifying that $\sigma_1.P = H_1(C_i||C_1||X_s||t_s)X_s$ and it proceeds only if the computation holds
- Step 3: First computes, $Y_d = x_d C_1$ and then DAP carries out data aggregation by partial decryption of the message from SM to OC as follows:

$$\begin{aligned} C_T^{agg} &= \sum_{i=1}^{n} (C_i - Y_d) \\ &= \sum_{i=1}^{n} (m_i + ID_s + r_\alpha X_\alpha) \end{aligned}$$

  which is the result of additive homomorphic operation on an encrypted data
- step 4: Then DAP computes a signature of the partially decrypted cipher-text, including both its private key and public key in the computation , $\sigma_2 = H_2(C_T^{agg}||C_2||C_3||X_d||t_d)x_d$.
- step 5: Now DAP forwards the partially decrypted message $M_2 = \{C_2, C_3, C_T^{agg}, X_d, t_d, \sigma_2\}$ to the OC.

(e) **Data Recovery**: Upon receipt of the partially decrypted message M_2 from DAP, the OC carries the following steps in the process of decrypting.

- step 1: Selects the timestamp t_α, and checks if $|t_\alpha - t_d| < \delta$ holds and then proceeds, otherwise it quits the process.
- step 2: Furthermore the OC checks the integrity of the partially decrypted ciphertext from the DAP by verifying, $\sigma_2.P = H_2(C_T^{agg}||C_2||C_3||X_d||t_d)X_d$.
- step 3: Afterwards, OC carries out user identification extraction by computing:

$$ID_s = C_3 \oplus C_2.x_\alpha$$

- step 4: Finally, the OC is able to securely extract the total electricity consumption reporting data for each user by carrying out the computation:

$$Cons_{Total} = C_T^{agg} - \sum_{i=1}^{n} (ID_s + x_\alpha C_2)$$

$$= \sum_{i=1}^{n} m_i + \sum_{i=1}^{n} ID_s + r_\alpha X_\alpha - \sum_{i=1}^{n} (ID_s + x_\alpha C_2)$$
$$= \sum_{i=1}^{n} (m_i + ID_s + r_\alpha X_\alpha) - \sum_{i=1}^{n} (ID_s + x_\alpha r_\alpha P)$$
$$= \sum_{i=1}^{n} m_i$$

It should be noticed that, if any checking in the scheme's steps fails to hold then the scheme is immediately aborted. A summary of message flow for the scheme between concerned entities is depicted in Table 2.

Now we can look at the correctness of the main computations of DAP and OC to verify he consistence of the messages.

DAP Computation Correctness:

$$\begin{aligned} Y_d &= x_d C_1 \\ &= x_d r_d P \\ &= r_d x_d P \\ &= r_d X_d \end{aligned}$$

The DAP computes partial decryption:

$$\begin{aligned} C_T^{agg} &= \sum_{i=1}^{n} (C_i - Y_d) \\ &= \sum_{i=1}^{n} (m_i + ID_s + r_\alpha X_\alpha + r_d X_d + k_i - x_d C_1) \\ &= \sum_{i=1}^{n} (m_i + ID_s + r_\alpha X_\alpha + r_d X_d + k_i - x_d r_d P) \\ &= \sum_{i=1}^{n} (m_i + ID_s + r_\alpha X_\alpha + r_d X_d + k_i - r_d x_d P) \\ &= \sum_{i=1}^{n} (m_i + ID_s + r_\alpha X_\alpha + k_i) \\ &= \sum_{i=1}^{n} (m_i + ID_s + r_\alpha X_\alpha) + \sum_{i=1}^{n} k_i \\ &= \sum_{i=1}^{n} (m_i + ID_s + r_\alpha X_\alpha) \end{aligned}$$

OC Computation Correctness:
Extraction of ID_s:

$$ID_s = C_3 \oplus C_2 x_\alpha$$

$$
\begin{aligned}
&= (r_\alpha X_\alpha \oplus ID_s) \oplus (r_\alpha P)x_\alpha \\
&= (r_\alpha X_\alpha \oplus ID_s) \oplus (r_\alpha X_\alpha) \\
&= r_\alpha X_\alpha \oplus ID_s \oplus r_\alpha X_\alpha \\
&= ID_s \oplus r_\alpha X_\alpha \oplus r_\alpha X_\alpha
\end{aligned}
$$

Electricity data extraction:

$$
\begin{aligned}
Cons_{Total} &= C_T^{agg} - \sum_{i=1}^{n}(ID_s + x_\alpha C_2) \\
&= \sum_{i=1}^{n}(m_i + ID_s + r_\alpha X_\alpha) - \sum_{i=1}^{n}(ID_s + x_\alpha C_2) \\
&= \sum_{i=1}^{n}(m_i + ID_s + r_\alpha X_\alpha) - \sum_{i=1}^{n}(ID_s + x_\alpha r_\alpha P) \\
&= \sum_{i=1}^{n} m_i + \sum_{i=1}^{n}(ID_s + r_\alpha X_\alpha) - \sum_{i=1}^{n}(ID_s + X_\alpha r_\alpha) \\
&= \sum_{i=1}^{n} m_i + \sum_{i=1}^{n}(ID_s + r_\alpha X_\alpha) - \sum_{i=1}^{n}(ID_s + X_\alpha r_\alpha) \\
&= \sum_{i=1}^{n} m_i
\end{aligned}
$$

The actual total consumption data $\sum_{i=1}^{n} e_i$ is extracted from $\sum_{i=1}^{n} m_i G$ by applying the Pollard lombda operation since $e_i = m_i G$.

Thus, this shows that the computations are consistently correct as the entities in the scheme carry out them.

4 Security Analysis and Performance Evaluation

We will first give formal security and privacy proofs by random oracle model of the proposed electricity power consumption aggregation scheme and then conduct a performance evaluation in terms of computational overhead and communication overhead to demonstrate the feasibility of the proposed scheme's merits.

4.1 Security Requirements Analysis

Here we will analyze the satisfaction of security requirements of the proposed scheme and also compare it with other related works.

1. **Authentication**: The messages plying between SM and DAP as well as the ones between DAP and OC are intrinsically authenticated by either party to assure the communication is with the rightful or intended entity. The

Table 2. LPPDA scheme message flow.

Smart meter (SM)	Data Aggregator (DAP)	Operation Center (OC)
SM Computes:		
$C_1 = r_d P$, s.t $r_d \in Z_q^*$		
$C_2 = r_\alpha P$, s.t $r_\alpha \in Z_q^*$		
$C_3 = r_\alpha X_\alpha \oplus ID_s$		
$C_i = m_i + ID_s + r_\alpha X_\alpha + r_d X_d + k_i$		
$\sigma_1 = H_1(C_i \| \| C_1 \| \| X_s \| \| t_s) x_s$		
	$\xrightarrow{M_1 = \{C_1, C_2, C_3, C_i, t_s, X_s, \sigma_1\}}$	
	DAP Data Aggregation	
	Checks if:	
	$\|t_d - t_s\| < \delta$	
	$\sigma_1.P = H_1(C_i \| \| C_1 \| \| X_s \| \| t_s) X_s$	
	Then DAP computes:	
	$Y_d = x_d C_1$	
	$C_T^{agg} = \sum_{i=1}^{n} (C_i - Y_d)$	
	$\sigma_2 = H_2(C_T^{agg} \| \| C_2 \| \| C_3 \| \| X_d \| \| t_d) x_d$	
		$\xrightarrow{M_2 = \{C_2, C_3, C_T^{agg}, X_d, t_d, \sigma_2\}}$
		OC Data Recovery
		Checks if:
		$\|t_\alpha - t_d\| < \delta$
		$\sigma_2.P = H_2(C_T^{agg} \| \| C_2 \| \| C_3 \| \| X_d \| \| t_d) X_d$
		Then OC computes:
		$ID_s = C_3 \oplus C_2.x_\alpha$
		$Cons = C_T^{agg} - \sum_{i=1}^{n} (ID_s + x_\alpha C_2)$

message, $\{C_1, C_2, C_3, C_i, X_s, t_s, \sigma_1\}$, is authenticated by DAP by checking that $\sigma_1.P = H_1(C_i||C_1||X_s||t_s)X_s$ in the message M_1 against the timestamp t_s. This checking additionally affirms the sender as legitimate since it works as a digital signature as σ_1 is generated by SM's private key and verified with its public key. Similarly, the receiver is authenticated before carrying out partial decryption by ensuring that it is able to calculate $Y_d = x_d C_1$, which is used in the partial decryption. The message is designed such that only the intended DAP, is the one able to compute $Y_d = x_d C_1$, since the computation requires a private key of DAP. Hence this entails the partially decrypted data C_T^{agg} sent to OC is an authenticated one. In turn upon extraction of user ID_s, C_T^{agg} is further used to calculate the total consumption report, which ensures authentication of the smart meter user, ID_s, to OC. Consequently the message itself and the source authentication is satisfied.

2. **Data Confidentiality**: In order to obtain the electricity consumption reporting data the OC must compute $Cons_{Total} = C_T^{agg} - \sum_{i=1}^{n} ID_s(x_\alpha C_2)$ of which is infeasible to be calculated by any malicious party without the specific knowledge of the private key x_α and the hidden identity ID_s. An attacker will have to solve an intractable problem of ECCDH to obtain x_α from the public key $X_\alpha = x_\alpha P$ as well as extracting ID_s from $C_3 = r_\alpha X_\alpha \oplus ID_s$, which is impossible. Therefore, the proposed scheme does meet the data confidentiality requirement.
3. **Data Integrity**: In the proposed scheme the ciphertext integrity is ensured by inclusion of verifiable signatures and freshness checking values. In the ciphertext M_1, σ_1 is used to check the integrity of the sent message to DAP

by verifying whether $\sigma_1.P = H_1(C_i||C_1||X_s||t_s)X_s$ holds. This subsequently checks the integrity of the sent message as well as the source . Likewise σ_2 upholds the integrity of the ciphertext M_2 from DAP to OC and if the verification of $\sigma_2.P = H_2(C_T^{agg}||C_2||C_3||X_d||t_d)X_d$ fails the session will be aborted. Since no attacker can generate a valid σ_1 associated with the ciphertext C_i as it is infeasible to calculate the components r_d, r_α, ID_s and m_i from the publicly accessible message M_1. Similarly, σ_2 cannot be forged to fake the authenticity of the message M_2, as it is impossible for an attacker to fabricate C_T^{agg} without knowledge of DAP's private key x_d. Thus, the proposed scheme provides message integrity.

4. **Consumer Privacy and Anonymity**: In the proposed scheme an attacker has no access to any consumer related information from the plying messages from SM to OC. The user related information in $C_3 = r_\alpha X_\alpha \oplus ID_s$ and $C_i = m_i + ID_s + r_\alpha X_\alpha + r_d X_d$ from SM to DAP is transmitted in concealed form and cannot be obtained by an attacker before solving an intractable ECCDH problem to obtain the ID_s from $C_3 = r_\alpha X_\alpha \oplus ID_s$. This is well known hard problem for an attacker to resolve, hence the proposed scheme LPPDA provides consumer privacy and anonymity befitting the wireless communication between SM and DAP, where eavesdropping can easily be done in the public channel.
5. **Attack Resilience**: The proposed LPPDA scheme has the merits to resist against well known attacks such as: replay attack, impersonation attack and man-in-the-middle attack.
   - *Replay attack*: User electricity data report from SM to OC for a particular time slot $i = 1, 2, \cdots, n$ is hashed in $\sigma_1 = H_1(C_i||C_1||X_s||t_s)x_s$ which comprises of fresh random numbers r_d and r_α for each session run. Similarly, OC verifies the freshness of M_2 by checking that $\sigma_2 = H_2(C_T^{agg}||C_2||C_3||X_s||t_d)x_d$ holds. Thus, the timestamp t_s and t_d checks against any replay attempts by an attacker on the messages M_1 and M_2 respectively. Therefore by this procedure, LPPDA could resist replay attack.
   - *Impersonation attack*: No attacker can produce the ciphertext $C_i = m_i + ID_s + r_\alpha X_\alpha + r_d X_d$ associated with the identity ID_s without knowledge of either random number r_d or the private key, x_d. Thus, an attacker has no idea of the user who is to be impersonated on. By this, LPPDA is able to resist any impersonation attack.
   - *Man-in-the-middle attack*: From the ciphertext M_2 obtained by OC originating from SM, only the OC could extract the identity ID_s that authenticates SM to OC in the process $ID_s = C_3 \oplus C_2.x_\alpha$. Furthermore, the integrity check components, $\sigma_1 = H_1(C_i||C_1||X_s||t_s)x_s$ and $\sigma_2 = H_2(C_T^{agg}||C_2||C_3||X_d||t_s)x_d$ are generated with sender's private key to be verified by its corresponding public key, meaning that no attacker can formulate a verifiable ciphertext impersonating ID_s to DAP and OC, because of lack of private keys of the victims. Therefore an attacker in between SM and DAP cannot generate valid C_i and σ_1 that can be verified. Similarly the messages from DAP to OC are secure from an attacker

in between as C_T^{agg} cannot be forged for a targeted identity ID_s as it is verified by σ_2.

Furthermore, the merits of the proposed LPPDA scheme are compared against other related schemes [44,46,48,61,65] with respect to security features satisfied, as shown in the Table 3. In the table the ✓ symbol shows that a particular security feature is satisfied while the ✗ symbol shows that the security feature is not satisfied. For easy representation in the table the following features: confidentiality, authentication, integrity, anonymity, replay attack, impersonation attack, internal attack and man-in-the-middle attack are denoted as, F1, F2, F3, F4, F5, F6, F7, F8 respectively. So clearly, it is evident from the Table 3, that LPPDA scheme has merits on satisfaction of security requirements over related works.

Table 3. Security comparison

Security features comparison								
Scheme	F1	F2	F3	F4	F5	F6	F7	F8
[44]	✓	✗	✗	✗	✓	✓	✗	✗
[46]	✓	✓	✓	✗	✓	✓	✗	✓
[48]	✓	✓	✓	✗	✓	✓	✓	✓
[65]	✓	✓	✓	✗	✓	✓	✓	✓
[61]	✓	✓	✓	✗	✓	✓	✗	✗
LPPDA	✓	✓	✓	✓	✓	✓	✓	✓

4.2 Performance Evaluation

In this section, performance analysis is carried out in comparison with relevant related schemes [44,46,48,61,65] based on computational cost required for SM, DAP and OC as well as the communication cost analysis of the channels SM to DAP and DAP to OC respectively are presented.

4.3 Computation Cost

The computational times variables used in LPPDA, for carrying out specific cryptographic operations are adapted from [48] which were simulated on MIRACL Crypto SDK, which is a multi-precision integer, rational arithmetic C/C++ library, [58] run on a 2.53 GHz i5 CPU, 4 GB RAM on a 64 bit windows 10 operating system. The experimentation used a 160 bit key length for security parameters chosen from G over F_p. The data of the average quantified running times is depicted in given Table 4, which was obtained after taking the averages of 1000 runs in the simulation.

Table 4. Estimated running times for different operations in milliseconds (ms) averaged after 1000 runs

Notations	Description of operation	Execution time
T_{sm-ecc}	Scalar multiplication in ECC	0.38
T_{DL}	Solving DL operetion in mod$p$	0.64
T_{mtp}	Map to a point hash function	3.58
T_{n^2}	Exponentiation in Z_{n^2}	2.02
T_p	Bilinear paring	10.31
T_{p-D}	Paillier public key decryption	11.82
T_{p-E}	Paillier public key encryption	9.89
T_{exp-p}	Exponentiation in p	0.13
T_m	Scalar multiplication in bilinear paring	1.42
T_n	Exponentiation in Z_n	0.58

In this regard the computation cost comparison will be done based on the quantification weights in Table 4, averaged after 1000 experiment runs in order to estimate time complexities for different operations. In this evaluation, lightweight operations like hash functions, hash chain, concatenation, point addition and XORing are disregarded, since they have negligible computation overload. So, our focus will be on heavier computation operations only such as: map-to-point hash function, bilinear parings, paillier public key encryption and paillier public key decryption operations among others as portrayed in the Table 5.

Table 5. Computed average running times of schemes for different operations in milliseconds

Scheme	SM	DAP	OC	Total time
[44]	$1T_m + (n+1)T_{n^2} + 1T_{mtp}$ $= 2.02n + 7.02ms$	$(w+1)T_p + 1T_m + (w+1)T_{mtp}$ $= 13.89w + 15.70ms$	$2T_p + 1T_{mtp}$ $= 36.04ms$	$13.89w + 2.02n + 58.74ms$
[46]	$3T_{exp-p} + 3T_m$ $= 4.65ms$	$(3w+1)T_{exp-p} + 3T_m$ $= 0.39w + 4.39ms$	$4T_{exp-p} + 3T_m$ $= 4.78ms$	$0.39w + 13.82ms$
[48]	$6T_{sm-ecc}$ $= 2.28ms$	$(2w+2)T_{sm-ecc}$ $= 0.76w + 0.76ms$	$4T_{sm-ecc} + 1T_{DL}$ $= 2.16ms$	$0.76w + 5.2ms$
[65]	$5T_{sm-ecc}$ $= 1.9ms$	$2T_p + (w+1)T_{sm-ecc}$ $= 0.38w + 21ms$	$1T_p + (2z+1)T_{sm-ecc} + 1T_{mtp}$ $= 4.78z + 14.26ms$	$= 0.38w + 4.78z + 37.16ms$
[61]	$3T_{exp-p} + 3T_m$ $= 4.65ms$	$(3w+1)T_{exp-p} + 3T_m$ $= 0.39w + 4.39ms$	$4T_{exp-p} + 3T_m$ $= 4.78ms$	$0.39w + 13.82ms$
LPPDA	$6T_{sm-ecc}$ $= 2.28ms$	wT_{sm-ecc} $= (1.14w)ms$	$(z+3)T_{sm-ecc}$ $= (0.76z)ms$	$= 0.76w + 0.38z + 3.68ms$

The computational cost is categorized in three levels according to the entity carrying out the operation which are the smart meters, data aggregation points and the operations center for the sake of clarity and fairness in the evaluation. Thus, we will evaluate the computational cost of the SM, DAP and OC individually before calculating the total computation incurred in the schemes [44,46,48,61,65] and the summary of the comparison is done in Table 5.

Firstly, we calculate computational cost required for a SM in different schemes. The computational cost required for SM in [44] is $(n + 1)$ exponentiation operations in Z_{n^2}, one map-to-point hash operation and one scalar multiplication in bilinear pairing. So the amount of the computation cost for a SM is $(n+1)T_{n^2} + 1T_m + 1T_{mtp} = 2.02 + 7n.02ms$. In schemes [46] and [61], there are three exponentiation operations in G_1 and three scalar multiplication in bilinear pairing. So the amount of computation needed for SM is $3T_{exp-p} + 3T_m = 4.65ms$. While in [48], SM needs six scalar multiplication which amounts to $6T_{sm-ecc} = 2.28ms$ and in [65] five scalar multiplication operations are needed with an overhead of $5T_{sm-ecc} = 1.9ms$. On the other hand in LPPDA, SM requires six scalar multiplication with the computational cost of $6T_{sm-ecc} = 2.28ms$. However, although [65] has the same SM computational cost as in the proposed scheme, in LPPDA scheme the computation of which, C_1, C_2 and C_3 can be pre-computated prior to the data reporting time and this further reduces the computational requirement. Thus, our design would require $3T_{sm-ecc} = 1.14ms$ for real-time computations, therefore it does not overload the SM computation overhead during the run time, besides it ensures provision of user anonymity, unlike in the other schemes.

Secondly, the computation cost for DAP in these schemes is such that in [44] there are $w+1$ bilinear pairing operations where w is the numbers of smart meters in a particular neighborhood, one scalar multiplication in bilinear pairing and $w+1$ map-to-point hash operations. Consequently, the DAP requires $(w+1)T_p + 1T_m + (w+1)T_{mtp} = 13.89w + 15.7ms$ computation cost. Whereas in [46] and [61], the schemes need $3w+1$ exponentiation operations in G_1 and three scalar multiplication operations in bilinear pairing and DAP's computation cost is $(3w+1)T_{exp-p} + 3T_m = 0.39w + 4.39ms$. In scheme [48], the DAP requires $(2w+2)$ scalar multiplication operations with a computation cost overhead of $(2w+2)T_{sm-ecc} = 0.76w + 0.76ms$. Whereas in [65], two bilinear pairing operations and $(w+1)$ scalar multiplication are needed so the computation cost overhead for DAP is $2T_p + (w+1)T_{sm-ecc} = 0.38w + 21ms$. On the other hand the DAP in the proposed LPPDA, requires $3w$ scalar multiplication operations and one exponentiation operation, thus the computation cost for DAP is $3wT_{sm-ecc} = (1.14w)ms$.

Thirdly, we analyze the computation cost required for OC in each scheme and we assume that there are z, DAPs in the wide area network where OC is in-charge. In Lu et' al [44] scheme, OC requires to two bilinear pairing operations, one paillier public key decryption operation and one map-to-point hash operation. The computation cost required in this case is $2T_p + 1T_{p-D} + 1T_{mtp} = 36.04$. Whereas in [46] and [61], the OC requires four exponentiation operations in G_1, three scalar multiplication operations in bilinear pairings, so the computation cost overhead is $4T_{exp-p} + 3T_m = 4.78ms$. In Ming et al. [48], on the other hand, OC requires four scalar multiplication operations and DL operation $mod\ p$, with the computational cost of $4T_{sm-ecc} + 1T_{DL} = 2.16ms$. In [65], the OC needs one bilinear pairing operation, $(2z+1)$ scalar multiplication operation and one map-to-point hash operation, resulting into the computation cost of $1T_p + (2z+1)T_{sm-ecc} + 1T_{mtp} = 4.78z + 14.26ms$. While in the proposed

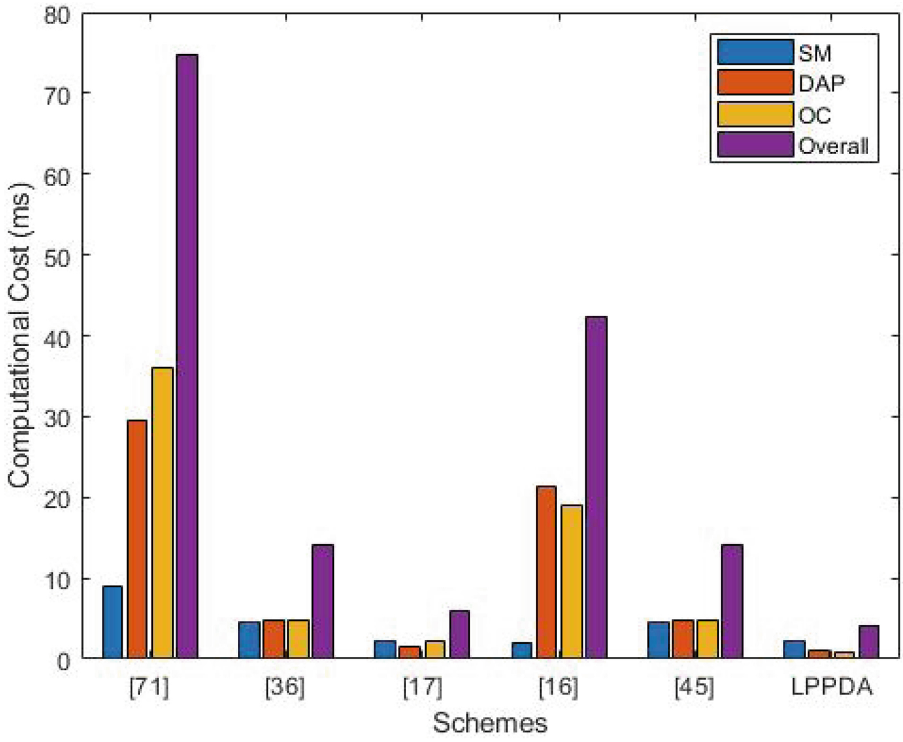

Fig. 2. Computation cost for the schemes.

LPPDA, the OC requires $2zT_{sm-ecc}$ scalar multiplication operations yielding to computation cost of $2zT_{sm-ecc} = 0.76zms$.

Thus, with reference to the computation loads of each entity reflected in Table 5, where w stands for the number of smart meters reporting to DAP and z stands for number of DAPs reporting to OC. The proposed LPPDA scheme has overall lower computation cost as depicted in Fig. 2, with both w and z equated to one for simplicity, even though [48] and [65] have a slightest edge for lower SM overhead. The computation results considers the incurred overhead per individual entity, that is SM, DAP and OC separately. The actual run time computation cost for SM is much lower to the tune of $1.14ms$ upon excluding pre-computated operations unlike the $2.28ms$ shown in the Fig. 2, of which is still lower than comparable schemes hence affirming the efficiency of the proposed LPPDA scheme.

4.4 Communication Cost

Now we will evaluate the communication cost of the proposed LPPDA scheme in relation to the schemes [44,46,48,61,65], however ECC based schemes are generally more efficient and saves more bandwidth since with reduced short length security parameter with high security achievement. Different mathematical structures have different lengths for the security parameters. The elements in G_1, G_T, G, Z_q^*, Z_n and Z_{n^2} respectively, have the lengths, 512 bits, 1024 bits, 160 bits, 160 bits, 1024 bits and 2048 bits. The length of a hash function is 160 bits whilst that for an identity and timestamp is 32 bits. The communication cost analysis between SM and DAP is given as follows. SM sends the message $\{C_i, \sigma_i, RA, U_i, TS\}$ to the DAP in [46], such that $C_i \in Z_{n^2}$, $\sigma_i \in G_1$ whereby RA, U_i and TS are all 32 bit long each. SM to DAP communication cost is therefore, $|C_i + \sigma_i + RA + U_i + TS| = 2048 + 512 + 32 + 32 + 32 = 2659$ bits. Whereas in scheme [44], the SM sends the message C_i of length $|C_i| = 1024$ bits. While in [48] et 'al, the message $\{C_{1,i}, C_{2,i}, ID_i, L_i, v_i, T\}$ is sent from SM to DAP with the overall communication load of $|C_{1,i} + C_{2,i} + ID_i + L_i + v_i + T| = 160 + 160 + 32 + 160 + 160 + 32 = 704$, where $C_{1,i}, C_{2,i}, L_i \in G$, $v_i \in Z_q^*$ and further, ID_i and T are both 32 bit identity and timestamp respectively. On the other hand, in the scheme [65], SM sends the message $R_{it}||S_{it}||t||ID_{U_i}||\sigma_{it}$ to corresponding DAP. By considering the same conditions for [65] elements in the group G has length of 160 bits. Thus, the communication bandwidth required for [65] is, $|R_{it} + S_{it} + t + ID_{U_i} + \sigma_{it}| = 160 + 160 + 32 + 32 + 160 = 544$ bits. In the scheme [61], SM sends the message $\{C_i, H_i\}$ to DAP with $C_i \in G_T$ and H_i a one way hash function. The computation cost requirement is therefore, $|C_i + H_i| = 1024 + 160 = 1184$ bits. On the on other hand, in the proposed LPPDA, SM sends the message $M_1 = \{C_1, C_2, C_3, C_i, X_s, t_s, \sigma_1\}$ to DAP all from the group G. So the required communication bandwidth is $|C_1 + C_2 + C_3 + C_i + X_s + t_s + \sigma_1| = 160 + 160 + 160 + 160 + 160 + 32 + 160 = 992$.

Lastly, the communication cost between DAP and OC in the concerned schemed is analyzed as follows. DAP sends the message $\{C, \sigma_g, RA, GW, TS\}$ to OC in [44], where $C \in Z_{n^2}$. $\sigma_g \in G_1$, and further, RA and TS are both 32 bits elements. Thus, the communication cost is $|C + \sigma_g + RA + GW + TS| = 2048 + 512 + 32 + 32 + 32 = 2659$ bits. The DAP in [46] sends C as the message to OC where $C \in G_T$, so the communication cost is $|C| = 1024$ bits. On the other hand, DAP sends the message $\{C_1, C_2, ID_{GW}, L_{GW}, v_{GW}, T\}$ to OC in [48], with $C_1, C_2, L_{GW} \in G$, $v_{GW} \in Z_q^*$ and further, ID_{GW} and T are both 32 bits long elements. Hence the communication cost amounts to $|C_1 + C_2 + ID_{GW} + L_{GW} + v_{GW} + T| = 160 + 160 + 32 + 160 + 160 + 32 = 704$ bits. In [65], the DAP sends the message $R_t||S_t||t||ID_{AG_j}||\sigma_{AG_j,t}$ to OC, so the communication cost is $|R_t + S_t + t + ID_{AG_j} + \sigma_{AG_j,t}| = 160 + 160 + 32 + 32 + 1024 = 1408$ bits. While in Tahir et al. [61], DAP sends the message, $\{C, H\}$ to OC with $C \in G_T$ and H as a 160 bits one-way hansh function, therefore yielding the communication cost of $|1024 + 160| = 1184$ bits. On the other hand, in the proposed LPPDA the DAP sends the message $M_2 = \{C_2, C_3, C_T^{agg}, X_d, \sigma_2, t_d\}$ to the OC which has the communication cost of $|C_2 + C_3 + C_T^{agg} + \sigma_2 + X_d + t_d| = 160 + 160 + 160 + 160 + 160 + 32 = 832$ bits.

Based on the comparison values of the concerned schemes summarized in Table 6, we obtain the graphical representation in Fig. 3 to show the performance of the proposed scheme.

Table 6. Communication cost of SM-DAP and DAP-OC transmissions

Scheme	SM-DAP	DAP-SM
[44]	1024	2659
[46]	2659	1024
[48]	704	704
[65]	544	1408
[61]	1184	1184
LPPDA	992	832

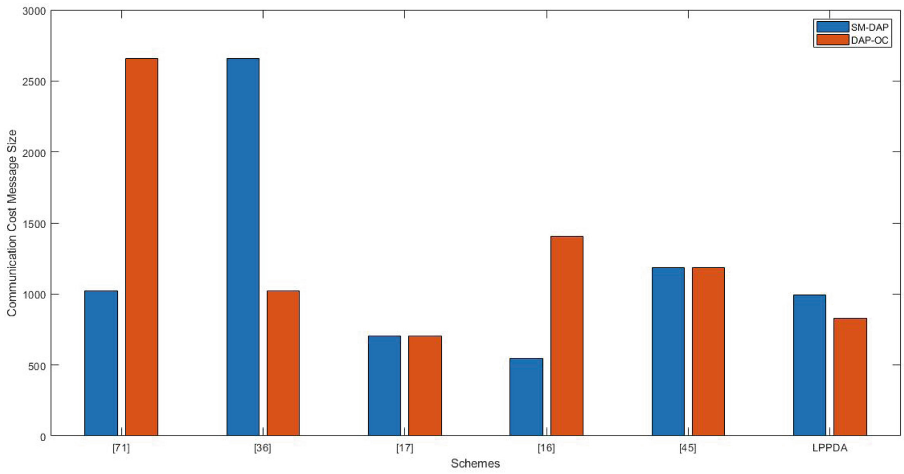

Fig. 3. Communication cost for the schemes.

Thus, from the analysis, the proposed LPPDA has an overall efficient communication cost to the related works as seen in Fig. 3 relative to the scheme in [48] in terms of SM-DAP communication segment cost while our scheme, LPPDA, has better merits overall the comparable schemes. In regards to the DAP-OC communication segment, our scheme has an efficient bandwidth load of 832 bits which providing desirable security requirements. Therefore, from the performance evaluation, in general our scheme is lightweight in both computation cost and communication cost, making it an ideal security scheme in a complex communication based network environment as that of smart grid.

Remark: By considering pre-computed operations, the overall communication and computation performance efficiency of LPPDA scheme surpasses that of relevant related works with a significant margin. The performance of the proposed work is achieved with a great advantage comparatively, as much as twice to the better performing scheme, while achieving user anonymity without sacrificing computation efficiency.

This work is on going and we will progressively devote on including other versatile security techniques that ensure a robust modern grid with its applications and functionalities. In our further works, multi-dimension data aggregation techniques will be considered, to ensure security, classification and uploading of different types of electrical appliance measurements data to the utility company. We consider incorporating a privacy-preserving data aggregation focusing on a block-chain smart grid network environment and its functionality requirements. We also would endeavor to employ smart grid security schemes adaptive to dynamic machine learning based techniques in optimal scheduling and billing.

5 Conclusion

In this paper, we proposed a lightweight privacy-preserving data aggregation scheme based on elliptic cryptography, for smart grid communications. The framework is a two leveled architecture with the residential area data aggregator acting as an intermediate party between the smart meter and the utility service company. Unlike many other researches the proposed scheme protects user's real-time power consumption and identity from privacy and security breaches in an efficient manner by using a ECCDH problem in the encryption algorithm and additive homomorphic encryption. The comparison analysis shows that the scheme is efficient on overall system performance as well as robust in security, hence ideal for an IoT enhanced smart grid system. Our future work, related to this discipline will focus on secure and efficient multidimensional data preservation in smart grid network environment. Aspects of versatile technologies such as; block-chain, lattice cryptography and machine learning techniques will be considered, in order to realize a modern smart grid.

Acknowledgments. This research was possible because of the financial support from the World Bank, at the ACEIoT (African Center of Excellence in Internet of Things) which resides at the College of Science and Technology of the University of Rwanda, KN Street Nyarugenge, P.O. Box 3900, Kigali.

References

1. Abbasinezhad-Mood, D., Nikooghadam, M.: Design and hardware implementation of a security-enhanced elliptic curve cryptography based lightweight authentication scheme for smart grid communications. Futur. Gener. Comput. Syst. **84**, 47–57 (2018)
2. Abdallah, A., Shen, X.S.: A lightweight lattice-based homomorphic privacy-preserving data aggregation scheme for smart grid. IEEE Trans. Smart Grid **9**(1), 396–405 (2016)
3. Akyol, B.A., Kirkham, H., Clements, S.L., Hadley, M.D.: A survey of wireless communications for the electric power system. Technical report, Pacific Northwest National Lab (PNNL), Richland, WA, United States (2010)
4. Ali, I., Khan, E., Sabir, S.: Privacy-preserving data aggregation in resource-constrained sensor nodes in internet of things: a review. Future Comput. Inform. J. **3**(1), 41–50 (2018)

5. Alvarez-Alvarado, M.S., Jayaweera, D.: Reliability-based smart-maintenance model for power system generators. IET Gener. Trans. Distrib. **14**(9), 1770–1780 (2020)
6. Badra, M., Zeadally, S.: Lightweight and efficient privacy-preserving data aggregation approach for the smart grid. Ad Hoc Netw. **64**, 32–40 (2017)
7. Bae, M., Kim, K., Kim, H.: Preserving privacy and efficiency in data communication and aggregation for AMI network. J. Netw. Comput. Appl. **59**, 333–344 (2016)
8. Bao, H., Lu, R.: Comment on "privacy-enhanced data aggregation scheme against internal attackers in smart grid". IEEE Trans. Industr. Inf. **12**(1), 2–5 (2015)
9. Bekara, C.: Security issues and challenges for the IoT-based smart grid. In: FNC/MobiSPC, pp. 532–537 (2014)
10. Boudia, O.R.M., Senouci, S.M., Feham, M.: Elliptic curve-based secure multidimensional aggregation for smart grid communications. IEEE Sens. J. **17**(23), 7750–7757 (2017)
11. Chim, T.W., Yiu, S.M., Hui, L.C., Li, V.O.: Pass: privacy-preserving authentication scheme for smart grid network. In: 2011 IEEE International Conference on Smart Grid Communications (SmartGridComm), pp. 196–201. IEEE (2011)
12. Chim, T.W., Yiu, S.M., Li, V.O., Hui, L.C., Zhong, J.: PRGA: privacy-preserving recording & gateway-assisted authentication of power usage information for smart grid. IEEE Trans. Dependable Secure Comput. **12**(1), 85–97 (2014)
13. Das, D., Rout, D.K.: Adaptive algorithm for optimal real-time pricing in cognitive radio enabled smart grid network. ETRI J. **42**, 585–595 (2020)
14. Deng, R., Xiao, G., Lu, R., Chen, J.: Fast distributed demand response with spatially and temporally coupled constraints in smart grid. IEEE Trans. Industr. Inf. **11**(6), 1597–1606 (2015)
15. Desai, S., Alhadad, R., Chilamkurti, N., Mahmood, A.: A survey of privacy preserving schemes in IoE enabled smart grid advanced metering infrastructure. Clust. Comput. **22**(1), 43–69 (2019). https://doi.org/10.1007/s10586-018-2820-9
16. Dong, X., Zhou, J., Alharbi, K., Lin, X., Cao, Z.: An elgamal-based efficient and privacy-preserving data aggregation scheme for smart grid. In: 2014 IEEE Global Communications Conference, pp. 4720–4725. IEEE (2014)
17. Dwork, C., Kenthapadi, K., McSherry, F., Mironov, I., Naor, M.: Our data, ourselves: privacy via distributed noise generation. In: Vaudenay, S. (ed.) EUROCRYPT 2006. LNCS, vol. 4004, pp. 486–503. Springer, Heidelberg (2006). https://doi.org/10.1007/11761679_29
18. Eissa, M.: New protection principle for smart grid with renewable energy sources integration using WiMAX centralized scheduling technology. Int. J. Electr. Power Energy Syst. **97**, 372–384 (2018)
19. Eissa, M., Ali, N.: Performance evaluation of IEEE 802.16 real time polling and unsolicited grant service scheduling for protecting transmission and subtransmission systems with multi-terminals. Electric Power Syst. Res. **167**, 48–57 (2019)
20. Ekanayake, J.B., Jenkins, N., Liyanage, K., Wu, J., Yokoyama, A.: Smart Grid: Technology and Applications. Wiley, Hoboken (2012)
21. Erkin, Z., Troncoso-Pastoriza, J.R., Lagendijk, R.L., Pérez-González, F.: Privacy-preserving data aggregation in smart metering systems: an overview. IEEE Signal Process. Mag. **30**(2), 75–86 (2013)
22. Farash, M.S., Attari, M.A.: A secure and efficient identity-based authenticated key exchange protocol for mobile client-server networks. J. Supercomput. **69**(1), 395–411 (2014). https://doi.org/10.1007/s11227-014-1170-5

23. Ferrag, M.A.: EPEC: an efficient privacy-preserving energy consumption scheme for smart grid communications. Telecommun. Syst. **66**(4), 671–688 (2017). https://doi.org/10.1007/s11235-017-0315-2
24. Ferrag, M.A., Ahmim, A.: Security solutions and applied cryptography in smart grid communications. IGI Global (2016)
25. Ferrag, M.A., Maglaras, L.A., Janicke, H., Jiang, J., Shu, L.: A systematic review of data protection and privacy preservation schemes for smart grid communications. Sustain. Urban Areas **38**, 806–835 (2018)
26. FitzPatrick, G.J., Wollman, D.A.: NIST interoperability framework and action plans. In: IEEE PES General Meeting, pp. 1–4. IEEE (2010)
27. Ford, V., Siraj, A., Rahman, M.A.: Secure and efficient protection of consumer privacy in advanced metering infrastructure supporting fine-grained data analysis. J. Comput. Syst. Sci. **83**(1), 84–100 (2017)
28. Garcia, F.D., Jacobs, B.: Privacy-friendly energy-metering via homomorphic encryption. In: Cuellar, J., Lopez, J., Barthe, G., Pretschner, A. (eds.) STM 2010. LNCS, vol. 6710, pp. 226–238. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-22444-7_15
29. Goel, S., Hong, Y.: Security challenges in smart grid implementation. In: Goel, S., Hong, Y. (eds.) Smart Grid Security. SC, pp. 1–39. Springer, London (2015). https://doi.org/10.1007/978-1-4471-6663-4_1
30. Gong, Y., Cai, Y., Guo, Y., Fang, Y.: A privacy-preserving scheme for incentive-based demand response in the smart grid. IEEE Trans. Smart Grid **7**(3), 1304–1313 (2015)
31. Gope, P., Sikdar, B.: Lightweight and privacy-friendly spatial data aggregation for secure power supply and demand management in smart grids. IEEE Trans. Inf. Forensics Secur. **14**(6), 1554–1566 (2018)
32. He, D., Kumar, N., Zeadally, S., Vinel, A., Yang, L.T.: Efficient and privacy-preserving data aggregation scheme for smart grid against internal adversaries. IEEE Trans. Smart Grid **8**(5), 2411–2419 (2017)
33. He, D., Zeadally, S., Wang, H., Liu, Q.: Lightweight data aggregation scheme against internal attackers in smart grid using elliptic curve cryptography. Wireless Communications and Mobile Computing 2017 (2017)
34. Hu, C., Huo, Y., Ma, L., Liu, H., Deng, S., Feng, L.: An attribute-based secure and scalable scheme for data communications in smart grids. In: Ma, L., Khreishah, A., Zhang, Y., Yan, M. (eds.) WASA 2017. LNCS, vol. 10251, pp. 469–482. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-60033-8_41
35. Humayed, A., Lin, J., Li, F., Luo, B.: Cyber-physical systems security-a survey. IEEE Internet Things J. **4**(6), 1802–1831 (2017)
36. Hur, J., Koo, D., Shin, Y.: Privacy-preserving smart metering with authentication in a smart grid. Appl. Sci. **5**(4), 1503–1527 (2015)
37. Jo, H.J., Kim, I.S., Lee, D.H.: Efficient and privacy-preserving metering protocols for smart grid systems. IEEE Trans. Smart Grid **7**(3), 1732–1742 (2015)
38. Koo, J., Lin, X., Bagchi, S.: RL-BLH: learning-based battery control for cost savings and privacy preservation for smart meters. In: 2017 47th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN), pp. 519–530. IEEE (2017)
39. Li, D., Aung, Z., Williams, J., Sanchez, A.: P3: privacy preservation protocol for automatic appliance control application in smart grid. IEEE Internet Things J. **1**(5), 414–429 (2014)
40. Li, S., Xue, K., Yang, Q., Hong, P.: PPMA: privacy-preserving multisubset data aggregation in smart grid. IEEE Trans. Industr. Inf. **14**(2), 462–471 (2017)

41. Liu, J.K., Yuen, T.H., Au, M.H., Susilo, W.: Improvements on an authentication scheme for vehicular sensor networks. Expert Syst. Appl. **41**(5), 2559–2564 (2014)
42. Liu, Y., Cheng, C., Gu, T., Jiang, T., Li, X.: A lightweight authenticated communication scheme for smart grid. IEEE Sens. J. **16**(3), 836–842 (2015)
43. Liu, Y., Guo, W., Fan, C.I., Chang, L., Cheng, C.: A practical privacy-preserving data aggregation (3PDA) scheme for smart grid. IEEE Trans. Industr. Inf. **15**(3), 1767–1774 (2018)
44. Lu, R., Alharbi, K., Lin, X., Huang, C.: A novel privacy-preserving set aggregation scheme for smart grid communications. In: 2015 IEEE Global Communications Conference (GLOBECOM), pp. 1–6. IEEE (2015)
45. Lu, R., Heung, K., Lashkari, A.H., Ghorbani, A.A.: A lightweight privacy-preserving data aggregation scheme for fog computing-enhanced IoT. IEEE Access **5**, 3302–3312 (2017)
46. Lu, R., Liang, X., Li, X., Lin, X., Shen, X.: EPPA: an efficient and privacy-preserving aggregation scheme for secure smart grid communications. IEEE Trans. Parallel Distrib. Syst. **23**(9), 1621–1631 (2012)
47. Mahmood, K., Chaudhry, S.A., Naqvi, H., Kumari, S., Li, X., Sangaiah, A.K.: An elliptic curve cryptography based lightweight authentication scheme for smart grid communication. Futur. Gener. Comput. Syst. **81**, 557–565 (2018)
48. Ming, Y., Zhang, X., Shen, X.: Efficient privacy-preserving multi-dimensional data aggregation scheme in smart grid. IEEE Access **7**, 32907–32921 (2019)
49. Molina-Markham, A., Shenoy, P., Fu, K., Cecchet, E., Irwin, D.: Private memoirs of a smart meter. In: Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems For Energy-efficiency in Building, pp. 61–66 (2010)
50. Morello, R., Mukhopadhyay, S.C., Liu, Z., Slomovitz, D., Samantaray, S.R.: Advances on sensing technologies for smart cities and power grids: a review. IEEE Sens. J. **17**(23), 7596–7610 (2017)
51. Paillier, P.: Public-key cryptosystems based on composite degree residuosity classes. In: Stern, J. (ed.) EUROCRYPT 1999. LNCS, vol. 1592, pp. 223–238. Springer, Heidelberg (1999). https://doi.org/10.1007/3-540-48910-X_16
52. Pawar, A., Rahane, S.: Opportunities and challenges of wireless communication technologies for smart grid applications. Int. J. Comput. Netw. Wirel. Mobile Commun. (IJCNWMC) **3**(1), 289–296 (2013)
53. Phiri, K.K., Kim, H.: Linear secret sharing scheme with reduced number of polynomials. Secur. Commun. Netw. (2019)
54. Qu, H., Shang, P., Lin, X.J., Sun, L.: Cryptanalysis of a privacy-preserving smart metering scheme using linkable anonymous credential. IACR Cryptol. ePrint Arch. **2015**, 1066 (2015)
55. Rivest, R.L.: Cryptography. In: Algorithms and Complexity, pp. 717–755. Elsevier (1990)
56. Savaglio, C., Ganzha, M., Paprzycki, M., Bădică, C., Ivanović, M., Fortino, G.: Agent-based internet of things: State-of-the-art and research challenges. Futur. Gener. Comput. Syst. **102**, 1038–1053 (2020)
57. Saxena, N., Choi, B.J., Lu, R.: Authentication and authorization scheme for various user roles and devices in smart grid. IEEE Trans. Inf. Forensics Secur. **11**(5), 907–921 (2015)
58. Scott, M.: Multiprecision integer and rational arithmetic cryptographic library (2003)
59. Shen, H., Zhang, M., Shen, J.: Efficient privacy-preserving cube-data aggregation scheme for smart grids. IEEE Trans. Inf. Forensics Secur. **12**(6), 1369–1381 (2017)

60. Sultan, S.: Privacy-preserving metering in smart grid for billing, operational metering, and incentive-based schemes: a survey. Comput. Secur. **84**, 148–165 (2019)
61. Tahir, M., Khan, A., Hameed, A., Alam, M., Khan, M.K., Jabeen, F.: Towards a set aggregation-based data integrity scheme for smart grids. Ann. Telecommun. **72**(9–10), 551–561 (2017). https://doi.org/10.1007/s12243-017-0602-7
62. Tan, S., De, D., Song, W.Z., Yang, J., Das, S.K.: Survey of security advances in smart grid: a data driven approach. IEEE Commun. Surv. Tutor. **19**(1), 397–422 (2017)
63. Tonyali, S., Akkaya, K., Saputro, N., Uluagac, A.S., Nojoumian, M.: Privacy-preserving protocols for secure and reliable data aggregation in IoT-enabled smart metering systems. Futur. Gener. Comput. Syst. **78**, 547–557 (2018)
64. Ustun, T.S., Khan, R.H., Hadbah, A., Kalam, A.: An adaptive microgrid protection scheme based on a wide-area smart grid communications network. In: 2013 IEEE Latin-America Conference on Communications, pp. 1–5. IEEE (2013)
65. Vahedi, E., Bayat, M., Pakravan, M.R., Aref, M.R.: A secure ECC-based privacy preserving data aggregation scheme for smart grids. Comput. Netw. **129**, 28–36 (2017)
66. Vaidya, B., Makrakis, D., Mouftah, H.: Provisioning substation-level authentication in the smart grid networks. In: 2011-MILCOM 2011 Military Communications Conference, pp. 1189–1194. IEEE (2011)
67. Vallent, T.F., Hanyurwimfura, D., Mikeka, C.: Efficient certificate-less aggregate signature scheme with conditional privacy-preservation for vehicular ad hoc networks enhanced smart grid system. Sensors **21**(9), 2900 (2021)
68. Wu, D., Zhou, C.: Fault-tolerant and scalable key management for smart grid. IEEE Trans. Smart Grid **2**(2), 375–381 (2011)
69. Xia, J., Wang, Y.: Secure key distribution for the smart grid. IEEE Trans. Smart Grid **3**(3), 1437–1443 (2012)
70. Yun, M., Yuxin, B.: Research on the architecture and key technology of internet of things (IoT) applied on smart grid. In: 2010 International Conference on Advances in Energy Engineering, pp. 69–72. IEEE (2010)
71. Zuo, X., Li, L., Peng, H., Luo, S., Yang, Y.: Privacy-preserving multidimensional data aggregation scheme without trusted authority in smart grid. IEEE Syst. J. **15**(1), 395–406 (2020)