The Scientific Temper (2024) Vol. 15 (4): 3214-3224

E-ISSN: 2231-6396, ISSN: 0976-8653

Doi: 10.58414/SCIENTIFICTEMPER.2024.15.4.31

https://scientifictemper.com/

RESEARCH ARTICLE

A smart grid data privacy-preserving aggregation approach with authentication

Manpreet Kaur*, Shweta Mishra

Abstract

Authentication of smart grid privacy-preserving aggregation addresses two of the key privacy and security issues of the smart grids: user data confidentiality and grid node communication safety. The proposed study elaborates on a new approach to data aggregation with authentication in smart grid systems for the safe and efficient exchange of information. The proposed solution would apply techniques, such as homomorphic encryption along with advanced cryptographic techniques, to calculate encrypted data without leaking sensitive information. Data and device integrity are more likely to be maintained when using better authentication techniques like blockchain and quantum key distribution (QKD). This dual layered aggregation with privacy-preserving combined with robust authentication can strengthen the smart grids against unauthorized access and data tampering, along with other cyber-attacks. The results show that the proposed approach for aggregation in smart meters is more accurate and useful in terms of data as compared to the conventional approaches. As far as mean relative error (MRE) is concerned, the MRE of the proposed layer model is 0.0007, which is substantially smaller than the differentially private model (0.0023) and Gaussian model (0.0058). The minimum MRE of the proposed model was achieved in the aggregator layer at 0.0029 compared with the corresponding differentially- private model's 0.0063 and Gaussian model's 0.0117. As the privacy parameter ε increases, noise levels drop precipitously from 14.137738 for ε = 0.1 to 0.282786 for ε = 5.0. The proposed methodology improves smart grid data aggregation with a balance between privacy and accuracy.

Keywords: Smart grid, Privacy-preserving aggregation, Cryptographic techniques, Homomorphic encryption, Cyber-attacks, Smart meters, Data privacy.

Introduction

Smart grids have become one of the most important and swiftly rising technological developments in modern energy distribution systems (Collier, 2016). In contrast to the traditional one-way electrical grid, where the utility company only communicates with its consumers and vice versa, smart grids start based on a very similar but more advanced two-way flow of information and energy (Alaba *et al*., 2017). This new technology allows for so much more effective and efficient delivery of energy, with real-time data analysis and much greater involvement with renewable energy sources (Alotaibi *et al*., 2020). It was created to help face growing energy management challenges, balance generation and consumption, minimize carbon emissions, and reduce power outages. However, as this technology advances, so do the challenges it presents, particularly in the areas of data privacy and security (see Figure 1) (Koo *et al*., 2017; Wang *et al*., 2017).

In smart grid systems, enormous amounts of data are generated and transmitted between nodes: smart meters, control centers, and utility providers (Wang *et al*., 2017; Guan *et al*., 2019). Such data contains sensitive information relating to the electricity usage patterns by the users, which could lead to gross privacy violations if compromised. Unauthorized access to such information might further spawn malicious activities like interfering with energy distribution, causing financial loss, and even power outages (El Mrabet *et al*., 2018). The significantly increasing role of data aggregation in optimizing energy distribution surfaces the need for robust mechanisms

Department of Computer Science and Applications, Desh Bhagat University, Mandi Gobindgarh, Punjab, India.

***Corresponding Author:** Manpreet Kaur, Department of Computer Science and Applications, Desh Bhagat University, Mandi Gobindgarh, Punjab, India, E-Mail: manpreetk209@gmail.com

How to cite this article: Kaur, M., Mishra, S. (2024). A smart grid data privacy-preserving aggregation approach with authentication. The Scientific Temper, **15**(4):3214-3224.

Doi: 10.58414/SCIENTIFICTEMPER.2024.15.4.31

Source of support: Nil

Conflict of interest: None.

Received: 10/10/2024 **Accepted:** 26/11/2024 **Published** : 20/12/2024

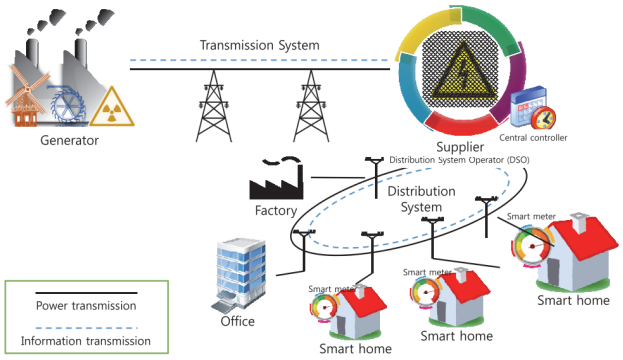

Figure 1: An overview of smart grid architecture (Koo *et al.*,2017)

that safeguard the private and secure exchange of data in smart grid networks (Zhang & Zhang, 2017).

Privacy-Preserving Aggregation Approaches

Despite the vulnerabilities of smart grid systems, there is a growing need for privacy-preserving aggregation techniques (Wang *et al.*, 2021). Privacy-preserving aggregation approaches present schemes that promise aggregated data transmitted between smart meters and utility providers to be anonymous while individual user data remains protected from disclosure (Abdallah & Shen, 2016). Determining efficient energy distribution by utility providers after analyzing aggregated data requires aggregated data analysis using privacy-preserving aggregation techniques by utility providers while keeping consumer data confidential (Tonyali *et al.*, 2018; Zhang *et al.*, 2016).

A simple solution could be to build cryptographic mechanisms allowing the points of aggregate data to be computed such that the secret for every user is his corresponding data (Kaaniche & Laurent, 2017). For example, homomorphic encryption allows operations on encrypted data without decrypting it, thus keeping each user's data private at every step of aggregation. Such privacy-preserving mechanisms enable smart grid operators to achieve consumer privacy even as they try to optimize the performance of respective grids (Othman *et al.*, 2015; Butun *et al.*, 2020).

Role of Authentication in Secure Smart Grid Communication

Besides privacy-preserving aggregation, another crucial aspect of securing communications within the smart grid systems is authentication (Gunduz & Das, 2020). This is because authentication ensures only that authorized devices and users can access or transmit data within the grid. This fact is very sensitive against unauthorized access, data manipulation, or even injection of false information into the system (Saxena *et al.*, 2015).

Several device and user authentications are relied on for authenticating smart grids. Digital signatures and public key infrastructure can be put into use for authentication purposes to check the integrity of the messages (Marino *et al.*, 2019). In this respect, smart grids implement the first layer of security protection against cyber-attacks and hence ensure that communications exchanged between smart meters, utility providers, and control centers are safe and reliable (Gunduz & Das, 2018; Kabalci, 2016).

Challenges and Vulnerabilities in Smart Grid Systems

Smart networks present significant problems despite their many benefits, including increased energy economy and the integration of renewable resources with improved stability (Ahmad & Zhang, 2021). The most significant challenge is related to the prospective breach of privacy because large amounts of data are obtained from smart meters and other devices (Abdalzaher *et al.*,2022). For this reason, smart grids are especially vulnerable to cyberattacks, as they rely on internet-based communication channels (Gunduz & Das, 2018). Hackers can intercept or manipulate data to find personal information, and this allows unauthorized people to gain access, steal identities and so forth (Dabrowski *et al.*, 2017).

One of the other main characteristics of smart grids is their decentralized nature. That is, hundreds of thousands of devices or nodes are involved in controlling the entire grid (Kulkarni *et al.*, 2019). This also increases the number of possible access routes because each device or node that has been interconnected in the network can become an entry route by hackers (Kulkarni *et al.*, 2019). There is no standardization of security protocols for each device, and therefore, this worsens the problem. This unregulated access to smart grid systems would eventually compromise the confidentiality of consumers but threaten the stability and reliability of the whole grid (Das & Zeadally, 2019).

Contributions

The proposed study has several important contributions to the field of smart grid security. First, it would come up with a novel approach to private aggregation, which would enable one to collect and analyze energy consumption data in a secure manner without compromising individual privacy. This approach would be designed to defend against numerous cyber threats, such as eavesdropping, data tampering, and unauthorized access.

An effective mechanism of authentication would be introduced to ensure confidentiality in the communication between nodes in the smart grid. Such an authentication mechanism can be designed to be scalable so that it can apply to any range of devices and configurations of smart grids. This study intended to provide a solution to the smart grids regarding privacy and security by integrally incorporating authentication with privacy-preserving aggregation. The case study with real-world applications would be demonstrated with practical applications of the technique using simulations.

Section 1 introduces smart grids, emphasizing data aggregation and privacy-security challenges while outlining the research's aim, scope, and contributions. Section 2 provides a literature review of smart grid technologies, identifying gaps in privacy-preserving methods and authentication mechanisms. Section 3 discusses the methodology, detailing cryptographic techniques and the experimental setup. Section 4 presents the results, comparing the proposed approach to existing methods. Finally, section 5 concludes the research and suggests future directions.

Literature Review

In this section, various related work based on smart grid data privacy-preserving aggregation is discussed below:

Singh and Kumar (2023) aimed to overcome privacy and security issues in the aggregation and classification of smart grid data. A cryptographic technique and machine learning-based model for the secure and privacy-preserving aggregation of smart grid data that facilitates the efficient processing and analysis of that data was proposed. Correct classification was ensured, but there was no unauthorized access to data during aggregation. The results indicated better security for the data and classification accuracy, hence proving that such a model works quite well in protecting smart grid communications.

Chang *et al*. (2023) proposed a realistic scheme that is privacy-preserving along with fault tolerance to address the privacy issues in smart grids. Cryptography techniques and mechanisms for fault tolerance were developed for secure aggregated data transmission inside the grid. Its methodology employed privacy-preserving protocols and fault recovery strategy. Demonstrating breach-free privacy and sustaining the performance of the grid during faults improved privacy protection along with system reliability.

Kserawi *et al*. (2022) proposed a dynamic differential private perturbatory approach in a fog-based environment to overcome the privacy concerns arising in the aggregation of smart grid data. A basic fog architecture was used in conjunction with differential privacy techniques to apply real-time differential-privacy data perturbation on input datasets. The outcomes showed better privacy protection to the users but preserved the data utility, making it an efficient, scalable solution for data aggregation in smart grids.

Ming *et al*., (2022) focused on the issues of privacy and fault tolerance in a smart grid data aggregation technique. It proposed the use of an efficient privacy-preserving data aggregation scheme via homomorphic encryption and mechanisms to achieve fault tolerance, thereby ensuring security and reliability. The methodology was based on making sure that once data is transmitted securely, it can be treated appropriately even when faults do occur. Results were compared among the data aggregation schemes, showing improvements in data privacy, security, and fault tolerance over existing schemes.

Fawaz Ahmad Kserawi (2021) addressed the problem of privacy-preserving data aggregation in smart power grids. The goal was to offer efficient data aggregation with protection assured for user privacy. Cryptographic primitives were used with homomorphic encryption to secure the aggregation of data. Indeed, in the results, it was found that data accuracy as well as communication efficiency within the smart grid system, could be maintained if user privacy was assured.

Mohammadali and Haghighi (2021) sought a method to tackle smart grid privacy issues within data aggregation by providing a fault-tolerant homomorphic encryption scheme. Based on methodology, several dimensions are used for the enhancement of security and fault tolerance in the aggregation of metering data. The result showed that the proposed scheme could maintain data privacy and address secure and efficient aggregation of metering data even in faulty conditions.

Khan *et al*. (2021) addressed privacy concerns and fault tolerance issues in the improved smart grids with fog support. A privacy-preserving data aggregation methodology was designed, with the aid of homomorphic encryption coupled with fault tolerance, to achieve the security of data in transmission and node failure tolerance as well. For this reason, the approach might improve data privacy and fault tolerance about the smart grid performance on reduced computation overhead and efficient communication.

Zuo *et al*. (2020) addressed the challenge of privacy-preserving multidimensional data aggregation in smart grids without depending on any trusted authority. The scheme used homomorphic encryption combined with random masking techniques to keep the data confidential. The proposed methodology might protect the user's privacy and allow the correctness of data aggregation. Results showed better privacy preservation and security than others, together with a lower computation overhead.

Fan *et al*., (2020) presented a decentralized scheme and applied blockchain technology that could ensure security in the aggregation process of smart grid data without revealing the details of the individual users through applying blockchain technology along with homomorphic encryption. The results pointed out that data privacy was improved yet maintained its efficiency in the aggregation process, thus providing a scalable solution for secure smart grid communication.

Guo *et al*. (2020) addressed the issues of mobile users' privacy and security in smart grids. An author proposed a privacy-preserving aggregation and authentication scheme using homomorphic encryption techniques together with bilinear pairing techniques. Thus, the methodology also maintained the confidentiality of data transmission without

disclosing individual user data. The outcomes showed the ensured protection of user privacy, as the scheme maintains system efficiency and security during the operation of the smart grid.

Based on the literature review, the following research gaps can be identified in the context of privacy-preserving data aggregation and security in smart grid systems:

While addressing privacy and fault tolerance, the study does not account for real-time performance and delays during fault recovery (Chang *et al*., 2023).

Improvements in privacy and fault tolerance are noted, but the research lacks consideration of communication overhead in large-scale smart grid environments (Ming *et al*., 2022).

Although privacy is maintained, the work does not explore the implications of integrating machine learning for enhanced efficiency (Kserawi, 2021).

The scheme protects privacy but lacks a detailed analysis of its impact on real-time data aggregation efficiency (Fan *et al*., 2020).

Research Objective

A privacy-preserving aggregation method with authentication for smart grid data might have the following research objectives:

To design and implement a robust data aggregation framework that ensures privacy preservation while maintaining the efficiency of smart grid operations.

To develop and integrate advanced authentication mechanisms, such as quantum key distribution (QKD) or blockchain, to secure communication between smart grid components.

To assess the scalability and adaptability of the proposed privacy-preserving aggregation approach across diverse smart grid architectures and operational scales.

To provide effective authentication and threat detection, the suggested approach's resistance against several cybersecurity risks, such as data breaches, illegal access, and tampering, requires evaluation.

Research Methodology

This section discusses the purpose, significance, dataset, and workflow charts and algorithms that are employed in the proposed methodology of the research.

Significance of the Proposed Approach

The proposed work aims to design a privacy-preserving aggregation technique for smart grids with considerations for very effective authentication mechanisms to negate the problems of data confidentiality and unauthorized access. To this end, it proposed the development of cryptographic techniques for secure data aggregation and then protocols of authentication to ascertain that nodes of a grid communicate with one another safely. Such a mechanism is necessary to make smart grids truly secure as they are starting to play an even more pivotal role in critical infrastructure. With the rising generation of data in smart grids, high-strength privacy and security measures can ensure such a case where no cyber-attack and causing energy distribution disruption takes place to ensure that the system works in a highly secure and efficient manner.

Dataset Description

The UMass Smart Home dataset (Barker *et al*., 2012) contains minute-by-minute energy use records from a few residential homes. It reports aggregate energy consumption in the home, as well as statistics on the use of about 20 individual appliances. It captures, over many months, the detailed trends in daily and seasonal consumption. It contains millions of data points, and the average is 1,440 data points that are generated by each household daily. The dataset might include abundant contextual information, for example, occupancy rates and weather forecasts; hence, it is the best dataset for the study and development of smart grid solutions that use privacy.

Techniques Used

This section defines the technique used in the present study to evaluate and enhance privacy preservation in smart grid systems:

Quantum key distribution

QKD is based on quantum cryptography and uses physics to protect the distribution of symmetric encryption keys (Sonko *et al*., 2024). The approach guarantees that any intrusion attempt can be discovered since seeing the quantum states utilized for transmission disrupts the connection. The confidentiality of the key distribution procedure is preserved since this disruption notifies the conversing parties of the existence of an interceptor (Khanna & Khanna, 2016). Figure 2 shows the procedure for distributing quantum keys.

In the context of smart grids, QKD can significantly enhance security by guaranteeing that the cryptographic keys used to secure communications between various grid components are resistant to eavesdropping (Kaur & Mishra, 2024). Integrating QKD into smart grids makes them more secure and resilient, which is especially important in the face of advanced cyber-attacks that might harm

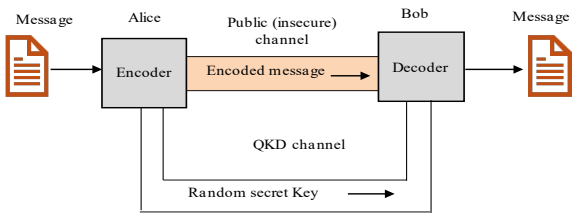

Figure 2: Quantum Key Distribution. https://www.drishtiias.com/daily-news-analysis/quantum-key-distribution/print_manually

vital infrastructure (Alshowkan *et al*., 2022). Enforcing the grid's defense mechanisms against unwanted access and data breaches, QKD is used in smart grid settings by creating a robust feedback system to continually monitor and authenticate machine-to-machine connections. By integrating these systems, researchers can safeguard critical infrastructure and set up a system to identify and respond to threats in real time, making sure the smart grid is safe and running well (Thakur *et al*., 2016).

Reference Control and Support System

The security and longevity of smart grids are greatly improved by utilizing it, which plays an essential role in detecting possible risks by comparing incoming data with established security criteria (Zibaeirad *et al*., 2024). At the central processing step, reference control and support system (RCISS) checks and analyzes the input data using QKD algorithms once it produces an alert (Singh, 2021). By doing so, one can be confident that no unauthorized or insecure data would be sent for examination. The RCISS has a feedback mechanism that can detect and respond to emerging dangers in real-time to keep the smart grid system stable and secure (Yu *et al*., 2024). This system's authentication and verification features help to forestall unwanted access and lessen the likelihood of harm caused by harmful actions.

GIS-based Emergency Alarm

Smart grid situational awareness and reaction skills are enhanced by an emergency alarm system that is based on GIS (Anevlavis, 2022). Smart grids might use geographic information system (GIS) technology to effectively map and monitor real-time data on electricity usage, outages, and other important occurrences (Liu *et al*., 2016). The GIS-based system can swiftly identify the site and provide comprehensive details to the command center once an emergency is detected, such as a power outage or unlawful access to data. Because of this, one could respond with pinpoint accuracy and speed, lessening the severity of the incident (Tomaszewski, 2020). Better decision-making and smart grid infrastructure resilience are both aided by the system's ability to aggregate and analyze geographical data.

High-Performance Computing Resource (HPC)

High-performance computing resources are essential for effectively managing the massive amounts of real-time data generated by smart power grids (Wang *et al*., 2018). It is essential to keep smart grids' operating efficiency and security under check, and HPC systems make that possible by efficiently computing and analyzing massive information. For instance, during the grid's threat grouping and sorting process, HPC resources analyze real-time sensor data and anticipate possible problems more quickly than in real time. With this capacity, risks might be quickly identified and mitigated, guaranteeing that the grid can operate reliably and continuously (Li & Yan, 2022). Power distribution optimization and grid resilience enhancement against cyber-attacks and other disturbances are both aided by HPC systems' capability for sophisticated modeling and simulation.

Proposed Methodology

The proposed workflow starts from the RCISS, feeding collected data into the system. From this point, the unit control processing technique QKD and a GIS-based emergency alarm decide whether there is a threat in the data or not. If no threat is established, then data enters a routine process and goes to output.

The system performs a more detailed scan of the threat if identified. Once confirmed, the threat is categorized using HPC resources, and an emergency alert is started to allow manual clearing of the threat. This stage of manual clearance has two possibilities: either threat elimination or the activation of restoration measures in case the threat persists.

In elimination cases, data restoration is followed using RCISS parameters. The mitigation process is then enforced if the threat is of a persisting nature or is to be mitigated; again, this more updated knowledge source is used. Once the threat has been neutralized, the system guarantees data restoration as per the predefined RCISS parameters, thus ensuring the integrity and safety of the data within the system. This cyclic approach enforces continuous checking and subsequent response followed by restoration. The flowchart (see Figure 3) describes a complete workflow of

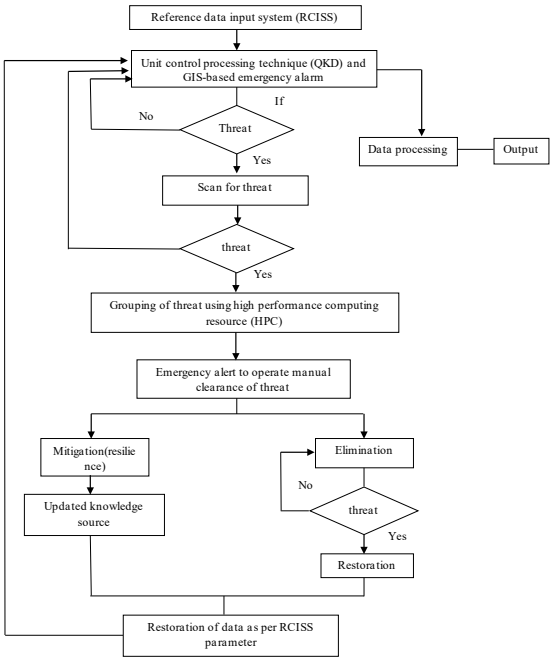

Figure 3: Flowchart of proposed methodology

the data processing and threat management process by using several advanced technologies.

Proposed algorithm

Proposed algorithm: Secure threat-resilient smart grid data aggregation algorithm

Step 1: Initial data input and reference system setup

Reference control and support system (RCISS) Initialization:

Let D be the dataset containing real-time grid metrics

$$D = \left\{ d_1, d_2, \ldots, d_n \right\}$$

Step 2: Unit control processing and initial threat detection

Quantum key distribution (QKD) and GIS-based emergency alarm:

- *Encryption*

Use QKD for secure communication. The key K is generated using

$$K = 2 \times A + \frac{log \frac{Q}{2}}{\log 2} + S\frac{Q}{2} + L$$

- *Threat detection*

For each data point $d_i \in$ D, check for threats T

$$T = \begin{cases} Yes \quad if \exists t_i \epsilon d_i \left(threat\ detected \right) \\ \\ No\ otherwise \end{cases}$$

Step 3: Threat scanning and verification

- *Scanning for threats*

If T=Yes, scan for the specific nature of the threat t_i:

$$Scan\left(t_i \right) \rightarrow Threat\ confirmed$$

Step 4: High-performance computing (HPC) resource allocation

- *Grouping of threats*

Utilize HPC resources to analyze and categorize threats

$$Group\left(t_i \right) \rightarrow Threat\ Group\ T_g$$

Step 5: Emergency alert and manual intervention

- *Generate emergency alert*

Notify operators for manual clearance:

$$Alert\left(T_g \right)$$

step 6: threat management pathways

- *Mitigation and resilience*

If mitigation is possible, apply resilience measures.

$$Mitigation\left(T_g \right) \rightarrow Resilience$$

- *Elimination and restoration*

If elimination is required, verify and neutralize the threat.

$$Eliminate\left(Tg \right) \rightarrow Check\ for\ remaining\ threats$$

If there are no further threats, proceed to restoration.

$$Restore \rightarrow Normal\ Operations$$

Step 7: Restoration of data and system integrity

- *Data restoration*

Restore data according to RCISS parameters:

$$D_{restored} = RCISS\left(D \right)$$

Step 8: Output and continuous monitoring

- *Final output*

Output processed data for its intended use:

$$Output\left(D_{restored} \right)$$

- *Continuous monitoring*

Implement a feedback loop for continuous threat detection.

$$Monitor \rightarrow Repeat\ from\ Step\ 2$$

Results

In this section, results are discussed with utilized evaluation parameters.

Evaluation Parameters

An evaluation of a privacy-preserving aggregation approach for smart grid data might be done using parameters such as MRE, cumulative distribution function (CDF), global active power (kW), window size, threat score (TS), noise level (Standard Deviation), and privacy loss (PL).

Mean relative error

Mean Relative Error compares forecasts to actual values. It is especially effective when relative precision is more crucial than absolute accuracy.

$$\boldsymbol{MRE} = \frac{1}{\boldsymbol{n}} \sum_{i=1}^{\boldsymbol{n}} \left| \frac{\boldsymbol{y_i} - \hat{\boldsymbol{y_i}}}{\boldsymbol{y_i}} \right| \qquad (1)$$

Where, n is number of observations, actual value is denoted by y_i, and predicted value by the model is presented by $\hat{y_i}$.

Cumulative distribution function

According to its probability distribution, the CDF calculates the chance that a real-valued random variable X is smaller than or equal to x.

$$F(x) = P(X \leq x) \quad (2)$$

Where, $F(x)$ is Cumulative distribution function, P shows Probability, and X presents Random variable.

Global active power

Global Active Power is the total power utilized by all grid-connected devices. It is essential for energy utilization and smart grid functioning.

Window size

Window size is the data analysis timeframe. Larger window widths minimize data noise and smooth out swings in smart grid data processing.

Threat score

TS estimates threat detection system efficacy in smart grid security binary categorization. It assesses the ratio of genuine positive threats to true positives, false negatives, and false positives.

$$TS = \frac{\text{True positives}}{\text{True positives} + \text{False negatives} + \text{False positives}} \quad (3)$$

Noise Level (Standard Deviation)

Noise Level measures data privacy-preserving noise. Differential Privacy masks data values with noise to preserve privacy. Privacy and data usefulness depend on this noise's standard deviation.

$$\sigma_{noise} = \frac{1}{\epsilon} \quad (4)$$

Privacy loss

PL estimates the privacy loss from privacy-preserving methods like Differential Privacy. It analyzes the differential privacy parameter (ϵ) and a modest chance (δ) of privacy promise violation.

$$PL = \delta + \epsilon \quad (5)$$

Performance Analysis

A thorough assessment of the smart grid data privacy-preserving aggregation approach is included in the outcome analysis via the comparison of different models and approaches.

The cumulative frequency of standardized global active power across original data, private model data, and gaussian model data is shown in Figure 4. The original data (blue) and gaussian model data (red) begin accumulation about -10 and reach near complete accumulation around 7. The private model data (green) accumulates about -5 and continues up to 15, showing a wider dispersion. The original and gaussian model data exhibit a sharp rise between 0 and 5, with a cumulative frequency near 1.0×10^6, while the Private Model Data exhibits greater variability, resulting in a wider range of cumulative values.

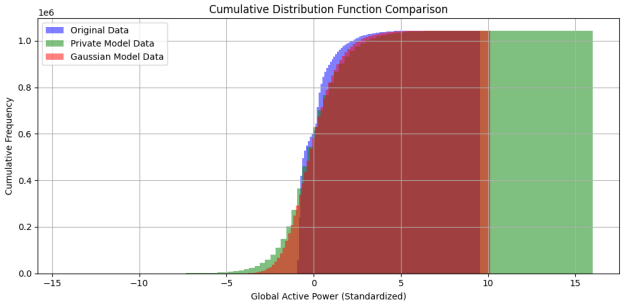

Figure 4: Cumulative distribution function (CDF) comparison graph

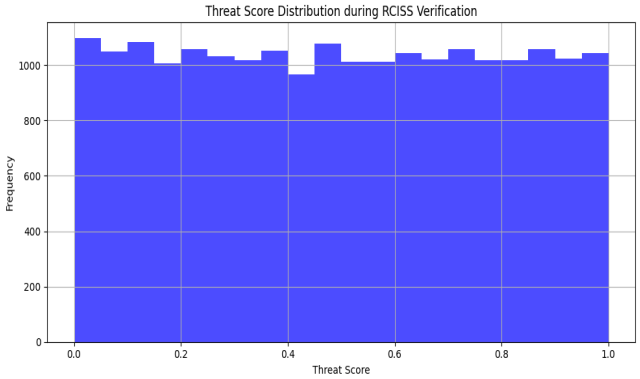

Figure 5: Threat score distribution during RCISS verification

The frequency distribution of threat ratings produced during RCISS verification is shown in Figure 5. With a frequency near to 1000 for each threat score range, the distribution seems uniform. During verification, threat ratings are uniformly distributed over the spectrum, demonstrating a consistent assessment technique that does not favor any scoring range.

In Figure 6, the noise level (standard deviation) reduces dramatically as epsilon goes from 0.1 to 5.0. Noise is high at 14.0 at epsilon = 0.1, implying good privacy but severe data distortion. At 0.5 epsilon, the noise level lowers to 3.5,

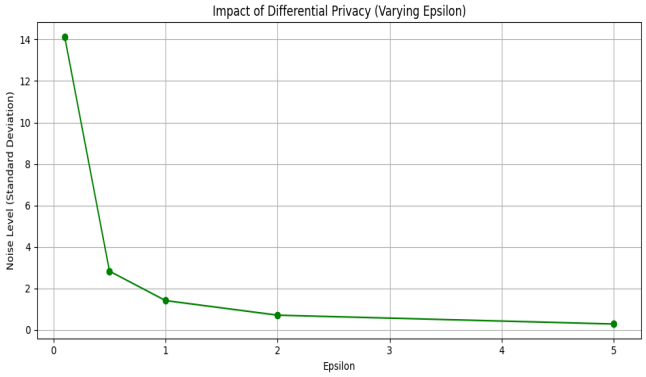

Figure 6: Impact of differential privacy (Varying Epsilon)

Table 1: Privacy budget analysis

Epsilon	*Noise level (Std Dev)*
0.1	14.137738
0.5	2.829226
1.0	1.412544
2.0	0.707682
5.0	0.282786

and at 1.0, it drops to 2.0. With 2.0 and 5.0 epsilon values, noise levels drop to 1.0 and 0.5, respectively, boosting data usefulness but diminishing privacy.

Table 1 shows how various epsilon values affect differential privacy noise (measured as standard deviation). Noise diminishes dramatically as epsilon grows, suggesting a privacy-data accuracy trade-off. At epsilon = 0.1, the noise level is 14.137738, which gives good privacy but distorts data. As the epsilon climbs to 0.5, noise lowers to 2.829226, and at 1.0, it drops to 1.412544. At 2.0 and 5.0 epsilon values, noise levels drop to 0.707682 and 0.282786, respectively, boosting data usefulness but lowering privacy.

Table 2 shows threat-score-based data chunk verification results. Each piece has a threat score between 0 and 1, with lower ratings suggesting fewer dangers. The table reveals that chunks 2, 3, and 6, with threat ratings of 0.262467, 0.030903, and 0.133767, were validated ("True"). However, chunks 1, 4, 5, 7, 8, 9, and 10 failed the verification procedure (marked "False"), suggesting that the system considered them more dangerous. This table shows that lower threat levels lead to successful verification and higher scores for rejection.

Figure 7 plots global active power versus time for original data, proposed model, and traditional Gaussian mechanism smart meter data aggregation. The dashed blue line represents the original data, the baseline. The green line for "proposed model" centralizes and smooths data better, following the original data with less noise, preserving data features and usefulness. In comparison, the orange dotted line indicating the traditional Gaussian mechanism distorts

Table 2: RCISS verification results

Chunk number	*Threat score*	*Verified*
1	0.949915	False
2	0.262467	True
3	0.030903	True
4	0.545557	False
5	0.926117	False
6	0.133767	True
7	0.714076	False
8	0.906133	False
9	0.550916	False
10	0.673351	False

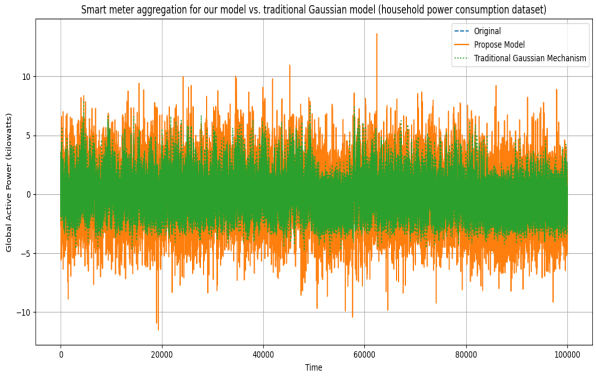

Figure 7: Smart meter data aggregation of various models

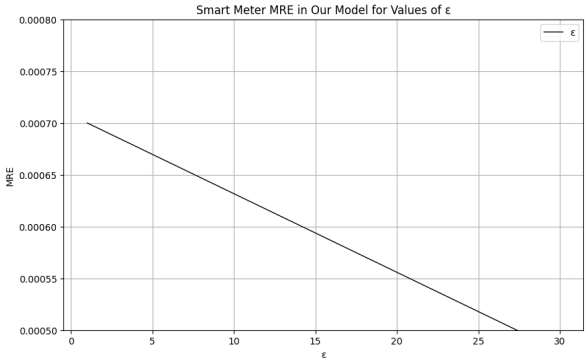

Figure 8: Variable ε values effect on MRE for proposed model

data more due to randomness and noise. The results show that the "Proposed Model" balances data privacy and utility, reduces noise, and handles outliers better than the traditional Gaussian mechanism, making it a more accurate and trustworthy smart meter data aggregation method.

In a smart meter data aggregation paradigm, Figure 8 shows the link between MRE and privacy parameter ϵ. At ϵ=0, the MRE is about 0.00070, showing increased error due to strict privacy regulations. When ϵ reaches 30, MRE decreases to 0.00050, indicating greater accuracy due to loosened privacy restrictions. Lower ϵ values increase privacy but increase MRE, whereas higher ϵ values reduce MRE and improve data accuracy while sacrificing privacy.

Comparative Analysis

Table 3 shows that the proposed model regularly beats the others with the lowest mean relative error (MRE), suggesting better data gathering. At the smart meter layer, the "proposed model" has an MRE of 0.0007, much lower than the differentially private model (0.0023) and gaussian model (0.0058) (Kserawi *et al.*, 2022). At the Aggregator layer, the "proposed model" has the lowest MRE of 0.0029, compared to 0.0063 for the differentially private model and 0.0117 for the gaussian model. These findings show that the "proposed model" is the most accurate and dependable alternative for data integrity during aggregation, whereas the Gaussian Model adds the greatest error and is less effective.

Table 3: Comparative analysis

Model	*Layers*	
	Smart meter	*Aggregators*
Proposed model	0.0007	0.0029
Differentially private model (Kserawi et al., 2022)	0.0023	0.0063
Gaussian model (Kserawi et al., 2022)	0.0058	0.0117

The research found that the smart grid data privacy-preserving aggregation approach balances privacy and utility. The approach protects data without affecting smart grid accuracy or performance by using differential privacy and homomorphic encryption. The suggested model regularly beats previous models in MRE, CDF, and noise level, making it a dependable and effective smart grid security and performance solution. These findings demonstrate the model's usefulness in real-world situations where privacy and accuracy are crucial.

Conclusion and Future Scope

An authentication-based smart grid data privacy-preserving aggregation approach addresses two of the most burning concerns about security and privacy in smart grids: how to enhance user-data secrecy while achieving safe communication between grid nodes. To accomplish safe and efficient interchange of data, this study propounds a new method of privacy-preserving data aggregation with authentication as a solution particularly for smart grid systems. The system can perform complex computations directly over the encrypted data, and the confidential information remains strictly confidential using sophisticated methods of cryptography, including homomorphic encryption, for privacy-preserving data aggregation. The devices and data possessed by the network are authenticated through an authentication process, making use of blockchain-based solutions as well as QKD. Smart grids are much more secure from cyberattacks and, data tampering, and illegal access if a two-layered approach that integrates strong authentication with privacy-preserving aggregation could applied.

The proposed privacy-preserving aggregation methodology offers better accuracy and utility in data as compared to the traditional techniques. The MRE of the proposed smart meter layer model is significantly reduced to 0.0007 as compared to the differentially private model (0.0023) and the Gaussian model (0.0058). Comparing the proposed model at the aggregator layer with the differentially private model (0.0063), and the Gaussian model (0.0117) shows the lowest value of MRE, which is 0.0029. From 14.137738 ($\varepsilon = 0.1$) to 0.282786 ($\varepsilon = 5.0$), the noise levels drop dramatically as the privacy parameter (ε) increases. The approach enhances data aggregation for smart grids by finding a happy medium between privacy and precision. More advanced technologies, such as blockchain and AI-driven threat detection systems, would be applied more thoroughly in future research for better real-time responsiveness, enhanced security, increased system resilience, and robust approach application of the method in evolving infrastructures of smart grids.

Acknowledgment

At the opening of my research paper, We would like to express my profound gratitude to everyone who has assisted us in this quest. We would like to express my heartfelt gratitude to our research supervisor and our principal for providing us with the opportunity to create this research paper on the topic "A smart grid data privacy-preserving aggregation approach with authentication." which allowed us to conduct extensive study and learn about many new things. We also express our heartfelt thanks to our parents and family members, who have always morally and financially supported us. Last but not least, our thanks go to all of my friends who provided excellent advice and direction for the completion of my research paper. Cooperation and constructive criticism were beneficial to them. Finally, we would like to thank everyone who has already been recognized.

References

Abdallah, A., & Shen, X. S. (2016). A lightweight lattice-based homomorphic privacy-preserving data aggregation scheme for smart grid. IEEE Transactions on Smart Grid, 9(1), 396-405.

Abdalzaher, M. S., Fouda, M. M., & Ibrahem, M. I. (2022). Data privacy preservation and security in smart metering systems. Energies, 15(19), 7419. https://doi.org/10.3390/en15197419

Ahmad, T., & Zhang, D. (2021). Using the Internet of things in smart energy systems and networks. Sustainable Cities and Society, 68, 102783. https://doi.org/10.1016/j.scs.2021.102783

Alaba, F. A., Othman, M., Hashem, I. A. T., & Alotaibi, F. (2017). Internet of Things security: A survey. Journal of Network and Computer Applications, 88, 10-28. https://doi.org/10.1016/j.jnca.2017.04.002

Alotaibi, I., Abido, M. A., Khalid, M., & Savkin, A. V. (2020). A comprehensive review of recent advances in smart grids: A sustainable future with renewable energy resources. Energies, 13(23), 6269. https://doi.org/10.3390/en13236269

Alshowkan, M., Evans, P. G., Starke, M., Earl, D., & Peters, N. A. (2022). Authentication of smart grid communications using quantum key distribution. Scientific Reports, 12(1), 12731. https://doi.org/10.1038/s41598-022-16090-w

Anevlavis, T. (2022). A mithrilian approach to safety and robustness of autonomous cyber-physical systems. University of California, Los Angeles.

Barker, S., Mishra, A., Irwin, D., Cecchet, E., Shenoy, P., & Albrecht, J. (2012). Smart*: An open data set and tools for enabling research in sustainable homes. SustKDD, August, 111(112), 108.

Butun, I., Lekidis, A., & dos Santos, D. R. (2020). Security and Privacy in Smart Grids: Challenges, Current Solutions and Future Opportunities. ICISSP, 10, 0009187307330741. DOI: 10.5220/0009187307330741

Chang, Y., Li, J., Lu, N., Shi, W., Su, Z., & Meng, W. (2023). Practical privacy-preserving scheme with fault tolerance for smart grids. IEEE Internet of Things Journal.

Collier, S. E. (2016). The emerging enernet: Convergence of the smart grid with the internet of things. IEEE Industry Applications Magazine, 23(2), 12-16.

Dabrowski, A., Ullrich, J., & Weippl, E. R. (2017, December). Grid shock: Coordinated load-changing attacks on power grids: The non-smart power grid is vulnerable to cyber attacks as well. In Proceedings of the 33rd Annual Computer Security Applications Conference (pp. 303-314). https://doi.org/10.1145/3134600.3134639

Das, A. K., & Zeadally, S. (2019). Data security in the smart grid environment. In Pathways to a smarter power system (pp. 371-395). Academic Press. https://doi.org/10.1016/B978-0-08-102592-5.00013-2

El Mrabet, Z., Kaabouch, N., El Ghazi, H., & El Ghazi, H. (2018). Cyber-security in smart grid: Survey and challenges. Computers & Electrical Engineering, 67, 469-482. https://doi.org/10.1016/j.compeleceng.2018.01.015

Fan, H., Liu, Y., & Zeng, Z. (2020). Decentralized privacy-preserving data aggregation scheme for smart grid based on blockchain. Sensors, 20(18), 5282. https://doi.org/10.3390/s20185282

Guan, Z., Zhang, Y., Zhu, L., Wu, L., & Yu, S. (2019). EFFECT: An efficient flexible privacy-preserving data aggregation scheme with authentication in smart grid. Science China Information Sciences, 62, 1-14. https://doi.org/10.1007/s11432-018-9451-y

Gunduz, M. Z., & Das, R. (2018, September). Analysis of cyber-attacks on smart grid applications. In 2018 international conference on artificial intelligence and data processing (IDAP) (pp. 1-5). IEEE.

Gunduz, M. Z., & Das, R. (2018, September). Analysis of cyber-attacks on smart grid applications. In 2018 international conference on artificial intelligence and data processing (IDAP) (pp. 1-5). IEEE.

Gunduz, M. Z., & Das, R. (2020). Cyber-security on smart grid: Threats and potential solutions. Computer networks, 169, 107094. https://doi.org/10.1016/j.comnet.2019.107094

Guo, F., Cao, Z., Liu, Z., & Cao, N. (2020). A privacy-preserving aggregation and authentication scheme towards mobile users in smart grid. Journal of Shanghai Jiaotong University (Science), 25, 37-43. https://doi.org/10.1007/s12204-019-2137-8

Kaaniche, N., & Laurent, M. (2017). Data security and privacy preservation in cloud storage environments based on cryptographic mechanisms. Computer Communications, 111, 120-141. https://doi.org/10.1016/j.comcom.2017.07.006

Kabalci, Y. (2016). A survey on smart metering and smart grid communication. Renewable and Sustainable Energy Reviews, 57, 302-318. https://doi.org/10.1016/j.rser.2015.12.114

Kaur, M., & Mishra, S. (2024). ENHANCING SECURITY AND RESILIENCE IN SMART GRIDS THROUGH QUANTUM KEY DISTRIBUTION: CHALLENGES AND OPPORTUNITIES." https://doi.org/10.30780/

Khan, H. M., Khan, A., Jabeen, F., & Rahman, A. U. (2021). Privacy preserving data aggregation with fault tolerance in fog-enabled smart grids. Sustainable Cities and Society, 64, 102522. https://doi.org/10.1016/j.scs.2020.102522

Khanna, V. K., & Khanna, V. K. (2016). Cyber security and confidentiality concerns with implants. Implantable Medical Electronics: Prosthetics, Drug Delivery, and Health Monitoring, 209-225. https://doi.org/10.1007/978-3-319-25448-7_11

Koo, D., Shin, Y., & Hur, J. (2017). Privacy-preserving aggregation and authentication of multi-source smart meters in a smart grid system. Applied Sciences, 7(10), 1007. https://doi.org/10.3390/app7101007

Kserawi, F. A. (2021). Privacy-preserving data aggregation in smart power grid systems (Master's thesis). http://hdl.handle.net/10576/21579

Kserawi, F., Al-Marri, S., & Malluhi, Q. (2022). Privacy-preserving fog aggregation of smart grid data using dynamic differentially-private data perturbation. IEEE Access, 10, 43159-43174.

Kulkarni, S., Gu, Q., Myers, E., Polepeddi, L., Lipták, S., Beyah, R., & Divan, D. (2019). Enabling a decentralized smart grid using autonomous edge control devices. IEEE Internet of Things Journal, 6(5), 7406-7419.

Kulkarni, S., Gu, Q., Myers, E., Polepeddi, L., Lipták, S., Beyah, R., & Divan, D. (2019). Enabling a decentralized smart grid using autonomous edge control devices. IEEE Internet of Things Journal, 6(5), 7406-7419.

Li, Y., & Yan, J. (2022). Cybersecurity of smart inverters in the smart grid: A survey. IEEE Transactions on Power Electronics, 38(2), 2364-2383.

Liu, Y. B., Liu, J. Y., Taylor, G., Liu, T. J., Gou, J., & Zhang, X. (2016). Situational awareness architecture for smart grids developed in accordance with dispatcher's thought process: a review. Frontiers of Information Technology & Electronic Engineering, 17(11), 1107-1121. https://doi.org/10.1631/FITEE.1601516

Marino, F., Moiso, C., & Petracca, M. (2019). PKIoT: A public key infrastructure for the Internet of Things. Transactions on Emerging Telecommunications Technologies, 30(10), e3681. https://doi.org/10.1002/ett.3681

Ming, Y., Li, Y., Zhao, Y., & Yang, P. (2022). Efficient Privacy-Preserving Data Aggregation Scheme with Fault Tolerance in Smart Grid. Security and Communication Networks, 2022(1), 5895176. https://doi.org/10.1155/2022/5895176

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

Othman, S. B., Bahattab, A. A., Trad, A., & Youssef, H. (2015). Confidentiality and integrity for data aggregation in WSN using homomorphic encryption. Wireless Personal Communications, 80, 867-889. https://doi.org/10.1007/s11277-014-2061-z

Saxena, N., Choi, B. J., & Lu, R. (2015). Authentication and authorization scheme for various user roles and devices in smart grid. IEEE transactions on Information forensics and security, 11(5), 907-921.

Singh, A. K., & Kumar, J. (2023). A secure and privacy-preserving data aggregation and classification model for smart grid. Multimedia Tools and Applications, 82(15), 22997-23015. https://doi.org/10.1007/s11042-023-14599-4

Singh, K. (2021). Management of smart grids: A review. International Journal for Research in Applied Science & Engineering Technology (IJRASET), 9(6), 1763-1769. https://doi.org/10.22214/ijraset.2021.36734.

Sonko, S., Ibekwe, K. I., Ilojianya, V. I., Etukudoh, E. A., & Fabuyide,

A. (2024). Quantum cryptography and US digital security: a comprehensive review: investigating the potential of quantum technologies in creating unbreakable encryption and their future in national security. Computer Science & IT Research Journal, 5(2), 390-414. https://doi.org/10.51594/csitrj.v5i2.790

Thakur, K., Ali, M. L., Jiang, N., & Qiu, M. (2016, April). Impact of cyber-attacks on critical infrastructure. In 2016 IEEE 2nd International Conference on Big Data Security on Cloud (BigDataSecurity), IEEE International Conference on High Performance and Smart Computing (HPSC), and IEEE International Conference on Intelligent Data and Security (IDS) (pp. 183-186). IEEE.

Tomaszewski, B. (2020). Geographic information systems (GIS) for disaster management. Routledge. https://doi.org/10.4324/9781351034869

Tonyali, S., Akkaya, K., Saputro, N., Uluagac, A. S., & Nojoumian, M. (2018). Privacy-preserving protocols for secure and reliable data aggregation in IoT-enabled smart metering systems. Future Generation Computer Systems, 78, 547-557. https://doi.org/10.1016/j.future.2017.04.031

Wang, J., Wu, L., Zeadally, S., Khan, M. K., & He, D. (2021). Privacy-preserving data aggregation against malicious data mining attack for IoT-enabled smart grid. ACM Transactions on Sensor Networks (TOSN), 17(3), 1-25. https://doi.org/10.1145/3440249

Wang, K., Du, M., Maharjan, S., & Sun, Y. (2017). Strategic honeypot game model for distributed denial of service attacks in the smart grid. IEEE Transactions on Smart Grid, 8(5), 2474-2482.

Wang, K., Du, M., Maharjan, S., & Sun, Y. (2017). Strategic honeypot game model for distributed denial of service attacks in the smart grid. IEEE Transactions on Smart Grid, 8(5), 2474-2482.

Wang, L., Ma, Y., Yan, J., Chang, V., & Zomaya, A. Y. (2018). pipsCloud: High performance cloud computing for remote sensing big data management and processing. Future Generation Computer Systems, 78, 353-368. https://doi.org/10.1016/j.future.2016.06.009

Yu, Z., Feng, J., Tang, S., Liu, Z., Yan, Y., & Luo, N. (2024). Disaster Intelligent Perception and Emergency Command of Power Grid (p. 373). Springer Nature. 10.1007/978-981-99-7236-4

Zhang, L., & Zhang, J. (2017). EPPRD: an efficient privacy-preserving power requirement and distribution aggregation scheme for a smart grid. Sensors, 17(8), 1814. https://doi.org/10.3390/s17081814

Zhang, L., Wang, X., Lu, J., Li, P., & Cai, Z. (2016). An efficient privacy preserving data aggregation approach for mobile sensing. Security and Communication Networks, 9(16), 3844-3853. https://doi.org/10.1002/sec.1546

Zibaeirad, A., Koleini, F., Bi, S., Hou, T., & Wang, T. (2024). A Comprehensive Survey on the Security of Smart Grid: Challenges, Mitigations, and Future Research Opportunities. arXiv preprint arXiv:2407.07966. https://doi.org/10.48550/arXiv.2407.07966

Zuo, X., Li, L., Peng, H., Luo, S., & Yang, Y. (2020). Privacy-preserving multidimensional data aggregation scheme without trusted authority in smart grid. IEEE Systems Journal, 15(1), 395-406.