Computers and Electrical Engineering 124 (2025) 110371

Contents lists available at ScienceDirect

Computers and Electrical Engineering

journal homepage: www.elsevier.com/locate/compeleceng

A comprehensive survey on privacy-preserving technologies for Smart Grids

Hafsa Bibi [a,*], Mehran Abolhasan [a], Justin Lipman [a], Mahrokh Abdollahi [b], Wei Ni [b]

[a] School of Electrical and Data Engineering, University of Technology Sydney, Sydney, New South Wales, 2007, Australia
[b] Commonwealth Scientific and Industrial Research Organisation (CSIRO), Marsfield, Sydney, New South Wales, 2122, Australia

ARTICLE INFO

Keywords:
Cryptography
Federated learning
Power consumption datasets
Privacy
Security
Smart Grids

ABSTRACT

Smart Grid (SG) systems greatly contribute to the reliability, efficiency, and sustainability of traditional power grids. Besides the many promising aspects of the SG, such as cheaper maintenance, more effective peak control, and wider energy markets, the growth in SG infrastructure has led to privacy and security issues because large-scale consumers' data is exchanged through heterogeneous devices connected via public communication networks. In this paper, we present a comprehensive overview of privacy-preserving schemes for SG systems. In particular, we examine and analyze privacy-preserving schemes published or developed in the context of SGs between 2018 and 2024. We discuss several privacy and security requirements for SG systems. Then, we present a novel taxonomy for various schemes proposed to address these issues, considering the privacy goal and Key Characteristics of each scheme. We then offer our recommendations for using these solutions practically in SGs. Additionally, we discuss and compare different publicly available energy consumption datasets. Finally, we investigate the open research challenges and recommendations for further research to ensure privacy in SG systems.

1. Introduction

The realization of smart cities depends on several key domains, including governance, healthcare, security, and energy. Among these, energy plays a crucial role in ensuring sustainability and integrating diverse stakeholders [1]. The transition from traditional power grids to Smart Grids (SGs) is essential for smart cities, as disruptions in electricity supply can halt critical urban utilities. SGs overcome the limitations of conventional grids by significantly enhancing efficiency, reliability, and stability [2]. They boost efficiency by optimizing power generation, minimizing energy wastage, and enabling demand-response mechanisms. Reliability is improved through self-healing capabilities, predictive maintenance, and the integration of distributed energy resources. Additionally, SGs enhance stability by regulating voltage and frequency, incorporating smart inverters, and ensuring a balanced supply–demand system. These advancements make SGs a more resilient, adaptive, and sustainable alternative to traditional power grids [3]. The primary SG attributes include (i) self-healing, (ii) distributed generation, (iii) improved quality of electricity, (iv) quick demand response, (v) user contribution, and (vi) effective resource management.

A core advantage of SGs is their bidirectional energy flow, which modernizes power systems from generation to consumption. SGs leverage millions of interconnected devices—deployed across power plants, distribution centers, and consumer premises to

* Correspondence to: University of Technology Sydney, 15 Broadway, Ultimo NSW 2007, Australia.
E-mail addresses: hafsa.bibi@student.uts.edu.au (H. Bibi), Mehran.Abolhasan@uts.edu.au (M. Abolhasan), Justin.Lipman@uts.edu.au (J. Lipman), Mahrokh.Abdollahilorestani@data61.csiro.au (M. Abdollahi), Wei.Ni@data61.csiro.au (W. Ni).

https://doi.org/10.1016/j.compeleceng.2025.110371
Received 1 August 2024; Received in revised form 20 March 2025; Accepted 14 April 2025
Available online 1 May 2025

monitor and analyze grid performance [4,5]. SG systems rely on specialized communication networks to respond to real-time events, requiring reliable, fast, and secure information flow between devices, applications, grid operators, and consumers. However, the growing scale and complexity of SGs introduce significant challenges in multi-source data management [6–9]. Frequent state measurements and control actions generate vast amounts of data transmitted over open networks, leading to concerns about energy consumption, bandwidth usage, and cyber-security threats, including False Data Injection (FDI) attacks.

Smart Meters (SMs), integral to SGs, measure and report energy consumption and grid operation data to Meter Data Management Units (MDMUs), typically at 15 min intervals. MDMUs process this data to enhance grid efficiency and manage consumer billing [10,11]. SMs also support local grid management by balancing stochastic power injections with flexible demand, reducing the need for expensive grid expansion. Advanced technologies such as Advanced Metering Infrastructure (AMI), Demand Response Programs (DRPs), Energy Storage Systems (ESSs), Renewable Energy Sources (RESs), and Electric Vehicles (EVs) further enhance SGs but also add complexity, necessitating novel solutions for emerging challenges [12,13].

1.1. Motivation

Despite their advantages, SGs raise significant privacy concerns due to the fine-grained energy data collected by SMs [14]. This data can reveal sensitive consumer information, including personal routines [15], appliance usage patterns [16], occupancy status [17], and even economic conditions [18]. Various threats pose risks to consumer data privacy, for instance, (i) reconstruction attacks can extract household activities [15], (ii) Non-Intrusive Load Monitoring (NILM) attacks can identify appliance usage [16], (iii) inference attacks can reveal personal routines and occupancy [17], (iv) linkage attacks combine datasets to infer economic conditions [18], and (v) data manipulation attacks can alter readings [19]. The widespread display, transmission, storage, and processing of such information heighten privacy concerns, limiting SG adoption.

Protecting individual energy consumption profiles while maintaining data usability and ensuring efficient grid operations is a critical challenge, particularly given the limited computational capabilities of resource-constrained SMs. This article is deeply inspired by previous research on privacy preservation strategies within SG networks. Privacy in this context encompasses both user privacy and data privacy. Numerous privacy-preserving techniques have been proposed to address these concerns [20]. However, given the rapid advancements in SG technology and privacy-preserving mechanisms, a comprehensive evaluation of recent research is essential. Specifically, we consider and evaluate privacy-preserving techniques published recently between 2018 and 2024. By presenting a structured and up-to-date analysis of privacy-preserving solutions for SGs, this study aims to assist researchers in designing more effective and secure energy data protection mechanisms while ensuring grid efficiency and reliability.

1.2. Our contributions

The following contributions are made in this paper:

- We provide a comprehensive review and comparative analysis of recent survey articles in the SG domain, highlighting advancements in SG applications, technologies, privacy, and security based on prior research.
- We explore the interplay between privacy and security within SG systems, providing a detailed classification of privacy attacks, including adversary models, attack surfaces, and underlying intentions.
- We provide a detailed taxonomy of privacy-preserving schemes for SG systems, focusing on publications from 2018 to 2024. These schemes are systematically categorized and summarized in comparative tables to facilitate easier analysis and understanding.
- We conduct a thorough review of several publicly available power consumption datasets, highlighting their key features and identifying their limitations through a multi-faceted analysis. Additionally, we pinpoint open research challenges and propose directions for future investigation based on our findings.

1.3. Organization of the paper

The rest of this paper is structured as follows: Section 2 delves into a comprehensive review of existing surveys and compares our survey with related surveys in the field. Following that, Section 3 outlines a comprehensive background of SG systems and intricate privacy and security challenges encountered by SG systems. Moving on to potential resolutions, Section 4 critically assesses privacy-preserving solutions proposed for SGs and intricately examines software-based and hardware-based privacy-preserving approaches tailored for SGs. We present the findings of our study in Section 5. For a more nuanced understanding, Section 6 offers an in-depth comparison of publicly available datasets utilized in various SG studies. Exploring open challenges and avenues for future research is presented in Section 7. Finally, Section 8 concludes the paper.

Table 1
Comparison of previous survey articles for SGs.

Ref.	Year	Research area	O	ATT	Cr	Pr	Agg	6GT	B	FL	Phy	KC	Rec	PD	OF
[21]	2016	Privacy and Security	✓	✓	✓	–	✓	–	–	–	∘	∘	–	–	✓
[20]	2017	Privacy and security	✓	∘	✓	–	✓	–	–	–	∘	∘	–	–	–
[22]	2018	Privacy and security	✓	–	∘	–	∘	–	–	–	∘	–	∘	–	∘
[23]	2018	Privacy and security	✓	✓	✓	✓	✓	✓	–	–	∘	✓	∘	–	✓
[24]	2018	Privacy and security, communication technologies	✓	✓	∘	–	∘	–	–	–	–	–	∘	–	✓
[25]	2018	Privacy and security	✓	✓	✓	∘	✓	–	–	–	✓	✓	✓	–	✓
[26]	2018	Privacy and security, communication technologies	∘	–	∘	∘	∘	–	–	✓	–	✓	✓	–	✓
[27]	2019	Privacy and security, communication technologies	✓	∘	∘	–	∘	–	–	✓	–	✓	✓	–	✓
[28]	2019	Privacy and security	✓	∘	✓	✓	✓	–	–	–	✓	✓	–	–	–
[29]	2019	Privacy and security, SG communications	✓	✓	✓	✓	✓	–	–	–	∘	✓	✓	–	✓
[30]	2019	Privacy and security	✓	∘	∘	✓	∘	–	–	–	✓	✓	✓	–	✓
[19]	2020	Privacy and security	✓	✓	∘	–	∘	–	–	✓	–	✓	–	–	✓
[31]	2020	Privacy and security	✓	✓	✓	✓	✓	–	–	∘	–	✓	–	–	✓
[32]	2020	Privacy and security	✓	–	∘	–	∘	–	–	✓	–	∘	–	–	∘
[33]	2020	Privacy and security	✓	∘	–	–	–	–	✓	–	–	∘	–	–	✓
[34]	2021	Communication technologies	✓	✓	✓	∘	∘	–	–	–	–	∘	–	–	✓
[35]	2021	Privacy and security	✓	✓	✓	–	–	–	–	–	∘	∘	–	–	∘
[36]	2021	Communication technologies	✓	✓	∘	–	∘	–	–	∘	∘	–	–	–	∘
[37]	2021	Communication technologies, energy trading	✓	✓	∘	–	–	✓	✓	✓	–	✓	∘	–	✓
[38]	2022	Privacy and security	✓	–	✓	✓	✓	✓	–	✓	–	✓	✓	–	∘
[39]	2022	Privacy and security, communication technologies	✓	✓	✓	✓	✓	–	✓	✓	✓	✓	∘	–	✓
[40]	2023	Privacy and security	✓	✓	✓	–	✓	–	✓	–	–	–	–	–	∘
[41]	2023	Privacy and security	∘	✓	∘	–	∘	–	–	✓	∘	∘	–	–	✓
This work	2024	Privacy and security	✓	✓	✓	✓	✓	✓	∘	✓	✓	✓	✓	✓	✓

O: Overview of SG Components and Communication Technologies, ATT: Privacy and Security Goals, Cr: Cryptographic Approaches, Pr: Data Perturbation Approaches, Agg: Data Aggregation Approaches, GT: Game Theory Approaches, B: Blockchain Approaches, FL: Federated Learning Approaches, Phy: Approaches based on Physical Equipment, KC: Key Characteristics, Rec: Recommendations, PD: Power Consumption Datasets, OF: Open Challenges and Future Directions
✓: Fully covered, ∘: Partially covered, -: Not covered.

2. Related works

There are several survey papers published in recent years in SG domain covering SG privacy, security, applications, and communications technologies. Table 1 summarizes and compares the key contributions of our survey paper with the existing SG survey articles from previous years on several aspects.

Han et al. [21] examined privacy-preserving techniques used in vehicle-to-grid networks, such as anonymous authentication, location and identity privacy, covert data aggregation, billing and payment privacy, and data publication. In order to highlight existing issues and provide solutions for unresolved ones, they investigated a variety of techniques, including homomorphic encryption, blind signature, and anonymity networks. Asghar et al. [20] focused on automated SMs, which are critical devices for monitoring energy use in near real-time and enabling different SG features. The article examined the applications of metering data, privacy regulations, and security solutions for privacy-preserving metering data distribution and management. It also examined current developments in privacy preservation in gathering meter data for operations, billing, and value-added services such as demand response.

Leszczyna. [22] discussed the significance of reliable information and communication in the SGs by highlighting the challenge of finding relevant standards for operators and stakeholders. To combine SG standards linked to cyber-security challenges, this study thoroughly evaluates 36 security articles and 11 privacy publications. In-depth explanations are provided regarding the study methodology, as well as the selection and assessment criteria for standards.

The work by Ferrag et al. [23] provides a thorough analysis of privacy-preserving communication methods for SGs. Additionally, 32 previous approaches are categorized into AMI, data aggregation, smart community of house gateways, marketing architecture, and vehicle-to-grid architecture. The study discusses privacy violations, defenses, and game-theoretic strategies for each strategy. It also examines current survey articles on standardization, security, communications, and applications for the SG. Based on the survey results, this article concludes with recommendations for additional research. Mocrii et al. [24] offered a thorough analysis of IoT-based smart homes, including a discussion on their definition, construction, and positioning within the SG. They also covered the aspects of smart home management systems, software solutions, communication technologies, and privacy and security concerns for SGs. The article identifies present issues, prospective remedies, and upcoming developments in smart home technology.

In a very comprehensive survey, Stellios et al. [25] studied the rapid development of IoT and its relevance to cyber-attacks. The authors emphasized how malicious attacks on IoT industries like SGs might be overlooked despite the fact that attacks on vital systems linked to IoT technologies are well-known. In this article, IoT-enabled cyber-attacks across many areas since 2010 are surveyed, with an emphasis on actual occurrences and proof-of-concept attacks. Its objectives include determining hidden attack vectors, assessing mitigation techniques for all application domains, and assessing today's threat landscape.

Mohammadi et al. [26] offered a thorough introduction to applying deep learning methods for analytics in the IoT domains, including SGs. They analyzed IoT data properties and described two main IoT analytics approaches: (i) IoT streaming data analytics and (ii) IoT big data analytics. Different architectures and methods are examined, along with the potential of deep learning in IoT data analytics. The study also examined attempts to conduct research on and implement deep learning in smart IoT devices, as well as deep learning implementation strategies in fog and cloud centers. Finally, the study discussed problems and potential paths for future research.

Hossain et al. [27] explored the use of machine learning and big data in the SG, emphasizing the IoT as a connection provider and a major source of data production. In the context of the SG, this study focused on using data analysis and machine learning for accurate load forecasting, cost-effectiveness, and decision-making. Additionally, the paper covers the crucial cybersecurity problem in the IoT-integrated SG and offers a thorough review of the research results, constraints, and potential future research in this area.

Desai et al. [28] studied the transformation of the traditional electrical network into an SG, made possible by the Internet of Everything (IoE), to handle the issues provided by renewable resources and rising energy consumption. The primary focus is on the AMI, which allows two-way communication between SMs and utilities but simultaneously increases the risk of cyberattacks and jeopardizes consumer privacy. The study gives an in-depth assessment of privacy-specific research in the IoE-enabled SGs, analyzing privacy challenges in AMI and suggesting future research topics.

Kumar et al. [29] studied the importance of smart metering infrastructure in SG networks, which improves energy infrastructure reliability, effectiveness, and conservation. This work highlights the security and privacy risks associated with coordinating a large number of disparate devices across public communication networks. Additionally, the paper presents a threat taxonomy, outlines security and privacy requirements, discusses proposed schemes to address these threats, and identifies open research issues for future investigation in SG metering networks.

Hassan et al. [30] provided an extensive overview of differential privacy methods tailored for Cyber-Physical Systems (CPSs). Specifically, they examine the utilization and integration of differential privacy across four primary CPS applications: energy systems, transportation systems, healthcare and medical systems, and the Industrial Internet of Things (IIoT). Additionally, they highlighted unresolved issues, hurdles, and prospective research avenues concerning the application of differential privacy techniques within CPSs.

In another study, the usage of big data in SG systems and the security issues brought on by FDI attacks are discussed by Cui et al. [19]. In addition to defining categories of FDI attacks and outlining data security standards, this survey gives a brief description of the SG architecture with its data sources. The following section of the study reviews state estimation, non-technical losses, and machine learning-based load forecasting methods for identifying FDI attacks. Finally, possible research trajectories are investigated, covering the detection of adversarial intrusions, decentralized and collaborative detection frameworks, privacy-preserving detection, and sophisticated machine learning methods.

Ferrag et al. [31] provided a classification and analysis of cybersecurity approaches for fog-based SG SCADA systems. They presented a taxonomy of security vulnerabilities addressed by authentication and privacy-preserved solutions, and categorized the Intrusion Detection Systems (IDS) on the basis of machine learning methods. Joudaki et al. [32] discussed the value of smart energy grids in mitigating the drawbacks of conventional power systems. While considering the SG sophistication and huge data production, the paper emphasizes the usage of deep learning techniques to improve privacy and security. It gives a summary of the SG architecture, introduces fundamental ideas in deep learning, and evaluates research papers that use deep learning-based techniques to protect user privacy and security in the SG.

Zhuang et al. [33] thoroughly analyzed the potential of blockchain technology to be used for cybersecurity in SGs. It offers insights into the theories, designs, and implementation methods for blockchain applications in the SG. The survey's objectives are to direct future research projects and act as a useful resource for researching blockchain for cybersecurity in the context of SGs.

Abrahamsen et al. [34] provided an extensive survey on communication technologies for SG. Their survey covered the communication requirements, network architectures, and physical layer technologies relevant to SG systems. Philips et al. [35]

analyzed the cyber security threats, attack models, and solutions in SG networks. In particular, they examined the lightweight Key Management systems (KMS) methods in the context of SGs.

Jha et al. [36] presented the communication technologies in the context of Cyber-Physical Systems (CPS) for SG. They discussed different communications standards and protocols, challenges including security, reliability, safety, and resilience, and identified research gaps in the SG-CPS domain. Aggarwal et al. [37] focused on how information and communication technology can be integrated with the SGs to improve energy management. They looked at the complications caused by elements like real-time monitoring, automated outage management, and renewable energy sources. The survey discusses enabling technologies, offers a taxonomy of methods utilized in energy trading mechanisms, and presents a four-layered energy trading framework for the SG. It concludes with suggestions for future research and reader insights. Abdalzaher et al. [38] reviewed the SM applications, including SGs and the best possible power flows. It draws attention to the necessity of SM privacy implementations that are more cost-effective and efficient. In order to preserve data privacy and reduce cyberattacks, which could ultimately have an impact on both human life and disaster management infrastructure, the article addresses the weaknesses of SMs and provides eleven trust models for increasing SM security.

Mirzaee et al. [39] discussed the bidirectional flow of energy and information in SGs, which is made possible by the integration of power networks and information and communication technology. The paper emphasizes the need to explore security and privacy procedures in order to secure customers' sensitive data and prevent potential breaches. It also covers the application of machine learning algorithms in SG components, as well as the vulnerabilities caused by adversarial machine learning attacks. The survey includes taxonomies and tables to help researchers understand the links between variables in the field, as well as an analysis of the implications of evolving technologies on SG security and privacy.

Ghiasi et al. [40] summarized the different decisive problem-solving strategies and defense mechanisms related to cybersecurity in SG systems. In this review, the authors discussed and analyzed the different cyber-attack models and then described cybersecurity solutions, such as quantum computing and blockchain, in the context of energy. Ortega et al. [41] discussed and analyzed the effects of DoS attacks in the SGs, and presented the survey of detecting and mitigating methods using reinforcement learning algorithms against DoS attacks in the SG context.

2.1. Comparison with existing surveys

The following key aspects define the novelty of this survey and set it apart from existing studies.

- Some existing studies have categorized security issues, real-time attacks, threats, IoT-enabled attacks, and security analytics in the context of CPSs [28–31,40]. We discuss the different privacy and security goals in separate sections, including adversaries, attack surfaces, and intentions in the energy domain.
- Since privacy and security in SGs is a vast domain. We surveyed the literature from current years (2018–2024) while keeping in mind the recent trends and advancements in privacy-preserving technologies. This aspect can help the research community to develop advanced mechanisms by having insights into emerging privacy-preserving solutions from recent years.
- Several studies review existing solutions for SG systems with a limited number of privacy features. Compared to other surveys, our survey provides a taxonomy of the state-of-the-art privacy-preserving schemes for SG systems and presents a comprehensive overview of these privacy-preserving approaches along with corresponding shortcomings.
- None of the related articles covers power consumption datasets, which are crucial for evaluating privacy countermeasures. In order to investigate and develop efficient solutions, choosing a database can be challenging. Our survey discuss and strengths and weaknesses of 18 publicly available power consumption datasets in detail. This section can help researchers choose an appropriate power consumption dataset for the design and evaluation of a model.
- On the basis of this detailed analysis and our findings, we provided our recommendations for the practical usefulness of different privacy-preserving technologies. This can aid researchers in narrowing down a research domain for developing privacy-preserving solutions for SG systems.
- Finally, this survey identifies open research challenges and future recommendations by integrating innovative research with adapted existing solutions. We believe that by the end of this study, researchers will have a deeper knowledge of the key aspects of security and privacy issues and their solutions in different SG domains.

3. SG systems and data privacy: Background

The Internet of Things (IoT) serves SG systems to carry out many network operations from generation to utilization of power by integrating the IoT devices, including sensors, metering devices, and actuators and by allowing required automation, connections and tracking of this wide range of devices. IoT enables collaborative and direct network connections to users and devices over various communication mechanisms, power equipment by means of unified IoT devices, and interaction to realize a two-way, real-time high-speed information exchange across multiple applications to improve the inclusive performance of SG. The three-layered IoT architecture can be applied in the SG paradigm as follows: for deploying different IoT smart devices to monitor equipment states at the perception layer, for gathering information from equipment via connected IoT smart devices at the network layer, and for supervising the SGs through the application interfaces underlying the application layer of IoT [42,43].

IoT integrates the power flow and distribution flow with information flow to allow the practical implementation of data sensing, collection and transmission, network construction and operation, maintenance, security monitoring and user interactions in an

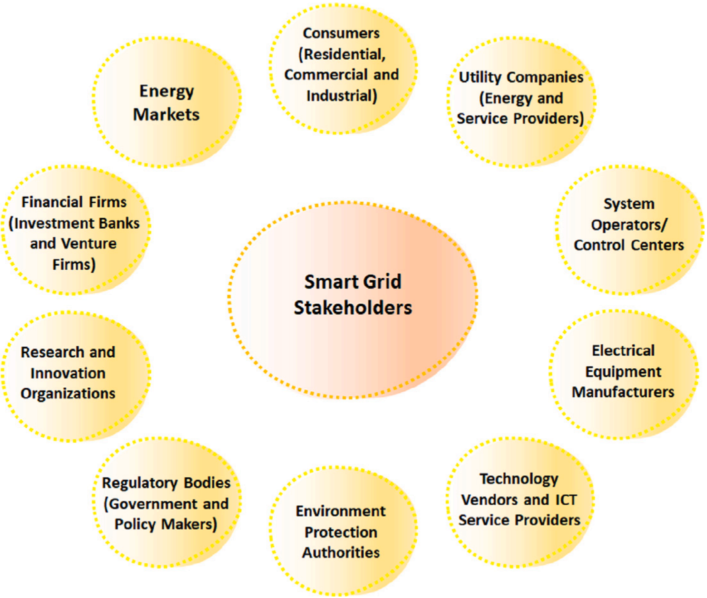
Smart Grid Stakeholders
Consumers (Residential, Commercial and Industrial)
Utility Companies (Energy and Service Providers)
System Operators/ Control Centers
Electrical Equipment Manufacturers
Technology Vendors and ICT Service Providers
Environment Protection Authorities
Regulatory Bodies (Government and Policy Makers)
Research and Innovation Organizations
Financial Firms (Investment Banks and Venture Firms)
Energy Markets

Fig. 1. SG Stakeholders.

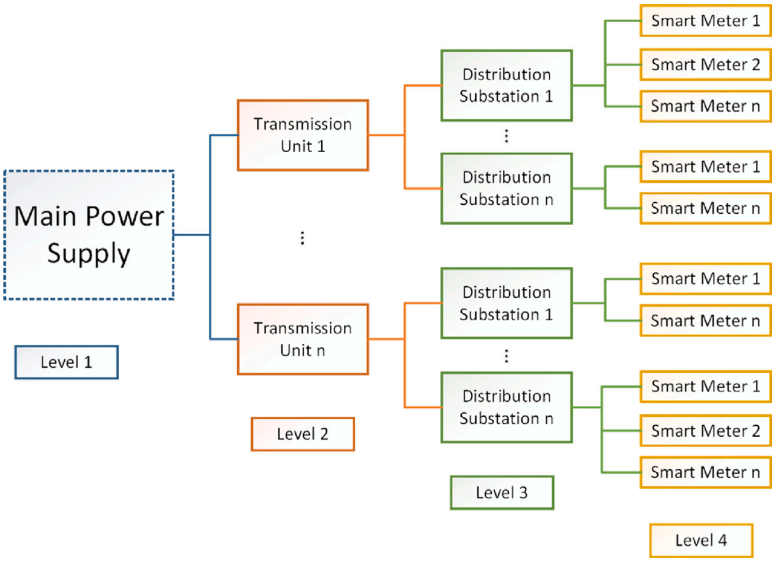

Fig. 2. SG Communication Hierarchy [29,46].

SG [44]. IoT strengthens the four power grid subsystems, described hereafter: In the power generation phase, IoT is applied to manage and monitor the consumption of energy, equipment, units, pollutants and gas discharge, energy production and consumption predictions, energy storage, distributed power plans, biomass power, wind power as well as photovoltaic power generation. In power transmission, IoT controls and supervises the transmission lines and substations with the guarantee of providing transmission tower security. In power distribution, IoT is used in distributed automation and management of operations. Lastly, in power utilization, IoT is functional for smart homes, electric vehicle charging and discharging, automatic meter reading, home appliances information collection, power load and demand detection and multi-network consumption [45].

3.1. SG components and technologies

SG incorporates several technologies, consumer solutions, policy, and regulatory drivers to ensure robust performance. A diverse group of entities are involved in SG systems, as depicted in Fig. 1.

Various stakeholders play distinct roles to ensure the efficient, secure, and sustainable operation of the grid. Here's a breakdown of each stakeholder's role:

- *Consumers (Residential, Commercial, and Industrial):* End-users of electricity who actively participate in energy management by adjusting their consumption patterns based on dynamic pricing, demand response programs, and through the adoption of energy-efficient and smart appliances.
- *Utility Companies (UCs) (Energy and Service Providers):* Generate, transmit, and distribute electricity to consumers. They are responsible for maintaining infrastructure, ensuring energy reliability, and implementing SG technologies to optimize operations. Distribution System Operator (DSO) manages the grid at the local level. They provide a medium to low voltage in distribution networks of a specific area and keep quality constraints at the limit set by the regulatory authority. On the other hand, the Transmission System Operator (TSO) caters to the high-voltage electrical transmission lines and balances demand and supply continuously. Finally, the regulatory authority audits the grid operations and establishes technical rules for the grid. Fig. 2 presents the communication hierarchy for SG networks as suggested by [29,46].
- *System Operators/Control Centers (CC):* Monitor and control grid operations in real-time. They manage load balancing, ensure grid security, and handle energy flow, facilitating efficient distribution and preventing outages.
- *Energy Markets:* Energy trading markets are also a significant part of the SG systems, allowing SGs to balance the demand and generation of energy to some extent. There are different types of energy markets, including the wholesale market, the retail market, and the balancing market. Energy Markets facilitate the buying and selling of electricity, allowing efficient energy trading and pricing. They help balance supply and demand, promoting economic incentives for renewable energy adoption and grid stability.
- *Electrical Equipment Manufacturers:* Produce and supply essential equipment for grid infrastructure, such as transformers, circuit breakers, and SMs. They also innovate in energy storage and grid automation technologies to enhance grid resilience.
- *Technology Vendors and ICT Service Providers:* Develop and provide communication, data management, and cyber-security solutions for the SG. They play a key role in integrating digital technologies that enable grid monitoring, control, and data analysis.
- *Environment Protection Authorities:* Ensure that the SG's operation complies with environmental regulations. They advocate for sustainable practices, reduced emissions, and the integration of renewable energy sources in the grid.
- *Regulatory Bodies (Government and Policy Makers):* Set policies and regulations that govern SG operations. They establish standards for security, consumer privacy, energy efficiency, and market operations to promote a fair and reliable energy system.
- *Research and Innovation Organizations:* Conduct research on new technologies, algorithms, and processes to improve grid efficiency, security, and sustainability. They drive innovation in renewable integration, energy storage, and demand-side management.
- *Financial Firms (Investment Banks and Venture Firms):* Provide funding for infrastructure upgrades, technology development, and expansion projects. They support the financial viability of SG initiatives and encourage private sector investment in green technologies.

Each of these stakeholders contributes to the advancement, sustainability, and resilience of the SG, working collectively to optimize energy production, distribution, and consumption in a way that benefits society and the environment.

3.2. Communication networks for SGs

A communication network is needed to exchange information in the SG paradigm. Traditionally, three network models are employed in the SG framework: Home Area Network (HAN), Neighborhood Area Network (NAN), and Building Area Network (BAN). The SG architecture also includes a central entity as CC and several cloud services [12,47]. The HAN consists of two additional networks, Local Area Network (LAN) and the Wide Area Network (WAN). On the other hand, BAN is responsible for connecting various networks within a building with higher communication rates. The NAN keeps a secure channel for communication between the UC and individual users. Moreover, gateways are placed in each network for the transmission of control information among these networks [48]. Fig. 3 presents a typical SG infrastructure. Communication networks facilitate data and command exchange among various devices, entities, and stakeholders. Here's an explanation of each type of network:

- **Home Area Network (HAN):** HANs are local networks within residential homes, connecting appliances, SMs, thermostats, home energy management systems, and sometimes electric vehicle chargers to monitor and control energy consumption at the household level. HAN provides connectivity within the home, data flows between the SM and various appliances, enabling users to receive real-time energy usage information, apply demand response, and manage energy consumption efficiently. The HAN also sends information to the BAN or NAN via gateways, allowing communication with external systems.
- **Building Area Network (BAN):** BANs cover larger buildings like commercial or industrial facilities, connecting energy management systems within these larger structures. BANs operate similarly to HANs but manage more extensive and often more complex equipment in non-residential settings. BANs include HVAC systems, lighting controls, industrial machines, and building automation systems. The BAN gathers data from various equipment within the building and sends this data to a gateway. Through the gateway, it transmits aggregated data to the NAN or directly to Utility Providers (UPs), enabling centralized monitoring, control, and optimization of energy usage across multiple buildings or facilities.

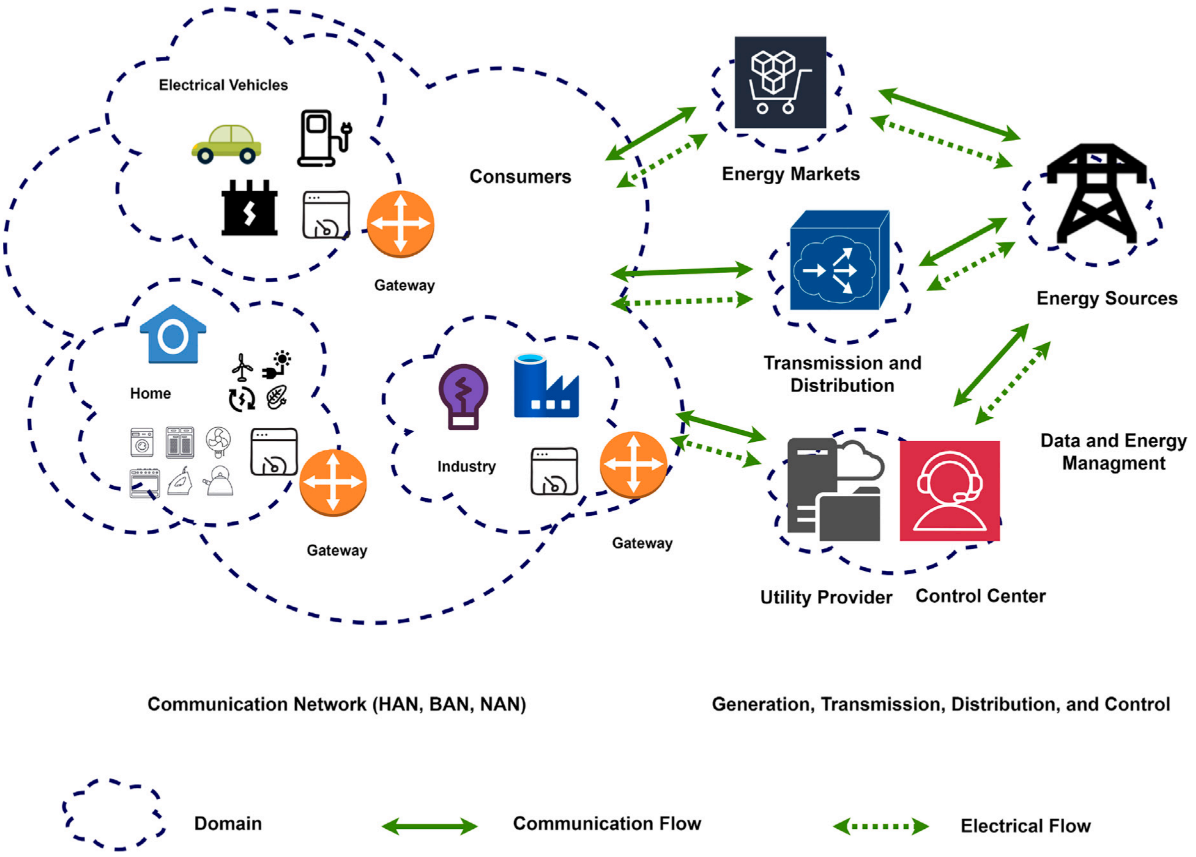

Fig. 3. A Typical SG Infrastructure Representing HAN, BAN, and NAN. The dashed lines denote the flow of power, and the solid lines denote the information flow.

- **Neighborhood Area Network (NAN):** NANs connect multiple HANs and BANs within a localized geographic area (*e.g.*, neighborhood or district). It acts as an intermediary network that gathers and aggregates data from various residential and commercial sources before forwarding it to the wider utility network. SMs, gateways, and other data concentrators collect information from HANs and BANs in the neighborhood. Data flows from individual HANs and BANs to the NAN through gateways. The NAN then forwards aggregated data to UPs, transmission, and distribution networks, allowing for broader grid management, demand response, and load balancing across larger areas.

Communication Flow in the SG can be described as follows.

- *Within HAN and BAN:* Communication flow occurs between smart devices and the local network's central control unit or energy management system, enabling device monitoring and control.
- *From HAN and BAN to NAN:* Communication flows from local networks (HAN/BAN) to the NAN through gateways, where data is aggregated for broader area management.
- *NAN to UPs and CCs:* NANs transmit aggregated data to UPs, CCs, and sometimes directly to transmission and distribution networks, facilitating comprehensive grid management. This ensures efficient energy distribution, load balancing, and integration of DERs within the SG.

3.3. Privacy and security in SGs

SG systems enable electrical infrastructure to transmit and distribute power along with information exchange over communications networks [49]. The primary objective of an SG is to ensure precise customer billing while efficiently managing and distributing electrical energy. Within the SG framework, an SM plays a pivotal role as the key entity. The deployment of SMs gives rise to concerns about meter tampering and consumer privacy, highlighting the necessity for legislating the use of these SMs. To hold the privacy regulations, specific properties such as integrity, confidentiality, authenticity, and availability are crucial in ensuring privacy [20,50].

The presence of an SM introduces the risk of data tampering, where an adversary can breach the security of the SM. Once compromised, accessing a cryptographic key becomes relatively easy. Exploiting a common vulnerability can compromise hundreds of SMs, potentially enabling the manipulation of real-time energy consumption data. Thus, developing an access control system is necessary to protect meters and ensure that stored information can only be used for billing purposes and related value-added services. While the primary role of an SM is to calculate accurate bills, frequent exchanges of consumer information with the UC may inadvertently disclose certain private information [51].

Since the energy usage from household appliances is typically collected by SMs after a 15 min period, subsequently, the SM generates an energy usage report and transmits it to the UC. However, there exists a potential risk wherein an adversary may illicitly observe the energy consumption patterns of a smart home, thereby enabling them to make predictions about the homeowner's lifestyle and routines. Consequently, this is a significant threat to the security and privacy of home automation systems. In order to safeguard privacy and security, power consumption information is encrypted at the individual appliances or prior to its transmission from the HAN.

The encrypted energy consumption data is then refined, and the CC obtains the energy consumption details from multiple SMs. Generally, each meter records its consumption reports individually, or an aggregation method is employed to combine the consumption records from all SMs within a specific sub-region or zone. Subsequently, a consolidated consumption summary is sent by the CC after the aggregation process [52]. This information serves as the basis for generating monthly bills and maintaining an overview of the region's electricity consumption. The CC continually collects consumption statistics from various devices.

To ensure the integrity of the data, the received consumption information is either encrypted or subjected to aggregation. This precautionary measure is implemented to detect any potential tampering by an intruder. By employing encryption or aggregation, any modifications made to the transmitted message can be readily identified [53].

3.3.1. Security goals

This section encompasses an explanation of the diverse security requirements associated with SG systems. Table 2 presents a detailed explanation of the various security goals related to the attack category in SGs.

- **Confidentiality:** Confidentiality is a crucial aspect of security that assesses whether specific data should be protected from unauthorized disclosure. While it is the least critical factor in terms of grid communications reliability, it holds significant importance for end consumers [54]. The confidentiality aspect is closely tied to the privacy of end users since consumption usage data can expose individuals' daily routines. Home appliances transmit their energy details to aggregators, SMs, or service providers, and these details might reflect a user's personal profile. Privacy and confidentiality are interconnected, and safeguarding messages from unauthorized access violates the privacy of home occupants. If the privacy of home users is compromised, confidentiality is automatically violated. Therefore, it is crucial to maintain the confidentiality of consumption usage data at utility servers. Various techniques, such as homomorphic encryption, blind signature, and in-network aggregation, ensure confidentiality in SG systems [55,56].
- **Integrity:** Smart appliances consistently transmit consumption patterns to an aggregator or SM at regular intervals. Later, the SM periodically forwards consumption data to utility servers through gateways. While aggregators and SMs may possess physical security, they remain susceptible to various attacks, such as MITM attacks, replay attacks, and alteration attacks [70]. Integrity ensures data and control commands remain unchanged or unaltered without proper authorization. Compromising integrity puts valuable information at risk and may lead to incorrect network management and control decisions. In terms of message transmission, integrity implies that the receiver receives the message sent by the sender without any modifications. Several schemes are employed to maintain data integrity in HANs, including digital signature, message digest, Message Authentication Code (MAC), and Hash-based Message Authentication Code (H-MAC) [55,71].
- **Availability:** Availability refers to the state in which data, applications, or systems are accessible to end-users whenever they are needed. However, availability can be compromised if an unauthorized individual impersonates an authorized user, gaining access to the system and causing network congestion [72]. The usage data from SMs in the HAN now allows utility suppliers to manage and balance bulk energy demand and supply effectively. Consequently, the reports generated by SMs are of significant importance in providing energy feedback. Ensuring SM data availability is imperative for SG reliability. The specific availability demands may differ depending on the applications being utilized. A protective relay, for example, necessitates a latency of 4 ms to detect faulty lines and circuits promptly. Consequently, ensuring data availability should be among the primary design objectives in SG systems [73,74].
- **Access Control:** The concept of access control refers to the capability of verifying whether SMs, gateways, and aggregators seeking access to a resource possess the appropriate privileges to do so. In the context of consumer usage data, various stakeholders participate, each driven by their specific application interests, such as demand supply and load management [75]. Nevertheless, it is essential for these stakeholders to implement sufficient authorization policies to ensure that customers' data cannot be accessed without proper permission [76].
- **Authentication:** The SG system consists of a vast number of entities, making it crucial to ensure the authenticity of each entity's claims. Authentication or identification serves as a logical approach to verifying the legitimacy and identity of entities, including end-users, SMs, and more. To ensure the integrity of SG metering applications, a robust authentication mechanism must be implemented, capable of identifying and rejecting malicious commands and connections [77]. Furthermore, authentication schemes ought to exhibit high efficiency and resilience against cyberattacks and incorporate multicast functionality within SG Systems [78].

Table 2
Different cyber-attacks in SG systems and their description.

Ref.	Attack classification	Description	Target layer	Security goals
Cui et al. [19], Riggs et al. [57]	FDI attack	Tampering with the real data and injecting the untrusted or false data packets into communication links	Application, transport, network and data link	Integrity
Tufail et al. [58]	Man-In-The-Middle (MITM) attack	Mitigating or altering data while it is being transmitted across the network when sensitive information have an unauthorized access	Session, network and data link	Integrity and confidentiality
Sakhnini et al. [59]	Buffer overflow attack	Overwriting buffer memory, crashing the system and exhausting the resources	Application and transport	Availability
Huseinovic et al. [60]	Denial of Service (DoS) attack	Stops user's requests from being served	Network	Availability
Zhang et al. [61], Lu et al. [62]	Flooding attack	Preventing users from accessing resources, leading to exhaustion	Application, transport, network and data link	Availability
El Mrabet et al. [63], Gunduz et al. [64]	Social engineering attack	Attacking individuals instead of machines or networks, stealing identity and violating the users' privacy, temporary or persistent harm to system	Application	Integrity and confidentiality
Kurt et al. [65]	Channel jamming attack	Jamming the network, interrupting and obstructing the transmission and reception of data, preventing one or multiple nodes from sending and receiving information due to communication collisions	Network, data link, and physical	Availability
Peng et al. [66], Bi et al. [67]	Eavesdropping attack	Capturing, monitoring, and analyzing the network traffic	Network and physical	Confidentiality
Huang et al. [68]	Session hijacking attack	Blocking the user from resources temporarily, stealing or predicting a valid token to unauthorized access to the web server	Session	Integrity and confidentiality
Ganguly et al. [69]	SM tampering attack	Modify the actual readings of SM leading to inaccurate usage and billing information	Physical	Integrity

- **Non-Repudiation:** Given the extensive scope of the SG, the non-repudiation property plays a crucial role in identifying instances where an individual has engaged in deceptive activities, such as energy theft by consumers, which means only to deny any involvement or responsibility for their actions afterward.
- **Auditing and Accountability:** Regular accountability and auditing are essential prerequisites to enhance the validation of security mechanisms within SG systems. While security breaches may occur, the implementation of periodic accountability procedures enables the identification of responsible parties behind such breaches. By holding the SG and its information systems accountable, accountability ensures that comprehensive analysis can be conducted while facilitating the traceability of actions and events [79].

3.3.2. Privacy goals

This section will specifically address communication privacy within the SG systems. We delve into the following privacy properties, which have been taken from [80–82] for discussion.

- **Anonymity:** Anonymity is related to maintaining the confidentiality of an end user's true identity. When sharing secret control signals or data, a device has the ability to hide its true identity from other devices [83]. In a HAN, appliances and SMs cannot identify other communicating devices. The objective of anonymity is to conceal one's identity during interactions between appliances, appliances and SMs, SMs and service providers, and service providers and appliances. Several techniques are utilized to achieve anonymity, including Password-based Anonymous Lightweight Key (PALK), Anonymous Secure Framework (ASF), and Threshold-Based Anonymous Identification (TAI) [84,85].
- **Pseudonymity**: Pseudonymity serves as an alternate identifier for a subject, distinct from their real name(s) [81]. In SG communications, numerous stakeholders may access power consumption data from SMs. Thus, an SM needs multiple identifiers, *i.e.*, pseudonyms. These identifiers are exclusively held by the designated entities engaging in information exchange with SMs.
- **Unlinkability:** In this scenario, the data or information lacks sufficient linkage between two entities within a system. From a broader perspective, unlinkability emerges as a critical property requiring implementation across various levels, including HAN (comprising SMs, sensors and home gateways) and data concentrators and substations situated in the cloud or SG servers [80,81].

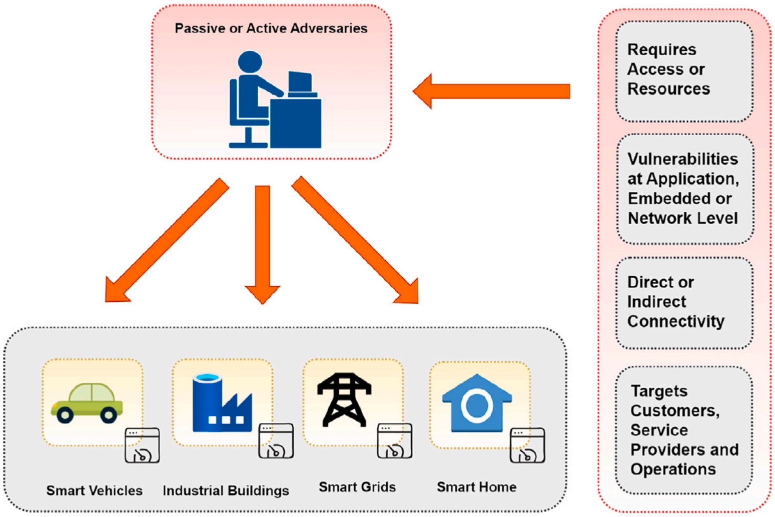
Passive or Active Adversaries
Smart Vehicles
Industrial Buildings
Smart Grids
Smart Home
Requires Access or Resources
Vulnerabilities at Application, Embedded or Network Level
Direct or Indirect Connectivity
Targets Customers, Service Providers and Operations

Fig. 4. SG Applications with Adversaries Existence.

- **Unobservability:** This property ensures that an external observer cannot discern whether the communication is occurring or not. In other words, an uninvolved entity, such as an outside observer, cannot adequately differentiate whether a specific entity (like an SM) has carried out particular messages or actions of interest, such as transmitting consumption, demand-bidding, payment, or logging in information [81].
- **Undetectability:** various entities like devices, machines, equipment, users and applications or their information/data could be targets for malicious users or adversaries seeking to detect or glean knowledge about them. Hence, it is imperative that these items or data remain undetectable to the adversary. Moreover, undetectability can be categorized into zero undetectability and maximum undetectability. For more detailed information, readers can refer to [82].

3.3.3. *Privacy attacks in SGs*

Research in SG vulnerabilities has become increasingly popular over the past few years; we discuss the privacy attacks in SG systems here. Fig. 4 illustrates SG applications in the presence of various adversaries. *Passive adversaries* are typically eavesdroppers who monitor communication channels, aiming to collect sensitive information without interfering directly with data or systems. Their goal is often to gather intelligence or observe the behavior of SG components. *Active adversaries* take a more aggressive approach, actively engaging in attacks to disrupt systems, inject false data, alter messages, or compromise resources. These adversaries may launch attacks like DoS, MITM attacks, or data manipulation to directly impact SG operations. Passive and active adversaries could leverage multiple vulnerabilities across the network, application, and embedded levels to achieve their objectives, such as data theft, service disruption, or unauthorized access.

Fig. 4 presents a list of potential characteristics of adversarial actions within the SG systems. These highlight the objectives and methods that adversaries may use:

- *Requires Access or Resources:* Attackers need access to certain devices, or communication channels to initiate an attack. Gaining access could involve physical attacks, social engineering, or exploiting software vulnerabilities to infiltrate SG networks.
- *Vulnerabilities at Application, Embedded, or Network Level:* SG components are susceptible to attacks at various layers. For example (i) application-level vulnerabilities might involve flaws in software applications or improper access controls, (ii) embedded vulnerabilities could target IoT devices and SMs, which often have limited processing power and may lack robust security mechanisms, and (iii) network-level vulnerabilities include issues in communication protocols or encryption weaknesses that allow attackers to intercept and manipulate data.
- *Direct or Indirect Connectivity:* Adversaries can exploit either direct connections (such as device-to-device communication) or indirect connections (such as communication through the internet or third-party networks) to achieve their goals.
- *Targets Customers, Service Providers, and Operations:* The attacks may impact a wide range of stakeholders, including: (i) customers by compromising their privacy and personal data, (ii) service providers by disrupting energy distribution, leading to service outages or financial losses, and (iii) operations by causing inefficiencies, equipment damage, or manipulation of data to distort real-time energy management.

We discuss different privacy-related attacks in the following section.

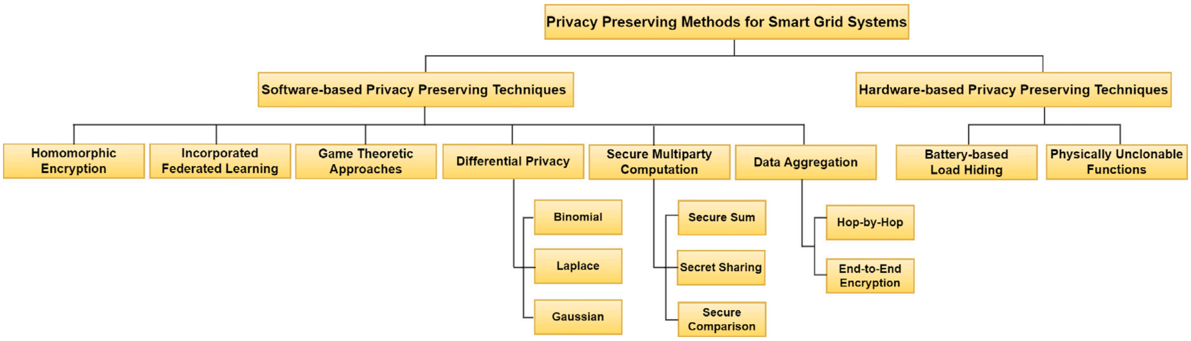
Privacy Preserving Methods for Smart Grid Systems
Software-based Privacy Preserving Techniques
Hardware-based Privacy Preserving Techniques
Homomorphic Encryption
Incorporated Federated Learning
Game Theoretic Approaches
Differential Privacy
Secure Multiparty Computation
Data Aggregation
Battery-based Load Hiding
Physically Unclonable Functions
Binomial
Laplace
Gaussian
Secure Sum
Secret Sharing
Secure Comparison
Hop-by-Hop
End-to-End Encryption

Fig. 5. Classification of Privacy Schemes for SG Systems.

- **Impersonation Attacks:** At any given moment, an adversary has the potential to intercept and capture data transmitted within the SG by other SMs, subsequently gaining access to the content and recovering information regarding energy consumption within a smart home. The SM's memory stores the on/off status of each device, and every 15 min, the appliances transmit their consumption data to the SM. If a compromised appliance impersonates another device, it can generate false readings for a certain period of time unless detected and rectified [86]. For instance, if an air conditioner masquerades as a light bulb or a fan, or vice versa, it can significantly impact the billing. Moreover, when some appliances impersonate the SM and instruct other appliances to report electricity consumption every 15 min, it can pose dangerous consequences, potentially leading to disasters or even electricity theft [66].
- **Eavesdropping:** SGs serve not only as a means of electricity supply between the grid and homes but also as communication channels enabling smart homes to interact with the SG. They facilitate the transmission of various control messages and allow for advanced power demand forecasting. However, if an adversary eavesdrops or gains unauthorized access to an SM, they can easily gather sensitive information about the homeowner's routine, living habits, lifestyle, and interests (including preferred TV channels) [86,87]. Additionally, they can determine when the homeowner is at work or at home. This kind of information compromise poses a significant threat to customer privacy and can be exploited to commit theft or other illicit activities.
- **Replay Attack:** Home automation systems and the SG maintain a continuous flow of communication, exchanging information regarding power consumption and predicting potential electricity demands. In the event of breaching an SM or appliance, an adversary would have the ability to access the consumption reports. This opens the possibility for the adversary to perform replay attacks, wherein an old consumption report is substituted for the current one. Furthermore, they can manipulate the supply–demand report or replay previous control messages [88]. For instance, if there is a request for additional power or an appliance is configured to operate in the off-peak period, replay attacks may modify the demand to lower power or cause the appliance to turn on simultaneously, resulting in inconvenience or disruption [89].
- **Alteration Attack:** Essentially, the alteration attack occurs when an appliance, HAN or SM has been compromised, enabling a malicious party to modify the consumption statistics or forge messages illegally. The consequences of such forged messages or altered power usage reports can be detrimental [90]. For instance, when a message is intended to set an oven's temperature to 120 °C but is altered to activate the water heater at the same temperature, it can result in personal injury or system failure, such as a short circuit. Additionally, forged consumption reports can lead to customers being billed for the electricity they did not actually consume, causing financial harm [91].
- **Message Modification Attack:** In SG, communication plays a crucial role in distinguishing it from conventional grids. However, if a malicious entity exists between the HAN and SG, it can manipulate the messages being transmitted to or from the HAN/SG. This malicious interference has the potential to erode trust between the involved entities, ultimately resulting in significant damage on both sides [92,93].
- **Energy Export/Import Attack:** The SG facilitates the implementation of distributed power generation, enabling consumers to install renewable power generation resources at their premises. This allows them to supply excess energy to the national grid while also being able to request additional energy resources from the grid when required. However, in the presence of an adversary, there is a risk of manipulative behavior [92]. For instance, the adversary may falsely demand energy imports from the grid when they are unnecessary while exporting energy to the grid even when required for domestic use. Similarly, if a plug-in electric vehicle unnecessarily draws energy from the grid during peak hours, it can lead to power shortages and necessitate load shedding.

4. Privacy-preserving solutions for SG systems

In this section, we investigate various privacy-preserving solutions developed and applied in the SG domain. The realization of these privacy-preserving schemes involves a number of considerations, for instance, the description of the system and communication model (*e.g.*, SG marketing architecture, home gateways community, and vehicle-to-grid architecture, etc.), targeted adversarial or attacker model (*e.g.*, data-based attacks, physical-based attacks, and key-based attacks, etc.), the explanation of the privacy concern (*e.g.*, the privacy of users' data, identity privacy, and location privacy, etc.), selection of approach (*e.g.*, cryptographic methods or non-cryptographic methods), explanation of primary steps of the designed method (*e.g.*, initialization, registration and aggregation process, etc.), and security analysis and performance evaluation criteria (*e.g.*, computational cost and communication cost, etc.).

Based on the above observations, SG privacy-preserving techniques may be categorized into multiple privacy-preserving models. So, we put forward a detailed taxonomy of these privacy schemes and present an overview of each category using references to contrast and compare them.

We broadly classify the SG privacy-preserving techniques into software and hardware domains. The software category includes solutions that preserve the privacy of the system through some logical algorithms and methods, while the hardware group of schemes incorporates privacy in the system with the help of some physical entities or devices.

Fig. 5 shows a comprehensive taxonomy of SG privacy-preserving approaches. It is worth noting that the individual classes in the figure do not necessarily demonstrate one technique exclusively. Indeed, a privacy-preserving mechanism reported in the existing literature possibly considers multiple concerns that we mentioned earlier in this section and implements combinations of them. The following section compares the software and hardware categories with the sub-classes.

4.1. Software privacy techniques

The first major category of privacy-preserving techniques for SG systems that we consider is software-based approaches. These approaches are developed over time and are independent of any external, tangible entities. In the following subsections, we discuss the key features of these methods in detail. Additionally, we present the positive and negative aspects of each category of privacy-preserving methods for SG systems as lessons learned at the end of each section.

4.1.1. Homomorphic encryption

Concerning energy consumption data privacy in SG systems, a significant part of the literature consists of employing cryptography and designing methodologies to limit information leakage [28]. These techniques are being designed to protect the privacy of SGs. To prevent the leakage of users' data through SM, most cryptographic privacy-preserving methods use homomorphic encryption [28,94]. On the other hand, many schemes use aggregation methods that combine the SM data before forwarding it to CC rather than sending individual reports from household data [95,96].

Numerous research efforts have been made to address the privacy concerns in SGs, and various privacy-aware mechanisms for SG environments have been suggested so far [20]. For instance, one straightforward approach to preserve the privacy of users is the use of encryption on individual SM profiles before their transmission. Homomorphic encryption is widely used to ensure data privacy in the aggregation process to protect data privacy because it allows for computational operations over encrypted data [97–100]. In homomorphic encryption, plaintext tasks can be converted into corresponding ciphertext tasks. The SM encrypts the raw data using homomorphic encryption and sends it to the aggregator, where additional operations can be performed on the encrypted data. Afterward, the aggregator combines the protected data coming from SMs and may decrypt the accumulated ciphertext [101,102].

Generally, an encryption scheme would be homomorphic to an operation "$*$" if it is compatible with the expression $E(m_1) * E(m_2) = E(m_1 * m_2)$, $\forall\ m_1, m_2 \in M$, here E represents the encryption mechanism and M represents the set of feasible messages. Informally, homomorphic encryption allows one to perform a computation over cipher texts to derive encrypted results so that if this cipher text is decrypted, the result obtained from plain texts will remain the same. Here, we discuss the homomorphic encryption-based privacy-preserving techniques from recent years in detail. Table 3 summarizes various Homomorphic Encryption-based Cryptographic methods for SG privacy.

Chen et al. [8] presented a Paillier homomorphic encrypted SM data aggregation scheme that enables the utility service provider to obtain the accumulative consumption readings of all SMs instead of getting them from individual SMs. This algorithm can also report different data elements contained within a message, making it easier for a supplier to execute variance analysis and one-way variance analysis on the data. The authors exhibited that their proposed method is secure and more cost-effective both for the user and the intermediate gateway. Chen et al. [103] extended their scheme to an elliptic-based scalable and multidimensional data aggregation scheme. The new scheme had an improvement over the preceding scheme with some additional features, such as the possibility of recording or deleting SMs in case of SM failure.

Wang et al. [104] examined the security risks associated with edge computing devices in SG environments and proposed a lightweight yet reliable algorithm for the authentication of edge nodes during initialization of communication with the aim of preventing information outflow and key loss. The proposed scheme employed the Elliptic Curve Encryption (ECC) method to encrypt the device's private information and then used timestamps to validate every session. In addition, the authors computed the random numbers to strengthen the security of their authentication system.

The robust features of the ECC algorithm restricted the attackers from deciphering ECC in polynomial time and thereby made them unable to retrieve information about random numbers and key values. By using the ProVerif tool, the authors proved that the proposed algorithm outperforms the existing techniques with regard to reliability and security, has lower computation and communication costs, and meets the cost requirements for larger-scale SG implementations.

Gai et al. [105] proposed an edge computing-based blockchain model for SG environments aimed at solving the two significant problems of SGs, privacy preservation and energy security, by combining blockchain with edge computing, transparent processes provided by the framework assisted in identifying improper energy consumption behaviors to avoid and eliminate energy-related attacks. The authors used Covert Channel Authorization (CCA) and group signatures to guarantee user authenticity. Further, they utilized smart contracts of the blockchain to construct an optimized security-aware model. A group of super nodes has also been initialized in blockchains, which are responsible for allocating available resources and using the voting method of blockchain to verify user identity. Models based on this approach had better decision-making capability in terms of security performance, energy consumption, and execution time.

Lyu et al. [106] attempted to attain a balanced trade-off in terms of utility and privacy and developed a method of fog-based data aggregation scheme. They utilized the Gaussian mechanism more expeditiously and distributed the noise generation mechanisms among the parties. Then, two-layered encryption is employed to enhance security and ensure the aggregator's obliviousness. In the first step, One-Time Pad (OTP) is applied to encrypt the individual noisy measurements to secure the aggregator's obliviousness. In contrast, in the second step, public-key encryption is implemented to perform authentication. These modifications prolonged the network's lifetime by reducing time delays, saving energy consumption, and mitigating bandwidth bottlenecks.

An energy data privacy-preserving with energy theft detection system for the SG has been proposed by Yao et al. [107]. To detect the abnormalities in metering data, the authors modified the CNN model and used it with a long-period pattern observation method. Additionally, the consumption data of users, as well as the number of honest users, are protected through Paillier homomorphic encryption. Their experimental results revealed that the proposed technique can detect anomalies accurately up to 92.67% with low communication and computational overhead compared to previous detection schemes.

In 2020, Sui et al. [108] designed a new identity-based tree model for secure data aggregation in SG systems. In this protocol, the SM can create signatures without running signing algorithms upon transmission. Due to the generation of homomorphic ciphers of SM data, privacy is preserved without a third party, resulting in reduced overall computation costs. Besides, the SMs execute finite iterations to create signatures for aggregated ciphers, and it requires a constant time to compute and verify aggregated signatures. Simulations and results of the proposed protocol indicated its superior performance as compared to other existing tree-based models.

As a means of ensuring privacy in SG communications, Ge et al. [109] designed a new scheme named Fine-Graded Data Analysis (FGDA). As compared to previous methods, the SM is fault tolerant in this approach and only needs to share the users' energy consumption information with higher entities only once. The proposed FGDA algorithm eliminated the use of bilinear maps and Pollard's lambda and enabled the CC to perform various statistical computations, reducing the communication overhead and achieving better communication efficiency. Zhang et al. [110] proposed a spatial and temporal aggregation scheme for preserving privacy for smart metering to protect individual customers' fine-grained utility usage information.

The method is resilient in various attack scenarios, including when the CC colludes to decrypt the data coming from gateways that were initially encrypted with homomorphic encryption.

Lessons Learned: The homomorphic encryption scheme can be based on several different approaches, such as Paillier encryption [111], Lattice-based encryption [55], Elliptic Curve encryption [112], and ElGamal encryption [113]. Homomorphic encryption-based approaches preserve users' data privacy. However, these encryption methods pose a significant computational load on resource-restrained SMs, which generally have limited processing capacity. Additionally, these encryption models require a trusted third party for their implementation. Since SMs transmit periodic and frequent data updated to aggregators, homomorphic encryption is a less feasible approach for ensuring privacy.

4.1.2. Differential privacy for preserving privacy in SGs

An alternative solution to cryptography for ensuring privacy for data aggregation in SG systems is the use of data perturbation [114–116]. Users or SMs perturb their readings by adding random noise before transmitting it to an aggregator. An example of this approach is the use of differential privacy [117,118]. Differential privacy makes it possible to conceal the information on energy consumption either with random noise and/or simply apply a particular algebraic operation on the fine-grained energy consumption data [119,120]. Differential privacy is an empirically proven framework for preserving privacy in various massive data aggregation frameworks for SGs [121].

The following section explains the preliminary knowledge for implementing differential privacy on a given data for ensuring information privacy.

- **Differential Privacy**: Differential privacy measures the privacy guarantees provided by an algorithm. Initially presented by Dwork in [122], motivated by the instinct that the output of a randomized mechanism applied on two neighboring datasets is approximately indistinguishable.
- ϵ**-Differential Privacy:** An algorithm is considered differentially private if addition or removal of a single record from the dataset does not substantially affect the output of that algorithm. Formally, a mechanism $\mathcal{M}$ is ϵ-differentially private if $\forall$ D_1 and $D_2 \in D_n$ differing in only one record, and for any set of outputs $R \subseteq Range(\mathcal{M})$, the differential privacy equation holds [122], $\Pr[\mathcal{M}(D_1) \in R] \leq e^{\epsilon} \Pr[\mathcal{M}(D_2) \in R]$, Where $Range(\mathcal{M})$ denotes the range of outputs for the query function, while ϵ is a non-negative parameter that controls the privacy level. Lower values of ϵ provide a stronger privacy guarantee. In differential privacy mechanisms, a random noise, typically Laplacian distribution noise, is employed to obfuscate raw data. In practice, ϵ-differential privacy is typically implemented using a core mechanism, such as the Laplace mechanism, which depends on the ℓ_1-sensitivity parameter which quantifies the change in output of the function with addition or removal of an individual's data from the dataset [123]. Mathematically, the ℓ_1 sensitivity of a given function f is represented as: $\Delta f = \max_{D_1,D_2\,:\,D_1 \sim D_2} \|f(D_1) - f(D_2)\|_1$.
- **Laplace Mechanism:** The Laplace mechanism uses ℓ_1 sensitivity to ensure differential privacy by adding noise computed from the Laplace distribution of function output. Given a function f and sensitivity Δf, the Laplace mechanism is given as: $\mathcal{M}(D) = f(D) + Lap\left(\frac{\Delta f}{\epsilon}\right)$. Here Lap($\frac{\Delta f}{\epsilon}$) represents the Laplace distribution with scale $\frac{\Delta f}{\epsilon}$, mean = 0, and ϵ is the privacy parameter.

Table 3
Summary of homomorphic encryption-based cryptographic methods used in privacy-preserving schemes for SGs.

Scheme	Approach used	Entities involved	Privacy goals	Key characteristics
Chen et al. [8]	Paillier homomorphic encryption-based SM data aggregation scheme	SM, aggregator, utility service provider	Privacy of SM, identity privacy, location privacy	Efficient in computational cost at user and aggregator
Chen et al. [103]	Elliptic curve encryption-based multidimensional data aggregation scheme	SM, aggregator, utility service provider	Privacy of SM, identity privacy, location privacy	Efficient in computational and communication cost, report or delete SM in case of failure
Xiao et al. [104]	Elliptic curve encryption-based identity authentication scheme	Edge nodes in SGs	Identity privacy, location privacy, authentication	Preventing information outflow and key loss, reliable and secure, low computational and communication cost
Gai et al. [105]	Covert Channel Authorization (CCA) and group signatures-based authentication scheme and blockchain	Edge nodes and super nodes	Identity privacy, location privacy, authentication	Improved security performance, execution time, and energy consumption
Lyu et al. [106]	Fog-based data aggregation scheme with two-layered encryption	SMs and fog nodes	Identity privacy, location privacy, authentication	Balance between privacy and utility, reduced time delays in the network, improved bandwidth utilization
Yao et al. [107]	Energy data privacy with energy theft detection system using CNN model and Paillier homomorphic encryption	SMs, aggregator, CC, and TTP	Privacy of users' data, confidentiality, integrity, and authentication	Anomaly detection with higher accuracy and utility, low communication and computational overhead
Sui et al. [108]	Identity-based tree model for secure hop-by-hop data aggregation	SMs, data collector, key generator	Privacy of users' data, integrity, and confidentiality	Reduced computational costs and transmission delays, prevents MITM, replay and masquerade attacks
Ge et al. [109]	Fine-Graded Data Analysis (FGDA) scheme for data privacy	SMs, local gateways, CC	Privacy of users' data, authentication, integrity, and confidentiality	Fault tolerant, reduced computational costs, and improved communication efficiency
Zhang et al. [110]	Spatial and temporal aggregation scheme with Paillier homomorphic encryption	SMs, local gateways, CC	Location privacy, identity privacy, integrity, and confidentiality	Prevention against internal and external attacks without additional computation and communication overheads

Differential privacy describes privacy without making any assumptions about the background knowledge of the adversaries. Therefore, differential privacy can resist the majority of privacy attacks, such as linkage attacks. Furthermore, differential privacy provides a proven privacy guarantee with a quantitative analysis of privacy disclosure risk through statistical probability models. However, in time-series data, where the data is usually correlated, implementation of differential privacy becomes a challenge to guarantee utility [124].

As a means of preventing adversaries from determining the classification of household appliances in SG, Ou et al. [125] combined the Singular Spectrum Analysis (SSA) with Local Differential Privacy (SSA-LDP) to prevent the adversaries from identifying SG classification for household appliances. Initially, the Fourier spectrum noise was combined with the geometric sum, which was proven to produce Laplace noise. The findings revealed that optimized SSA-LDP achieves a higher level of privacy while still delivering a higher level of data utility as compared to other privacy approaches. Additionally, the optimized SSA-LDP can also be employed for a variety of time-series datasets for privacy preservation besides the SG applications. Table 4 summarizes various Differential Private Approaches for SG privacy.

Fernandez et al. [126] focused on obtaining privacy-preserved load forecasts in dynamic power generation demand scenarios and investigated how federated learning can be combined with some privacy-preserving mechanisms, including secure data aggregation and differential privacy. For this study, the authors simulated a large residential dataset and analyzed the effect of different privacy-preserving solutions and federated learning frameworks. Their results revealed that combining federated learning with data aggregation and perturbation techniques may acquire high prediction accuracy and near-complete level privacy. Moreover, they identified that these combinations allow a higher level of information exchange with the assurance of preserving the privacy of both forecasting models and local consumption data.

Table 4
Summary of differential private approaches used in privacy-preserving schemes for SGs.

Scheme	Approach used	Privacy goal	Implementation context	Key characteristics
Ou et al. [125]	Singular Spectrum Analysis (SSA) with Local Differential Privacy (SSA-LDP)	Privacy of users' data, identity privacy	Fourier spectrum noise addition; decomposition of SSA	Higher data utility with privacy, universal model for time series data
Fernandez et al. [126]	Differential privacy and federated learning for privacy-preserved load forecasts in dynamic power demand scenarios	Privacy of users' consumption data, identity privacy	Simulation of large residential dataset, combining federated learning with secure aggregation and differential privacy	High prediction accuracy with absolute privacy
Zheng et al.[127]	Decentralized Differentially Private (DDP) method with random permutation algorithm	Privacy of SM data, identity privacy, location privacy	Addition of Laplace noise at clients; permutation to hide the sensitive information	Decentralized approach; Protection against NILM attacks
Wu et al. [128]	Secure and Efficient Multifunctional Data Aggregation (SEMDA) scheme, differential privacy for privacy preservation	Privacy of users' data, confidentiality, integrity, and authentication	System initialization; gathering of data; aggregation at edge; aggregation at cloud	Fine-grained, robust and multifunctional, low computational and communication overheads
Hassan et al. [129]	Differential Privacy-based Load Monitoring (DPLM) approach, differential privacy for privacy preservation	Privacy of users' data, confidentiality, and integrity	Consumption data collection; aggregation; noise addition; peak load value restriction	Real-time privacy protection, limited peak values
Gough et al. [130]	Differential privacy for SM data protection with diverse privacy preferences	Privacy of consumers, authentication	Customer privacy preference identification; noise addition; Shapley value, Nucleolus and Vickrey–Clark–Groves (VCG) cost allocation to different classes of customers	Privacy of customers in line with personal preferences, no additional loss or cost at the retailer, scalable

Taking another step ahead, Zheng et al. [127] presented a Decentralized Differentially Private (DDP) method by joining the differential privacy with a random permutation method to protect the SM data measurements and, in that way, uphold the power usage patterns. This privacy-preserving mechanism introduced the Laplace noise in the gathered power measurements at the customer's end in a distributed way, and then it permuted this power usage sequence by separating the measuring and releasing time points of meter data. This leverages differential privacy over aggregated power data and erases the sensitive characteristics of power usage and the operating state of the connected appliances to ensure privacy from curious participants and other malicious adversaries in SG networks. As part of the study, the DDP framework was evaluated by using the Electricity Consumption and Occupancy dataset based on NILM attack protection and in terms of aggregation and accumulative errors analysis.

Wu et al. [128] proposed a Secure and Efficient Multifunctional Data Aggregation (SEMDA) scheme without involving trusted third parties in edge-enhanced systems. The authors employed lightweight cryptography algorithms to make their data aggregation approach fine-grained, robust, and multifunctional. Moreover, they introduced differential privacy in their proposed SEMDA to preserve privacy. An honest-but-curious model is adopted to analyze the security of SEMDA, whose outcome demonstrated that the SEMDA method performs well in achieving authentication, privacy, integrity, and confidentiality. Meanwhile, comprehensive experimental results revealed that SEMDA is more effective concerning computational and communication overheads.

Hassan et al. [129] integrated differential privacy with renewable energy resources and proposed a Differential Privacy-based Load Monitoring (DPLM) approach to preserve the privacy of consumers in real-time. This method masks the load data of users in a way that would prevent the utility from being able to estimate the consumption of renewable energy resources and SM. Conversely, it provides grid utilities with adequate information to calculate demand response and load forecasts. The authors compared the DPLM scheme with the Gaussian Noise Differential Privacy (GNDP) approach and revealed that SG users' data privacy can be protected by efficiently adding noise with a smaller tolerance of only 1.5%.

Gough et al. [130] utilized differential privacy to protect the SM data with diverse privacy preferences of consumers. The authors investigated the impacts of their proposed algorithm in terms of distribution grid operations from both consumer cost and energy retailer perspectives. Further, they studied the additional costs, power quality variations and losses to the grid by using an AC Optimal Power Flow (OPF) model. In addition, several cooperative game theory-based cost allocation procedures were employed to divide the extra costs among participants in an efficient, unbiased, and equitable way. Generally, the proposed approach preserved

Table 5
Summary of SMC approaches used in privacy-preserving schemes for SGs.

Scheme	Approach used	System model components	Privacy goal	Key characteristics
Wagh et al. [135]	Privacy preservation using Shamir's secret sharing scheme	SM, aggregators, UC	Privacy of SM, identity privacy, location privacy	Resilient at a single point of failure
Cheng et al. [136]	Multi-party household load scheduling framework using homomorphic encryption	HEMS, aggregator, DSO	Privacy of users' data, identity privacy, location privacy	Reduced peak loads for the aggregator, no increase in electricity costs for the customers
Palacios et al. [137]	Symmetric and asymmetric key encryption and random additive shares-based aggregation scheme	Home agents, aggregators, DSO/TSO	Privacy of users' energy data, identity privacy, location privacy	Optimized computational costs
Khan et al. [138]	Fog-Enabled Secure Multiparty Computation (FESMPC) data aggregation mechanism	SM, Fog nodes, CC	Privacy of SM, identity privacy, location privacy	Fault-tolerant, robust against FDI attacks, improved communication and computational costs

the privacy of customers while respecting their personal preferences and contributed with no significant losses and additional costs on the energy retailer side.

Lessons Learned: With regard to differential privacy, the main advantages are its comparatively low cost and analysis, its sequential presentation, and straightforward implementation, demonstrating its strength and usefulness in real-world SG applications. These methods can be efficient; however, current differential privacy-based solutions lack the ability to effectively stabilize the trade-off between achieving data privacy and maintaining the utility at the same time [131]. When the user consumption profiles are obfuscated individually at SMs, the accumulated noise variance in aggregated results increases linearly over time. Consequently, it degrades the data utility of aggregated results and can lead to accuracy loss in different grid operations. In differential privacy, there is no guarantee of complete privacy. An adversary can still use the generalization of datasets to draw statistical inferences on user behavior. However, if more users are introduced to a dataset, the impact of individual users may be reduced, and therefore, it would be harder to make statistical inferences about an individual with a larger dataset.

4.1.3. Secure multiparty computation for preserving privacy in SGs

Secure Multiparty Computation (SMC) is an alternative to trusted third-party frameworks. SMC is a general primitive in cryptography that empowers various distributed parties to compute a function jointly without disclosing their private information [132]. SMC ensures that after a multiparty computation, the participants can learn nothing except their own inputs and the results. Therefore, intended data should be deduced only from inputs and results. Initially, SMC was developed to address Yao's Millionaire Problem [133], where any two parties may determine who is wealthier without revealing their real assets. Yao's two-party mechanism has been extended to multi-party in [134], and from then, numerous studies have focused on both theory as well as application-based SMC protocols to address security and privacy concerns in emerging technologies such as IoT, cloud computing, and mobile computing. For privacy-preserving, SMC is appropriate when:

- The functions are available and known.
- The computation task is distributed across multiple owners of data.
- Each owner of data has a privacy requirement.

Accordingly, SMC is expressed as : Assume N parties $\{P_1, \dots, P_N\}$ with N databases $\{D_1, \dots, D_N\}$. Here, SMC will determine a joint function $f(D_1, \dots, D_N)$ by preserving the private input information D_i of each party. Primarily, SMC is implemented through homomorphic encryption. Table 5 summarizes various SMC-based schemes for SG privacy that appeared in recent literature.

Wagh et al. [135] employed Shamir's Secret Sharing Scheme (SSSS) to restrict the electrical UC and dedicated aggregation devices from linking the aggregated measurements to a particular SM. The authors conducted a feasibility study to evaluate the scheme with regard to security. The proposed mechanism was capable enough to offload the aggregation jobs from the electrical utility, distribute trustworthiness among multiple available dedicated aggregation devices, and make the framework resilient to the single point of failure.

A growing number of distribution networks are incorporating renewable energy sources. In order to improve system reliability, demand-side management has become increasingly important. Home Energy Management Systems (HEMS) are currently capable of decentralized real-time DSM. Due to competition and privacy concerns, it may be difficult to coordinate these HEMSs. When HEMSs shift their loads without coordination, additional peaks may appear at the local aggregators.

To ensure encrypted data is shared between HEMSs, a multi-party framework for scheduling household loads based on HE was proposed by Cheng et al. [136].

Table 6
Summary of game theoretic approaches used in privacy-preserving schemes for SGs.

Scheme	Approach used	Privacy goal	Application scenario	Key characteristics
Chung et al. [139]	A model-free approach with non-cooperative stochastic game	Privacy of users' data	Intelligent management of load consumption profiles of home appliances	Reduction in power grid operation cost and household electricity cost
Wang et al. [140]	Nash Bargaining theory, cooperative game without need of a central entity	Privacy of consumers	Peer-to-Peer (P2P) marketing architecture for energy trading	Direct energy trading between prosumers, enhanced security, eliminated the need of third party
Doan et al. [141]	A game theory-based double auction mechanism for energy trading with Stackelberg equilibrium	Privacy of consumers' bids and queries	P2P marketing architecture for energy trading	Buyers are able to set their energy demand according to varying prices, Maximize social welfare of participants

In this study, tree-based deep Q-networks were developed as a reinforcement learning method for solving decentralized demand-side management in real-time. Studies have found that this model of data sharing performed better than conventional demand-side management in terms of reducing peak loads for the aggregators while providing no increase in electricity costs for the customers. Additionally, the proposed reinforcement learning method protects the privacy of users in comparison with reinforcement learning methods that do not preserve privacy.

To address the problem of high computation costs of cryptographic techniques for privacy preservation, Palacios et al. [137] proposed a symmetric and asymmetric key encryption and random additive shares-based aggregation algorithm. The authors compared their proposed method with Additive Homomorphic Encryption (AHE) schemes and advanced MPC mechanisms. Their findings revealed that generic encryption techniques like homomorphic increase computation and communication costs significantly at the client nodes. In contrast, although MPC methods offer better resilience to large networks, they have higher communication costs among users. In these conditions, the random additive shares method is the most optimized solution for demand-side management with better performance, a straightforward flow of information, and the flexibility of adding redundant intermediate parties for greater resilience.

Khan et al. [138] proposed a privacy-aware and secure Fog-Enabled Secure Multiparty Computation (FESMPC) data aggregation mechanism for SG systems. This secure data aggregation scheme is privacy-preserving, fault-tolerant, and robust against collusive FDI attacks. Paillier homomorphic cryptography is applied to ensure user data privacy during the aggregation phase.

Concurrently, an extended Shamir's secret scheme is employed to deal with collusive attacks. The authors conducted an in-depth security analysis to demonstrate that their proposed solution achieves data privacy, fault tolerance, and source authentication and deals effectively with FDI attacks. Furthermore, extensive experimental evaluations revealed that the proposed mechanism outperforms the other existing state-of-the-art frameworks in terms of computational costs of encryption, decryption, and aggregation, as well as overall communication costs.

Lessons Learned: SMC is mainly implemented by homomorphic encryption. A significant advantage of SMC-based methods is that no trusted third parties can see the private information of users because SMC eliminates the need to trust an external party for data security and broker exchanges. Participants never share data beyond their local network. Additionally, SMC reduces the trade-off between privacy and utility of users' data, as it does not mask or drop data features to preserve its privacy. The entire set of features may be incorporated into data analysis, which can help maintain SG services and operations with high accuracy and precision without compromising privacy. Along with these benefits, there are several limitations associated with SMC-based schemes: (i) high computational costs, which slow down the execution due to the generation of random numbers; (ii) high communication costs incurred by players, as secret sharing requires connectivity across all parties, which may increase communication costs relative to plain text compute and, (iii) Vulnerable to attacks from colluding parties.

4.1.4. Game theoretic approaches used in privacy-preserving schemes for SGs

Game theory was developed as a way to encourage cooperation among parties to arrive at a fair outcome in the system. Game theory is, therefore, a sophisticated problem-solving subset associated with intelligent optimization methods. Game theory models depict a competition between groups of participants who either engage cooperatively or aggressively to maximize their payoffs/outcomes based on the specific set of strategies adopted by progressive players.

The key definitions of game parameters can be summarized as follows [142,143]:

- A game refers to a strategic alliance between cooperative or antagonistic interests where the limitations and rewards for participation are considered for the outcome.
- A player is an integral part of any game, and in the game, each player, denoted by i, is expected to act rationally, indicated by A_i. A player may represent a human, a node, or a group of players N.

- The payoff is defined by a reward or penalty given to a player for their actions during the game provided by $u_i : A \longrightarrow R$, which determines the outcome for ith player, specified by the actions of participating players $A = \times_{i \in N} A_i$, where "×" corresponds to a Cartesian product.
- The term strategy refers to an action plan employed during a game so that a player can take part in a well-planned game $\{N, (A), (u_i)\}$.

Game theory can be applied in the SG security sector to detect malicious nodes, reduce foreign invasions, and reveal self-centered nodes that overburden the entire system. In general, Nash Equilibrium (NE) is an intelligent way of decision-making in game theory which implies that a player is likely to succeed and acquire the desired outcome without changing their original strategy. In the NE, it is assumed that the strategy of each player is optimal relative to other players' decisions. NE represents the optimal action profile of players, $a^* \in A$, from the perspective of an individual player, $i \in N$, it cannot gain reward if radically deviating and opting for another action plan [144]. The process can be expressed as a utility function, $u_i(a_i^*, a^*\_i) \geq u_i(a_i, a^*\_i)\ \forall\ a_i \in A$, where a_i is the strategy of player i and a_i are the strategies of all other players excluding i.

Table 6 presents different game theory-based concepts used in preserving privacy in SG systems.

Since the electricity bills and household appliances' energy consumption are very dynamic, intelligent management of load consumption profiles of home appliances becomes challenging. Chung et al. [139] presented a model-free approach for households, capable of working even with a limited amount of data about uncertain conditions. In particular, a non-cooperative stochastic game is employed for modeling the interactions between the power grid and the households, where the voltage price is considered a stochastic variable. Then, a distributed deep reinforcement learning-based method is adopted to find the NE for their stochastic game. The authors claimed that their method is also capable of preserving the privacy of households. For experiments and assessment of the proposed method, they used the load consumption data of around 1000 households from a real-world dataset of Pecan Street Incorporation [145]. The results revealed that the power grid operation cost and household electricity cost can be minimized with the reduction of average load variance and peak-to-average ratio.

Wang et al. [140] proposed a decentralized P2P marketing architecture for energy trading which not only allowed energy trading between prosumers directly through distribution networks but also enhanced the security and efficiency of the system and eliminated the need for an intermediate entity. The authors applied the Nash Bargaining theory to break down P2P energy sharing problem into two sub-problems: (i) power flow optimization and (ii) payment bargaining optimization. Additionally, the authors adopted a decentralized optimization algorithm, the Alternating Direction Method of Multiplier (ADMM), to preserve privacy in the proposed P2P energy trading model. To evaluate the usefulness of the proposed energy trading framework, a modified 33-bus distribution system-based case study has been performed, which validated the proposed system's effectiveness in loss reduction, enhanced accuracy, and voltage security.

Doan et al. [141] studied the P2P energy trading among multiple prosumers based on a double auction game theory approach. During the purchase process, the buyer can adjust the level of energy to be purchased to match the changing electricity price to maximize its benefit. An auctioneer oversees the game, while the seller refrains from taking part in the event, but eventually reaches the maximum welfare level. This method benefits all the game players by maximizing their reward and protecting their private information, including their bids and queries. The authors further investigated incentive compatibility and individual rationality properties of the proposed auction method at Stackelberg equilibrium. Moreover, they implemented their system using blockchain to demonstrate its applicability in real-time P2P energy trading.

Lessons Learned: Game theory offers privacy protection during data sharing in different SG domains, such as in distributed generation integration, P2P energy trading and demand response programs, along with reduced total cost, improved grid performance and fewer power losses. Furthermore, it encourages the prosumers to participate in energy trading, facilitating P2P trading models and saving energy. Conversely, there may be some possible limitations associated with the game-theoretic models. Modeling and analyzing the complexity of interactions among numerous players in different SG applications is a challenging task. Additionally, the design and development of game theory models may entail significant computing resources and specialized knowledge.

4.1.5. *Privacy preserving though incorporating federated learning*

Federated learning has emerged recently in machine learning, where the global machine learning model is offloaded to devices and trained locally instead of transmitting the input data obtained from local sources to a remote server and then training the machine learning models globally [77,146]. In a federated learning environment, only parameters of local nodes are sent to a central server for global machine learning model learning and analysis [147]. Consequently, federated learning offers a great opportunity to solve the security and privacy issues of SG systems [148].

The traditional way of uploading the power consumption data from SM to the central cloud, followed by executing machine learning algorithms for different services, entails several disadvantages, such as exposing private and sensitive data of customers to potential adversaries and increased latency [149]. Federated learning, due to its inherent characteristics of training the machine learning model at the local nodes and then transmitting only the resulting model parameters to a central device for further training, can effectively address these concerns of SG systems [150,151]. The rest of the section describes the applications of federated learning in preserving SG privacy, which have been designed in recent years. Tables 7 summarize various integrated federated learning approaches for SG privacy.

As SG environments are much more sensitive, have massive data and fragmentation concerns, and the introduction of Multi-Party Computation (MPC) to such systems increases security risks, Yin et al. [152] presented a federated learning approach to design a rigorous data collaboration framework. This paradigm, named FDC, fulfills the requirements of SMC in SG environments.

In particular, the mechanism involved three components: (i) a private data center, (ii) a public data center, and (iii) a blockchain. The private data center was responsible for collecting, storing, managing, and registering the IoT data. The authors used this private data center to perform required computations on local data and eliminated the need to move this data to other resources outside this private cloud. The public data center is comprised of a distributed and heterogeneous data storage system that is designed to facilitate fragmented and scattered data through cryptography, thereby ensuring the internal security of the public cloud. Hence, the private data center intended to perform distinct tasks from those in the public data center. In the next phase, a data access authentication system based on blockchain and a token scheme was devised to realize efficient and adaptable access control for participants with variable incentives. Moreover, to improve the usefulness of data authentication, a unified and linear authorization logic approach was adopted to map the identities of data owners and users. Finally, a blockchain with a distributed consensus paradigm and a distributed ledger was employed to maintain secure records of behaviors for realizing multiparty privacy preservation.

Wen et al. [153] came up with a distributed federated learning privacy-preserving framework for energy theft sensing, namely, FedDetect. The framework considered a federated learning model consisting of a CC, a data center, and several detection stations. In this framework, to preserve the privacy of local consumers, Local Differential Privacy (LDP) is employed at consumer data before sending it to the detection stations, which restricts the access of detection centers to raw data. The authors have also designed a secure algorithm to enable model training to facilitate the detection stations passing encrypted trained parameters to centralized control. Next, the homomorphic encryption method was applied at the data center to secure aggregated model parameters, and the updated parameters were sequentially returned to participating detection stations. Besides, the cutting-edge deep learning-based Temporal Convolutional Network (TCN) model has been extended for successful energy theft detection. For a substantial security verification of the proposed framework, comprehensive data-driven simulations were carried out over a real-world energy consumption dataset in this work. Results demonstrated the proposed federated learning-based approach's high energy theft detection accuracy with a reduced computation overhead.

Su et al. [154] developed an edge and cloud-assisted federated learning mechanism to accelerate the private sharing of data within the SG network. The authors formulated the two optimization strategies, including optimal local data training strategy and optimal user payoff strategy with Non-Independently and Identically Distributed (non-IID) data, to enhance both the privacy preservation of users and communication efficiency. These optimal strategies were employed using deep reinforcement learning for multidimensional private user data and massive state space. Firstly, each energy service provider applies a Deep Q-network (DQN) to acquire the optimal payoff function for users to balance the problem of expenditure and accuracy loss in federated learning. Secondly, each user implements the DQN to obtain the optimal training plan for the local model relating to data size and data quality to amplify its payment in the dynamic environment. The effectiveness of the proposed scheme was evaluated through wide-ranging simulations, and it was found that the proposed model can efficiently work with the training of a high-quality local model on the user side, develop optimum strategies for all the participants, and consequently reduce task delays.

Wang et al. [112] designed a federated deep reinforcement learning-based anomaly detection algorithm to accomplish anomaly detection at the user level with the aim of mitigating inconsistencies between anomaly detection centers. The authors defined a privacy outflow degree and constructed a universal model for user anomaly detection without revealing their user identity, thus preserving privacy with generalized anomaly detection. The technique performs anomaly detection first on anomaly detection centers to find abnormalities and to improve the user anomaly detection accuracy. Moreover, an appeal mechanism has also been introduced to recover the non-anomaly of misinterpreted nodes to increase detection accuracy. Experimental results illustrated that the proposed federated learning-based anomaly detection model precisely detected abnormal nodes with enhanced throughput and minimal latency and thereby can be utilized in the Industrial IoT paradigm.

Lu et al. [155] incorporated federated learning with permission blockchains and guaranteed privacy-preserved data sharing over various distributed participants involved in industrial IoT setup. The author employed differential privacy while training models in federated learning to reduce the risk of leaks in the data-sharing process. They first devised a blockchain-permitted secure data-sharing structure for multiple distributed participants and then incorporated privacy-aware federated learning to frame a machine learning algorithm consistent with the data-sharing problem, thereby maintaining the privacy of user data by exchanging the trained model parameters rather than revealing the original information. Finally, they modified the consensus stage of permission blockchain with the integration of federated learning and utilized consensus computations to speed up the federated training of models. Numerical results demonstrated that the proposed data-sharing mechanism achieved comparatively high accuracy, efficiency, and security when simulated over a real-world dataset.

The utility centers in an SG system analyze the information extracted from SM data to determine the socio-demographic features and needs of consumers so that they can deliver them diverse services. However, SM data is gathered and owned by various retail entities in the retail market that may hesitate to share this information. Accordingly, Wang et al. [156] employed federated learning and put forward a distributed method for the identification of energy consumers' characteristics along with maintaining the privacy level of the retailers. For feature extraction from SM data, authors exploited the Principal Component Analysis (PCA) and then trained an Artificial Neural Network (ANN) in the federated learning environment with various weighted averaging approaches to traverse the SM records and socio-demographic features of customers. The authors verified the performance of the proposed federated learning method through case studies on the Irish Commission for Energy Regulation (CER) [163] with the scenarios of both balanced and unbalanced datasets.

While determining NILM, local data owners may get inadequate information and may breach the privacy of customers. Wang et al. [157] introduced a federated learning-based NILM method called Fed-NILM. This scheme distributes local model parameters among different data owners instead of sharing the local load data. Then, appropriately weighted parameters are added in an average function to obtain the global NILM model. Experiments on two energy consumption datasets are performed in this work

Table 7
Summary of privacy-preserving schemes for SGs through federated learning.

Scheme	System model components	Privacy approach	Privacy goal	Steps overview	Key characteristics
Yin et al. [152]	A private data center, a public data center, and a blockchain	Multi-Party Computation (MPC) and federated learning	Privacy of users and data owners, location privacy, and identity privacy	Local data processing at a private cloud; encrypted data storage at a public cloud	Multiparty privacy preservation with an assurance of integrity
Wen et al. [153]	A data center, a CC, and several detection stations	Distributed federated learning, temporal convolutional networks for model training	Privacy of customers, identity privacy, and energy theft detection	LDP at consumer data; model training at detection centers; secure aggregation at central server	Higher energy theft detection accuracy with reduced computation overhead
Su et al. [154]	SGs with edge nodes and a cloud	Deep reinforcement learning for multidimensional user data and massive state space	Privacy of users and communication efficiency	DQN at energy service provider for user payoff; DQN at user for model training	Optimal user payoff strategy with non-IID data, reduced task delays
Wang et al. [112]	Global, regional, and local anomaly detection centers	Federated deep reinforcement learning-empowered anomaly detection algorithm	Anomaly detection at the user level, identity privacy	Anomaly detection on detection center; an appeal mechanism in case of misinterpreted detection	Increased anomaly detection accuracy, enhanced throughput, and reduced latency
Lu et al. [155]	Secure data sharing in distributed participants	Federated learning with permissioned blockchain for distributed data sharing of parties	Privacy of users, data-sharing in distributed participants	Data-sharing in blockchain transactions	High accuracy, efficiency, and security
Wang et al. [156]	SG with advanced metering infrastructure	Federated learning for privacy of model parameters	Privacy of retailers, privacy of customers, guarantees integrity	Feature extraction; model training	Mapping of customers' preferences for a balanced demand and response
Wang et al. [157]	Local data owners	Privacy of data owners, privacy of customers	NILM model in federated learning settings	Distribution of model parameters among local data owners	Improved scalability and convergence
Lee et al. [158]	Building energy management system, energy storage, and global server	Distributed deep reinforcement learning framework	Privacy of consumers, location privacy, energy scheduling of smart buildings	Generate and broadcast global model to the LBEMS agents	Optimal scheduling of energy consumption of smart buildings and charging and discharging
Lin et al. [159]	Smart household connected to retailers	A hybrid model combining the LSTM and CNN in the federated learning environment	Privacy of retailers, household characteristic identification	Local model training at retailers; delay compensation in global model aggregation	Improved model accuracy and model training speed
Gao et al. [160]	Smart residential buildings and home agents	PriResi, a privacy-aware federated learning load forecasting system	Privacy of users, residential data privacy, load forecasting	Local training edge; broadcast updates among home agents; gradient selection	High forecasting accuracy with reduced communication overhead
Fekri et al. [161]	Smart metering data in residential buildings	Federated learning for load forecasting on machine learning algorithms, FedSGD and FedAVG comparison	Privacy of consumers, residential data privacy, location privacy, load forecasting	Local training of data at the SMs without sharing of data; simulations for FedSGD and FedAVG	Higher precision in load prediction with FedAVG as compared to FedSGD and its localized and centralized models
Zhang et al. [162]	Energy management in Integrated Energy Microgrids (IEM)	Appliance load forecasting using CNN-Attention-LSTM models in federated learning	Privacy of consumers, location privacy, load forecasting, FDI attack	Features extraction using CNN-Attention-LSTM; training of forecasting models	Higher precision in load forecasting,FDI attacks prevention with FedAdagrad

to examine the generalization ability of the Fed-NILM algorithm. Additionally, the authors have compared the proposed Fed-NILM with the centralized Fed-NILM model and localized Fed-NILM model to exhibit superior performance in terms of convergence and scalability.

Lee et al. [158] utilized federated reinforcement learning for privacy-preserved energy monitoring of Shared Energy Storage Systems (SESS) in heterogeneous smart building scenarios. This distributed deep reinforcement learning framework consisted of various Local Building Energy Management Systems (LBEMSs) and a Global Server (GS). Privacy-preserved power scheduling of smart buildings associated with different SESS, is realized in a way that the LBEMS agents send only a random and selective part of the trained neural network model to the GS instead of sending consumers' actual energy consumption data. In their framework, the GS is responsible for executing two processes: (i) to generate and broadcast the energy consumption global model to the LBEMS agents so that they can maintain their models locally and (ii) to train the SESS agent's energy charging and discharging schedule from the utility center and to the smart buildings. The authors conducted the simulation using three smart buildings connected to one SESS with solar photovoltaic systems to show the optimal performance of the proposed scheme in scheduling the charge and discharge time of the SESS and the best possible energy utilization of ventilators, air conditioners and heater in diverse buildings with privacy-preservation of buildings' local electricity consumption data at the same time.

Lin et al. [159] introduced a deep learning model in the federated learning environment for household characteristics identification. In this framework, firstly, a Long Short-Term Memory (LSTM) neural network and a Convolutional Neural Network (CNN) are combined to design a hybrid model for spatial–temporal feature extraction from the load profiles, and this hybrid model is then applied to the distributed federated learning environment. Models are trained locally by each retailer, and only the resulting parameters are shared with other retailers to construct a global model, thereby preserving privacy. Moreover, to enhance the model accuracy and training speed, the authors employed a delay compensation approach in asynchronous stochastic gradient descent. This mechanism updates the global parameters without waiting for additional retailer nodes before aggregation of the global model, which accelerates the training process. Finally, Taylor expansion is used to cope with the gradient delay problem of previous asynchronous methods.

To ensure the uninterrupted performance of smart home appliances along with preserving the privacy of consumers' data, Gao et al. [160] proposed a distributed federated learning mechanism for the neighborhood and examined the load forecasting scenario in residential buildings for analysis. The authors presented PriResi, a privacy-aware load forecasting system that is highly communication-efficient and solves the collaborative training problem in residential buildings without the need for a cloud service. Firstly, a distributed federated learning framework enables users to train all the consumption data on edge locally and broadcast the model parameters updates among smart home agents in each residential area. Secondly, a gradient selection method is adopted, which reduces the number and frequency of gradient aggregations and gradient broadcasts, respectively, to acquire high forecasting results with reduced communication overhead.

Fekri et al. [161] investigated the limitations associated with machine learning load forecasting models and devised a computationally efficient federated learning strategy for load forecasting in SM data. Their scheme trains a mutual model for all the participants, eliminating the need for sharing local data among participating SMs. For analysis, the authors examined two surrogate federated learning approaches: (i) FedAVG, which performs several steps of gradient descent before aggregating updates at the central server, and (ii) FedSGD, which needs one iteration of gradient descent at the client side before combining updates. Since the residential consumption data is highly diversified in nature and these load profiles make it more challenging to train a single model across varying consumers, their results demonstrated that FedAVG outperforms the FedSGD in terms of accuracy and also requires reduced communication rounds. Additionally, FedAVG acquires improved accuracy when compared to localized and centralized versions of model training on SM data.

Integrated Energy Microgrids (IEM) are emerging as an important method of utilizing energy that alleviates environmental and economic stress. IEM planning and operation scheduling is made easier through multi-energy load forecasting, which is integral to energy demand forecasting. By using CNN-Attention-LSTM models based on federated learning, Zhang et al. [162] proposed a method for forecasting the multi-energy load of IEMs that increases data diversity while improving model generalization. For extracting features, CNN-Attention-LSTM is used as the global model. Federated learning allows IEMs to train forecast models distributedly without sharing information locally. Using different learning approaches, this study examined the performance of different central, local, and federated models (FedAdam, FedAdagrad, FedAvg, and FedYogi). Additionally, the consequences of FDI attacks are also examined. FedAdagrad reported the best prediction performance of all the federated models. Its accuracy was similar to the central model, but its precision was higher than that of individual models. Additionally, FedAdagrad can maintain stability in the event of FDI attacks.

Lessons Learned: From the above discussion, we have learned that a federated learning framework has great potential in dealing with the challenges of SG applications, such as preserving the privacy of consumers, resource management, and processing large volumes of data of SGs, low latency and communication cost when transferring large amounts of data rapidly to central device model training [164], real-time data analysis, customized decision making based on individual needs and geographical locations, as well as maintaining the diversity of data [165]. However, to maximize the benefits of federated learning in preserving SG privacy, several challenges have yet to be considered, such as machine learning model poisoning at local devices, detection of malicious computing devices that may affect the accuracy of resulting models, and excessive computing resources required at local devices.

Table 8
Summary of privacy-preserving data aggregation schemes for SGs.

Scheme	Approach used	Privacy goals	Steps overview	Key characteristics
Su et al. [166]	Lightweight and Communication Efficient Data Aggregation (LCEDA) protocol for Smart household connected to retailers	Privacy of users, forward secrecy	SMs form an aggregation zone; masking value share updates to minimize complexity	Scalability, data utility
Saleem et al. [167]	Fog-Enabled Secure Data Aggregation (FESDA)	Privacy of users' data, integrity, and confidentiality	SM data encryption; SM data aggregation; SM data decryption, fault-tolerant aggregation	Prevents FDI Attacks, reduced communication and computation costs
Zhu et al. [168]	Fog-based data aggregation scheme with identity verification	Identity privacy, authentication	Blind signatures and short randomizable signatures for certification; data aggregation	Solves billing issues in SM, anonymous authentication
Zuo et al. [169]	Privacy-preserving multidimensional data aggregation	Privacy of users' data, identity privacy, confidentiality, and integrity	Targeted multidimensional user data for aggregation; ECElGamal cryptosystem for encryption	Reduced computational and communication overhead, centralized approach
Singh et al. [170]	Deep learning-based privacy-aware data aggregation scheme with blockchain integration	Integrity and authentication, irregular electricity usage detection	Data pre-processing at client nodes; data aggregation at cloud	Reliable data aggregation, reduced computational overhead, enhanced SM manipulation detection — Centralized approach
Qiu et al. [171]	Universally Composable Meter Reading Aggregation (UCMRA) for SM data	Privacy of SM data, confidentiality, and integrity	Superincreasing sequence to structure the multidimensional data; encrypt the structured data; data aggregation directly on ciphertext	Minimizes the risk of MITM attack, comparatively low communication and computational overhead
Wang et al. [172]	Homomorphic encryption and blockchain-based distributed data aggregation scheme	Privacy of SM data	Blockchain provides data storage system; homomorphic cryptography over the blockchain for secure aggregation	Data tampering resistance in SMs, secure against eavesdropping, enhanced cost-efficiency and robustness
Mohammad Ali et al. [173]	Novel Homomorphic Privacy-Preserving Protocol (NHP3) for advanced metering infrastructure of SG systems	Privacy of SM data	System initialization; user report generation and aggregation; security report reading	Multi-dimensional data aggregation, reduced computational cost, fault-tolerant, batch verification at global server and aggregator

4.1.6. Data aggregation

The final approach in software privacy-preserving schemes is data aggregation. Data aggregation, which also addresses the privacy issue, takes information from several sources and combines it into comprehensive information. Aggregation is performed to accumulate and concatenate data packets from several entities using average or sum functions. The aggregation process decreases data transmission; however, privacy concerns could arise because aggregation directly accesses plain-text data. In this solution, the power measurements are less accurate as more data are aggregated, thereby degrading their positioning.

When data aggregation is implemented, another privacy issue arises: individual participating SM can view in-between plain-text aggregated results that have been routed through them. The reason for this is that intermediate SMs have the authority to perform accumulated data decryption to implement mathematical operations used in aggregation that cannot be executed over encrypted data. However, this issue can be resolved by incorporating homomorphic encryption to provide reliable in-network encryption for privacy protection. The power usage measurements from respective SMs are encrypted using a semantic encryption algorithm. Meanwhile, aggregation functions are enabled by permitting the algebraic operations of plain data to be performed in the cipher domain. The primary downside of this data aggregation approach is its relatively high computational overhead. Table 8 compares various data aggregation schemes for SGs based on their main attributes.

To cope with the increased computation cost of public key homomorphic encryption communication overhead of masking methods, Su et al. [166] suggested a Lightweight and Communication Efficient Data Aggregation (LCEDA) protocol for secure data aggregation in SG systems. First off, the LCEDA method enables SMs to form an aggregation zone at reduced computation and communication costs. Next, LCEDA performs masking value share updates to minimize complexity and achieve forward security of consumption data. In addition, LCEDA is scalable; it allows SMs to dynamically enroll and leave the system to mitigate the migration and malfunction of SMs. Extensive security analysis and evaluations demonstrated the LCEDA outperforms the other existing methods in terms of preserving individual data privacy and ensuring big data utility.

The Fog-Enabled Secure Data Aggregation (FESDA) approach by Saleem et al. [167] invoked four algorithms. The first algorithm encrypts the SM consumption data by means of Paillier encryption for the sake of privacy, the second algorithm aggregates the metering data, the third algorithm performs decryption, and finally, the fourth algorithm carries out aggregation with fault-tolerance. In addition, the proposed FESDA scheme prevents FDI attacks by screening the misleading information inserted by external adversaries. Performance evaluation of the proposed scheme demonstrated that FESDA reduces the communication cost to half and achieves good results in terms of aggregation and decryption compared to other fog-based privacy-preserving aggregation schemes.

Zhu et al. [168] introduced a fog computing-based architecture and its classic applications for SGs, then examined the privacy and security issues and presented an extensive review of the privacy preservation and authentication requirements in fog-based SG contracts. Furthermore, they put forward a privacy-aware identity verification and data aggregation strategy along with its competency analysis. Specifically, they utilized blind signatures and short randomizable signatures; blind signatures for anonymous certification in a subjective way, followed by the Paillier homomorphic cryptosystem to aggregate the SM measurements. Finally, fog nodes were operated to figure out billing problems effectively once anonymous authentication was completed.

In another study, Zuo et al. [169] offered a privacy-preserving scheme to aggregate multidimensional user data in SG to retain the identity of the user and the privacy of data. The proposed mechanism incorporated the following primary characteristics: (i) feasibility: it targeted the multidimensional data for aggregation; (ii) security: it protected the identity of consumers and privacy of shared data; (iii) robustness: it performed regular operations even in the scenarios where any SM turned off or faulty; (iv) high efficiency: it implemented ECElGamal cryptosystem and delivered an improved computation cost. A random oracle mode was utilized to prove the reliability of the proposed framework. It is revealed from the modeling results that this scheme considerably improved outcomes in computational and communication overhead, which matched the SG requirements.

Singh et al. [170] contributed with homomorphic encryption and deep learning-based privacy-aware data aggregation scheme to strengthen the prediction model precision by reducing the adverse impact of flash workload. Integrity and authenticity were considered privacy measures for given data, and a mutual security framework was provided at the user and cloud layers to help prevent both insider and rival adversaries' attacks. In the next phase, a cloud-based data aggregation mechanism with the integration of blockchain has been employed to share the workload of SMs with less computational power and keep a record of the transactions for security purposes. The authors introduced efficient and reliable data compilation and pre-processing procedures that were used to discard or improve the low-quality data gathered from different Home Area Networks (HANs) before being involved in big data analysis. Based on the performance evaluation and security analysis, the authors claimed that their proposed framework ensured a reliable aggregation of data with significantly reduced computational overhead. Particularly, the proposed system delivered 80% enhanced accuracy as compared to the previous approaches in identifying SM abnormalities along with 20% to 80%, lower computation cost than traditional techniques.

Qiu et al. [171] proposed an SM data aggregation technique called the Universally Composable Meter Reading Aggregation (UCMRA). It was an extension of the Efficient and Privacy-Preserving Aggregation (EPPA) algorithm by Lu et al. [111], targeted at minimizing the threat of the MITM attack. The authors utilized the universally composable symbolic analysis for the security analysis of the proposed UCMRA scheme. Moreover, their results showed that the general functionality of the proposed UCMRA is comparable to the EPPA scheme.

Wang et al. [172] introduced a blockchain and homomorphically encrypted data aggregation scheme for SG systems. This combination improved the privacy-preserving and security level for SM consumption data aggregation. In this method, blockchain with a consensus mechanism provides a data storage system that can deal with trust issues, and enhance cost-efficiency and robustness. The homomorphic cryptography over the blockchain performs data encryption for secure aggregation and transmission without disclosing the original data.

Mohammad Ali et al. [173] focused on customers' privacy concerns in advanced metering infrastructure of SG systems and proposed a Novel Homomorphic Privacy-Preserving Protocol (NHP3). The NHP3 possesses a number of characteristics: (i) it successfully performed multi-dimensional data aggregation with reduced computational cost, (ii) it has fault tolerance properties, and (iii) it is able to do batch verification both at the global server and intermediate aggregation nodes. Their proposed method is fault-tolerant and delivers the same outcomes even if a malicious aggregator or gateway is present in the system. Any compromised SM is not allowed to access other consumers' consumption data. Moreover, this protocol prevents the curious global server from inferring any user data, even if it captures data packets coming from SM to intermediate gateways. Finally, the authors carried out a comprehensive analysis to support the superior performance of their proposed protocol in terms of communication and computation cost.

Lessons Learned: Privacy-preserving schemes based on data aggregation are highly reliable and computationally advantageous. SG efficiency is crucial because electricity is consumed for the execution of these algorithms, therefore there is a tradeoff between energy consumption and privacy, which must be considered when deploying such algorithms. Additionally, a low computational overhead is necessary because of the limited computing power of SMs. To counteract this, SM manufacturers can install comparatively high processing units in SMs, potentially increasing their costs.

Our research and analysis found that data aggregation is a prominent category of schemes to preserve data privacy in SG systems, and the overhead required to set up and maintain the scheme framework is of medium-high complexity. Using ideas from other fields, such as big data, mobile computing and edge computing, potential research could be directed towards further improving data aggregation overhead.

Table 9
Summary of privacy-preserving schemes for SGs using battery-based load hiding approaches.

Scheme	Implementation strategy	Privacy goals	Scenario/ Phases	Key characteristics
Giaconi et al. [177]	Privacy protection through adding battery noise to SM data	Privacy of SM data	Scenario of the presence of both a rechargeable battery and a renewable energy resource	Reduced information leakage rate, requires more battery capacity to ensure privacy
Cho et al. [178]	Privacy with rechargeable batteries in multi-user scenario	Privacy of users' data	Combine power supply of multiple users; derive upper bound and lower bounds for information leakage rate	Privacy leakage can be reduced by a factor with increased cooperative users, requires an increased battery capacity with increased users
Li et al. [179]	Smart metering system with rechargeable batteries	Privacy of users' data	Series reduction to the optimal battery charging policy; determining the information leakage rate	Optimized charging policy and information leakage rate in IID load
Elkazaz et al. [180]	Hierarchical and distributed Home Battery Storage System (HBSS) for P2P energy trading	Integrity and confidentiality	Daily energy cost optimization; P2P energy sharing with reduced operational costs	Significant reduction in annual household energy costs
Liu et al. [181]	Reinforcement Learning with battery-based Intermittently Differential Privacy (RL-IDP) scheme for household power systems	Privacy of SM data	IDP mechanism to ensure privacy through battery noise; RL to maintain battery conditions and cost saving	Maintained a better privacy protection level along with reduced cost of implementing differential privacy

Table 10
Summary of Physically Unclonable Functions (PUF) Used in privacy-preserving schemes for SGs.

Scheme	Approach used	Entities involved	Privacy goals	Key characteristics
Gope et al. [182]	Key agreement authentication scheme	SM, aggregator, utility service provider	Privacy of SM, physical security, authentication	Suitable for SMs having low computational power — Susceptible to epithermal leakage attacks, lack in ensuring session key secrecy and backward secrecy
Tahavori et al. [183]	Key agreement authentication scheme	SM, aggregator, utility service provider	Privacy of SM, physical security, authentication	Robust against physical tampering, lightweight

4.2. Hardware-based privacy techniques

The second major category we consider in this article for preserving privacy in the SG systems is hardware-based privacy techniques. The subsequent sub-sections will discuss the main features of these techniques.

4.2.1. Battery-based Load Hiding (BLH) approaches

Battery-based Load Hiding (BLH) is a well-known method that employs rechargeable batteries to partially meet energy demands to manipulate and hide actual SM readings. The following factors can affect the relative performance of BLH algorithms: (i) Privacy level: Several measures can be used to quantify the privacy level available with charging protocols. These measures generally consider original and obfuscated load data as input and return a value that indicates the degree of privacy achieved as output, (ii) Storage properties: Charging and discharging rate, battery capacity, and round trip time typically represent inputs for BLH algorithms, (iii) Data characteristics: Load profile length, discretization rate, household location and individual power consumption behaviors may impact the efficiency of BLH algorithms. A number of BLH algorithms have recently been developed, including Non-Intrusive Load Leveling (NILL) [174], Best Effort (BE) [175] and Lazy Stepping (LS) [176]. In Table 9, we summarize all the countermeasures that employ additional batteries to hide and protect the SM data to ensure privacy.

Giaconi et al. [177] explored how privacy in SM data can be preserved in conjunction with both a rechargeable battery and a renewable energy resource to obscure the consumers' energy usage data partially. Here, privacy is determined by the leakage rate of data, which represents the proportion of shared data between actual user energy consumption and energy requests from the grid, which the SM transmits to the UP. Their results showed that the leakage rate of information can be decreased by increasing the accessibility to renewable energy sources, and more energy storage is required to effectively utilize the stored energy consumption information to achieve improved privacy.

Cho et al. [178] examined the impact of user cooperation on SM data privacy in the presence of rechargeable batteries. Giaconi et al. considered the one-user per battery; at the same time, the authors in this work looked at the multi-user perspective, where their findings revealed that the privacy leakage could be reduced by approximately up to 1/N with N added users. However, this approach requires an increased battery capacity with increased users. Additionally, it is worth noting that a loose bound is insufficient to protect privacy when more than three users are involved.

Li et al. [179] developed a rechargeable battery-powered smart metering system for the purpose of partially obscuring the user's energy demand. The authors made the following assumption in their model: (i) user energy demand pattern can be presented as a discrete Markov process, (ii) the battery can ensure idealistic charge management, and (iii) the privacy level can be evaluated by comparing the leakage rate of user load input and the battery output. The authors performed a series of reductions on the formulated problem of deriving the optimal battery charge policy and remodeled it in the Markov Decision Process (MDP). In their study, an example of IID load is considered to explicitly express the leakage rate and define the optimal charging policy. This optimal charging policy intuitively preserves an invariance property of the state. Finally, the authors presented two opposing views on acquiring optimality using dynamic programming and information-theoretic arguments.

A decentralized and hierarchical energy management system is presented by Elkazaz et al. [180] that allows prosumers to trade energy in a decentralized manner. Implementing a decentralized Home Battery Storage System (HBSS) combined with shiftable home appliances in a distributed manner can help reduce household energy expenditures further than they would be if each house were operated individually (not as a community member). Three levels are included in the hierarchical system: Selective, P2P, and home level. The lower layer is used to optimize the daily energy costs of each household. A P2P energy-sharing algorithm is then applied to improve the results by selecting pairs of houses with greater cost savings through joint optimization. A sensitivity analysis was performed to determine how the proposed management system's size and efficiency affect the livelihood of stakeholders. The size of the PV power output, the size of the home battery storage, and the household's average consumption per year, are assessed. Results indicated that when operated as part of a community, the proposed system resulted in further reductions in household energy expenses (up to 8.96%), compared to when operated individually in contrast to when operated independently (*i.e.*, not participating in a community).

Liu et al. [181] integrated reinforcement learning with Intermittently Differential Privacy (IDP) and proposed a battery-based RL-IDP scheme enabling a privacy-aware household power system. First, the battery-based IDP mechanism incorporates differential privacy in accordance with the charging and discharging capacity of the battery. A reinforcement learning-based algorithm is implemented to save cost and maintain the battery energy level to generate an acceptable amount of noise. Their experimental results demonstrated the better performance of the RL-IDP scheme in terms of privacy protection and cost-saving.

Lessons Learned: Regarding data privacy, the majority of approaches that use rechargeable batteries or renewable energies come under hardware privacy techniques. Rather than using electricity directly for household appliances, these approaches need battery charging to preserve privacy. More specifically, these techniques hide the individual consumption events in home load signatures and use rechargeable batteries to compensate for the difference between a constant SM load profile and the actual energy consumption of consumers.

Unfortunately, these schemes may compromise the privacy of consumers because of battery capacity limitations and variable charging/discharge rates. Batteries may overcharge or become extremely low in charge to maintain constant loads, and the load-change information may be leaked when there is an obvious change in load.

Moreover, these schemes are potentially vulnerable to attacks that can reveal appliance events. These types of attacks utilize peak load moderation and reduction algorithms to facilitate load shifting in SM load [175]. As a result, BLH methods restrict the functionality of the SG systems to deliver appliance-level energy management, as they are primarily focusing on protecting against NILM attacks, which can track individual appliance usage by analyzing energy consumption profiles through load signature files [184,185].

Additional batteries are expensive; their lifetime is limited and even more reduced due to repeated charging and discharging, and they require significant setup and maintenance costs. This makes it less than ideal for both the UC and the customer.

Furthermore, renewable energy is dissipated when batteries are sufficiently charged, or the energy demand is less than the available energy. Researchers have demonstrated that batteries always have a certain amount of energy loss. It is, therefore, suboptimal to use batteries since the SG's primary goal is to save energy.

4.2.2. Physically Unclonable Functions (PUF) approaches

The second type of hardware-based approach for preserving privacy in SG systems is the use of Physically Unclonable Functions (PUF). PUF features a chip in SMs that can be programmed to generate a private key for encryption and provide protection against internal adversaries. PUF-based devices have a low manufacturing cost and can support hardware-based integrity and authentication mechanisms that are resilient against impersonation attacks. In this method, two SMs are aware of their public key, yet each SM obtains its own private key. This way, a more secure private key is generated from hardware without human interaction [186]. Consequently, PUF achieves consumer privacy through the one-way functions embedded in the physical design of the SM. PUFs use randomly generated signatures in accordance with their complex and unique physical properties. A major characteristic of PUFs is that they are unpredictable and unclonable, thereby preventing internal users in the network from breaching privacy. Table 10 summarizes various Physically Unclonable Functions (PUF)approaches for SG privacy.

Gope et al. [182] introduced a privacy-sensitive key agreement authentication scheme for more secure communication between SM and service providers. This method of the key agreement was based on a PUF and one-way hashing and, thus, was suitable for low computational SMs. The scheme attempted to ensure reliable communications and physical security for SMs.

However, according to Tahavori et al. [183], Gope's proposed mechanism was indeed susceptible to epithermal leakage attack under some adversarial models, such as Canetti and Krawczyk (CK), and thereby lacking in ensuring backward secrecy and session key secrecy. The authors then came up with an enhanced PUF-based key agreement technique to secure end-to-end communications between the SM and the service provider. In experiments, their scheme was proven to be lightweight and robust against physical tamper attacks and secret leakage attacks in SMs.

Lessons Learned: The PUF has evolved into an indispensable privacy and security primitive because it creates a unique fingerprint for each SM, and is built to be robust to reverse engineering. Software and hardware blocks can be protected using authentication protocols devised by the user. However, due to its inherent nature, it suffers from low reliability and is susceptible to both modeling and physical attacks. Accordingly, PUF requires particular methods and processing components for it to be robust and unique, and further research is needed to find ways to fix its inherent flaws so that it can be used as a trusted primitive.

5. Comprehensive overview of privacy-preserving methods in SGs: Benefits and drawbacks

We conducted in-depth research and analysis, focusing on privacy concerns within SGs and exploring potential solutions. Our objective is to assist researchers in both academic and industry settings to enhance privacy within SGs in a more practical manner. Through our comprehensive exploration, this study reveals key insights into privacy concerns within SGs. The following section presents a summary of the proposed privacy-preserving solutions and outlines their pros and cons derived from the findings of this study. We offer our recommendations for the practical implementation of these solutions in SG environments. Table 11 provides a comprehensive overview of the pros and cons of various privacy-preserving solutions for SG systems. Our insights are then condensed into a dedicated "recommendation" column, where we classify each method as "low", "medium", or "high" recommended based on its effectiveness, providing a concise evaluation of their practical suitability.

5.1. Cryptography

Cryptography is the most common and widely used approach for preserving privacy in SG systems. In recent years, researchers have primarily focused on implementing homomorphic-based encryption methods due to their practicability. These methods enable the execution of computational operations over encrypted data. However, at the same time, they impose a computational load on resource-constrained SMs and often rely on a trusted third party.

- **High Computational Overhead:** Cryptographic algorithms, especially advanced techniques like homomorphic encryption and public-key cryptography, are computationally intensive. SMs and other devices in SG systems often have limited computational resources, memory, and processing power, making it difficult for them to perform complex cryptographic operations without affecting the SG's performance.
- **Latency and Delay:** Cryptographic processing introduces additional latency, which can slow down data transmission and decision-making in SG systems that require real-time or near-real-time responses. This delay is problematic for applications that need instantaneous data for grid stability, demand response, and other critical operations.
- **Communication Overhead:** Cryptographic schemes require exchanging of keys, certificates, and sometimes large encrypted data packets, which increases communication overhead. In SG systems, where data is frequently transmitted between devices and central controllers, this can lead to network congestion and inefficient bandwidth usage that can affect the real-time SG operations.
- **Key Management Complexity:** Managing cryptographic keys across a distributed network of devices in SG systems is challenging. Keys must be securely distributed, stored, and periodically refreshed to maintain security, which requires sophisticated key management infrastructure. In large-scale SG systems, keys can be compromised which can render the entire cryptographic scheme vulnerable.
- **Risk of Single Points of Failure:** When practically employed in SG systems, cryptographic techniques often rely on central authorities or trusted third parties for key distribution and management. These entities can become single points of failure in SGs. If the central authority is compromised, the security of the entire SG system could be at risk.
- **Vulnerability to Quantum Computing:** Emerging technologies, particularly quantum computing, pose significant threats to conventional cryptographic algorithms implemented in SG domain. Many algorithms, including ECC, could be broken by quantum computers, necessitating new quantum-resistant cryptographic protocols and imposing additional computational burdens on resource-constrained SG devices.
- **Privacy and Security Trade-offs:** Cryptography focuses on data security (*i.e.*, confidentiality, integrity, and authenticity), it does not inherently address privacy concerns. For instance, in SGs, while encrypted data is secure, metadata such as timing, volume, and frequency of communication between entities such as SMs and CC can still reveal sensitive information about user behavior and energy consumption patterns. Techniques like differential privacy, which focuses directly on privacy rather than security, can sometimes be more effective for preserving privacy in SG systems.
- **Device Constraints and Cost:** Implementing cryptographic solutions for preserving users' privacy in SG systems often requires additional hardware support to handle encryption and decryption operations securely and efficiently. Adding this hardware to low-cost devices, like SMs, increases production costs and may be impractical for widespread SG deployment.

5.2. SMC

To eliminate the need for a trusted third party and leverage the benefits of cryptography, SMC is a primitive method that predominantly utilizes homomorphic encryption. It enables distributed computation of tasks, aiming to enhance the accuracy of data for various SG services and operations. However, SMC comes with high communication costs for execution.

- **No Need for a Trusted Third Party:** Unlike many privacy-preserving techniques, SMC does not require a centralized trusted authority to collect or process data in SGs. Instead, trust is distributed among participating parties such as SMs, lowering the risk of single-point failures or data leaks due to a compromised central authority in large-scale SGs.
- **Improved Security Against Collusion:** SMC protocols can be designed to be secure against a certain number of colluding parties (*i.e.*, a subset of SMs that may conspire to reveal others' inputs). Some protocols, such as threshold SMC, provide security as long as fewer than a specific threshold number of SMs collude.
- **High Computational Overhead:** SMC protocols are computationally intensive, often requiring multiple rounds of communication and significant processing power to ensure users' privacy in SGs. This is a challenge for resource-constrained SMs and real-time decision-making in SG systems.
- **Communication Overhead:** SMC typically involves a substantial amount of communication between SMs or CCs. For example, in large-scale SG networks, each party may need to exchange multiple messages with others, leading to high communication costs and network congestion in SGs.
- **Limited Scalability:** As the number of SMs increases, the computational and communication demands of SMC grow substantially, making it challenging to scale for large networks in SGs to handle complex computations. This is particularly obvious in SG configurations involving millions of SMs.
- **Complexity of Implementation:** Implementation of SMC protocols is technically challenging in SG environment, requiring specialized knowledge in cryptography and distributed computing. The complexity and cost of implementing SMC in large-scale SG systems can be barriers due to a lack of expertise or resources.
- **Susceptibility to Network Failures:** Since SMC relies on continuous communication between parties, it can be disrupted by network failures or interruptions. This can affect the reliability and availability of SG services relying on SMC in environments with unstable network connectivity.

5.3. Data perturbation

Data perturbation techniques have emerged as an optimal primitive due to their straightforward implementation. Data can be perturbed either by adding manual noise distributions to input data or by introducing noise from hardware components, such as a rechargeable battery. However, the main drawback of data perturbation is that it diminishes data utility and may impact the real-time accuracy of SG services, such as billing.

- **Simplicity and Reduced Computational Load:** Perturbation methods, such as adding random noise, are relatively straightforward to implement in SGs, as compared to cryptographic approaches. Perturbation generally requires less computational power than encryption, making it suitable for resource-constrained SMs.
- **Customization for Privacy Levels:** Perturbation levels can be adjusted to control the trade-off between data utility and privacy, allowing for flexible privacy settings based on user or application needs. For example, in differential privacy, the privacy parameter, epsilon ϵ, allows for tuning the level of privacy, which can help balance the need for privacy against data accuracy requirements.
- **Strong Privacy Guarantee:** Data Perturbation such as differential privacy offers a robust, formalized privacy guarantee in SGs that limits the risk of exposing individual information, contributing to enhance the reliability of SGs, even when attackers have additional external knowledge.
- **Impact on Data Utility:** Higher levels of perturbation degrade data quality in SGs, making it difficult to balance privacy and data utility. Critical insights of energy consumption data may be lost if the data is overly perturbed. Differential privacy introduces noise based on the privacy budget, which can significantly affect data utility, especially in applications that require high precision, like fault detection or load forecasting in SGs.
- **Vulnerability to Sophisticated Attacks:** Advanced data mining and machine learning techniques can sometimes bypass perturbation in SGs, especially if attackers have access to auxiliary information. When multiple analyses are conducted over the same data, privacy loss accumulates, which may eventually require higher noise or reduced data usability. This is particularly challenging in SGs where continuous data is collected and precise operations such as billing and pricing are required.

5.4. Game theory

Game theory provides privacy protection during data sharing in P2P energy trading and demand response programs. It encourages prosumers through incentives and rewards, motivating their active participation in energy conservation and improving grid performance with fewer power losses.

- **Dynamic Interaction Modeling:** Game theory can model and predict the behavior of various participating entities in the SGs, including attackers and users, allowing for adaptive privacy mechanisms based on real-time behaviors of energy consumers.

- **Incentivize Privacy-Preserving Actions:** Game-theoretic approaches can provide incentives for users to engage in privacy-preserving behaviors *e.g.*, cooperating with certain privacy protocols or data-sharing agreements while deterring malicious actions.
- **Adaptable to Multi-Agent Systems:** SGs involve multiple stakeholders with competing interests. Game theory is well-suited for scenarios such as P2P energy trading where multiple participants interact and need to reach a balanced solution without a single trusted authority.
- **Higher Complexity:** Modeling all possible interactions and incentives among stakeholders can be computationally intensive and challenging, especially in real-time applications within the SGs. Creating effective incentive structures and strategies for privacy requires a deep understanding of each participant's goals, motivations, and interactions, making game theory complex to design and deploy in SG systems.
- **Limited Privacy Guarantees:** Game theory does not inherently provide formal privacy guarantees like differential privacy. Instead, it relies on incentive structures and strategies, which may be less robust against advanced or adaptive attacks in SG domains such as energy trading markets.

5.5. Data aggregation

Data aggregation is highly reliable and a prominent category of schemes to preserve data privacy in SG systems, which can be combined with other methods, such as cryptography, data perturbation, and federated learning. The overhead required to set up and maintain the scheme framework is of medium-high complexity.

- **Reduced Individual Exposure:** Aggregation hides individual consumption patterns by combining data from multiple users into a single, collective energy consumption pattern. This makes it more difficult to identify or profile a single household's behavior. Aggregated data can still support many SG operations, such as demand response, load forecasting, and overall grid optimization, without exposing individual users' data.
- **Improved Privacy with Minimal Complexity:** Compared to advanced techniques like differential privacy or homomorphic encryption, data aggregation is relatively simple to implement in SGs and can offer privacy improvements without significant computational or storage requirements.
- **Compatibility with Legacy Systems:** Aggregation can be applied without extensive changes to the SG infrastructure, making it compatible with existing systems and protocols. This can facilitate quicker adoption in SG environments.
- **Loss of Data Granularity:** Aggregation inherently reduces the level of details available in the energy consumption data. For tasks requiring precise, real-time monitoring or control such as anomaly detection or individualized billing, aggregated data may not provide sufficient insights.
- **Privacy Vulnerability to Disaggregation Attacks:** In some SG scenarios, attackers can perform disaggregation attacks to reconstruct individual user data from aggregated data, especially if they have side information. This is a major limitation, as aggregation alone may not provide strong privacy guarantees.
- **Limited Utility for Personalized Services:** Aggregated data is less useful for applications that require individual-level insights, such as personalized energy-saving recommendations, fault detection for specific devices, or household-specific billing. Aggregation may limit the ability to deliver these personalized services, which is a substantial aspect of modern SGs.
- **Lack of Formal Privacy Guarantees:** Unlike differential privacy, which provides mathematical assurances about privacy levels, aggregation does not inherently guarantee users' privacy. The privacy benefits are implicit rather than rigorously quantified, reducing the robustness of SGS against advanced attacks or data inference.

5.6. Federated learning

Federated learning has proven to be beneficial in resource management, processing large volumes of data with low latency and communication costs. It can facilitate real-time SG services, such as power consumption forecasting, to reduce peak loads with higher accuracy, ultimately contributing to the maintenance of grid stability.

- **Enhanced Data Privacy:** Federated learning allows individual SMs to keep raw data locally, which reduces the risk of exposing sensitive data, such as household energy usage patterns. Only less sensitive model updates are shared with CCs or UPs, enhancing privacy protection of energy consumers.
- **Personalized Model Adaptation:** Federated learning can enable personalized models by allowing SMs to train locally based on their specific consumption patterns. This can improve the accuracy and effectiveness of services like demand forecasting and personalized recommendations, which would otherwise require individual data processing.
- **Real-Time Processing Capability:** By allowing computations to happen at the edge, federated learning supports real-time data processing in SGd, which is beneficial for time-sensitive applications like fault detection, dynamic pricing, or grid balancing.
- **Vulnerable to Inference Attacks:** Although federated learning reduces the need to centralize data, it does not eliminate privacy risks entirely. Attackers in SGs can still perform inference attacks on the shared model updates from SMs, potentially extracting sensitive information about household behaviors from the gradient data.
- **Device Constraints:** SMs and other edge devices in SGs may have limited computational resources, battery life, and storage. These constraints can hinder the performance of federated learning, especially for complex models requiring substantial processing power.

Table 11
Benefits and challenges of using different privacy methods in SG Systems and our recommendations based on findings.

Privacy method	Benefits	Limitations	Recommendation
Homomorphic Encryption	Allows for computational operations over encrypted data	Computational load on resource-constrained SMs, need trusted third party	Low — While homomorphic encryption ensures strong data security, its high computational demands make it impractical for real-time SG operations. It is best suited for offline processing tasks, such as secure energy billing and privacy-preserving analytics, where computational resources are less constrained
Secure Multi-Party Computation	Eliminates the need for a trusted third party reduces trade-off between data utility and privacy	Communication overhead due to secret sharing, vulnerable to attacks from colluding parties	Medium — Suitable for decentralized SG systems, such as distributed energy trading and demand-side management, where trustless computation is required. However, researchers must address high communication overhead and develop efficient protocols for scalable implementation
Differential Privacy	Simpler computations, able to interactively facilitate machine learning models	Trade-off between privacy and data utility, need trusted third party	High — Highly recommended for privacy-preserving SM data aggregation, demand response programs, and AI-driven energy forecasting. Researchers should fine-tune privacy budgets to balance data utility and security, ensuring practical implementation in real-time grid operations
Game Theory	Interaction between nodes as a cooperative or non-cooperative game, rationality among participants	Complex, comparatively hard to implement	Medium — Suitable for optimizing energy trading, load balancing, and cooperative decision-making in SGs. However, its complexity requires advanced optimization techniques and real-time implementation strategies to ensure effective deployment
Federated Learning	Reduced communication cost and latency, efficient and timely decision-making due to processing at edge level	Need of optimization algorithms, data heterogeneity, device constraints	High — Strongly recommended for privacy-preserving energy consumption prediction and decentralized grid intelligence. Researchers should address communication constraints and model heterogeneity to ensure smooth deployment in resource-constrained edge devices *i.e.*, SMs
Data Aggregation	Aids in gathering information from dissimilar, heterogeneous, and large number of sources	Requires encryption, additional overheads	Medium — Well-suited for SG environments where real-time energy usage data is collected from diverse sources. Optimization of encryption techniques and lightweight aggregation protocols is essential for efficient performance
Battery-based Load Hiding	No need for extra energy for processing contributes to meeting energy demands	Low battery capacity and life, costly to install additional batteries	Low — While effective for masking energy consumption patterns, its dependence on battery capacity limits its long-term viability in SGs. Future research should focus on enhancing energy storage efficiency and reducing implementation costs
Physically Unclonable Functions	Low manufacturing cost, high integrity	Low reliability, can suffer from tampering attacks, requires specific units for processing	Low — PUFs can be leveraged for lightweight and cost-effective hardware-based authentication in SG devices, such as SMs and IoT sensors. However, improvements in tamper resistance and reliability are necessary for wider adoption in grid security applications

- **Data Heterogeneity:** The data generated by different devices in an SG is often non-iid and consumption patterns may vary significantly across households. This heterogeneity can make it challenging to train a global model effectively and may lead to poor generalization.
- **Higher Communication Costs for Frequent Updates:** While federated learning reduces the frequency and size of data transfers from SMs to other entities in SGs, it still requires regular communication of model updates. In cases where frequent updates are needed, this can create significant communication overhead, especially if the SG system involves many devices.
- **Complexity in Model Aggregation:** Aggregating updates from diverse edge devices such as SMs in SG environments is complex, where devices may have varied data distributions, processing capabilities, and network connectivity. This may affect the efficiency and effectiveness of the training process.

5.7. Use of additional hardware

The use of additional hardware components may help achieve privacy for users' data to some extent; however, they are less optimal solutions due to their installation and maintenance costs, scalability issues, and concerns related to energy wastage.

- **Additional Backup Power:** A battery installed for privacy can also function as a backup power source, providing resilience during power outages. This can improve the reliability of the power supply for the household, offering a dual purpose for the investment.
- **Environmental Benefits:** Batteries that store renewable energy *e.g.*, solar or wind, and use it strategically contribute to environmental sustainability by reducing dependency on grid power and potentially smoothing grid demand.
- **High Installation and Maintenance Costs:** Batteries are expensive to install and maintain. The upfront cost of purchasing and installing a battery system can be prohibitive for many households, and maintenance costs can add to the long-term expense.
- **Limited Battery Life and Capacity:** Batteries have limited capacity, which means they can only mask load patterns up to a certain point. Large or sustained loads may exceed the battery's capacity, leaving parts of the load profile exposed. Additionally, battery degradation over time reduces efficiency and increases the need for replacement.
- **Impact on Household Energy Bills:** The additional energy usage due to battery inefficiencies and potential time-of-use rate changes could increase energy bills. Additionally, some households might end up using grid power to charge the battery, which could result in additional costs if not managed properly.
- **Space and Installation Constraints:** Not all households have the space or infrastructure to install batteries, especially in densely populated urban areas or older buildings with limited electrical capacity. This can limit the accessibility of battery-based privacy solutions.

6. Publicly available power consumption datasets

Due to privacy concerns and marketing policies, most power companies do not disclose their electricity demand and generation data. Consequently, the lack of access to load data has been problematic for researchers to conduct research in the SG domain. However, several load datasets are publicly available and have been used in a variety of studies. There are several datasets available in the literature, and each has its own characteristics. In order to investigate and develop solutions for energy efficiency issues, choosing a database can be challenging.

6.1. Power consumption data analysis steps

Fig. 6 presents a general framework for acquiring and analyzing power consumption data within SGs and outlines, the associated components essential for pre-processing, data analysis, and result interpretation. This framework serves as a versatile model applicable for conducting experiments in diverse SG environments.

6.1.1. Data collection

The first step involves gathering various types of data relevant to energy usage. This includes:

- **Power Information of Home Appliances:** Detailed energy consumption data of individual home appliances (*e.g.*, washing machine, refrigerator, fan, oven) that can provide insights into household energy patterns. Recent studies emphasize the importance of disaggregating appliance-level data for better demand-side management and user engagement.
- **Climate Data:** Information on temperature, humidity, and other climatic factors influence energy demand, particularly for heating and cooling systems. Studies show that incorporating climate data can significantly enhance the accuracy of energy consumption forecasts and load profiling.
- **Renewable Energy Resources Data:** Data on the availability and usage of renewable energy sources (*e.g.*, solar or wind energy). The data from renewable energy sources is crucial for understanding the potential integration of renewables into household energy consumption. Integrating renewable data with household energy data can lead to more sustainable and cost-effective energy management solutions.

6.1.2. Data storage

Collected data is stored in the structured database for easy access and processing. Efficient data storage systems are critical, especially with the increasing volume of data generated by SMs and IoT devices. The importance of scalable and secure data storage in energy management systems is discussed in recent works.

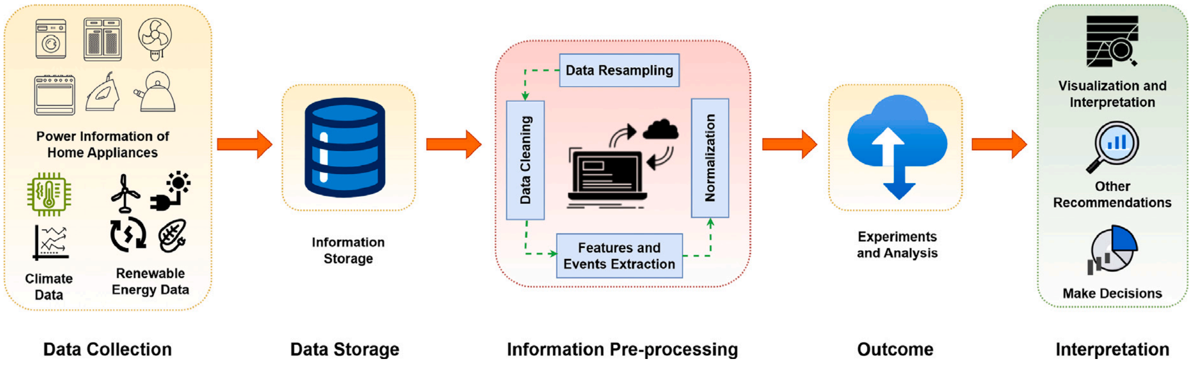

Fig. 6. General Framework for Power Consumption Data Analysis in SGs and its Essential Components.

6.1.3. Information pre-processing

The stored data undergoes pre-processing to prepare it for analysis [187]. This includes:

- **Data Cleaning:** Cleaning data by removing noise, errors, and inconsistencies. Techniques such as outlier detection and noise reduction are commonly applied in energy datasets.
- **Data Resampling:** Adjusting data frequency (*e.g.*, hourly, daily) for consistency across different data sources is essential for a unified analysis framework.
- **Normalization:** Standardizing data to ensure comparability across different sources and variables is particularly relevant in heterogeneous SG data environments.
- **Features and Events Extraction:** Identifying important features within the data, such as peaks in energy usage or unusual consumption patterns, helps in understanding demand trends. Techniques like event detection and feature extraction have been shown to improve load forecasting accuracy.

6.1.4. Experiments and analysis outcome

After pre-processing, the data is used in experiments and analysis. This stage helps in deriving insights, such as energy usage patterns, peak demand times, or the impact of specific appliances. Machine learning models and statistical methods are commonly employed to analyze energy data for load forecasting and anomaly detection.

6.1.5. Interpretation of results

The final stage involves interpreting the outcomes of the analysis. Visualization techniques have been shown to be effective in identifying consumption trends and promoting energy-saving behaviors. Insights or suggestions based on the analysis highlight the potential of personalized recommendations for improving household energy efficiency. Furthermore, these insights can be utilized to make informed decisions that can lead to improved energy efficiency and cost savings. Decision support systems based on energy analytics have proven effective in SGs.

6.2. Power consumption datasets

This section presents the main features of 18 power consumption datasets selected for this study. These datasets have been gathered in real-time and used by researchers to develop and evaluate privacy-preserving solutions for different SG domains. This work contributes an in-depth comparison between different datasets in accordance with various specifications. These specifications include the release date, country, collection period, number of observed houses, sampling rate, key features, and their applications. Power characteristics of various widely used open-load datasets found in the literature are presented in Table 12. HES [188], UMas Smart [189], SustData [190], and Dataport [191] are large-scale datasets. HES, UMas Smart, and Dataport analyzed energy consumption patterns down to the minute level, while SustSata measured power usage over very short intervals, *i.e.*, in seconds. Each of these datasets repositories contains appliance-level energy consumption measurements over a long time period.

For instance, HES and UMas Smart comprise energy data records for a duration of 12 months, whereas SustData includes energy usage patterns for a period of 1144 d. The Dataport repository has characteristics similar to the UMas Smart dataset. Primarily, it captures energy consumption data for more than 1200 households at the same sampling intervals as the UMas Smart dataset for a longer time period, *i.e.*, over four years.

Tracebase dataset [192] provides power usage profiles of a variety of devices collected at 1-second intervals. It is useful to use this dataset to evaluate energy efficiency; however, it cannot be utilized for energy disaggregation, appliance recognition or preference detection. Trasebase records information about 43 different appliances, each with a variety of recordings from several households over several days. Moreover, the date and time are also provided, along with the active and normalized power of appliances.

Aside from that, Electricity Consumption Benchmarks (ECB) [193], released in 2014, presents energy consumption statistics for 25 households from Victoria, Australia. This database used two years of aggregated consumption patterns and 30 min sampling rates to extract consumption patterns for individual appliances.

A detailed analysis of electricity consumption patterns of individual appliances in Austria and Italy was presented in Green Dataset (GREEND) [194] to describe detailed data collected through an experimental campaign at a sampling rate of 1 Hz. During the data collection process, eight households were observed, with each containing at least nine distinct devices. The device-level power consumption patterns are gathered for a duration of six months.

In the Netherlands, the Dutch Residential Energy Dataset (DRED) [195] was launched to collect data about energy usage, occupancy patterns, and environmental conditions in a pilot house. Sensors measure electricity consumption at the appliance level and aggregated consumption. Accordingly, 12 types of residential appliances are sampled at 1 min intervals, and aggregated consumption is recorded at the 1 Hz sampling rate.

In Reference Energy Disaggregation Dataset (REDD) [196], more than 10 kHz sampling frequency is used to capture energy consumption records. Comparatively, the observations are recorded only for a relatively short time period. In addition, data from six households are recorded in REDD, where energy consumption is monitored periodically at 0.5 Hz for a total of 24 devices.

London Carbon Project dataset [197] contains the energy consumption patterns of 5,567 households from London at a resolution of 30 min. The readings are collected from November 2011 to February 2014. Only electrical consumption appears to be associated with the data from the devices.

In [198], the authors published the MEULPv.1 and MEULPv.2 datasets. MEULPv.1 offers energy consumption data from 12 households in Canada, gathered at 1 min intervals for both appliance-specific and aggregated levels. During data collection for MEULPv.1, 8 different appliances were monitored. On the other hand, MEULPv.2 presents 1 year of data from 23 households, monitored at 1 min intervals, detailing both aggregated and appliance-based consumptions.

The RAE database has been introduced in [199], gathering data at a 1 Hz frequency. RAE serves as the first version of a power consumption repository, encompassing 1 Hz readings for sub-metered and aggregated levels of 2 households. In addition to power data, it integrates indoor temperature and humidity measurements of a house as well.

DISEC [200] provides a collection of diverse data from 19 accommodations within an Indian faculty housing complex over a period of 284 days. Various features, including power and weather, are gathered at 30-second sampling rate and subsequently aggregated into 15 min, 30 min, and 60 min intervals. Additionally, weather changes were continuously recorded from atmospheric conditions of nearby station readings.

BLOND [201] captured continuous energy consumption data and offers current and voltage readings for both appliance and aggregated levels. This dataset includes data from 53 devices spanning 16 appliance categories. Specifically, BLOND consists of two primary repositories: (i) BLOND-50, which includes consumption data captured at 50 kSps sampling rate for group circuits and 64 kSps for individual devices, and (ii) BLOND-250, which encompasses measurements over a 50-day timeframe collected with sampling rates of 250 kSps at for aggregated data and 50 kSps for appliance-specific data.

HUE [202] presents long-term power consumption profiles of 5 households at 1-h sampling frequency. It collects appliance-level consumption records from house 1 over a time period of two years with 1 1 min resolution rate, while records from house 2 over a time period of one year with a 1 Hz sampling rate. This information is referenced in [203].

The Residential Electricity End-use Demand Dataset of Costa Rica (REEDD-CR) [204] comprises data gathered from 51 households in Costa Rica. This dataset encompasses both branch circuit and aggregated measurements for each household, recorded at 1 min intervals for a minimum of one week. The measurements were collected from various locations across the country. Moreover, REEDD-CR incorporates a dataset featuring 197 load signatures, representing the energy usage patterns of 8 high-power demanding appliances: stove, lighting, refrigerator, dryer, water heating, microwave, air conditioning and washing machine. Every load signature is characterized by seven consumption and demand features, including peak and average power, average daily energy and events, day and night usage and time of use. The dataset also incorporates individual measurements for each appliance, which were utilized in calculating these load signatures.

The dataset from a domestic building in the UK [205] offers detailed power consumption data at the appliance level along with ambient climate conditions represented both in a time series and in a set of 2D images generated using Gramian Angular Fields (GAF). Its significance lies in two key aspects: (i) offering the researchers a dataset with a combination of appliance-level readings along with respective surrounding environmental measurements and (ii) presenting energy record summaries in the form of 2D images, facilitating the extraction of novel insights through machine learning and data visualization techniques. The heterogeneous dataset encompasses various parameters such as power consumption, current, voltage, occupancy, indoor temperature and humidity.

A 26-question-based survey was conducted in Greece, which revealed 188 data points from 104 households spanning various time periods. Each reporting point consists of attributes categorized into four groups: (i) household information such as residence type and its properties, (ii) socio-economic features of the occupants, including information on the number and types of occupants, total income and employment status, (iii) energy-related behaviors of the occupants and (iv) location of households to estimate climate conditions for the specified time period. Consequently, a set of corresponding features was extracted from the raw attributes [206].

6.2.1. *Characteristics comparison and discussion*

In this section, we compare existing power consumption datasets in order to obtain representative outputs and relevant interpretations. Each dataset possesses various properties, which play an important role in developing energy-efficient solutions. Table 12 compares the investigated characteristics of the aforementioned power consumption databases. In the following bullet points, we present some of the observations we have made for the above-mentioned datasets. In particular, we discuss the strengths and weaknesses of the datasets.

Table 12
Characteristics comparison of publicly available energy consumption datasets.

No.	Dataset	Year	Country	Period	Houses	Sample rate	Features	Applications
1	REDD [196]	2011	Massachusetts, USA	119 d	6	3 s	Current, voltage, active power	Energy saving, energy disaggregation
2	GREEND [194]	2015	Italy and Australia	6 months	8	1 s	Active power	Energy saving, energy disaggregation
3	HES [188]	2011	England, UK	1 year	26	10 min	Active power, temperature	Energy saving, energy disaggregation, anomaly detection
4	DRED [195]	2015	Netherlands	6 months	1	1 s	Active power, temperature, humidity, weather	Energy saving, occupancy detection, preference detection
5	UMas Smart [189]	2012	UK	3 months	3	1 s	Active power, apparent power	Energy saving, energy disaggregation
6	Tracebase [192]	2012	Germany	–	15	1-8 s	Active power, normalized power	Energy saving, energy disaggregation
7	Dataport [191]	2012	USA	4 years	1200	1 min	Active power	Energy saving
8	ECB [193]	2014	Australia	2 years	25	1 s	Active power	Energy disaggregation, demand prediction
9	SustData [190]	2014	Portugal	1144 d	50	2 s and 10 s	Current, voltage, active power, reactive power	Energy saving
10	Low Carbon London [197]	2011	UK	3 years	5567	30 min	Active power, temperature	Energy saving, energy disaggregation
11	MEULPv2 [198]	2017	Canada	1 year	23	1 min	Active power	Energy saving
12	RAE [199]	2018	Canada	72 d	1	1 Hz	Active power, reactive power, occupancy, frequency	Energy saving, demand prediction
13	DISEC [200]	2018	India	284 d	19	30 s, 15, 30, 60 min	Active power, weather	Energy saving, energy disaggregation
14	BLOND [201]	2018	Germany	213 d	5567	6.4 kSps	Active power, current, voltage	Energy saving, appliance recognition
15	HUE [202]	2019	Canada, Columbia	3 years	5	1 hr	Active power	Energy saving, energy disaggregation
16	REEDD-CR [204]	2022	Costa Rica	1 week	51	1 min	Average power, peak power	Energy saving, energy disaggregation
17	Domestic Building Energy Consumption [205]	2022	UK	7 months	1	1 min	Current, voltage, temperature, humidity	Energy saving, appliance recognition
18	Electricity Consumption of Residential Installations [206]	2023	Greece	1 month	104	1 min	Average power, weather, occupancy	Energy saving, appliance recognition

- The most extensive datasets both in relation to the time period and length, are UMas Smart, HES, Low Carbon Project, and SustData. Alternatively, HES has a very short observing period and comparatively high sampling frequency, *i.e.*, 2 min. The same applies to UMas Smart, where readings are recorded at 1 min sampling rate. Due to this, these datasets are insufficient for energy disaggregation, as individual devices and events will be difficult to distinguish. These two repositories, in contrast, provide details about observed homes, including the building type, building size, the number of rooms and occupants. Further, SustData uses only a sampling rate of 8s, which is still insufficient to conduct real-time analysis.
- In some datasets, *e.g.*, REDD, relatively few residential buildings are monitored at a very high frequency. Consequently, this fulfills the energy disaggregation requirements since high-frequency collection allows the extraction of comprehensive characteristics that capture dynamic behaviors.
- Majority of datasets are obtained from the USA and European regions, with 120 V and 230 V, respectively. Table 12 indicates that existing datasets are from different countries located on different continents, including America, Europe, Australia, and Asia. Therefore, these real-world databases were compiled in various climate zones, including humid regions (UMas Smart and REDD), marine West Coast climate (HES and Tracebase), and arid regions (ECB).
- The type and number of appliances and observed buildings significantly limit the usability of datasets. Statistical evaluations require a large number of appliances and houses. In this case, UMas Smart, Dataport, HES, and TraceBase are the most appropriate repositories. Moreover, some datasets monitor houses over multiple time intervals, making them difficult to compare and impractical. An important operator in analyzing the complexity of domestic equipment usage is the context in which it is used throughout the day. In this way, a real-world experiment would be carried out in homes, laboratories, and offices, rather than simulated environments.
- Some datasets, such as UMas Smart and DRED, include electrical parameters, namely current, voltage, frequency, and active and reactive power values. The majority of studied datasets lacked the ability to capture external variables, such as temperature, weather conditions, and humidity, which could have an impact on energy consumption.
- Recently collected datasets from Uk, Greece and Costa Rica offer appliance-level data and detailed socio-economic features, which can help in appliance modeling, energy demand model feedback, demand disaggregation testing, and enhancement in conducting experiments regarding electricity supply–demand studies with disaggregated and realistic data.
- Annotated power consumption datasets are not readily available to train anomaly detection models, where normal and anomalous data variables are labeled clearly. Particularly, almost all the investigated datasets lack in including labels identifying normal or abnormal power consumption, making them suitable only for training unsupervised anomaly detection algorithms without annotations.
- In most existing datasets, privacy and security considerations are not carefully addressed. The reason for this is that conventional meters require physical access and record energy consumption patterns for longer periods of time, *i.e.*, they are not capable of real-time monitoring.

7. Open research challenges and future directions

We will conclude our overview of privacy-preserving schemes for SG systems by looking at some open challenges and discussing them from the perspective of future research directions.

- **Allocation Network Usage Charges and Loss:** The utility-owned distribution network plays a significant role in aiding in the actual supply of electricity in P2P energy trading. Fees for using these corridors must be paid by each participant to the corresponding utility. Therefore, UCs may later use the accrued pricing to help with network maintenance or improvement. Additionally, line losses caused by power transfers throughout the network must be addressed. This can be done by adding extra power equivalent to the peer's losses or by adding costs proportional to the losses that happens during the power transfer. As a result, P2P trading must consider network utilization and loss charges using a smart P2P pricing system. This distribution ought to be determined by the volume of energy exchanged and the particular network resources used during trading.
- **Scalability and Participation of Customers:** To enhance P2P energy trading, it is essential to promote increased prosumer and consumer involvement in local energy sharing. Currently, the level of motivation and scalability determines the extent of client participation in P2P trading. P2P pricing mechanisms should be created so that participants enjoy advantages over Peer-2-Grid (P2G) trading at the very least. It is crucial to give people a variety of options for purchasing and selling energy. This can be accomplished by creating a scalable P2P system. A P2P energy trading platform must have the capacity to handle the increased amount of energy and financial transactions while also being able to accommodate more participants.
- **Privacy of Participants in P2P Trading:** Information is shared between the customer and a central entity in P2G energy trading. In contrast, sensitive information is repeatedly transferred across numerous peers in P2P energy trading, some of whom may not know one another. As a result, maintaining anonymity should be of utmost importance for P2P energy trading platforms. If this issue is ignored, there may be unfavorable privacy violations, which would decrease client engagement on P2P energy trading platforms.
- **Data Privacy in Behind the Meter Energy Storage Systems:** Owners of Energy Storage Systems (ESSs), particularly prosumers, frequently show an eagerness to enter into agreements with regional utilities to make money by offering ancillary services. The fact that it provides access to additional storage space and enables cost savings on expenditures is another reason why the utilities see it as useful. However, the main problem is establishing a stable bidirectional network to permit data transfer between the utilities and end-users. Several authors assumed that there can be an uninterrupted communication

channel between utilities and end-users. For instance, if ESS owners do not set their devices to discharge when the operator gives them the demand signal or does not receive it, the theoretical capacity of ESSs becomes useless. One strategy is to provide access to utility for ESS controls, which improves service quality but also increases the risk of cyber-attacks. Customers are also concerned that third parties might acquire their specific energy consumption data. Considerable hazards include cyber-attacks by nefarious third parties, such as sending misleading information to ESS owners or stealing user information.

- **Incentives to Encourage Energy Storage Systems Installation:** A pricing approach that allows for tariff flexibility is required to produce a fair price indicator that appropriately reflects power system costs in the retail markets for electricity. With this strategy, there is a great motivation to promote behind-the-meter investments. The interests of each network participant must be considered when developing tariffs. Moreover, it is crucial to set up a compensation scheme to motivate ESS owners to utilize their storage capacities. Currently, some countries are employing smart tariffs; however, their adoption in some areas is still incomplete and requires further examination. For example, the Netherlands Authority for Consumers and Markets (ACM) has initiated several pilot projects to explore the feasibility of implementing smart tariffs later on. The capacity of ESS has recently expanded, but it is projected that this trend will pick up speed with the advent of creative pricing models designed to achieve maximum ESS efficiency for all stakeholders, including utilities and owners.
- **Cyber Security Threats Detection in IoT-based SG Systems:** The risk of security vulnerabilities and cyber-attacks increases when relying on open internet connections to manage and monitor various physical devices and systems. Exploiting cutting-edge analytics methods like deep learning and machine learning to identify suspicious activities in these platforms remains a relatively unexplored area of study.
- **Data Privacy and Security in IoT-based SG Systems:** IoT-enabled SGs create a large amount of data, which raises the possibility of data exploitation and theft [207]. Additionally, the convergence of multiple system resources brought by interoperability may make data more vulnerable. Static and dynamic data may experience these problems. For instance, hostile actors are capable of changing SM data to change energy usage readings and reduce pricing. Additionally, such actors may be able to alter the operational data for the IoT-enabled SGs, providing a risk of monetary loss or serious deterioration to the utilities and assets of the SGs. Additionally, adversaries can determine whether clients are away from their residences using information on customer consumption. Furthermore, the storage and processing limitations of IoT-enabled SG equipment have limited their capacity to carry out complex and time-consuming traditional security methods. To protect data privacy and confidentiality, it is essential to improve existing methods. These solutions must be flexible to IoT devices with limited resources while also being scalable to support the widespread deployment of IoT devices across the SG. In addition, utilities need to put customer data privacy first, making sure that access to that data is only given with the consumers' express consent.
- **Authentication Schemes for IoT-based SG Systems:** It is vital to use reliable and effective authentication mechanisms for IoT-based SG devices with constrained resources to protect communications without adding to the constraints. By employing these techniques, only approved devices or users will be able to access resources or carry out particular tasks according to the number of access privileges they have been allowed. However, ensuring the dependability and clarity of such schemes continues to be a challenging research problem in IoT-enabled SG systems.
- **Lack of Datasets Related to an SG Environment:** The confidentiality of actual power system data frequently restricts its access. As a result, user-synthetic datasets are frequently used for testing and training. These datasets could add bias to the data selection process and may not include all of the qualities that could be associated with actual data collected from networks. The current public datasets are insufficient for the deployment of an IDS in an SG context because they exclude certain communication overheads that are frequently seen in SG communication protocols. Datasets like BATADAL, SWaT, and WADI mostly address the SCADA system within the electricity grid, but they exclude the unique characteristics connected to grid measurements, protocols, communication infrastructure, and advanced devices like smart inverters and intelligent electronic devices.
- **Maintaining Stability of Grids by Prosumers Energy Consumption Scheduling:** The ability of the current grid network to adequately handle the expanding use of green energy and the rising number of prosumers is raising serious concerns. The grid architecture also confronts several difficulties as a result of the erratic nature of renewable energy resources, the diverse prosumer energy usage patterns, and the charging habits of electric vehicles. Peaks in the demand load curve result from huge amounts of energy being unmanageably added to or removed from the system. These peaks typically occur in the morning and evening when there is a significant energy demand. As a result, there are imbalances in the grid network, which affect voltage stability and drive up operational costs. To ensure that the grid's capacity is properly managed, efficient scheduling of energy generation and consumption is necessary. This entails coordinating massive grid power drawn and optimizing renewable power input into the grid. Implementing a system that can schedule customers' real-time energy usage while taking into consideration their usage habits and power consumption/generation is necessary to overcome the difficulty of grid balancing. Therefore, there is a need to place more attention on scheduling energy use in a way that prioritizes grid stability while preserving security and privacy.
- **Use of Reinforcement Learning for Cyber-attacks Mitigation:** Most studies on reinforcement learning in the SG mainly aimed at solving various optimization problems in the control plane. There are not many studies, though, that particularly examine the application of reinforcement learning for countering attacks on the SG. There is a noteworthy lack of comprehensive strategies for dealing with cyber-attacks such as DoS in both the physical layer and the communication network. There has not been much research done on how DoS assaults affect the SG or how to create real-time mitigation techniques within realistic simulations. The time-consuming convergence and training requirements of deep reinforcement learning algorithms are one of the main causes of the dearth of studies on SG cyber-security. The majority of scientific studies have focused on

carrying out simulations on small scales and investigating particular attacks, in which the impact of the cyber-attack is mostly simulated and examined at the control plane. However, it becomes vital to use sophisticated co-simulation systems that can concurrently mimic both the communication and physical parts of the grid to thoroughly analyze the consequences of DoS assaults in the SG.

- **Intrusion Prevention Systems for Physical and Network Layer:** Studies, designs, and evaluations of Intrusion Prevention Systems (IPS) that are specifically adapted to the special needs of the SG are of special attention. These IPS solutions ought to be carefully evaluated in a co-simulated environment that closely mimics the dynamics of the SG. The co-simulation environment must be able to perform complex cyber-attacks that have an impact on both domains to examine the most effective defense mechanisms capable of mitigating attacks at the physical layer as well as the communication network layer. This makes it possible to analyze defense tactics in-depth in the setting of plausible attack scenarios.
- **Use of Game Theory for SG Privacy:** Market design and energy trading in SGs and microgrids have been evaluated using cooperative and non-cooperative game theory. This extension is partly due to the ability to develop price structures and incentive schemes using analyses that are suitable for the special properties of SGs and microgrids. According to a literature review, coalitional games are used in cooperative approaches while the Stackelberg game is frequently used in non-cooperative ways. Cooperative games can potentially encourage participants to work together on tactics that advance fairness and sustainability in obtaining desired results. It is important to recognize that game-theoretic methods can be difficult and data-intensive, especially when optimization depends on human behavior and erratic renewable resource generation.
- **Data Privacy and Security in Federated Learning:** Data privacy and security are major considerations when data is shared. However, federated learning techniques provide a solution by enabling cooperative model training across remote devices without sharing data with a central server. Federated learning is a decentralized machine learning type with significant advantages for SGs. Even during cooperative model training, local energy data can be kept private and need not be shared. Developed by Google in 2016, Federated learning has drawn much interest from researchers who want to address privacy issues linked to data sharing. While efforts have been made to develop incentive systems employing federated learning methods for various industries, only a small number of research have specifically examined its use in the context of SGs. However, given the benefits of federated learning, particularly in terms of privacy, this topic has a great deal of promise for academics to create prompt remedies to important problems in the SG sector.
- **Protection Against NILM Attacks**: Sensitive information about specific consumers and their personal lives is routinely obtained in SGs. This demonstrates the importance of reliable authentication, authorization, and secrecy methods in the SG context. To protect customer privacy, it is crucial to limit access to client data only to authorized parties with express authorization. Customers' names, addresses, and details of their energy use are all included in this sensitive data. UPs rely on SMs to provide accurate and time-specific readings of power consumption. These readings are necessary for proper billing and efficient grid management. Nevertheless, there is a possibility for more extensive applications of the measurement data gathered by SMs. Although risks are involved, analyzing usage patterns can help maximize power utilization. A lot of useful information about specific customers is contained in this data. Non-intrusive Appliance Load Monitoring (NALM) solutions use energy readings to gather thorough data on appliance consumption. Making predictions about whether a person is home or away and identifying the appliances being used are both made possible through data analysis and usage pattern analysis. Law enforcement, tax officials, and insurance companies are just a few of the organizations that may find this information of interest. The National Institute of Standards and Technology (NIST) has acknowledged that although SGs have the advantage of allowing for the collection of more detailed data from SMs and other devices, this advantage simultaneously poses the greatest privacy danger.

8. Conclusion

This study presented a review of recent advances in privacy-preserving solutions for SG systems developed between 2018 and 2024. Survey articles from recent years covering SG applications, technologies, communication frameworks, privacy, and security have been analyzed and compared based on key parameters. Furthermore, critical privacy and security requirements in SGs have been outlined, major privacy leakage attacks have been identified, and various privacy-preserving techniques have been categorized. A side-by-side tabular comparison of existing privacy-preserving solutions in SG technologies has been provided to offer a comprehensive overview. Additionally, publicly available energy consumption datasets have been examined, with key characteristics highlighted and recommendations suggested for improving data utility while preserving privacy.

The findings indicate a growing need for the development of robust and efficient privacy-preserving mechanisms in SG domains, particularly in energy trading markets, microgrids, and other power technologies. Several open research challenges have been identified, including privacy risks in P2P energy trading, the integration of renewable energy resources, cybersecurity in IoT-based SGs, and the application of reinforcement learning and game theory. Furthermore, the mitigation of privacy threats from NILM requires further exploration. Addressing these challenges will be crucial for ensuring secure and privacy-aware SG implementations in the future.

CRediT authorship contribution statement

Hafsa Bibi: Conceptualization, Formal analysis, Writing – original draft. **Mehran Abolhasan:** Conceptualization, Supervision. **Justin Lipman:** Supervision. **Mahrokh Abdollahi:** Writing – original draft, Supervision. **Wei Ni:** Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

All authors reviewed the draft and approved the final version of the manuscript.

Data availability

No data was used for the research described in the article.

References

[1] Atasoy T, Akınç HE, Erçin Ö. An analysis on smart grid applications and grid integration of renewable energy systems in smart cities. In: 2015 international conference on renewable energy research and applications. IEEE; 2015, p. 547–50.

[2] Tonyali S, Cakmak O, Akkaya K, Mahmoud MM, Guvenc I. Secure data obfuscation scheme to enable privacy-preserving state estimation in smart grid AMI networks. IEEE Internet Things J 2015;3(5):709–19.

[3] Curiale M. From smart grids to smart city. In: 2014 Saudi arabia smart grid conference. SASG, IEEE; 2014, p. 1–9.

[4] Gungor VC, Lu B, Hancke GP. Opportunities and challenges of wireless sensor networks in smart grid. IEEE Trans Ind Electron 2010;57(10):3557–64.

[5] Liu H, Ning H, Zhang Y, Xiong Q, Yang LT. Role-dependent privacy preservation for secure V2G networks in the smart grid. IEEE Trans Inf Forensics Secur 2013;9(2):208–20.

[6] Boudia ORM, Senouci SM, Feham M. Elliptic curve-based secure multidimensional aggregation for smart grid communications. IEEE Sensors J 2017;17(23):7750–7.

[7] Shen H, Zhang M, Shen J. Efficient privacy-preserving cube-data aggregation scheme for smart grids. IEEE Trans Inf Forensics Secur 2017;12(6):1369–81.

[8] Chen Y, Martínez-Ortega J-F, Castillejo P, López L. A homomorphic-based multiple data aggregation scheme for smart grid. IEEE Sensors J 2019;19(10):3921–9.

[9] Lang B, Wang J, Cao Z. Multidimensional data tight aggregation and fine-grained access control in smart grid. J Inf Secur Appl 2018;40:156–65.

[10] Chen X, Wen H, Ni W, Zhang S, Wang X, Xu S, et al. Distributed online optimization of edge computing with mixed power supply of renewable energy and smart grid. IEEE Trans Commun 2021;70(1):389–403.

[11] Yang Z, Zhu H, Yin C, Xie Z, Chen W, Chen C. Lightweight privacy-enhanced secure data sharing scheme for smart grid. Peer-to-Peer Netw Appl 2024;1–13.

[12] Ferrag MA, Maglaras LA, Janicke H, Jiang J, Shu L. Authentication protocols for internet of things: a comprehensive survey. Secur Commun Networks 2017;2017.

[13] Liu J, Xiao Y, Li S, Liang W, Chen CP. Cyber security and privacy issues in smart grids. IEEE Commun Surv & Tutorials 2012;14(4):981–97.

[14] Wang Y, Chen Q, Kang C, Xia Q. Clustering of electricity consumption behavior dynamics toward big data applications. IEEE Trans Smart Grid 2016;7(5):2437–47.

[15] Zhou K, Yang S. Understanding household energy consumption behavior: The contribution of energy big data analytics. Renew Sustain Energy Rev 2016;56:810–9.

[16] Wang Z, Zheng G. Residential appliances identification and monitoring by a nonintrusive method. IEEE Trans Smart Grid 2011;3(1):80–92.

[17] Chen D, Barker S, Subbaswamy A, Irwin D, Shenoy P. Non-intrusive occupancy monitoring using smart meters. In: Proceedings of the 5th ACM workshop on embedded systems for energy-efficient buildings. 2013, p. 1–8.

[18] Rouf I, Mustafa H, Xu M, Xu W, Miller R, Gruteser M. Neighborhood watch: Security and privacy analysis of automatic meter reading systems. In: Proceedings of the 2012 ACM conference on computer and communications security. 2012, p. 462–73.

[19] Cui L, Qu Y, Gao L, Xie G, Yu S. Detecting false data attacks using machine learning techniques in smart grid: A survey. J Netw Comput Appl 2020;170:102808.

[20] Asghar MR, Dán G, Miorandi D, Chlamtac I. Smart meter data privacy: A survey. IEEE Commun Surv & Tutorials 2017;19(4):2820–35.

[21] Han W, Xiao Y. Privacy preservation for V2G networks in smart grid: A survey. Comput Commun 2016;91:17–28.

[22] Leszczyna R. Cybersecurity and privacy in standards for smart grids–A comprehensive survey. Comput Stand Interfaces 2018;56:62–73.

[23] Ferrag MA, Maglaras LA, Janicke H, Jiang J, Shu L. A systematic review of data protection and privacy preservation schemes for smart grid communications. Sustain Cities Soc 2018;38:806–35.

[24] Mocrii D, Chen Y, Musilek P. IoT-based smart homes: A review of system architecture, software, communications, privacy and security. Internet Things 2018;1:81–98.

[25] Stellios I, Kotzanikolaou P, Psarakis M, Alcaraz C, Lopez J. A survey of iot-enabled cyberattacks: Assessing attack paths to critical infrastructures and services. IEEE Commun Surv & Tutorials 2018;20(4):3453–95.

[26] Mohammadi M, Al-Fuqaha A, Sorour S, Guizani M. Deep learning for IoT big data and streaming analytics: A survey. IEEE Commun Surv & Tutorials 2018;20(4):2923–60.

[27] Hossain E, Khan I, Un-Noor F, Sikander SS, Sunny MSH. Application of big data and machine learning in smart grid, and associated security concerns: A review. Ieee Access 2019;7:13960–88.

[28] Desai S, Alhadad R, Chilamkurti N, Mahmood A. A survey of privacy preserving schemes in IoE enabled smart grid advanced metering infrastructure. Clust Comput 2019;22:43–69.

[29] Kumar P, Lin Y, Bai G, Paverd A, Dong JS, Martin A. Smart grid metering networks: A survey on security, privacy and open research issues. IEEE Commun Surv & Tutorials 2019;21(3):2886–927.

[30] Hassan MU, Rehmani MH, Kotagiri R, Zhang J, Chen J. Differential privacy for renewable energy resources based smart metering. J Parallel Distrib Comput 2019;131:69–80.

[31] Ferrag MA, Babaghayou M, Yazici MA. Cyber security for fog-based smart grid SCADA systems: Solutions and challenges. J Inf Secur Appl 2020;52:102500.

[32] Joudaki M, Zadeh PT, Olfati HR, Deris S. A survey on deep learning methods for security and privacy in smart grid. In: 2020 15th international conference on protection and automation of power systems. IPAPS, IEEE; 2020, p. 153–9.

[33] Zhuang P, Zamir T, Liang H. Blockchain for cybersecurity in smart grid: A comprehensive survey. IEEE Trans Ind Informatics 2020;17(1):3–19.

[34] Abrahamsen FE, Ai Y, Cheffena M. Communication technologies for smart grid: A comprehensive survey. Sensors 2021;21(23):8087.

[35] Philips A, Jayakumar J, Lydia M. A review on cyber security in metering infrastructure of smart grids. Comput Methods Data Eng: Proc ICMDE 2020, Vol 1 2021;117–32.

[36] Jha AV, Appasani B, Ghazali AN, Pattanayak P, Gurjar DS, Kabalci E, et al. Smart grid cyber-physical systems: communication technologies, standards and challenges. Wirel Netw 2021;27:2595–613.

[37] Aggarwal S, Kumar N, Tanwar S, Alazab M. A survey on energy trading in the smart grid: Taxonomy, research challenges and solutions. IEEE Access 2021;9:116231–53.

[38] Abdalzaher MS, Fouda MM, Ibrahem MI. Data privacy preservation and security in smart metering systems. Energies 2022;15(19):7419.

[39] Mirzaee PH, Shojafar M, Cruickshank H, Tafazolli R. Smart grid security and privacy: From conventional to machine learning issues (threats and countermeasures). IEEE Access 2022;10:52922–54.

[40] Ghiasi M, Niknam T, Wang Z, Mehrandezh M, Dehghani M, Ghadimi N. A comprehensive review of cyber-attacks and defense mechanisms for improving security in smart grid energy systems: Past, present and future. Electr Power Syst Res 2023;215:108975.

[41] Ortega-Fernandez I, Liberati F. A review of denial of service attack and mitigation in the smart grid using reinforcement learning. Energies 2023;16(2):635.

[42] Al-Fuqaha A, Guizani M, Mohammadi M, Aledhari M, Ayyash M. Internet of things: A survey on enabling technologies, protocols, and applications. IEEE Commun Surv & Tutorials 2015;17(4):2347–76.

[43] Yaqoob I, Ahmed E, Hashem IAT, Ahmed AIA, Gani A, Imran M, et al. Internet of things architecture: Recent advances, taxonomy, requirements, and open challenges. IEEE Wirel Commun 2017;24(3):10–6.

[44] Zaveri MA, Pandey SK, Kumar JS. Collaborative service oriented smart grid using the Internet of Things. In: 2016 international conference on communication and signal processing. IEEE; 2016, p. 1716–22.

[45] Saleem Y, Crespi N, Rehmani MH, Copeland R. Internet of things-aided smart grid: technologies, architectures, applications, prototypes, and future research directions. IEEE Access 2019;7:62962–3003.

[46] Viswanatham VM, Chari A, Saritha V. Region-based group and hierarchical key management for secure smart grid communications. Int J Smart Grid Green Commun 2016;1(1):50–61.

[47] Ferrag MA. EPEC: An efficient privacy-preserving energy consumption scheme for smart grid communications. Telecommun Syst 2017;66(4):671–88.

[48] Hu S, Chen X, Ni W, Wang X, Hossain E. Modeling and analysis of energy harvesting and smart grid-powered wireless communication networks: A contemporary survey. IEEE Trans Green Commun Netw 2020;4(2):461–96.

[49] Amine Ferrag M, Maglaras LA, Janicke H, Jiang J. Authentication protocols for Internet of Things: A comprehensive survey. 2016, arXiv e-prints arXiv–1612.

[50] Walgama S, Hasinthara U, Herath A, Daranagama K, Kumarawadu S. An optimal electrical energy management scheme for future smart homes. In: 2020 IEEE 8th international conference on smart energy grid engineering. IEEE; 2020, p. 137–41.

[51] Paukstadt U. A survey of smart energy services for private households. 2019.

[52] Shen H, Liu Y, Xia Z, Zhang M. An efficient aggregation scheme resisting on malicious data mining attacks for smart grid. Inform Sci 2020;526:289–300.

[53] Liu Y, Guo W, Fan C-I, Chang L, Cheng C. A practical privacy-preserving data aggregation (3PDA) scheme for smart grid. IEEE Trans Ind Informatics 2018;15(3):1767–74.

[54] Fan X, Gong G. Security challenges in smart-grid metering and control systems. Technol Innov Manag Rev 2013;3(7).

[55] Abdallah A, Shen XS. A lightweight lattice-based homomorphic privacy-preserving data aggregation scheme for smart grid. IEEE Trans Smart Grid 2016;9(1):396–405.

[56] Kong W, Shen J, Vijayakumar P, Cho Y, Chang V. A practical group blind signature scheme for privacy protection in smart grid. J Parallel Distrib Comput 2020;136:29–39.

[57] Riggs H, Tufail S, Khan M, Parvez I, Sarwat AI. Detection of false data injection of pv production. In: 2021 IEEE green technologies conference. IEEE; 2021, p. 7–12.

[58] Tufail S, Parvez I, Batool S, Sarwat A. A survey on cybersecurity challenges, detection, and mitigation techniques for the smart grid. Energies 2021;14(18):5894.

[59] Sakhnini J, Karimipour H, Dehghantanha A, Parizi RM, Srivastava G. Security aspects of Internet of Things aided smart grids: A bibliometric survey. Internet Things 2021;14:100111.

[60] Huseinovic A, Mrdovic S, Bicakci K, Uludag S. A taxonomy of the emerging denial-of-service attacks in the smart grid and countermeasures. In: 2018 26th telecommunications forum. IEEE; 2018, p. 1–4.

[61] Zhang F, Mahler M, Li Q. Flooding attacks against secure time-critical communications in the power grid. In: 2017 IEEE international conference on smart grid communications. IEEE; 2017, p. 449–54.

[62] Lu Z, Lu X, Wang W, Wang C. Review and evaluation of security threats on the communication networks in the smart grid. In: 2010-milcom 2010 military communications conference. IEEE; 2010, p. 1830–5.

[63] El Mrabet Z, Kaabouch N, El Ghazi H, El Ghazi H. Cyber-security in smart grid: Survey and challenges. Comput Electr Eng 2018;67:469–82.

[64] Gunduz MZ, Das R. Cyber-security on smart grid: Threats and potential solutions. Comput Netw 2020;169:107094.

[65] Kurt MN, Yılmaz Y, Wang X. Real-time detection of hybrid and stealthy cyber-attacks in smart grid. IEEE Trans Inf Forensics Secur 2018;14(2):498–513.

[66] Peng C, Sun H, Yang M, Wang Y-L. A survey on security communication and control for smart grids under malicious cyber attacks. IEEE Trans Syst Man, Cybern: Syst 2019;49(8):1554–69.

[67] Bi S, Li K, Hu S, Ni W, Wang C, Wang X. Detection and mitigation of position spoofing attacks on cooperative UAV swarm formations. 2023, arXiv preprint arXiv:2312.03787.

[68] Huang X, Qin Z, Liu H. A survey on power grid cyber security: From component-wise vulnerability assessment to system-wide impact analysis. IEEE Access 2018;6:69023–35.

[69] Ganguly P, Nasipuri M, Dutta S. A novel approach for detecting and mitigating the energy theft issues in the smart metering infrastructure. Technol Econ Smart Grids Sustain Energy 2018;3(1):13.

[70] Ge L, Yu W, Moulema P, Xu G, Griffith D, Golmie N. Detecting data integrity attacks in smart grid. Secur Priv Cyber-Phys Syst: Found Princ Appl 2017;281–303.

[71] Yip S-C, Wong K, Phan RC-W, Tan S-W, Ku I, Hew W-P. A privacy-preserving and cheat-resilient electricity consumption reporting scheme for smart grids. In: 2014 international conference on computer, information and telecommunication systems. IEEE; 2014, p. 1–5.

[72] Kim M. A survey on guaranteeing availability in smart grid communications. In: 2012 14th international conference on advanced communication technology. IEEE; 2012, p. 314–7.

[73] Du D, Li X, Li W, Chen R, Fei M, Wu L. ADMM-based distributed state estimation of smart grid under data deception and denial of service attacks. IEEE Trans Syst Man, Cybern: Syst 2019;49(8):1698–711.

[74] Wang K, Du M, Maharjan S, Sun Y. Strategic honeypot game model for distributed denial of service attacks in the smart grid. IEEE Trans Smart Grid 2017;8(5):2474–82.

[75] Raza MA, Abolhasan M, Lipman J, Shariati N, Ni W, Jamalipour A. Statistical learning-based adaptive network access for the industrial internet-of-things. IEEE Internet Things J 2023.

[76] Yan Y, Qian Y, Sharif H, Tipper D. A survey on cyber security for smart grid communications. IEEE Commun Surv & Tutorials 2012;14(4):998–1010.

[77] Zheng J, Li K, Mhaisen N, Ni W, Tovar E, Guizani M. Exploring deep-reinforcement-learning-assisted federated learning for online resource allocation in privacy-preserving EdgeIoT. IEEE Internet Things J 2022;9(21):21099–110.
[78] Wang W, Lu Z. Cyber security in the smart grid: Survey and challenges. Comput Netw 2013;57(5):1344–71.
[79] Liu J, Xiao Y, Gao J. Achieving accountability in smart grid. IEEE Syst J 2013;8(2):493–508.
[80] Nicanfar H, Talebifard P, Alasaad A, Leung VC. Enhanced network coding to maintain privacy in smart grid communication. IEEE Trans Emerg Top Comput 2013;1(2):286–96.
[81] Pfitzmann A, Hansen M. A terminology for talking about privacy by data minimization: Anonymity, unlinkability, undetectability, unobservability, pseudonymity, and identity management. Germany: Dresden; 2010.
[82] Paverd A, Martin A, Brown I. Modelling and automatically analysing privacy properties for honest-but-curious adversaries. Tech. Rep, 2014.
[83] Sui Z, Niedermeier M, et al. TAI: A threshold-based anonymous identification scheme for demand-response in smart grids. IEEE Trans Smart Grid 2016;9(4):3496–506.
[84] Kumar P, Braeken A, Gurtov A, Iinatti J, Ha PH. Anonymous secure framework in connected smart home environments. IEEE Trans Inf Forensics Secur 2017;12(4):968–79.
[85] Khan AA, Kumar V, Ahmad M, Rana S, Mishra D. PALK: Password-based anonymous lightweight key agreement framework for smart grid. Int J Electr Power Energy Syst 2020;121:106121.
[86] Mahmud R, Vallakati R, Mukherjee A, Ranganathan P, Nejadpak A. A survey on smart grid metering infrastructures: Threats and solutions. In: 2015 IEEE international conference on electro/information technology. IEEE; 2015, p. 386–91.
[87] Duan L, Sun Y, Ni W, Ding W, Liu J, Wang W. Attacks against cross-chain systems and defense approaches: A contemporary survey. IEEE/CAA J Autom Sin 2023;10(8):1647–67.
[88] Altaf T, Wang X, Ni W, Liu RP, Braun R. NE-GConv: A lightweight node edge graph convolutional network for intrusion detection. Comput Secur 2023;130:103285.
[89] Kumar P, Gurtov A, Sain M, Martin A, Ha PH. Lightweight authentication and key agreement for smart metering in smart energy networks. IEEE Trans Smart Grid 2018;10(4):4349–59.
[90] Duan L, Yang L, Liu C, Ni W, Wang W. A new smart contract anomaly detection method by fusing opcode and source code features for blockchain services. IEEE Trans Netw Serv Manag 2023.
[91] Depuru SSSR, Wang L, Devabhaktuni V, Gudi N. Smart meters for power grid—Challenges, issues, advantages and status. In: 2011 IEEE/PES power systems conference and exposition. IEEE; 2011, p. 1–7.
[92] Diao F, Zhang F, Cheng X. A privacy-preserving smart metering scheme using linkable anonymous credential. IEEE Trans Smart Grid 2014;6(1):461–7.
[93] Altaf T, Wang X, Ni W, Yu G, Liu RP, Braun R. A new concatenated multigraph neural network for IoT intrusion detection. Internet Things 2023;22:100818.
[94] Bao H, Lu R. A lightweight data aggregation scheme achieving privacy preservation and data integrity with differential privacy and fault tolerance. Peer-To-Peer Netw Appl 2017;10:106–21.
[95] Tahir M, Khan A, Hameed A, Alam M, Khan MK, Jabeen F. Towards a set aggregation-based data integrity scheme for smart grids. Ann Telecommun 2017;72:551–61.
[96] Zhang J, Li H, Liu X, Luo Y, Chen F, Wang H, et al. On efficient and robust anonymization for privacy protection on massive streaming categorical information. IEEE Trans Dependable Secur Comput 2015;14(5):507–20.
[97] Li F, Luo B, Liu P. Secure information aggregation for smart grids using homomorphic encryption. In: 2010 first IEEE international conference on smart grid communications. IEEE; 2010, p. 327–32.
[98] Chim TW, Yiu S-M, Li VO, Hui LC, Zhong J. PRGA: Privacy-preserving recording & gateway-assisted authentication of power usage information for smart grid. IEEE Trans Dependable Secur Comput 2014;12(1):85–97.
[99] Chen L, Lu R, Cao Z. PDAFT: A privacy-preserving data aggregation scheme with fault tolerance for smart grid communications. Peer-To-Peer Netw Appl 2015;8:1122–32.
[100] Liu Y, Zhou C, Li L, Su L, Zhang Y. Fragile states metric system: An assessment model considering climate change. Sustainability 2018;10(6):1767.
[101] Zhang J, Zhang W, Wei X, Liu H. EPri-MDAS: An efficient privacy-preserving multiple data aggregation scheme without trusted authority for fog-based smart grid. High-Confid Comput 2024;100226.
[102] Zhang J, Wei J. PFDAM: Privacy-preserving fine-grained data aggregation scheme supporting multi-functionality in smart grid. IEEE Internet Things J 2024.
[103] Chen Y, Martínez-Ortega J-F, Castillejo P, López L. An elliptic curve-based scalable data aggregation scheme for smart grid. IEEE Syst J 2019;14(2):2066–77.
[104] Xiao L, Cai J, Qiu M, Liu M. A secure identity authentication protocol for edge data in smart grid environment. In: 2021 8th IEEE international conference on cyber security and cloud computing (cSCloud)/2021 7th IEEE international conference on edge computing and scalable cloud (edgeCom). IEEE; 2021, p. 188–93.
[105] Gai K, Wu Y, Zhu L, Xu L, Zhang Y. Permissioned blockchain and edge computing empowered privacy-preserving smart grid networks. IEEE Internet Things J 2019;6(5):7992–8004.
[106] Lyu L, Nandakumar K, Rubinstein B, Jin J, Bedo J, Palaniswami M. PPFA: Privacy preserving fog-enabled aggregation in smart grid. IEEE Trans Ind Informatics 2018;14(8):3733–44.
[107] Yao D, Wen M, Liang X, Fu Z, Zhang K, Yang B. Energy theft detection with energy privacy preservation in the smart grid. IEEE Internet Things J 2019;6(5):7659–69.
[108] Sui Z, de Meer H. An efficient signcryption protocol for hop-by-hop data aggregations in smart grids. IEEE J Sel Areas Commun 2019;38(1):132–40.
[109] Ge S, Zeng P, Lu R, Choo K-KR. FGDA: Fine-grained data analysis in privacy-preserving smart grid communications. Peer-To-Peer Netw Appl 2018;11:966–78.
[110] Zhang L, Zhang J, Hu YH. A privacy-preserving distributed smart metering temporal and spatial aggregation scheme. IEEE Access 2019;7:28372–82.
[111] 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 2012;23(9):1621–31.
[112] Wang X, Garg S, Lin H, Hu J, Kaddoum G, Piran MJ, et al. Toward accurate anomaly detection in Industrial Internet of Things using hierarchical federated learning. IEEE Internet Things J 2021;9(10):7110–9.
[113] 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. IEEE; 2014, p. 4720–5.
[114] Gope P, Sikdar B. An efficient data aggregation scheme for privacy-friendly dynamic pricing-based billing and demand-response management in smart grids. IEEE Internet Things J 2018;5(4):3126–35.
[115] Gong X, Hua Q-S, Qian L, Yu D, Jin H. Communication-efficient and privacy-preserving data aggregation without trusted authority. In: IEEE INFOCOM 2018-IEEE conference on computer communications. IEEE; 2018, p. 1250–8.
[116] Jia W, Zhu H, Cao Z, Dong X, Xiao C. Human-factor-aware privacy-preserving aggregation in smart grid. IEEE Syst J 2013;8(2):598–607.
[117] Dwork C, Kenthapadi K, McSherry F, Mironov I, Naor M. Our data, ourselves: Privacy via distributed noise generation. In: Advances in cryptology-EUROCRYPT 2006: 24th annual international conference on the theory and applications of cryptographic techniques, St. Petersburg, Russia, May 28-June 1, 2006. proceedings 25. Springer; 2006, p. 486–503.

[118] Bao H, Lu R. Ddpft: Secure data aggregation scheme with differential privacy and fault tolerance. In: 2015 IEEE international conference on communications. IEEE; 2015, p. 7240–5.

[119] Borden AR, Molzahn DK, Ramanathan P, Lesieutre BC. Confidentiality-preserving optimal power flow for cloud computing. In: 2012 50th annual allerton conference on communication, control, and computing. IEEE; 2012, p. 1300–7.

[120] Afrin S, Mishra S. An anonymized authentication framework for smart metering data privacy. In: 2016 IEEE power & energy society innovative smart grid technologies conference. IEEE; 2016, p. 1–5.

[121] Dwork C, McSherry F, Nissim K, Smith A. Calibrating noise to sensitivity in private data analysis. In: Theory of cryptography: third theory of cryptography conference, TCC 2006, New york, NY, USA, March 4-7, 2006. proceedings 3. Springer; 2006, p. 265–84.

[122] Dwork C, Roth A, et al. The algorithmic foundations of differential privacy. Found Trends® Theor Comput Sci 2014;9(3–4):211–407.

[123] Eibl G, Engel D. Differential privacy for real smart metering data. Comput Science-Res Dev 2017;32:173–82.

[124] El Ouadrhiri A, Abdelhadi A. Differential privacy for deep and federated learning: A survey. IEEE Access 2022;10:22359–80.

[125] Ou L, Qin Z, Liao S, Li T, Zhang D. Singular spectrum analysis for local differential privacy of classifications in the smart grid. IEEE Internet Things J 2020;7(6):5246–55.

[126] Fernández JD, Menci SP, Lee CM, Rieger A, Fridgen G. Privacy-preserving federated learning for residential short-term load forecasting. Appl Energy 2022;326:119915.

[127] Zheng Z, Wang T, Bashir AK, Alazab M, Mumtaz S, Wang X. A decentralized mechanism based on differential privacy for privacy-preserving computation in smart grid. IEEE Trans Comput 2021;71(11):2915–26.

[128] Wu Q, Zhou F, Xu J, Wang Q, Feng D. Secure and efficient multifunctional data aggregation without trusted authority in edge-enhanced IoT. J Inf Secur Appl 2022;69:103270.

[129] Hassan MU, Rehmani MH, Kotagiri R, Zhang J, Chen J. Differential privacy for renewable energy resources based smart metering. J Parallel Distrib Comput 2019;131:69–80.

[130] Gough MB, Santos SF, AlSkaif T, Javadi MS, Castro R, Catalão JP. Preserving privacy of smart meter data in a smart grid environment. IEEE Trans Ind Informatics 2021;18(1):707–18.

[131] Yuan X, Ni W, Ding M, Wei K, Li J, Poor HV. Amplitude-varying perturbation for balancing privacy and utility in federated learning. IEEE Trans Inf Forensics Secur 2023;18:1884–97.

[132] Zhao C, Zhao S, Zhao M, Chen Z, Gao C-Z, Li H, et al. Secure multi-party computation: theory, practice and applications. Inform Sci 2019;476:357–72.

[133] Yao AC. Protocols for secure computations. In: 23rd annual symposium on foundations of computer science. IEEE; 1982, p. 160–4.

[134] Micali S, Goldreich O, Wigderson A. How to play any mental game. In: Proceedings of the nineteenth ACM symp. on theory of computing. ACM New York, NY, USA; 1987, p. 218–29.

[135] Wagh GS, Gupta S, Mishra S. A distributed privacy preserving framework for the smart grid. In: 2020 IEEE power & energy society innovative smart grid technologies conference. IEEE; 2020, p. 1–5.

[136] Cheng L, Zang H, Wei Z, Sun G. Secure multi-party household load scheduling framework for real-time demand-side management. IEEE Trans Sustain Energy 2022;14(1):602–12.

[137] Palacios-Garcia EJ, Carpent X, Bos JW, Deconinck G. Efficient privacy-preserving aggregation for demand side management of residential loads. Appl Energy 2022;328:120112.

[138] Khan HM, Khan A, Jabeen F, Anjum A, Jeon G. Fog-enabled secure multiparty computation based aggregation scheme in smart grid. Comput Electr Eng 2021;94:107358.

[139] Chung H-M, Maharjan S, Zhang Y, Eliassen F. Distributed deep reinforcement learning for intelligent load scheduling in residential smart grids. IEEE Trans Ind Informatics 2020;17(4):2752–63.

[140] Wang L, Zhou Q, Xiong Z, Zhu Z, Jiang C, Xu R, et al. Security constrained decentralized peer-to-peer transactive energy trading in distribution systems. CSEE J Power Energy Syst 2021;8(1):188–97.

[141] Doan HT, Cho J, Kim D. Peer-to-peer energy trading in smart grid through blockchain: A double auction-based game theoretic approach. Ieee Access 2021;9:49206–18.

[142] Wang B, Wu Y, Liu KR. Game theory for cognitive radio networks: An overview. Comput Netw 2010;54(14):2537–61.

[143] Roy S, Ellis C, Shiva S, Dasgupta D, Shandilya V, Wu Q. A survey of game theory as applied to network security. In: 2010 43rd hawaii international conference on system sciences. IEEE; 2010, p. 1–10.

[144] Yang F, Zhou X, Jia G, Zhang Q. A non-cooperative game approach for intrusion detection in smartphone systems. In: 2010 8th annual communication networks and services research conference. IEEE; 2010, p. 146–51.

[145] Pecan Street Inc. Pecan Street Dataport. 2023, URL https://www.pecanstreet.org/dataport/. (Accessed 14 December 2023).

[146] Pham Q-V, Zeng M, Ruby R, Huynh-The T, Hwang W-J. UAV communications for sustainable federated learning. IEEE Trans Veh Technol 2021;70(4):3944–8.

[147] Singh P, Singh MK, Singh R, Singh N. Federated learning: Challenges, methods, and future directions. In: Federated learning for IoT applications. Springer; 2022, p. 199–214.

[148] Manzoor HU, Jafri A, Zoha A. Adaptive single-layer aggregation framework for energy-efficient and privacy-preserving load forecasting in heterogeneous federated smart grids. Internet Things 2024;28:101376.

[149] Kumari A, Gupta R, Tanwar S, Tyagi S, Kumar N. When blockchain meets smart grid: Secure energy trading in demand response management. IEEE Netw 2020;34(5):299–305.

[150] Yu G, Wang X, Sun C, Wang Q, Yu P, Ni W, Liu RP. Ironforge: An open, secure, fair, decentralized federated learning. IEEE Trans Neural Networks Learn Syst 2023.

[151] Hu S, Yuan X, Ni W, Wang X, Hossain E, Poor HV. OFDMA-f ^{2} L: Federated learning with flexible aggregation over an OFDMA air interface. 2023, arXiv preprint arXiv:2311.15141.

[152] Yin B, Yin H, Wu Y, Jiang Z. FDC: A secure federated deep learning mechanism for data collaborations in the Internet of Things. IEEE Internet Things J 2020;7(7):6348–59.

[153] Wen M, Xie R, Lu K, Wang L, Zhang K. FedDetect: a novel privacy-preserving federated learning framework for energy theft detection in smart grid. IEEE Internet Things J 2021;9(8):6069–80.

[154] Su Z, Wang Y, Luan TH, Zhang N, Li F, Chen T, Cao H. Secure and efficient federated learning for smart grid with edge-cloud collaboration. IEEE Trans Ind Informatics 2021;18(2):1333–44.

[155] Lu Y, Huang X, Dai Y, Maharjan S, Zhang Y. Blockchain and federated learning for privacy-preserved data sharing in industrial IoT. IEEE Trans Ind Informatics 2019;16(6):4177–86.

[156] Wang Y, Bennani IL, Liu X, Sun M, Zhou Y. Electricity consumer characteristics identification: A federated learning approach. IEEE Trans Smart Grid 2021;12(4):3637–47.

[157] Wang H, Si C, Liu G, Zhao J, Wen F, Xue Y. Fed-NILM: A federated learning-based non-intrusive load monitoring method for privacy-protection. Energy Convers Econ 2022;3(2):51–60.

[158] Lee S, Xie L, Choi D-H. Privacy-preserving energy management of a shared energy storage system for smart buildings: A federated deep reinforcement learning approach. Sensors 2021;21(14):4898.

[159] Lin J, Ma J, Zhu J. Privacy-preserving household characteristic identification with federated learning method. IEEE Trans Smart Grid 2021;13(2):1088–99.

[160] Gao J, Wang W, Liu Z, Billah MFRM, Campbell B. Decentralized federated learning framework for the neighborhood: a case study on residential building load forecasting. In: Proceedings of the 19th ACM conference on embedded networked sensor systems. 2021, p. 453–9.

[161] Fekri MN, Grolinger K, Mir S. Distributed load forecasting using smart meter data: Federated learning with Recurrent Neural Networks. Int J Electr Power Energy Syst 2022;137:107669.

[162] Zhang G, Zhu S, Bai X. Federated learning-based multi-energy load forecasting method using CNN-Attention-LSTM model. Sustainability 2022;14(19):12843.

[163] Commission for Energy Regulation (CER). CER smart metering project - Electricity customer behaviour trial, 2009–2010 [dataset]. 2012, Irish Social Science Data Archive. SN: 0012-00 URL https://www.ucd.ie/issda/data/commissionforenergyregulationcer/.

[164] Yu X, Xiao B, Ni W, Wang X. Optimal adaptive power control for over-the-air federated edge learning under fading channels. IEEE Trans Commun 2023.

[165] Li K, Zheng J, Yuan X, Ni W, Akan OB, Poor HV. Data-agnostic model poisoning against federated learning: A graph autoencoder approach. 2023, arXiv preprint arXiv:2311.18498.

[166] Su Y, Li Y, Li J, Zhang K. LCEDA: Lightweight and communication-efficient data aggregation scheme for smart grid. IEEE Internet Things J 2021;8(20):15639–48.

[167] Saleem A, Khan A, Malik SUR, Pervaiz H, Malik H, Alam M, et al. FESDA: Fog-enabled secure data aggregation in smart grid IoT network. IEEE Internet Things J 2019;7(7):6132–42.

[168] Zhu L, Li M, Zhang Z, Xu C, Zhang R, Du X, et al. Privacy-preserving authentication and data aggregation for fog-based smart grid. IEEE Commun Mag 2019;57(6):80–5.

[169] 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 2020;15(1):395–406.

[170] Singh P, Masud M, Hossain MS, Kaur A. Blockchain and homomorphic encryption-based privacy-preserving data aggregation model in smart grid. Comput Electr Eng 2021;93:107209.

[171] Qiu H, Zhang Z, Wang W, Zhang R, Zhou Y, Zhu L. Meter reading aggregation scheme with universally symbolic analysis for smart grid. Chin J Electron 2019;28(3):577–84.

[172] Wang Y, Luo F, Dong Z, Tong Z, Qiao Y. Distributed meter data aggregation framework based on blockchain and homomorphic encryption. IET Cyber-Phys Syst: Theory & Appl 2019;4(1):30–7.

[173] Mohammadali A, Haghighi MS. A privacy-preserving homomorphic scheme with multiple dimensions and fault tolerance for metering data aggregation in smart grid. IEEE Trans Smart Grid 2021;12(6):5212–20.

[174] McLaughlin S, McDaniel P, Aiello W. Protecting consumer privacy from electric load monitoring. In: Proceedings of the 18th ACM conference on computer and communications security. 2011, p. 87–98.

[175] Kalogridis G, Efthymiou C, Denic SZ, Lewis TA, Cepeda R. Privacy for smart meters: Towards undetectable appliance load signatures. In: 2010 first IEEE international conference on smart grid communications. IEEE; 2010, p. 232–7.

[176] Yang W, Li N, Qi Y, Qardaji W, McLaughlin S, McDaniel P. Minimizing private data disclosures in the smart grid. In: Proceedings of the 2012 ACM conference on computer and communications security. 2012, p. 415–27.

[177] Giaconi G, Gündüz D, Poor HV. Smart meter privacy with renewable energy and an energy storage device. IEEE Trans Inf Forensics Secur 2017;13(1):129–42.

[178] Cho K-H, Lee S-H, Khisti A. Effect of user cooperation on smart meter privacy with rechargeable batteries. IEEE Signal Process Lett 2019;26(7):971–5.

[179] Li S, Khisti A, Mahajan A. Information-theoretic privacy for smart metering systems with a rechargeable battery. IEEE Trans Inform Theory 2018;64(5):3679–95.

[180] Elkazaz M, Sumner M, Thomas D. A hierarchical and decentralized energy management system for peer-to-peer energy trading. Appl Energy 2021;291:116766.

[181] Liu X, Wang H, Chen G, Zhou B, ur Rehman A. Intermittently differential privacy in smart meters via rechargeable batteries. Electr Power Syst Res 2021;199:107410.

[182] Gope P, Sikdar B. Privacy-aware authenticated key agreement scheme for secure smart grid communication. IEEE Trans Smart Grid 2018;10(4):3953–62.

[183] Tahavori M, Moazami F. Lightweight and secure PUF-based authenticated key agreement scheme for smart grid. Peer-To-Peer Netw Appl 2020;13:1616–28.

[184] Wichakool W, Remscrim Z, Orji UA, Leeb SB. Smart metering of variable power loads. IEEE Trans Smart Grid 2014;6(1):189–98.

[185] Lin Y-H, Tsai M-S. An advanced home energy management system facilitated by nonintrusive load monitoring with automated multiobjective power scheduling. IEEE Trans Smart Grid 2015;6(4):1839–51.

[186] Rostampour S, Bagheri N, Ghavami B, Bendavid Y, Kumari S, Martin H, et al. Using a privacy-enhanced authentication process to secure IOT-based smart grid infrastructures. J Supercomput 2024;80(2):1668–93.

[187] Aurangzeb M, Wang Y, Iqbal S, Naveed A, Ahmed Z, Alenezi M, et al. Enhancing cybersecurity in smart grids: Deep black box adversarial attacks and quantum voting ensemble models for blockchain privacy-preserving storage. Energy Rep 2024;11:2493–515.

[188] Terry N, Palmer J. Household electricity survey. 2012, p. 1–31, UK Data Archive Study.

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

[190] Pereira L, Quintal F, Gonçalves R, Nunes NJ. Sustdata: A public dataset for ict4s electric energy research. In: ICT for sustainability 2014. Atlantis Press; 2014, p. 359–68.

[191] Parson O, Fisher G, Hersey A, Batra N, Kelly J, Singh A, et al. Dataport and NILMTK: A building data set designed for non-intrusive load monitoring. In: 2015 ieee global conference on signal and information processing. IEEE; 2015, p. 210–4.

[192] Reinhardt A, Morar O, Santini S, Zöller S, Steinmetz R. Cbfr: Bloom filter routing with gradual forgetting for tree-structured wireless sensor networks with mobile nodes. In: 2012 IEEE international symposium on a world of wireless, mobile and multimedia networks. IEEE; 2012, p. 1–9.

[193] Australian energy regulator; Electricity consumption benchmarks. 2014, URL http://www.energymadeeasy.gov.au.

[194] Monacchi A, Egarter D, Elmenreich W, D'Alessandro S, Tonello AM. GREEND: An energy consumption dataset of households in Italy and Austria. In: 2014 IEEE international conference on smart grid communications. IEEE; 2014, p. 511–6.

[195] Uttama Nambi AS, Reyes Lua A, Prasad VR. Loced: Location-aware energy disaggregation framework. In: Proceedings of the 2nd acm international conference on embedded systems for energy-efficient built environments. 2015, p. 45–54.

[196] Kolter JZ, Johnson MJ. REDD: A public data set for energy disaggregation research. In: Workshop on data mining applications in sustainability, vol. 25, Citeseer; 2011, p. 59–62.

[197] Druckman A, Jackson T. Household energy consumption in the UK: A highly geographically and socio-economically disaggregated model. Energy Policy 2008;36(8):3177–92.

[198] Johnson G, Beausoleil-Morrison I. Electrical-end-use data from 23 houses sampled each minute for simulating micro-generation systems. Appl Therm Eng 2017;114:1449–56.

[199] Makonin S, Wang ZJ, Tumpach C. RAE: The rainforest automation energy dataset for smart grid meter data analysis. Data 2018;3(1):8.

[200] Chen VL, Delmas MA, Locke SL, Singh A. Dataset on information strategies for energy conservation: A field experiment in India. Data Brief 2018;16:713–6.

[201] Kriechbaumer T, Jacobsen H-A. BLOND, a building-level office environment dataset of typical electrical appliances. Sci Data 2018;5(1):1–14.

[202] Makonin S. HUE: The hourly usage of energy dataset for buildings in British Columbia. Simon Fraser University; 2018.

[203] Kelly JP, James MA. Radiographic outcomes of hemiepiphyseal stapling for distal radius deformity due to multiple hereditary exostoses. J Pediatr Orthop 2016;36(1):42–7.

[204] Angulo-Paniagua J, Victor-Gallardo L, Alfaro-Corrales I, Quirós-Tortós J. REEDD-CR: Residential electricity end-use demand dataset from Costa Rican households. Data Brief 2023;46:108829.

[205] Alsalemi A, Amira A, Malekmohamadi H, Diao K. Novel domestic building energy consumption dataset: 1D timeseries and 2D gramian angular fields representation. Data Brief 2023;47:108985.

[206] Mischos S, Gkalinikis NV, Manolopoulou A, Dalagdi E, Zaikis D, Lazaridis A, et al. Household electricity consumption in Greece: A dataset based on socio-economic features. Data Brief 2023;48:109232.

[207] Zhao Z, Liu G, Liu Y. Practical privacy-preserving electricity theft detection for smart grid. IEEE Trans Smart Grid 2024.

Hafsa Bibi is currently a PhD candidate at School of Electrical and Data Engineering, University of Technology Sydney. She received her B.S. degree in Telecommunication and Networking from COMSATS University Islamabad, Pakistan, in 2016, and the M.S. degree in Computer Science from the University of Engineering and Technology Taxila, Pakistan, in 2021. Her research interests include Smart Grid Technologies, Differential Privacy, Game Theory and Optical networks.

Prof. Mehran Abolhasan completed a B.E in Computer Engineering and PhD in Telecommunications on 1999 and 2003 respectively at the University of Wollongong. Prof. Abolhasan has over 20 years of experience in R&D and serving in various research leadership roles. Some of these previous roles include serving as the Director of Research programs for the Faculty of Engineering and IT, Deputy Head of School for Research at the School of Electrical and Data Engineering and Lab Director for Telecommunication and IT Research Institute at University of Wollongong. He is currently the leader of intelligent Networks and Applications Lab within the Global Big Data Technology Centre at the Faculty of Engineering and IT at University of Technology Sydney.

Prof. Abolhasan has authored over 180 international publications and has won over seven million dollars in research funding. His current research Interests include: 5G/6G Wireless Networks, Cybersecurity, Software Defined Networking, Tactile Internet, Intelligent Transportation Systems (ITS), Internet of Things (IoT), Wireless Mesh, Wireless Body Area Networks and Sensor networks.

Dr. Justin Lipman fosters innovation in connected technologies and thrives at the intersection of academia and industry. As both Industry Associate Professor and Director of the Cyber Digital Centre (CDC) at the University of Technology Sydney, he leverages over 12 years of expertise leading R&D at Intel and Alcatel. Previously at University of Technology Sydney, he was Director of the RF and Communications Technologies Lab and served as Deputy Chief Scientist of the Food Agility Cooperative Research Centre. He has secured over $35M in research funding, driving innovation across cybersecurity, RF, IoT, digital agriculture, smart cities, and data privacy. He actively shapes future connected systems through standards development. As a dedicated bridge builder, he fuels impactful industry collaborations, translating research into real-world solutions and positioning himself at the forefront of innovation. He holds 24 U.S. patents and has published over 100 peer-reviewed articles in top-tier conferences and journals.

Dr. Mahrokh Abdollahi received the M.S. and Ph.D. degrees in electrical and data engineering from the University of Technology, Sydney, NSW, Australia, in 2016 and 2022, respectively. She is currently a Postdoctoral Fellow with the Privacy Technology Research Group at Data61, CSIRO. Her research interests include computer vision, deep learning, foundation models, privacy-preserving technologies, and cutting-edge advancements in 5G/6G networks and beyond.

Dr. Wei Ni received the B.E. and Ph.D. degrees in Electronic Engineering from Fudan University, Shanghai, China, in 2000 and 2005, respectively. He is a Principal Research Scientist at CSIRO, Sydney, Australia, and a Conjoint Professor at the University of New South Wales. He is also an Adjunct Professor at the University of Technology Sydney and an Honorary Professor at Macquarie University. He serves as a Technical Expert at Standards Australia in support of the ISO standardization of AI and Big Data. He was a Postdoctoral Research Fellow at Shanghai Jiaotong University from 2005 to 2008; Deputy Project Manager at Bell Labs, Alcatel/Alcatel-Lucent from 2005 to 2008; and Senior Researcher at Devices R&D, Nokia from 2008 to 2009. He has co-authored one book, ten book chapters, more than 300 journal papers, more than 100 conference papers, 26 patents, ten standard proposals accepted by IEEE, and three technical contributions accepted by ISO. His research interests include 6G security and privacy, machine learning, stochastic optimization, and their applications to system efficiency, security, and integrity.

Dr. Ni has been an Editor for IEEE Transactions on Wireless Communications since 2018, an Editor for IEEE Transactions on Vehicular Technology since 2022, and an Editor for IEEE Transactions on Information Forensics and Security and IEEE Communications Surveys and Tutorials since 2024. He served first as the Secretary, then the Vice-Chair and Chair of the IEEE VTS NSW Chapter from 2015 to 2022, Track Chair for VTC-Spring 2017, Track Co-chair for IEEE VTC-Spring 2016, Publication Chair for BodyNet 2015, and Student Travel Grant Chair for WPMC 2014.