frontiers | Research Topics

Planning and operation of integrated energy systems With deep integration of pervasive industrial internet-of-things

Edited by
Fengji Luo, Yunfei Mu, Gaoqi Liang, Yongxi Zhang and Linfeng Yang

Published in
Frontiers in Energy Research

ISSN 1664-8714
ISBN 978-2-83251-433-7
DOI 10.3389/978-2-83251-433-7

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Planning and operation of integrated energy systems with deep integration of pervasive industrial internet-of-things

Topic editors

Fengji Luo — The University of Sydney, Australia
Yunfei Mu — Tianjin University, China
Gaoqi Liang — The Chinese University of Hong Kong, Shenzhen, China
Yongxi Zhang — Changsha University of Science and Technology, China
Linfeng Yang — Guangxi University, China

Citation

Luo, F., Mu, Y., Liang, G., Zhang, Y., Yang, L., eds. (2023). *Planning and operation of integrated energy systems with deep integration of pervasive industrial internet-of-things*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83251-433-7

Table of contents

frontiers | Frontiers in Energy Research

TYPE Editorial
PUBLISHED 11 January 2023
DOI 10.3389/fenrg.2022.1121574

OPEN ACCESS

EDITED AND REVIEWED BY
ZhaoYang Dong,
Nanyang Technological University,
Singapore

*CORRESPONDENCE
Fengji Luo,
fengji.luo@sydney.edu.au
Yunfei Mu,
yunfeimu@tju.edu.cn

SPECIALTY SECTION
This article was submitted
to Smart Grids,
a section of the journal
Frontiers in Energy Research

RECEIVED 12 December 2022
ACCEPTED 27 December 2022
PUBLISHED 11 January 2023

CITATION
Luo F, Mu Y, Liang G, Zhang Y and Yang L (2023), Editorial: Planning and operation of integrated energy systems with deep integration of pervasive industrial internet-of-things.
Front. Energy Res. 10:1121574.
doi: 10.3389/fenrg.2022.1121574

Editorial: Planning and operation of integrated energy systems with deep integration of pervasive industrial internet-of-things

Fengji Luo[1]*, Yunfei Mu[2]*, Gaoqi Liang[3], Yongxi Zhang[4] and Linfeng Yang[5]

[1]School of Civil Engineering, The University of Sydney, Darlington, NSW, Australia, [2]School of Electrical and Information Engineering, Tianjin University, Tianjin, China, [3]School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, China, [4]School of Electrical and Information Engineering, Changsha University of Science and Technology, Changsha, China, [5]School of Computer Electronics and Information, Guangxi University, Nanning, China

KEYWORDS

smart grid, internet of things, energy system, power engineering, power system

Editorial on the Research Topic
Planning and operation of integrated energy systems with deep integration of pervasive industrial internet-of-things

Introduction

Modern energy systems have been evolving toward complex cyber-physical systems. As with many other industrial systems, modern energy systems are characterized by the pervasive integration of Internet-of-Things (IoT) devices. The IoT devices include smart meters, phasor measurement units, and sensors installed in demand-side devices (e.g., appliances). Deep penetration of IoT facilities enables the different stakeholders in energy systems to monitor the system's status on a granular level; in the meantime, this also imposes significant challenges and complexities on the development of planning and operation strategies to optimize the energy system's efficiency. With this context in mind, this Special Article Research Topic has been set up to solicit the most recent and original contributions to the planning and operation of integrated energy systems with deep integration of industrial IoT.

Ten papers are collected in this special article Research Topic, covering several important problems in industrial IoT-supported management technologies for integrated energy systems. The articles are categorized as follows.

Architectural design and analysis of IoT-enabled infrastructure for energy systems

The architectural design of integrated IoT and energy platforms plays an important role in modern energy systems. Two papers in this Research Topic are devoted to studying this aspect. Liu et al. provide a comprehensive review of the application of edge computing in the Ubiquitous Power Internet of Things (UPIoT). The article introduces the concept of

UPIoT and edge computing, investigates the architectural design of edge computing platforms for UPIoT, and discusses the potential multi-fold challenges. Shen et al. propose a Narrow Band Internet of Things (NB-IoT)-based demand-side data management framework for smart grades. The framework provides encrypted data transmission, management, and intelligent analysis services to support upper-level energy applications.

Advanced demand-side management

The pervasive IoT devices enable fine-grained energy data aggregation, and this creates new opportunities to develop data-driven demand-side management techniques. Three papers are collected in this direction, providing solutions for utilizing fundamental IoT facilities to improve demand-side energy efficiency.

Yang et al. propose an industrial IoT-enabled low-carbon demand response (LCDR) scheme. The scheme analyses the carbon cost on both the generation and demand sides; allowing the demand-side energy resources to be managed in an environmental-economic dispatch problem. Hou et al. study the visualization technique for the security risk analysis of vehicle networks. They develop a dynamic attack graph generation method to identify and visually display the security risks caused by the vulnerabilities in an Internet-of-vehicle system. This technique can help the system operator gain a better situational awareness of the vehicle network's security and perform better risk management. Zheng et al. develop a data-driven electricity theft detection technique that can effectively identify the abnormal energy metering data collected by edge computing devices.

Forecasting for renewable energy and power load

One direct advantage the pervasive IoT devices can provide to energy systems is more accurate forecasting for renewable power generation and power consumption. Two papers are collected on this Research Topic. Wang et al. propose a probabilistic forecasting method for industrial loads. The method utilizes a convolutional long-short-term memory network and a mixture density network to predict the probability density of the load. Liu et al. postulate a new wind power forecasting technique that uses super-resolution perception technology to detect errors in historical meteorological and wind power data collected by industrial IoT devices. The method will then correct any errors, recover the data from low- to high-frequency, and generate wind power predictions. These forecasting techniques can be used to assist upper-level energy system applications.

Planning and control of smart grids

IoT facilities have been playing an increasingly important role in optimizing the operation of smart grids. Three papers are collected in this special article Research Topic, which proposes innovative methods for planning and controlling the energy resources in smart grids. Cai et al. propose a planning model for grid-integrated energy storage systems (ESSs). The model determines the installation location and capacity of ESSs in the grid to minimize the workload in high-voltage transmission network reconfiguration, thereby mitigating transmission congestion. Zhao et al. develop a cluster partition-based zonal voltage control technique for power distribution networks with highly penetrated photovoltaics (PVs). The method partitions a distribution network into multiple clusters, and it designs algorithms to optimize the voltages in the clusters. Zhang et al. propose a control technique for solid oxide fuel cell (SOFC) hybrid direct-current microgrids. The technique analyzes the thermal and electrical characteristics of the microgrid; based on this, it regulates the current and voltage of the SOFC system to realize high efficiency and steady-state thermal safety of the system.

We hope the papers collected in this special article Research Topic can provide useful references to researchers and engineers and can advance knowledge in power and energy systems and ubiquitous computing.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

frontiers in Energy Research

REVIEW
published: 22 February 2022
doi: 10.3389/fenrg.2022.850252

Edge Computing Application, Architecture, and Challenges in Ubiquitous Power Internet of Things

Dongqi Liu[1], Haolan Liang[1,2], Xiangjun Zeng[1], Qiong Zhang[1], Zidong Zhang[1] and Minhong Li[3]*

[1]School of Electrical and Information Engineering, Changsha University of Science and Technology, Changsha, China, [2]Hunan Institute of Engineering, Xiangtan, China, [3]Shenzhen Power Supply Bureau Co., Ltd., Shenzhen, China

The Ubiquitous Power Internet of Things (UPIoT) is a concrete manifestation of the Internet of things (IoT) in the power industry, which is a deep integration of the interconnected power network and communication network, realizing full perception of the system status and full business penetration in all links of power production, transmission, and consumption. The introduction of edge computing in UPIoT fully meets the requirements of rapid response, real-time perception, and to some extent, privacy protection. However, there is currently no comprehensive investigation on the application of edge computing technology in UPIoT. First, this paper introduces the development background and construction of UPIoT and its technical architecture. Then the challenges faced by UPIoT in the process of construction are analyzed. Furthermore, the paper elaborates on the functions and features of edge computing, proposes that the support of edge computing technology can solve the challenges of efficient, fast, and secure processing of massive edge data faced by the traditional cloud-based centralized big data processing technology of UPIoT, and analyzes the architecture of the edge computing-assisted UPIoT. For the three typical scenarios of UPIoT, namely power monitoring system, smart energy system and power metering system, the edge computing architecture of the three scenarios are analyzed, and the specific application methods and roles played by edge computing in the three scenarios are also elaborated. Finally, we discuss the challenges of edge computing in UPIoT, in terms of policy challenges, market challenges, and technical challenges, as well as outline the outlooks of the technical challenges.

Keywords: power internet of things, UPIoT, edge computing, edge computing applications, edge computing architecture, smart grid

OPEN ACCESS

Edited by:
Fengji Luo,
The University of Sydney, Australia

Reviewed by:
Mingtao Feng,
Xidian University, China
Zhiqiang Miao,
Hunan University, China

****Correspondence:***
Dongqi Liu
liudongqi@csust.edu.cn

Specialty section:
This article was submitted to Smart Grids, a section of the journal Frontiers in Energy Research

Received: *07 January 2022*
Accepted: *27 January 2022*
Published: *22 February 2022*

Citation:
Liu D, Liang H, Zeng X, Zhang Q, Zhang Z and Li M (2022) Edge Computing Application, Architecture, and Challenges in Ubiquitous Power Internet of Things. Front. Energy Res. 10:850252. doi: 10.3389/fenrg.2022.850252

1 INTRODUCTION

The combination of the energy revolution with the digital revolution has led to the development of the fourth industrial revolution. With the application of these new technologies, such as IoT, edge computing, 5G communication, and artificial intelligence (AI) in the power system, the power system is being promoted to become intelligent, digital, and networked. The aim is for these modern technologies to break through the bottleneck of power development and realize 100% renewable electricity. Meanwhile, the development trends of the power grid are to adapt to the diverse needs of "new loads," improving the flexibility and flexibility of the power grid, and opening up the blue ocean

of the digital economy (state grid corporation of China, 2019a). In 2019, the state grid corporation of China set the goal of constructing the UPIoT, which integrates the power network and communication networks through various information sensing technologies, intelligent collection technologies, big data technologies, and other modern technologies to connect massive power-related entities, realizing comprehensive intelligent perception, identification, and management of power equipment, information interaction, and data sharing, as well as rapid response to demand.

With the continuous construction and development of UPIoT, a large number of intelligent terminals and devices are accessed, so the computing model with traditional cloud computing as the core is no longer effective in real-time transmission, computation, and storage for the process of the billions or trillions of data that are generated by heterogeneous massive intelligent terminals. Edge computing, a new computing model proposed under the rapid development of the Internet of Things, artificial intelligence, big data, and cloud computing, which is an open platform that uses network, computing, storage, and application core capabilities as a whole on the side close to the physical environment or data source, and the edge computing platform is deployed in the network measurement close to the data source to provide the nearest end service nearby, so as to get faster network service response and meet the basic needs of the industry in terms of real-time services, application intelligence, security, and privacy protection (Bai et al., 2020). Therefore, some services from the original cloud can be allocated to the edge side of the network for processing, so as to meet the real-time requirements of various tasks while ensuring overall system performance (Sharma and Wang, 2017; Fu et al., 2018; Zhang et al., 2018; Maier and Ebrahimzadeh, 2019; Song et al., 2019; Xu et al., 2019). Especially for supporting ubiquitous IoT, the processing will be done at the local edge computing layer, based on edge computing technology that can localize the computation, analysis, and control to provide a faster response to users without handing over responsibility to the cloud, thus enhancing the processing efficiency and reducing the data processing load in the cloud. Therefore, edge computing technology is naturally similar to UPIoT in terms of agile connectivity, computation, topology, real-time services, data optimization, application intelligence, security, and privacy protection, and can well support the construction of UPIoT.

There is much discussion in the literature of edge computing but very little work has been done towards the edge computing technology that is applied in the UPIoT. For edge computing, some surveys have studied basic characteristics, research challenges, and opportunities of different edge computing paradigms (Shi et al., 2016; Hu et al., 2015; Varghese et al., 2016; Dustdar et al., 2019; Caprolu et al., 2019). The concept of edge computing has been extended to the wider IoT. Scholars have discussed the application of edge computing in some IoT fields (Yu et al., 2017). conducted a survey to examine how edge computing can enhance the implementation and the performance of IoT, and compared the performance of the different loT applications that are based on the EC and cloud computing architectures. Alrowaily and Lu (2018) reviewed the concepts, features, security, and application of edge-computing-enabled IoT as well as its security features in the data-driven world. Porambage et al. (2018) surveyed multiaccess edge computing, and they presented a holistic overview of this paradigm in relation to IoT. The integration of multiaccess edge computing into IoT applications and their synergies are also analyzed and discussed. Pan and McElhannon (2018) investigated the key rationale, the efforts, the key enabling technologies, and typical IoT applications benefiting from edge cloud. Omoniwa et al. (2019) presented a survey on EC-based IoT literature in the period from 2008 to 2018, including services, enabling technologies, and open research issues, and briefly displayed how the EC-IoT was applied in real-life cyber-physical systems, such as the intelligent transportation system or smart grid. Qiu et al. (2020) introduced the concept of industrial IoT (IIoT), and presented the research progress and future architecture of the EC-assisted IIoT. As can be seen above, most articles mainly elaborate on the architecture, key technologies, advantages, and challenges of EC-assisted IoT, and there are few implementation plans and deployments for the specific application scenarios of edge computing. However, for the power grids, it simply explains the application and role of edge computing in the smart grid, and it does not specifically cover the various levels of the power generation, transmission, and distribution of power grids, and there is currently no comprehensive investigation on the application of edge computing technology in UPIoT.

This paper focuses on edge computing in UPIoT and combs through many research achievements concerning edge computing in UPIoT, discussing the architecture of edge computing in UPIoT. Then, the application of the three scenarios is explored, namely power monitoring system, smart energy system, and advanced metering infrastructure; meanwhile, the advantages of applying edge computing in the three scenarios of the UPIoT, in terms of data privacy protection, security, and communication time, are analyzed. Finally, we elaborate on the challenges and future directions for the application of edge computing in UPIoT.

2 AUTONOMOUS UBIQUITOUS INTERNET OF THINGS IN ELECTRICITY

Promoting the construction of UPIoT is an important initiative to realize the energy Internet, which is a strategic deployment for the development of the world economy and the upgrading of the world energy infrastructure. UPIoT is essentially a kind of Internet of things, a specific expression and application of Ubiquitous Internet of Things in the power industry. (the state grid corporation of China, 2019b). Around each link of the electric power system, modern information technologies and advanced communication technologies such as mobile Internet and artificial intelligence are fully applied to realize the interconnection of all things and human-computer interaction in each link of the electric power system (state grid corporation of

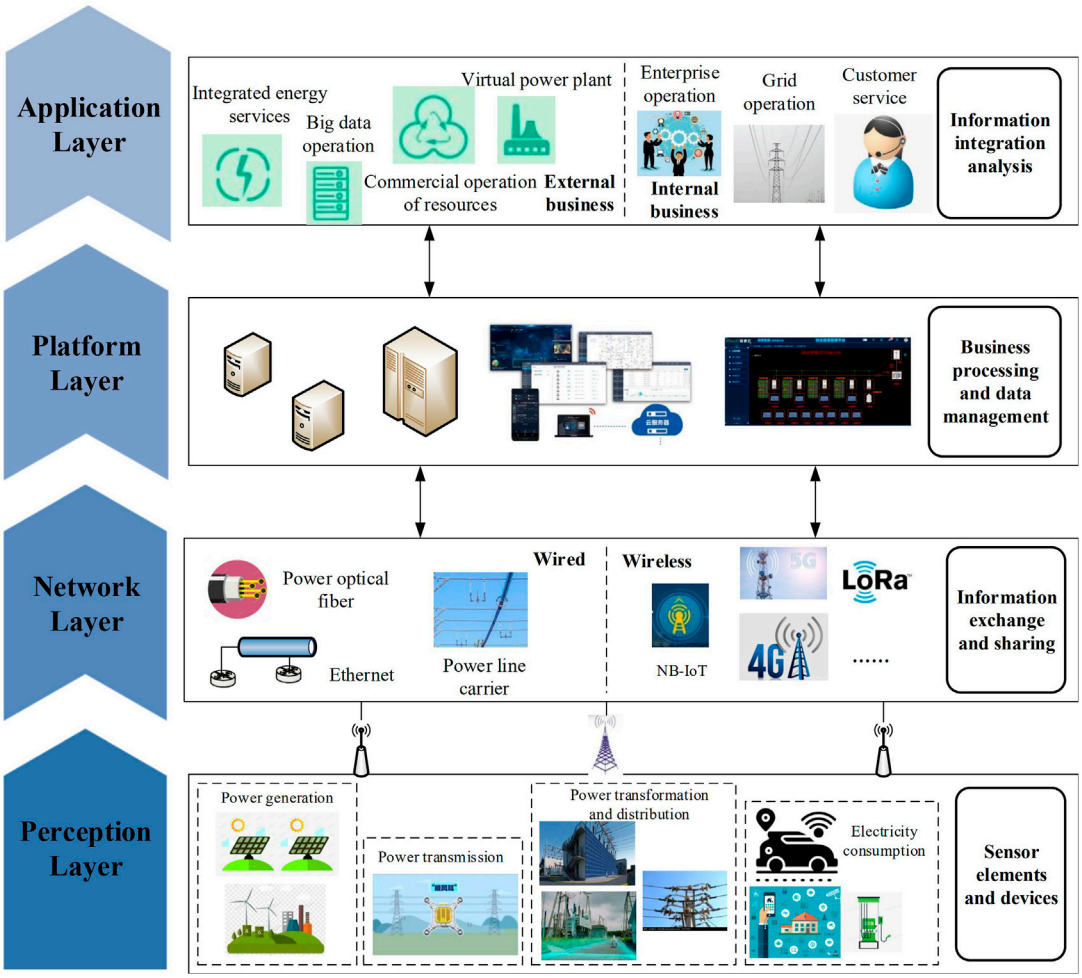


FIGURE 1 | The architecture of UPIoT.

China, 2020). It is a new form of network with a deep integration of traditional industrial technology and IoT technologies, and it is a concrete manifestation of the IoT in the power industry. Overall, it is a smart service system with state awareness, efficient information processing, and convenient and flexible application features. The architecture of UPIoT is shown in **Figure 1**.

Like IoT, UPIoT also includes a sensing layer, a network layer, a platform layer (management layer), and an application layer.

2.1 Perception Layer (Terminal Layer)

Smart sensing is the core foundation technology of the perception layer of UPIoT (Jiang et al., 2019b). Smart sensing devices are used to collect, monitor, and sense real-time data from end-of-grid operation equipment. The data flow of the power grid operation runs through the whole chain of power generation, power transformation, transmission, distribution, and consumption. For the primary measurement equipment of power system transmission and distribution, part of it is the basic data of the equipment, and the other part is the real-time operation status data of the equipment (Wang et al., 2019). The real-time data is monitored and sensed by the corresponding intelligent monitoring equipment, such as infrared thermometers, monitoring cameras, or inspection robots. For the secondary equipment on the distribution side, such as relay protection devices and electrical energy monitoring equipment, they can collect the system and equipment operation data. The most typical smart sensor of electricity consumption side equipment is the smart meter, which can sense the basic electricity consumption of users in real time and transmit the user's energy consumption and electric energy trading data to the system platform.

2.2 Network Layer

This is used to achieve efficient and secure data transmission over a wide area between the sensing layer (terminal layer) and the platform layer (Li Z. S. et al., 2018; Wang, 2019). In order to meet the access of different types of sensors, it may include different types of network compositions such as mobile communication, limited IoT, and local area network. They have specific

communication protocols and specifications and have strong extensibility. With the rapid development of 5G technology in China, 5G communication technology will be gradually applied in electric power IoT. 5G communication has the characteristics of wide wireless coverage, short transmission delay, system security engagement, and high transmission rate, which is the core technology to realize and build electric power IoT.

2.3 Platform Layer

The data transmitted by the network layer is stored and managed through a unified data center, meaning some of the data can also be shared across departments, subjects, and even industries. Using cloud computing technology, deep learning, big data analysis, and other core technologies, it realizes efficient data processing and IoT management. The platform layer is the basis for realizing advanced applications, and is dedicated to improving collaborative computing and real-time response to meet users' power supply needs and business response.

2.4 Application Layer

The application layer is the core goal of UPIoT, which provides control and decision support for users, electricity sellers, and grid operators. Through the interconnection of power and information, the application layer realizes the production and operation of the grid, operation management, and related energy services and breeds new business models and emerging businesses. These new models include micro-grid operation and management, operation and management of electric vehicles, campus energy management, and system multi-dimensional resource management.

The shape of UPIoT will gradually evolve as key technologies, such as sensor technology, communication technology, and cloud/edge computing technology, are deepened in the smart grid.

3 ARCHITECTURE OF EDGE COMPUTING FOR UBIQUITOUS POWER INTERNET OF THINGS

3.1 Edge Computing for Ubiquitous Power Internet of Things

Many new challenges are introduced by the development of the UPIoT. Although current cloud computing has been applied in the UPIoT, it still cannot meet the challenges that we summarize as follows (Chen et al., 2019): 1)low latency requirement, 2)Network bandwidth constraints, 3)Restricted equipment resources, 4)Uninterrupted connection and interaction with the cloud center, and 5)Privacy protection and data security.

In order to overcome these challenges and realize the construction of UPIoT, three key technologies in IoT are important components of its overall architecture: sensor technology, communication technology, and edge computing technology. Among them, edge computing is the important carrier for constructing it, and it is the most important core technology to support the UPIoT in realizing real-time response, short-period data analysis, various types of edge intelligent services, and so on.

Edge computing refers to a new computing model that analyzes and processes a portion of data using the computing, storage, and network resources distributed on the paths between data sources and the cloud computing center (Shi et al., 2016). Edge computing focuses on real-time, short-period data analysis, close to the device side, and it is better able to support local business real-time analysis and intelligent processing. Meanwhile, it has features such as distributed, low latency, high efficiency, and relieves traffic pressure, and it is more efficient and secure compared to simple cloud computing. The main advantages of edge computing applied to the UPIoT are as follows (Industrial Internet Consortium:Edge Computing Task Group, 2018; Du et al., 2021):

3.1.1 Improved System Performance

With edge computing, in addition to collecting and transmitting data to the cloud platform, data collected at the edge can be analyzed and processed in milliseconds. For example, the advanced measurement system in the smart grid, where the user electricity consumption data collected by smart meters is uploaded to the edge intelligent fusion terminal, without uploading tens of thousands of data to the cloud platform for processing. The computing software of the edge platform can analyze these data in real-time and only upload the fused or processed data to the cloud center, which greatly reduces the communication broadband, shortens the data transmission time, and improves the overall performance of the system.

3.1.2 Protected Data Security and Privacy

Cloud platform service providers give customers a comprehensive system of centralized data security protection solutions. However, once centralizing stored data get leaked, it will lead to serious consequences. Edge computing migrates computing closer to the device, avoiding the need to upload data to the cloud, which greatly reduces the risk of private data being compromised or corrupted during transmission, and also reduces the security risk of cyber attacks to the cloud center that result in leakage of all stored data stored.

3.1.3 Reduced Operational Costs

Since cloud computing requires uploading data to the cloud center for processing, cloud computing possesses the characteristics of data migration, bandwidth, and latency of cloud computing, which makes it very expensive to use cloud computing, whereas edge computing can significantly reduce operational costs by reducing the amount of data upload, thereby reducing data migration, bandwidth requirements, and latency.

3.2 Architecture of Edge Computing in the Ubiquitous Power Internet of Things

Thousands of power terminal devices and sensing nodes access the smart grid in a variety of ways, and can sense or control the power grid. These nodes are usually organized or self-organized

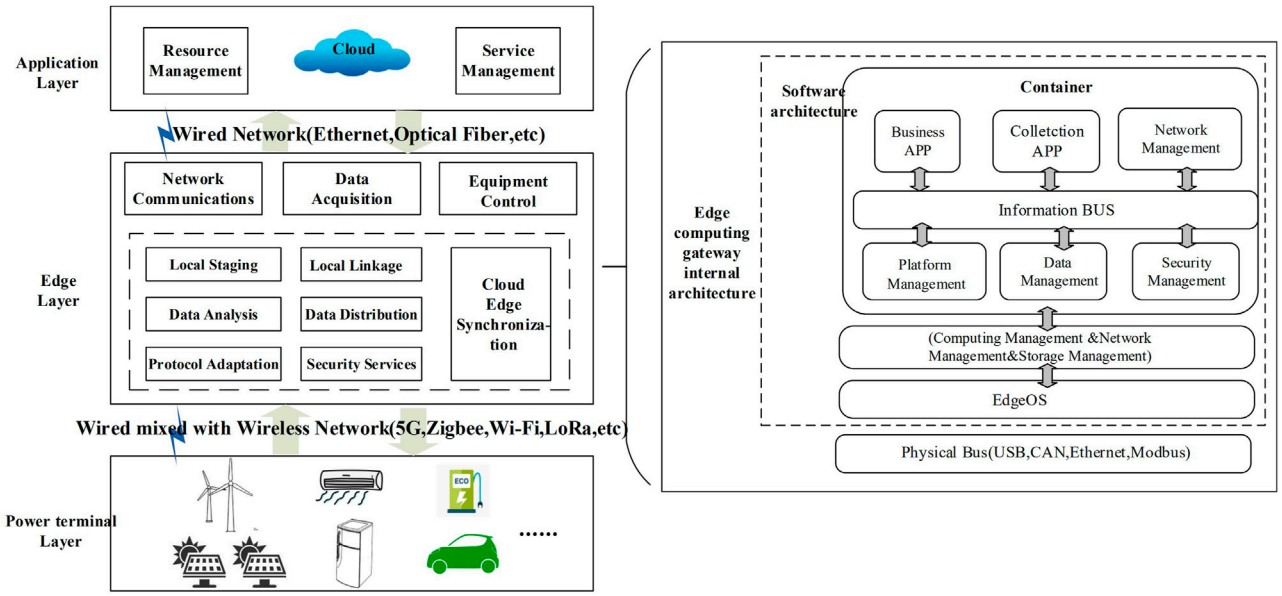


FIGURE 2 | Proposed reference architecture of edge computing architecture and the inner architecture of the edge computing gateway in UPIoT.

into various clusters to form edge networks (Cornel–Cristian et al., 2019). By deploying edge computing into the UPIoT, it can locally process massive heterogenous data and the acquisition of signals, solving the problem of fast response and centralized service, and reducing cloud pressure and communication overhead. The edge computing reference paradigm in UPIoT is as shown in **Figure 2**. The reference paradigm consists of device layer, edge layer, and cloud application layer; its architecture comprehensively describes the characteristics of UPIoT and edge computing. This paper clearly elaborates the inner architecture of the edge computing gateway in the edge layer. In edge computing gateway, which has certain computation resources that can provide a chance to offload part of the workload from the cloud, the edge not only requests service and content from the cloud, but also performs the computing tasks from the cloud. Meanwhile, the gateway can process amounts of data from various terminals, smart devices, and end users of the power terminal layer, and provides a distributed information computing service with large volumes of data and fast responses.

Edge can perform computing offloading, data storage, caching, and processing, as well as distribute request and delivery service from cloud to user. Under such circumstances, this mode can satisfy the demand of rapid responses required by devices and users in the smart grids, The internal architecture of the edge computing gateway is shown in **Figure 2** (Liu et al.,2019); its software system consists of a host operating system and multiple containers. APPs are running inside the container systems; the cloud server and the edge computing device perform data interaction instructions for uplink and downlink, and the terminal node uploads data to the edge computing device, various software applications run on edge computing devices. By encapsulating the application functions as APPs that are loaded in the container, this can provide support for advanced applications of smart grids such as edge autonomy of smart substations, intelligent online monitoring, rapid request distribution, and service delivery.

The application of edge computing in the IoT has attracted a great deal of research. Some typical IoT scenarios are applied, such as smart transportation, smart healthcare, the smart home, and the smart building. We roughly analyze how to apply edge computing in these four scenarios, and then elaborate on the current related articles on the application of edge computing to smart grids. For example, applying edge computing technology to smart transportation, mobile edge computing (MEC) puts the mobile base stations at the edge of the network, and the mobile base stations are deployed in a decentralized manner, providing servers for applications in base stations close to the edge of the network, allowing data to be processed as close to the vehicle and road sensors as possible, thus reducing the round-trip time for data. The server-side application of mobile edge computing can obtain local messages directly from the vehicle and road sensor applications, identify high-risk data and sensitive information that needs to be transmitted in near real-time through algorithm analysis, and send early warning messages directly to other vehicles in the area to facilitate early decision making by drivers to allow nearby vehicles to avoid hazards, slow down, or change routes (Li and Liu, 2017). It provides relatively real-time driving parameters setting guidelines for different vehicles, and the cloud service platform obtains the historical data from the area edge servers in different regions and applies these to the upper-level applications such as vehicle scheduling across regions, violation monitoring, and traffic map construction and updating. For smart healthcare, with the improvement of living standards, the demand for high quality medical and health care is increasing (Rahmani et al., 2018). Healthcare can also be

aided by edge computing. Edge computing applied to smart healthcare systems comprehensively improves global healthcare, and at the same time, brings safer, timely, and effective medical assistance to patients in different regions. In the smart medical system, edge computing and 5G technology are combined in the system to achieve real-time patient information acquisition by IoT devices, ensuring low latency and real-time computing through high-quality network transmission and computing power, realizing remote diagnosis, remote surgery, and emergency rescue, providing flexible and personalized medical services for patients, and meeting the patient's ability to send data about their health condition without leaving home. For example, MEC can help health advisers to assist their patients, independent of their geographical location. MEC enables smartphones to collect patient physiological information, such as pulse rate and body temperature, from smart sensors and send it to the cloud server. Health advisers who have access to the cloud server can immediately diagnose patients and assist them accordingly (Stantchev et al., 2015). Furthermore, the smart home system is the hot application under the IOT, through computer and communication technologies, the system produces an amount of sensitive data locally, and the data collected by the sensor is transmitted to the decision-making unit, which calculates the appropriate control signals to achieve the predetermined goal. The gateway is applied as the edge that is between the home device and the cloud, through internet services, such as Bluetooth,Wi-Fi, and home LAN, to realize locally controled kinds of smart home devices and remotely operated home devices. Therefore, the edge gateway form the home edge network and the intelligent home service platform, reducing uplink data transmission, service cloud platform load, and responding to user needs with ultra-low latency (Trimananda et al., 2018). Moreover, for the smart building, smart building control systems consist of wireless sensors that are deployed in different parts of buildings. Sensors are responsible for monitoring and controlling building environments, such as temperature, gas level, or humidity. In a smart building environment, sensors installed with edge computing are capable of sharing information and become reactive to any abnormal situation. These sensors can maintain the building atmosphere on the basis of collective information received from other wireless nodes. For example, if humidity is detected in the building, MEC can react and perform actions to increase the air in the building and blow out the moisture. As one of the IoT applications, the smart grid is a cyber physical system covering various smart devices. For instance, Tencent Cloud and Pengmai Energy Technology took edge computing into account and released the overall architecture of energy IoT solution in Cloud Tencent and Energy IoT Pengmai (2018). (Okay., 2016).proposed a fog computing-based smart grid model, and presented an example scenario, the smart homes, in terms of latency and security, of the advantages of the model. Sun H. Y. et al. (2019) introduced an edge computing technology for power distribution internet of things (PD-loT), and provided the architecture of edge computing for PD-loT, meanwhile, they analyzed the internal and external interaction mechanism of the data center construction under this architecture and the cloud-side collaboration mechanism based on the data center. Furthermore, they have analyzed the application of edge computing in the typical service of power distribution "orderly charging of electric vehicles." Li B. et al. (2018) discussed the application of edge computing in demand response, and analyzed the application of edge computing in specific scenarios in the field of power supply and demand, such as home energy gateways, non-intrusive load monitoring, and orderly power management. Gong et al. (2018) proposed a packet transport network (PTN) physical architecture model of active distribution network (ADN) based on edge computation, and constructed a cyber physical system (CPS) management and control model of ADN based on edge computation. Kumar et al. (2016) proposed a generalized architecture for data management based on vehicular delay-tolerant network (VDTN) using edge computing for smart gird, and also proposed an energy efficient virtual machine migration utilizing load forecasting. Zahoor et al. (2018) proposed a three-layered framework named cloud–fog-based smart grid, analyzed the edge computing layer close to the consumers' region that performs effective management of the network resource with low latency, and considered two scenarios for performance evaluation of their cloud–fog-based smart grid model. Liu et al. (2019) presented an architecture of edge computing-driven autonomous ubiquitous IoT in Electricity, which is based on the edge computing architecture and virtual synchronization technology, and introduced the applications that may be deployed in the edge computing gateways. Chen et al. (2019) introduced the services of IoT-based smart grid supported by edge computing, and proposed an architecture introducing edge computing into IoT-based smart grid, as well as presented the three scenarios of the smart grid. Long et al. (2020) discussed the advantages of edge computing technology used in the demand-side management of power consumption in the smart grid. Khan et al. (2020) highlighted the role of edge computing in realizing the vision of smart cities, and reviewed the state-of-the-art literature focusing on edge computing applications in smart cities, including smart transportation, smart health-care, smart grid, and smart farming.

Following an investigation, we found that there are only a few surveys discussing the current status of the applications of edge computing in UPIoT, since it is too new to have attracted too many people's attention. This paper will analyze the architecture and application of edge computing in three scenarios of power IoT, including distribution network automation monitoring system, smart energy system, and power metering system, as well as analyze the advantages of introducing edge computing in the three scenarios. It provides a significant reference for follow-up researchers, designers, and beginners.

4 APPLICATION OF EDGE COMPUTING IN UBIQUITOUS POWER INTERNET OF THINGS

As shown in **Figure 3** (Chen et al., 2019), applying edge computing technology, IoT technology, and 5G

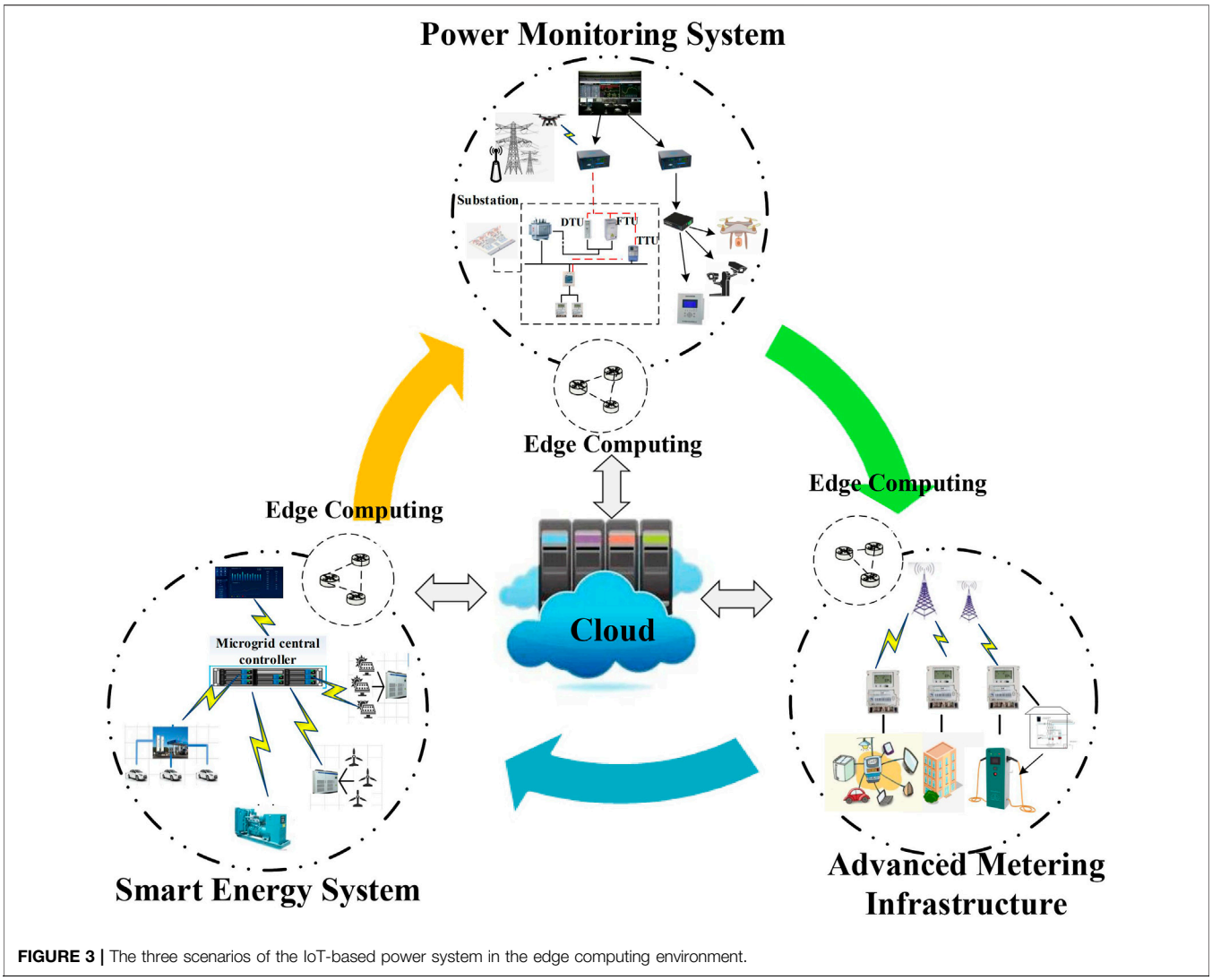


FIGURE 3 | The three scenarios of the IoT-based power system in the edge computing environment.

communication technology in three major scenarios, improves the performance of the power system and makes the power system more intelligent and automated.

4.1 Power Monitoring System

The power monitoring system is composed of the control center at all levels, such as substations, power line surveillance, and so on (Tao et al., 2017). The system applies modern control technology, visualization technology, modern communication technology, and Internet of things technology to intelligently monitor power equipment hot spots, power, and environment. It intelligently analyzes data, and realizes comprehensive visualization display and intelligent linkage alarm, at the same time, it effectively assists power equipment informatization, overhaul, and operation. Overall, it serves for smart grid overhaul, operation, and whole life cycle management (Li, 2019). As shown in **Figure 4** it mainly describes the two service applications of transmission lines and the intelligent substation in the power monitoring system based on edge computing.

4.1.1 Transmission Line Monitoring

On the one hand, in order to realize the intelligent inspection of the transmission line, the unmanned aerial vehicles obtain the image, video, and other data of the surrounding environment of the power transmission through the fixed-focus camera in the pan-tilt. The front-end edge computing detection module recognizes and locates the acquired video stream and the condition of the overhead line. The application of edge computing device nodes in the power line realizes the autonomous, procedural, and standardized collection of image information in the process of the transmission line equipment. Meanwhile, the acquired images and video image data are processed and analyzed locally, reducing invalid information sent back to the cloud platform, and improving the effectiveness of image data information. This method of

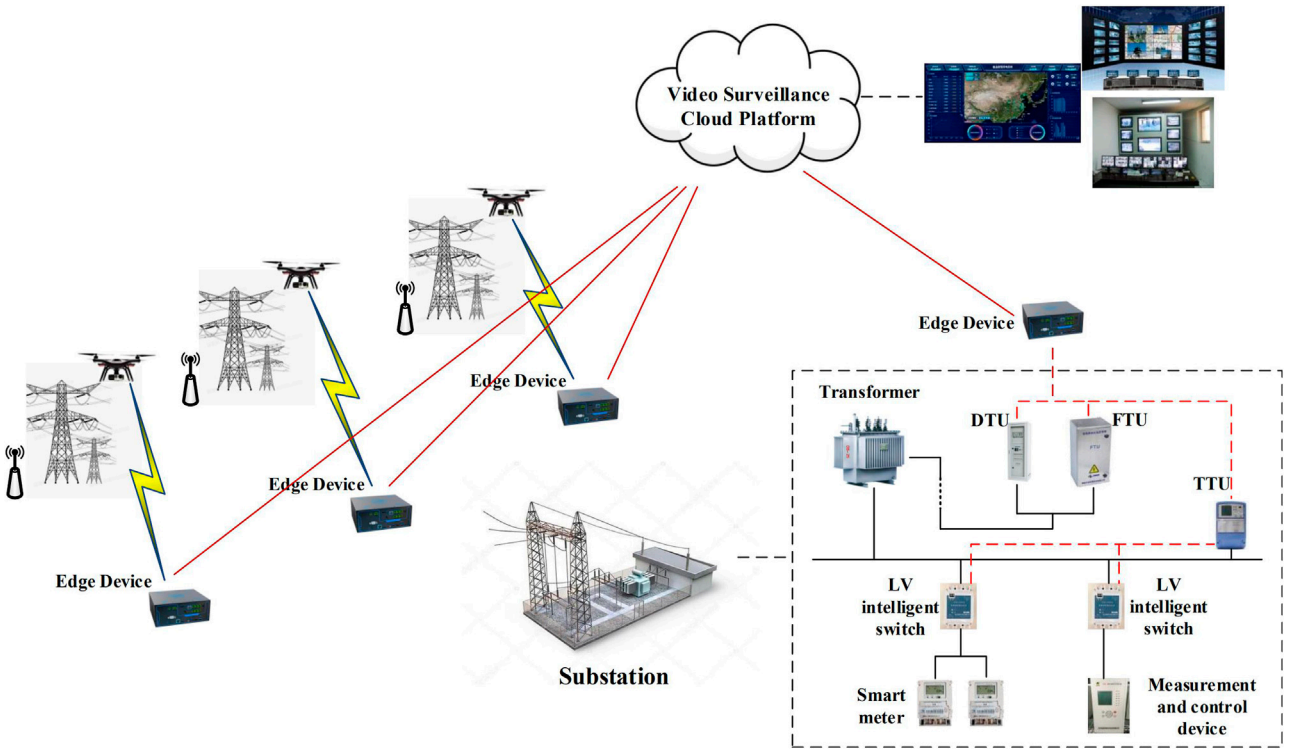


FIGURE 4 | The architecture diagram of edge computing for power distribution system.

utilizing the edge computing detection module that reduces the pressure on broadband reduces the load pressure on the cloud platform, and reduces the complexity of subsequent defect diagnosis. Furthermore, the manual participation in inspection operations is reduced, and the efficiency is improved by 5–7 times. On the other hand, it is neccesary to improve the intelligent level of transmission line monitoring and early warning, operation and maintenance, and overhaul. By deploying PMU, FTU, and other sensing units to collect local information, the data types include second-level switch position signals, fault signals, and minute-level signals. Equipment status data, such as electrical information and connector temperature, partial discharge signal, tower status, and channel status, support normalized operations, such as line inspection and troubleshooting. Among them, electrical information, switch status, and local analysis results will be uploaded to the master station layer as remote information, which realize real-time online monitoring and state detection and improve the intelligent level of environmental monitoring and early warning of important transmission channels, as well as realize intelligent operation, maintenance, and overhaul of the power grid.

4.1.2 Intelligent Monitoring of Substation

With the continuous improvement of the IEC 61850 standard and the continuous improvement and deepening of key technologies such as electronic voltage transformers, current transformers, MU, intelligent terminals, and process layers, as well as the continuous construction and deployment of UPIoT, he substation tends to be digital and intelligent (Tang et al., 2021). The system of the substation monitoring system is divided into process layer, bay layer, and station control layer. Substations need to monitor a huge number of secondary equipment such as lines, switch, and circuit breakers, and collect the electrical quantities of the lines and the on-off status of control equipment through the process-level intelligent integrated device of the terminal, then upload them to the GOOSE/SV network. SV messages and GOOSE messages are standard message formats in the IEC61850 communication protocol, they are the line state value and the switch state value, respectively. These messages are transmitted from the process layer to the bay layer, and the bay layer performs line stability control and relay according to the message content. The protection action is performed, and then protection is uploaded and information is relayed to the station control layer in MMS messages, so as to realize the analysis and processing of the monitoring data. Finally, the remote control host at the station control layer transmits it to the dispatch center through the communication protocol in the monitoring system, and waits for the analysis result and task dispatch of the dispatch center. However, both station domain protection and line protection of substations have extremely low delay requirements, and the impact of factors such as message analysis, communication congestion, and network packet loss will bring harm and hidden dangers to the operation of the power grid. By deploying the edge computing platform at the station

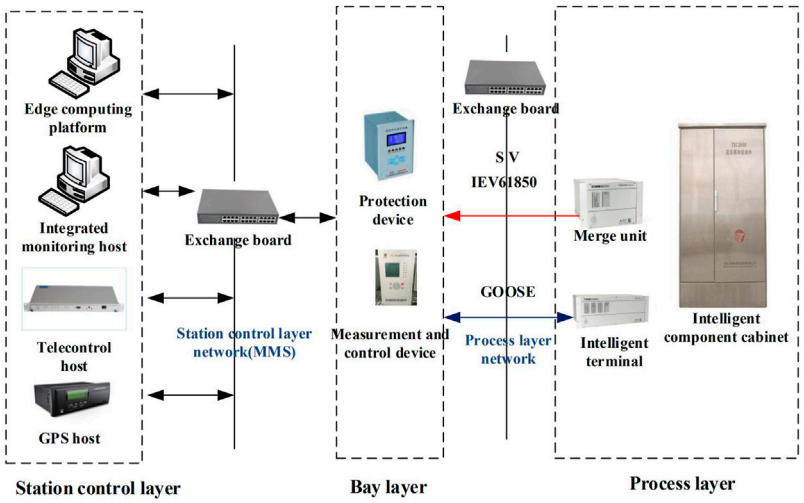


FIGURE 5 | The architecture of edge computing in the integrated automation system of substation.

control layer, the platform process, and storing part of the information on-site by capturing and parsing MMS messages, this would reduce the burden on telecontrol communications while ensuring the quality of application services (Bai et al., 2020). The architecture of edge computing in the integrated automation system of substation as shown in **Figure 5**.

4.2 Smart Energy Systems

The smart energy system is an integrated management system, which is made up of a distributed generator, energy storage devices, flexible loads, and energy conversion devices. The integrated energy management platform coordinates the electrical energy interactions in the power network and uses a microgrid central controller, a distributed power grid connection interface device, and an intelligent control terminal to implement the basic functions of the smart energy system.

The edge computing architecture for the smart energy system is shown in **Figure 6**. The architecture consists of three layers: device layer, edge layer, and cloud layer. The cloud layer takes the cloud platform as the core and provides various cloud services. For different scales, the cloud layer can deploy the public cloud, private cloud, or hybrid cloud. The equipment layer consists of various types of power devices, including uncontrolled distributed power sources such as photovoltaic and wind turbines, controlled distributed power sources such as diesel generators and power conversion devices such as inverters, energy storage devices such as electric vehicle charging piles and batteries, and various types of loads. The edge layer is the core of the entire architecture; it consists of edge gateway, edge platform, and edge services. The edge layer provides computing, storage, application deployment, and other functions at the edge side near the data source of the device. The edge gateway is the core device in the edge computing architecture, which collects the operation data of distributed power supplies, loads, power conversion devices, and energy storage devices in real-time, and then uploads them to the edge platform. Under the coordination of the edge platform, each edge gateway executes the control commands derived from the calculation results at the edge side to control the dispatchable power devices (Xu et al., 2020).

We will analyze the typical applications and the related advantages brought by the architecture based on the EC-IoT smart energy systems, including identification of malicious behavior of electricity consumption and Real-time perception of distributed energy power generation status, efficient data processing, and fast reactive voltage response.

4.2.1 Identification of Malicious Behavior of Electricity Consumption

With the massive access of a large number of distributed energy sources and terminals containing power electronic equipment, the terminal equipment in the smart energy system is vulnerable to permission attacks, data storage and encryption attacks, vulnerability threats, and remote control (Lei et al., 2020). These risks will lead to abnormal activities of the terminal's feedback data, allowing the micro-grid central controller to collect wrong information and then make wrong decision-making activities, causing the local or even the entire system of the micro-grid to collapse (Komninos et al., 2014; Lei et al., 2020). In order to monitor malicious behaviors of users online in real time, the smart energy system uses the microgrid central controller, distributed energy grid-connected interface devices, and charging piles, as edge computing modules to form an edge computing platform. For example, the edge computing module establishes the power generation behavior of each device with the characteristics of power and time through the distributed energy generation collected in real time. The edge computing module uses electricity consumption and time as parameters for the user's electricity consumption behavior in the microgrid, and then establishes the user's electricity consumption behavior pattern and constructs the database for users. Through the edge module, which can extract power generation and power consumption data

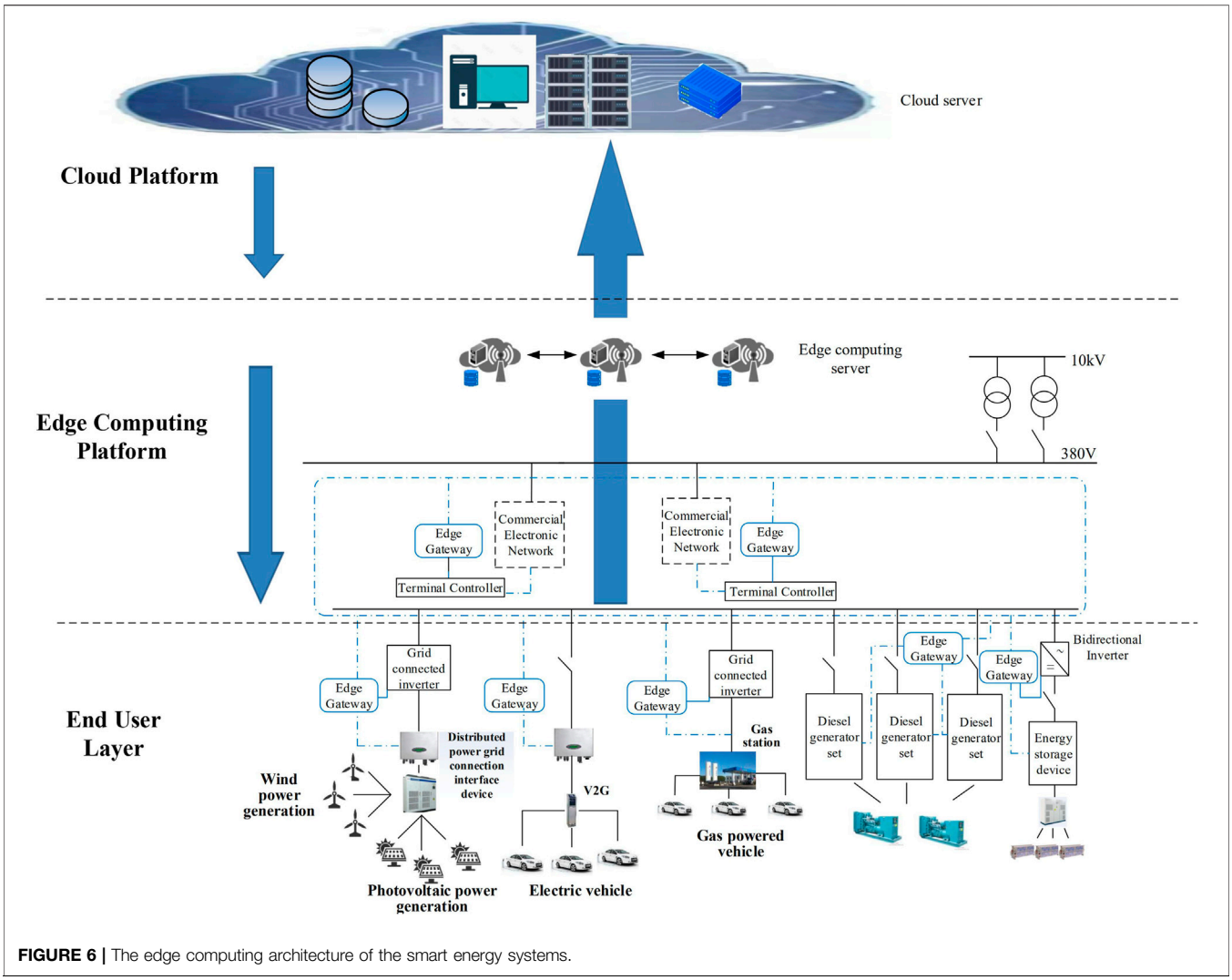


FIGURE 6 | The edge computing architecture of the smart energy systems.

anytime and anywhere, then realizes data interaction with the main station, identifies the characteristics of power generation and consumption, and realizes the power balance of smart energy and the identification of malicious behavior (Mao, 2020).

4.2.2 Real-Time Perception of Distributed Energy Power Generation Status, Efficient Data Processing, and Fast Reactive Voltage Response

By installing the edge computing equipment on the grid-connected inverter side of photovoltaic and other power generation equipment, the output electrical quantity and fault information of the inverter in real-time can be monitored and the main data can be uploaded to the integrated energy management system. On the one hand, edge computing devices can collect data from grid-connected inverters, box transformers, and combiner boxes in real time. On the other hand, the architecture of the edge computing module is composed of hardware and software, and has embedded control software. When the grid voltage and frequency collected in real time fluctuate abnormally, there is no need to wait for the control instructions of the cloud platform. The control algorithm adaptively controls the power output of the inverter, and quickly responds to the grid voltage and frequency, so as to realize the comprehensive perception of the state of the power generation unit and efficient data processing (Sun et al., 2021), and quickly support the safe and stable operation of the grid.

4.3 Advanced Metering Infrastructure

Advanced metering infrastructure (AMI) consists of smart meters, data concentrators, data centers, and communication networks. AMI is interconnected with the communication network to achieve two-way communication of power data. In the AMI, the smart meters are uploading their power usage information to the data concentrators through wired and wireless communication in the Neighborhood Area Network (NAN). And then the data center actively requests power data from data concentrators through the wide area network (WAN), or data concentrators pass through the WAN at a preset time interval, and they centrally upload power consumption data to the data center, then the data center distributes electricity price

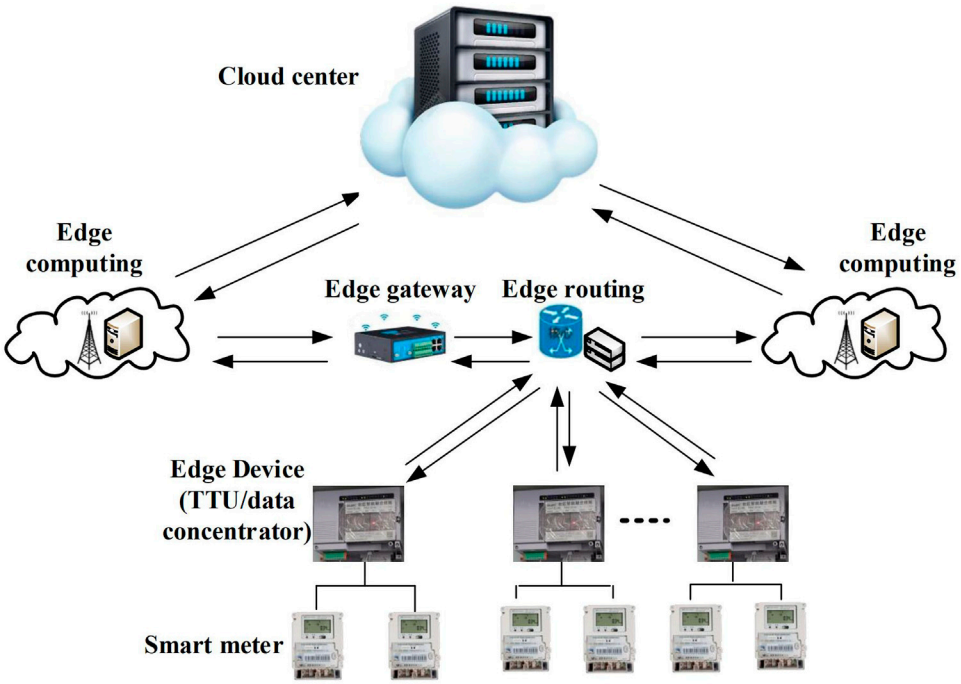


FIGURE 7 | The edge computing architecture of the AMI.

information to users and implements related measures such as load management, demand response, and meter control commands to improve customer service (Liang et al., 2021).

With the rapid development of the smart grid, the data generated by smart meters and other power terminal devices has exploded. Facing the computing demand of massive data, traditional cloud computing solutions face huge challenges in transmission bandwidth, transmission delay, data storage, and real-time response. The introduction of edge computing technology and the introduction of edge computing modules in the data concentrator constitute the AMI edge computing framework, as shown in **Figure 7**. The data collected by the terminal is processed locally in the concentrator, and only the calculation results are uploaded to the cloud, thus reducing the network burden, lowering transmission costs, and meeting users' real-time response needs, etc.

We will present the typical applications and the related advantages brought by the architecture based on the EC-IoT advanced metering infrastructure, including real-time power forecasting and efficient abnormal detection.

4.3.1 Real-Time Power Consumption Prediction

By adding an edge computing module under the current power concentrator, the massive power metering data is passed to the concentrator, and the edge module is an edge device. Based on the computing power of the edge module, the combination of deep learning and machine learning methods is adopted, such as online learning combined with machine learning algorithms (GBDT, XGBOOST, Linear Regression, etc.), which is used to train the time-series consumption data, then realize the real-time prediction of metering data and real-time feedback of consumption bills and the reasonable strategy of the power consumption to users that meet the demand response. Since the data from the massive power terminals are transferred to the edge device, they are no longer transferred to the cloud, and processed directly by the edge device's own computing resources, which greatly improves the data processing efficiency. In addition, for the new power data generated every moment, the online learning method can be used to quickly complete training and make predictions, which greatly improves the real-time data processing (Chen et al., 2019).

4.3.2 Efficiently Abnormal Detection

The traditional anomaly detection method relies on cloud computing technology to analyze and process all collected data in the data center. A distributed detection method based on edge computing is proposed, which sets up edge node detectors on the edge side of the grid to collect, store, and detect data directly instead of the original central processor. The edge computing-based detection method transforms the traditional centralized detection into a distributed detection method. Combined with the deep learning approach, the anomaly detection model is constructed, and the training process is separated from the edge nodes and placed at the central node to complete. The distributed detection model of the power metering system based on edge computing is shown in **Figure 8**. At the edge, the edge node concentrator is responsible for collecting and storing smart meters and related data around them, uploading the processed real-time data and the stored related historical data to the central node, which can perform local processing and detection based on

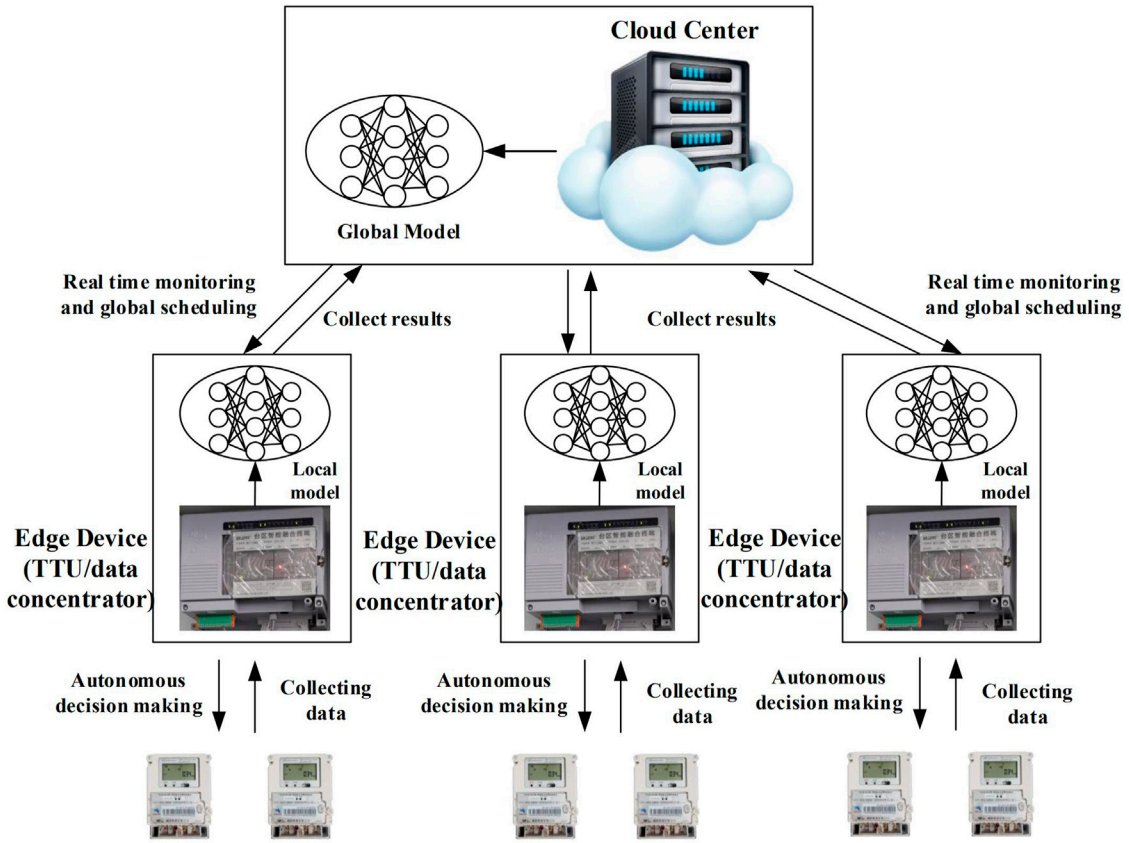


FIGURE 8 | The distributed detection model of the power metering system based on edge computing.

the data they monitor and collect, and sending the trained anomaly detection model down to the edge nodes. In the cloud center, it undertakes the training task of the detection model, and after the model is trained, it is then downlinked to the corresponding edge node. The edge node can realize offline detection, which can realize more efficient and safe monitoring of the power grid and reduce the pressure on the cloud center.

5 CHALLENGES AND OPPORTUNITY OF THE EDGE COMPUTING IN THE UBIQUITOUS POWER INTERNET OF THINGS

The UPIoT seeks to achieve the information-physical-social integration of power energy systems, however, the construction of the UPIoT is still in its infancy at this stage. Although edge computing technology has been gradually applied to various aspects of the power grid, there are still huge challenges to the application of edge computing in the UPIoT. These challenges come partly from the limitations of edge computing technology itself and partly from the potential constraints that exist when edge computing is combined and applied with UPIoT. This section analyzes the market, policy, and technology risks of edge computing in UPIoT applications, and discusses the technical challenges and corresponding outlook for UPIoT based on edge computing.

5.1 Market Challenge

Edge computing enables UPIoT. Constructing UPIoT requires the installation of massive amounts of edge devices, and power operators update traditional solutions to upgrade power infrastructure according to emerging development needs, but it is difficult to quantify its effect on power production, business capacity improvement, and cost reduction, and the corresponding business model is still in the exploration stage. Therefore, the development of the edge computing that is applied in the UPIoT will be faced with market and economic challenges, such as the possibility of investing hundreds of billions of dollars and how to develop operational and business models. Will the return on investment achieve the expected results? And how will the financial crisis be dealt with on the impact of grid planning, investment and operation? All these questions require grid companies, operators, governments, and others to think deeply.

5.2 Policy Challenge

Edge computing realizes the interconnection of power equipment and sensors through the integration of wireless networks, mobile center networks, Internet, and other communication networks.

The collection of massive power terminal data must be transmitted to the edge platform through communication, and then the data is processed and analyzed on the edge platform, so it is essential to ensure the security of the edge side network. On the one hand, it is necessary to ensure that data can be reliably and quickly transmitted from the remote end to the platform, and on the other hand, it is necessary to prevent data leakage and resist cyber attacks. Therefore, how to develop new information security protection technologies and systems is a difficult point for edge computing to be applied in the UPIoT, and there is a lack of a unified standard system in the power industry. In Europe and the United States, the development of smart grids is relatively advanced, and the development of security and stability standards is more detailed and comprehensive, For example, the European Telecommunications Standards Institute (ETSI) has defined typical service scenarios and engineering implementation guidelines for edge computing (European Telecommunications Standards Institute, 2016), so the application of edge computing can be promoted in the EU and other countries. However, in China, the security protection of the power secondary system follows the overall protection principle of "security zone, network dedicated, horizontal isolation, and vertical authentication" to ensure the security of the power data network. On the one hand, the standard system related to edge computing is not complete; on the other hand, due to the huge scale of China's power industry and the relatively conservative policy implementation, it is difficult to apply a new technology immediately. Therefore, in order to realize ubiquitous IOT intelligent sensing, the introduction of edge computing will inevitably lead to network and information security issues, and the relevant specific policies are not yet perfect, so it is difficult to apply and implement edge computing perfectly. Therefore, there are corresponding policy risks for edge computing to be applied to UPIoT, which require in-depth thinking by countries all over the world.

5.3 Technical Challenge

There are many distinctive technical challenges in UPIoT edge computing environments, including management and processing of massive heterogeneous power terminal devices accessing edge side, data offloading and load balancing, edge intelligence, edge network security, data sharing security, and privacy protection.

We summarize these technical challenges in detail and outline potential research directions.

5.3.1 Management and Processing of Massive Heterogeneous Power Terminal Devices Accessing Edge Side

Due to the stage of construction and real-time operation of the power system, a large amount of historical and real-time data, such as control, monitoring, and metering, will be accumulated on the power IoT platform, which constitutes a multi-source heterogeneous data source for the platform layer of the UPIoT. Heterogeneous data sources arise from a variety of power devices, owing to the variety of power devices, the lack of unified access to the edge node standards, and the edge network center node itself which is more difficult to expand. However, moving the location of the edge computing center will bring a lot of delay. Furthermore, in the edge computing network, there are not only static end-devices (e.g., sensors or video cameras), but also dynamic ones such as UAVs and electrical vehicles, making the device management even more challenging. Meanwhile, since different devices have huge differences in hardware configurations and software functions, and the corresponding data computing, storage, and communication capabilities vary, this also offers a challenge (Qiu et al., 2020). Therefore, there is an issue in management of massive heterogeneous power devices accessing the edge nodes, and the edge network must be programmable to support application-specific requirements of edge terminal devices. For this challenge, SDN and NFV technology realize the management and control of multi-source heterogeneous terminal equipment, as well as the scheduling and routing of data flows. SDN and NFV are two of the latest technologies designed to introduce flexibility in network management and orchestration. SDN is mainly characterized by the decoupling of the control plane from the data plane; it provides programmability for network application development and supports the new technology in a unified manner (Bera et al., 2017). SDN and NFV are applied for the edge network. It will make the network more flexible and programmable (Han et al., 2018). Many researchers have carried out certain discussions and research on the application of software-defined network technology in the power Internet of Things, and have achieved relevant results (Yang et al., 2017; Wang et al., 2015; Zhong et al., 2021), However, the combination of edge computing with SDN, the application of which in the UPIoT is in its infancy, will become a trend and is worth in-depth research in the future. Furthermore, utilizing deep learning techniques on the edge computing platform to assist in edge network control for the large-scale heterogenous power terminal devices is also an area of focus research for studies. At the same time, with the development of 5G networks and its application to the power Internet of Things, in the future, we can study the application of 5G technology to UPIoT-assisted edge computing for edge terminal network management; the edge node management issues will be easy to handle in a 5G network (Kumareshan and Poongodi, 2016). Specifically, 5G core networks can collect status information of various nodes regularly (e.g., node location, resource use, task list, and adjacent nodes), monitor and update node management information and strategies, and optimize other strategies such as data processing strategy and network protection strategy according to the collected information. The combination of edge computing with 5G that adapt to the challenge of management and control in the unified access of massive heterogeneous terminals in the UPIoT, is worth exploring in depth in the future. This is only one side. On the other hand, for a large amount of multi-source heterogeneous data, the computing power is deployed at the edge nodes to deal with heterogeneous data. However, in the actual operation of the UPIoT, on the one hand, the algorithm iteration speed is very fast, and multiple versions of calculation programs cannot be stored locally for a long time. On the other hand, the interface types of terminal devices vary greatly due to the influence of operating environment, communication network, and other factors.

There are both wired communication interfaces, such as Ethernet, PLC, RS485/232, and wireless communication interfaces, such as 4G/5G public network. There are both traditional wireless business terminals, such as RTU, PTU, and TTU, and intelligent business terminals. Therefore, it is difficult to process a large amount of multi-source heterogeneous data, which affects the efficiency and performance of edge computing. For this challenge, the technologies of UPIoT microservices, algorithm subscription, and container VM technology are worth studying in depth, which can be carried out to improve the flexibility of algorithm deployment, and then promote agile development and rapid iteration of edge computing in the UPIoT.

5.3.2 Realizing Edge Autonomy to Meet the Business Demand for Real-Time Response

With the construction and development of the UPIoT, the amount of power equipment has increased sharply, and the deployment area is also relatively wide. The on-site installation environment is not only complicated but also diverse, resulting in a sharp increase in the workload of installation, commissioning, management, and maintenance of power devices. Although edge computing is introduced, in current UPIoT systems based on edge computing, edge devices can only perform lightweight computing tasks. To enable edge devices and edge servers to perform more complex tasks with a higher data processing performance and lower latency, edge computing combined with artificial intelligence, big data analysis, and other technologies that are applied in the UPIoT assisted edge computing, which make edge devices and servers intelligent. However, edge equipment resources and computing capabilities are limited, it is difficult to achieve lean management, and there exists a challenge of realizing edge autonomy and Intelligence to meet the business demand for real-time response to improve the quality of service (QoS) and quality of experience (QoE).

Many scholars propose some ideas for coping with challenges from two aspects: the edge device itself and the edge model architecture. In terms of edge computing devices, this could be done by adding AI processor modules or re-designing the intelligent chip to improve the computing power of edge devices. In the design of chip architecture, the aim would be to support the edge computation paradigm and facilitate AI models (e.g., DNNs, CNN, etc) acceleration on the resource-limited IoT devices. Meanwhile, customized AI processors are developed to be better suited to specific edge devices and usage scenarios. For example, developing customized power chips and AI edge computing processors (Zhao et al., 2021) to meet the requirements of massive data processing and calculations. However, there exists a great challenge in developing devices more suited to edge AI and realizing the theoretical complementarity of edge computing and AI (Qiu et al., 2020). At present, in support of the intelligence of the edge computing for the UPIoT, the reconstruct of the edge device is not researched in-depth since some emerging technologies and architectures are under development. On the other hand, artificial intelligence models have been deployed at the edge layer. Although machine learning can be used to enhance the intelligence of edge devices, the high complexity of deep learning in many machine learning methods leads to the relatively high difficulty of the deployment. Due to the computing power of edge devices being generally weak, it is necessary to compress and simplify the model, and optimize the model architecture to adapt to the edge system to improve the processing performance. On the one hand, the technology of model compression for the edge device is worth researching since many approaches, such as model compression, conditional computation, and algorithm synchronization, are proposed to improve the efficiency of training and inference of deep AI models that are utilized in the edge. On the other hand, the architecture of the deep learning model is needed to optimize for the UPIoT assisted edge computing, As we know, the goal of the continuous development and construction of the UPIoT is realizing the transition from "collection + centralized control" to "collection + control + regional autonomy," and gradually shifting from "vertical closure" to "horizontal openness," The research of cloud-edge collaboration technology has become a developmental trend of edge computing in the UPIoT. Taking smart substation as an example, it will evolve from three stages as an important infrastructure for the construction of UPIoT: edge interconnection to edge intelligence, and then to edge autonomy. Utilizing the cloud-edge collaborative model for the substation, which remains in the initial stage, needs more efforts to be made, and is an effective way to realize the edge of substation autonomy, and can be the focus point for researchers that need to think and study in the next step.

5.3.3 Edge Network, Edge Node Security Protection, and Data Privacy Protection

Although applying edge computing to the UPIoT can bring the advantages of reduced transmitted data volume on the network, communication delay, computational costs, and enhanced flexibility, a large number of terminals accessed to the edge layer inevitably increase data interfaces that may be used as a springboard to attack edge nodes. On the other hand, edge nodes are close to the edge, and the network protection is weak. In the edge computing scenario, edge computing networks are distributed, scalable, and heterogeneous, the security measures of edge servers are weaker than traditional cloud servers, and they are vulnerable to attacks from malicious nodes in the network. Traditional security protection methods cannot satisfy the protection requirements of edge computing, as the security risks cannot be fully considered at the start of the design. Moreover, the integration of various technologies has also intensified the security threats related to data, networks, and applications. In UPIOT based on edge computing, there are challenges of edge network, edge node security protection, and data privacy protection. In response to this challenge, blockchain is being considered as a disruptive technology by academicians and industries that offers potential solutions to solve the security and privacy issues of edge computing networks and devices (Kang et al., 2019; Frey et al., 2019). The incorporation of blockchain and edge computing into a single framework, and then combining the attribute-based access control model, will make it possible to have reliable access and control over the network, storage, and distributed computational resources at the edge. Blockchain technology can also improve the security of the EC-assisted IoT paradigm as it permits only trusted IoT

devices/nodes to interact with each other. On the one hand, blockchain-based trusted data management schemes (e.g., lightweight consortium blockchain) for cooperative authentication, authorization, and privacy preserving could be developed (Gai et al., 2019; Mao et al., 2020; Wang et al., 2020), meanwhile, utilizing blockchain to form a security mechanisms for edge nodes/devices could ensure the security and credibility of edge nodes in the UPIoT. These two proposed methods can not only prevent edge nodes from being attacked and lead to data privacy of the problem of leakage, but also ensure the integrity and security of the data sharing using the external network; therefore, it ensures the credibility of regional terminal computing tasks in the edge layer, and then returns the correct calculation results to the cloud and terminal users safely and reliably. Overall, these are potential research directions for studies, and need more research and developments. On the other hand, the combination of blockchain and machine learning methods (such as federated learning, deep learning, reinforcement learning, and deep reinforcement learning) to enhance the access control, secure storage, and privacy -preserving data of EC-assisted UPIoT, can be adopted to detect abnormal behavior, such as energy theft with energy privacy protection in the smart gird (Yao et al., 2019), which is an emerging field and a promising direction for research.

5.3.4 Edge Computation Offloading and Load Balancing Realize Demand Response

UPIoT is a promising solution to meet the increasing electricity demand of modern cities, while challenges face the real-time processing and analysis of huge data collected by the power terminal devices due to the limited computing capability of the devices and long distance transmission from the cloud center. Edge computing enables power terminal tasks to be offloaded on the edge side for *in-situ* processing. It reduces communication delays and energy consumption. However, the power consumption, computing power, and storage space of edge-layer servers are also limited. Computation offloading and load balancing are great challenges in UPIoT systems based on edge computing.

For the challenge of the data offloading, scholars mostly combine artificial intelligence technologies (Lin et al., 2019; Sun W. et al., 2019) (such as deep reinforcement learning (Dinh et al., 2018; He et al., 2018; Dai et al., 2019; Luo et al., 2019; Min et al., 2019; Zhang et al., 2019a; Zhang et al., 2019b), such as markov chain decision-making, game theory, Lyapunov optimization, machine learning and so on) to improve the performance of computing offloading schemes. However, according to the characteristics of UPIoT, to find the optimal balance between energy consumption, delay, amount of data, bandwidth, it is necessary to design the computation offloading strategy according to the data volume, task type, and equipment capabilities (Jiang et al., 2019a; Pan et al., 2020), which is worthy of in-depth consideration, and is the outline for future studies by researchers at this stage. Furthermore, another challenge is considering how to allocate resources reasonably after making the data offloading decision, that is, the problem of where the resources are distributed. In other words, there is a new problem in that the data offloading scheme may lead to the overload of some edge devices, then creating another challenge of load balancing. Load balancing based on specific characteristics and scenarios of the UPIoT combining new technologies is a significant research direction. Considering the scale and frequency of scheduling are significantly larger, it is necessary to improve the existing load balancing algorithms to adapt the characters of the UPIoT edge system. A typical method depends on the NFV and SDN integrated edge cloud platform to orchestrate the resources to fulfill the offloaded tasks from the battery-constrained edge terminal devices. In UPIoT based on edge computing, using the task data and loading data from the equipment and edge servers, combined hierarchically with AI, a load balancing service can set up the load balancing scheme based on a machine learning model for each layer. In addition, SDN can be utilized to conduct load balancing scheduling from the global perspective of the edge network. To minimize the complexity of scheduling and routing, emerging SDN technology will have a significant impact on the routing scheme and communication mode of edge network, and brings more comprehensive and in-depth routing schemes for edge computing in UPIoT (Kaur et al., 2018; Li X. et al., 2018; Nayak et al., 2018; Al-Hubaishi et al., 2019). At present, some scholars propose to use SDN to establish a grid edge computing model, with minimum delay as the goal orientation, and use deep reinforcement learning to reasonably schedule and allocate computing resources (Shang et al., 2021). However, there is very little research on resource allocation strategies that target how to balance time delay and energy consumption, which is a direction worthy of research. On the other hand, Machine learning algorithms based on resource allocation in edge-cloud architecture, edge-edge architecture, could also solve the problem of load balancing. For example, based on the edge-cloud architecture, the computing tasks of the terminal cannot be completely offloaded to the edge side for execution, so some tasks are offloaded to the remote cloud server for calculation, and the result will be first returned to the edge server, and finally back to the terminal device. If a client's requirement is more critical, it will be handled by the cloud; otherwise, servicing is done by the edge. However, the reality of how edge computing and cloud computing can work together efficiently and seamlessly is a significant research direction (Li et al., 2020). When considering the computation offloading in edge computing, it is necessary to consider the gaming and cooperation between the edge and the cloud for task scheduling and collaboration.

6 CONCLUSION

Edge computing integrates network, computing, and storage on the edge of the network. The introduction of edge computing can solve the problems of cloud computing architecture facing the UPIoT, which is unable to handle massive heterogeneous data, communication delays, high computing pressure, data privacy leakage, and difficulty in satisfying user demand response and other issues. First, this article introduces the edge computing technology and the framework of the UPIoT, and gives the architecture of the combination of edge computing and the UPIoT and the internal architecture of the edge computing layer. Moreover, one of the contributions of this paper is to analyze the technical application of edge computing in the three power Internet of Things scenarios: power monitoring system,

smart energy system, and power metering system. It also gives the architecture of the edge computing in the three scenarios. Furthermore, the major contribution is putting forward the policy challenges, market challenges, and technical challenges of the application of edge computing in UPIoT, meanwhile, the technical challenges and outlooks in four major areas are analyzed in detail. This paper aims to obtain more attention from other researchers in edge computing in the UPIoT, and make power industry development more rapid and convenient.

AUTHOR CONTRIBUTIONS

DL conceived the overall structure and framework of the article. HL conceived the outline of the manuscript. QZ and ZZ wrote the manuscript and generated the figures. XZ and ML helped perform the manuscript with constructive discussions. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported in part by the National Natural Science Foundation of China (Grant No. 52177068), the Research Project in Hunan Province Education Department (Grant No. 21C0577), Postgraduate Research and Innovation Project of Hunan Province, China (Grant No. CX20210791) and in part by Key Laboratory of Renewable Energy Electric-Technology of Hunan Province (Changsha University of Science and Technology).

REFERENCES

Al-Hubaishi, M., Çeken, C., and Al-Shaikhli, A. (2019). A Novel Energy-Aware Routing Mechanism for SDN-Enabled WSAN. *Int. J. Commun. Syst.* 32 (17), e3724. doi:10.1002/dac.3724

Alrowaily, M., and Lu, Z. (2018). *Secure Edge Computing in IoT Systems: Review and Case studies[Conference Presentation]*."in 2018 IEEE/ACM Symposium on Edge Computing (SEC), Seattle, WA, USA, 440–444. doi:10.1109/SEC.2018.00060

Bai, Y. Y., Huang, Y. H., Chen, S. Y., Zhang, J., Li, B. Q., and Wang, F. Y. (2020). Cloud-edge Intelligence: Status Quo and Future Prospective of Edge Computing Approaches and Applications in Power System Operation and Control. *Acta Automatica Sinica* 46 (3), 397. doi:10.16383/j.aas.2020.y000001

Bera, S., Misra, S., and Vasilakos, A. V. (2017). Software-Defined Networking for Internet of Things: A Survey. *IEEE Internet Things J.* 4 (6), 1994–2008. doi:10.1109/jiot.2017.2746186

Caprolu, M., Di Pietro, R., and Lombardi, S. (2019). *Edge Computing Perspectives: Architectures, Technologies, and Open Security issues[Conference Presentation]*. Milan, Italy: IEEE International Conference on Edge Computing EDGE, 116–123. doi:10.1109/EDGE.2019.00035

Chen, S., Wen, H., Wu, J., lei, W., Hou, W., Liu, W., et al. (2019). Internet of Things Based Smart Grids Supported by Intelligent Edge Computing. *IEEE Access* 7, 74089–74102. doi:10.1109/ACCESS.2019.2920488

Cloud Tencent and Energy IoT Pengmai (2018). [Online]. Available: https://cloud.tencent.com/solution/energy_iot.

Cornel - Cristian, A., Gabriel, T., Calin-Arhip, M., and Zamfirescu, A. (2019). *Smart Grid Integration of IoT.[Conference Presentation] 2019 54th International Universities Power Engineering Conference (UPEC)*. Bucharest, Romania, 1–5. doi:10.1109/UPEC.2019.8893551

Dai, Y., Xu, D., Maharjan, S., Chen, Z., He, Q., and Zhang, Y. (2019). Blockchain and Deep Reinforcement Learning Empowered Intelligent 5G beyond. *IEEE Netw.* 33 (3), 10–17. doi:10.1109/mnet.2019.1800376

Dinh, T. Q., La, Q. D., Quek, T. Q. S., and Shin, H. (2018). Learning for Computation Offloading in mobile Edge Computing. *IEEE Trans. Commun.* 66 (12), 6353–6367. doi:10.1109/tcomm.2018.2866572

Du, Y., Zhang, Z. Y., and Zhao, Y. (2021). Application Review of Edge Computing in Smart Grid. *Hubei Electric Power(in Chinese)* 45 (3), 72. doi:10.19308/j.hep.2021.03.012

Dustdar, S., Avasalcai, C., and Murturi, I. (2019). *Invited Paper: Edge and Fog Computing: Vision and Research Challenges*. San Francisco, CA, USA: IEEE International Conference Service-Oriented System EngineeringSOSE, 9609–9696. doi:10.1109/SOSE.2019.00023

European Telecommunications Standards Institute (2016). *Mobile Edge computing(MEC)framework and Reference Architecture:GS MEC 003 V1.1.1*. Sophia Antipolis, France: [Standard]FranceNiceETSI.

Frey, M., Gundogan, C., Kietzmann, P., Lenders, M., Petersen, H., Schmidt, T. C., et al. (2019). *Security for the Industrial IoT: The Case for Information-Centric networking[Conference Presentation]*. IEEE 5th World Forum Internet Things (WF-IoT), Limerick, Ireland, 424–429. doi:10.1109/WF-IoT.2019.8767183

Fu, J.-S., Liu, Y., Chao, H.-C., Bhargava, B. K., and Zhang, Z.-J. (2018). Secure Data Storage and Searching for Industrial IoT by Integrating Fog Computing and Cloud Computing. *IEEE Trans. Ind. Inf.* 14 (10), 4519–4528. doi:10.1109/tii.2018.2793350

Gai, K., Wu, Y., Zhu, L., Xu, L., and Zhang, Y. (2019). Permissioned Blockchain and Edge Computing Empowered Privacy-Preserving Smart Grid Networks. *IEEE Internet Things J.* 6 (5), 7992–8004. doi:10.1109/jiot.2019.2904303

Gong, G. J., Luo, A. Q., Chen, Z. M., Luan, J. Z., An, X. N., Wang, X. B., et al. (2018). Cyber Physical System of Active Distribution Network Based on Edge Computing. *Power Syst. Technol.* 42 (10), 3128. doi:10.13335/j.1000-3673.pst.2018.0119

Han, Y. Q., HeLou, Y. Q. F. D., and WangGuo, Y. B. C. X. (2018). Analysis and Application of SDN Based Dynamic Optimal Route Strategy for Cyber Layer in Cascading Failures. *cyber-physical Power systemPower Syst. Technol.* 42 (8), 2620

He, Y., Zhao, N., and Yin, H. (2018). Integrated Networking, Caching, and Computing for Connected Vehicles: A Deep Reinforcement Learning Approach. *IEEE Trans. Veh. Technol.* 67 (1), 44–55. doi:10.1109/tvt.2017.2760281

Hu, Y. C., Patel, M., Sabella, D., Sprecher, N., and Young, V. (2015). Mobile Edge Computing-A Key Technology towards 5G. *ETSI White Paper* 11 (11), 1

Industrial Internet Consortium:Edge Computing Task Group (2018). *Introduction to Edge Computing in IIoT*. https://production.iiconsortium.org/pdf/Introduction_to_Edge_Computing_in_IIoT_2018-06-18.pdf.(Accessed November 03, 2018).

Jiang, C., Cheng, X., Gao, H., Zhou, X., and Wan, J. (2019a). Toward Computation Offloading in Edge Computing: A Survey. *IEEE Access* 7, 131543–131558. doi:10.1109/ACCESS.2019.2938660

Jiang, X. C., Liu, Y. D., Fu, X. F., Xu, P., Wang, S. Q., and Sheng, G. H. (2019b). Construction Ideas and Development Trends of Transmission and Distribution Equipment of the Ubiquitous Power Internet of Things. *High Voltage Engineering(in Chinese)* 45 (5), 1345.

Kang, J., Yu, R., Huang, X., Wu, M., Maharjan, S., Xie, S., et al. (2019). Blockchain for Secure and Efficient Data Sharing in Vehicular Edge Computing and Networks. *IEEE Internet Things J.* 6 (3), 4660–4670. doi:10.1109/jiot.2018.2875542

Kaur, K., Garg, S., Aujla, G. S., Kumar, N., Rodrigues, J. J. P. C., and Guizani, M. (2018). Edge Computing in the Industrial Internet of Things Environment: Software-Defined-Networks-Based Edge-Cloud Interplay. *IEEE Commun. Mag.* 56 (2), 44–51. doi:10.1109/mcom.2018.1700622

Khan, L. U., Yaqoob, I., Tran, N. H., Kazmi, S. M. A., Dang, T. N., and Hong, C. S. (2020). Edge-Computing-Enabled Smart Cities: A Comprehensive Survey. *IEEE Internet Things J.* 7 (10), 10200–10232. doi:10.1109/jiot.2020.2987070

Komninos, N., Philippou, E., and Pitsillides, A. (2014). Survey in Smart Grid and Smart Home Security: Issues, Challenges and Countermeasures. *IEEE Commun. Surv. Tutorials* 16 (4), 1933–1954. doi:10.1109/comst.2014.2320093

Kumar, N., Zeadally, S., and Rodrigues, J. J. P. C. (2016). Vehicular Delay-Tolerant Networks for Smart Grid Data Management Using mobile Edge Computing. *IEEE Commun. Mag.* 54 (10), 60–66. doi:10.1109/mcom.2016.7588230

Kumareshan, N., and Poongodi, P. (2016). Dynamic Mobility Management Architecture to Improve Quality of Experience (QOE) in Wireless Networks."in 2016 10th International Conference on Intelligent Systems and Control (ISCO) CoimbatoreIndia. [Conference presentation].

Lei, W., Xu, A., Lin, H., Hou, W., Zhang, Y., Jiang, Y., et al. (2020). On-line Detection of Malicious Activities Based on Edge Computing in Micro-grid System. *J. Phys. Conf. Ser.* 1646 (1), 012131–012136. doi:10.1088/1742-6596/1646/1/012131

Li, H. (2019). *Research on Intelligent Electric Power Monitoring System Based on Internet of Things Technology*. Jilin, China: Jilin university. [master's thesis].

Li, Q. H., Zhang, Y. J., Chen, J. Q., Yi, Y. Q., and He, F. L. (2020). Development Patterns and Challenges of Ubiquitous Power Internet of Things. *Automation Electirc Power Syst.* 44 (1), 13. doi:10.7500/AEPS20190726001

Li, Z. Z., and Liu, J. X. (2017). Application of mobile Edge Computing in Internet of Vehicles. *Mod. Sci. &Technology Telecommunications* 47 (3), 37

Li, B., Jia, B. C., Chao, W. Z., Tian, S. M., QiSun, B. Y., Zhu, W. Y., et al. (2018). Prospect of Application of Edge Computing in the Field of Supply and Demand. *Power Syst. Technol.* 51 (11), 154. doi:10.11930/j.issn.1004-9649.201712051

Li, X., Li, D., Wan, J., Liu, C., and Imran, M. (2018). Adaptive Transmission Optimization in SDN-Based Industrial Internet of Things with Edge Computing. *IEEE Internet Things J.* 5 (3), 1351–1360. doi:10.1109/jiot.2018.2797187

Li, Z. S., Xie, R. C., Sun, L., and Huang, T. (2018). *A Survey of mobile Edge Computing Telecommunication Science(in Chinese)*, 34.

Liang, H., Ye, C., Zhou, Y., and Yang, H. (2021). *Anomaly Detection Based on Edge ComputingFramework for AMI[Conference Presentation]*. China: IEEE International conference on Electrical Engineering and Mechatronics TechnologyQingdao, 385–390. doi:10.1109/ICEEMT52412.2021.9601888

Lin, C.-C., Deng, D.-J., Chih, Y.-L., and Chiu, H.-T. (2019). Smart Manufacturing Scheduling with Edge Computing Using Multiclass Deep Q Network. *IEEE Trans. Ind. Inf.* 15 (7), 4276–4284. doi:10.1109/tii.2019.2908210

Liu, D., Zeng, X., and Wang, Y. (2019). *Edge-Computing-Driven Autonomous Ubiquitous Internet of Things in Electricity: Architecture and Challenges [Conference Presentation]*. Changsha, China: The 3rd IEEE conference on Energy Internet and Energy System Integration. doi:10.1109/EI247390.2019.9062125

Long, Z., Wen, B., Chen, X., Wu, J., and Zhang, J. (2020). *Demand-side Ubiquitous Electric Power Internet of Things: Architecture, Functionalities and Technologies[Conference Presentation]. 2020 Chinese Control and Decision Conference (CCDC 2020)*. Hefei, China. doi:10.1109/CCDC49329.2020.9164391

Luo, Y., Zeng, M., and Jiang, H. (2019). Learning to Tradeoff between Energy Efficiency and Delay in Energy Harvesting-Powered D2D Communication: A Distributed Experience-Sharing Algorithm. *IEEE Internet Things J.* 6 (3), 5585–5594. doi:10.1109/jiot.2019.2903873

Maier, M., and Ebrahimzadeh, A. (2019). Towards Immersive Tactile Internet Experiences: Low-Latency WiFi Enhanced mobile Networks with Edge Intelligence. *J. Opt. Commun. Networking* 11 (4), B10–B25. doi:10.1364/jocn.11.000b10

Mao, Z. F., Wang, X. C., Jain, D. K., Khan, H., Hongmin, G., and Zhen, W. (2020). A Blockchain-Based Trusted Data Management Scheme in Edge Computing. *IEEE Transaction Ind. Inf.* 16 (3), 2013. doi:10.1109/TII.2019.2933482

Mao, Z. Q. (2020). *Research on Smart Energy Distribution Optimization Scheme Based on Edge Computing*. [Chengdu(China)]: University of electronic science and technology of China. [Master thesis].

Min, M., Xiao, L., Chen, Y., Cheng, P., Wu, D., and Zhuang, W. (2019). Learning-based Computation Offloading for IoT Devices with Energy Harvesting. *IEEE Trans. Veh. Technol.* 68 (2), 1930–1941. doi:10.1109/tvt.2018.2890685

Nayak, N. G., Durr, F., and Rothermel, K. (2018). Incremental Flow Scheduling and Routing in Time-Sensitive Software-Defined Networks. *IEEE Trans. Ind. Inf.* 14 (5), 2066–2075. doi:10.1109/tii.2017.2782235

Okay, F. Y., and Ozdemir, S. (2016). Tunisia, 1–6. doi:10.1109/ISNCC.2016.7746062A Fog Computing Based Smart Grid Model2016 International Symposium on Networks Computers and Communications(ISNCC)Yasmine Hammamet

Omoniwa, B., Hussain, R., Javed, M. A., Bouk, S. H., and Malik, S. A. (2019). Fog/edge Computing-Based IoT (FECIoT): Architecture, Applications, and Research Issues. *IEEE Internet Things J.* 6 (3), 4118–4149. doi:10.1109/jiot.2018.2875544

Pan, J., and McElhannon, J. (2018). Future Edge Cloud and Edge Computing for Internet of Things Applications. *IEEE Internet Things J.* 5 (1), 439–449. doi:10.1109/jiot.2017.2767608

Pan, X., Jiang, A., and Wang, H. (2020). Edge-cloud Computing Application, Architecture, and Challenges in Ubiquitous Power Internet of Things Demand Response. *J. Renew. Sust. Energ.* 12, 062702. Available. doi:10.1063/5.0014059

Porambage, P., Okwuibe, J., Liyanage, M., Ylianttila, M., and Taleb, T. (2018). Survey on Multi-Access Edge Computing for Internet of Things Realization. *IEEE Commun. Surv. Tutorials* 20 (4), 2961–2991. doi:10.1109/COMST.2018.2849509

Qiu, T., Chi, J., Zhou, X., Ning, Z., Atiquzzaman, M., and Wu, D. O. (2020). Edge Computing in Industrial Internet of Things: Architecture, Advances and Challenges. *IEEE Commun. Surv. Tutorials* 22 (4), 2462–2488. doi:10.1109/COMST.2020.3009103

Rahmani, A. M., Gia, T. N., Negash, B., Anzanpour, A., Azimi, I., Jiang, M., et al. (2018). Exploiting Smart E-Health Gateways at the Edge of Healthcare Internet-Of-Things: A Fog Computing Approach. *Future Generation Comput. Syst.* 78, 641–658. doi:10.1016/j.future.2017.02.014

Shang, L., Cai., S,, Cui, J. B., Ji, C. H., Cui, K. J., and Li, B. G. (2021). SDN-based MEC Resource Allocation of a Power Grid. *Power Syst. Prot. Control(in Chinese)* 49 (20), 136. doi:10.19783/j.cnki.pspc.201623

Sharma, S. K., and Wang, X. (2017). Live Data Analytics with Collaborative Edge and Cloud Processing in Wireless IoT Networks. *IEEE Access* 5, 4621–4635. doi:10.1109/access.2017.2682640

Shi, W., Cao, J., Zhang, Q., Li, Y., and Xu, L. (2016). Edge Computing: Vision and Challenges. *IEEE Internet Things J.* 3 (5), 637–646. doi:10.1109/jiot.2016.2579198

Song, C., Zhang, M., Zhan, Y. Y., Wang, D. S., Guan, L. Y., Liu, W., et al. (2019). Hierarchical Edge Cloud Enabling Network Slicing for 5G Optical Fronthaul. *J. Opt. Commun. Networking* 11 (4), B60–B70. doi:10.1364/jocn.11.000b60

Stantchev, V., Barnawi, A., Ghulam, S., Schubert, J., and Tamm, G. (2015). Smart Items, Fog and Cloud Computing as Enablers of Servitization in Healthcare. *Sensors Transducers* 185 (2), 121–128.

State Grid Corporation of China (2019b). *Full Deploys of the Construction of Ubiquitous Power Internet of Things*. http://h5ip.cn/ozIF (Accessed April 16, 2019).

State Grid Corporation of China (2019a). *The Outline of the Construction of the Ubiquitous Power Internet of Things Official Release*. http://www.chinasmartgrid.com.cn/news/20190311/632172.shtml (Accessed March 11, 2019).

State Grid Corporation of China (2020). *Ubiquitous Power Internet of Things Key Construction Task Outline in 2020*. https://shupeidian.bjx.com.cn/html/20191211/1027517.shtml (Accessed December 2, 2019).

Sun, W. W., He, G. Q., Liu, C., Zhao, W. R., and Qi, L. W. (2021). Research and Application of Key Technologies for Edge Computing Equipment Used in Photovoltaic Power Generation under Background of IoT. *Electric Power Automation Equipment* 41 (7), 38–43.

Sun, H. Y., Zang, J. C., Wang, P., Lin, J. Y., Guo, C., and Chen, L. (2019). Edge Computation Technology Based on Distribution Internet of Things. *Power Syst. Technol.* 43 (12), 4314. doi:10.13335/j.1000-3673.pst.2019.1750

Sun, W., Liu, J., and Yue, Y. (2019). AI-enhanced Offloading in Edge Computing: When Machine Learning Meets Industrial IoT. *IEEE Netw.* 33 (5), 68–74. doi:10.1109/mnet.001.1800510

Tang, B., Li, T., Li, B., Chen, L. Z., Yang, M., and Liu, S. Y. (2021). Analysis and Display of Intelligent Deployment and Control Data in Substations Based on Edge Computing. *Eng. Des.* 40 (2), 79

Tao, H., Zhou, J., and Liu, S. (2017). A Survey of Network Security Situation Awareness in Power Monitoring System."in IEEE Conference on Energy Internet and Energy System Integration (EI2). Beijing,china. doi:10.1109/EI2.2017.8245487

Trimananda, R., Younis, A., Wang, B., Xu, B., Demsky, B., and Xu, G. Vigilia: Securing Smart home Edge Computing."in 2018 Third ACM/IEEE Symposium on Edge Computing. Seattle, WA, USA(2018), 74–89.

Varghese, B., Wang, N., Barbhuiya, S., Kilpatrick, P., and Nikolopoulos, D. S. (2016). *Challenges and Opportunities in Edge computing[Conference Presentation]*. New York, USA: IEEE International Conference Smart Cloud, 20–26. doi:10.1109/SmartCloud.2016.18

Wang, J., Wu, L. B., Wu, L., Choo, K.-K. R., and He, D. (2020). Blockchain-based Anonymous Authentication with Key Management for Smart Grid Edge Computing Infrastructure. *IEEE Trans. Ind. Inf*, 16, 1984–1992. doi:10.1109/tii.2019.2936278

Wang, J. Y., LiuWu, C. J. M., Zhang, X. J., and Yu, H. (2015). Research of Software Defined Service Resource Equitable Allocation Technology of Power WAN. *Power Syst. Technol.* 39 (5), 1425. doi:10.13335/j.1000-3673.pst.2015.05.038

Wang, Y., Chen, Q. X., Zhang, N., Feng, C., Teng, F., Sun, M. Y., et al. (2019). Fusion of the 5G Communication and the Ubiquitous Electric Internet of Things: Application Analysis and Research Prospects. *Power Syst. Technology(in Chinese)* 43 (5), 1575–1585. doi:10.13335/j.1000-3673.pst.2019.0635

Wang, Z. H. (2019). Researches and Applications of Safety Program for Power Internet of Things Relying on Cloud-Computing Services. *Tech. Automation Applications(in Chinese)* 38 (12), 93

Xu, C. K., Wu, C. H., Shu, J., Yang, Z. Y., and Zhang, H. R. (2020). An Edge Computing Architecture of Distributed Generation Microgrid and its Application. *Adv. New Renew. Energ.* 8 (5), 383. doi:10.3969/j.issn.2095-560X.2020.05.006

Xu, X., Li, D., Dai, Z., Li, S., and Chen, X. (2019). A Heuristic Offloading Method for Deep Learning Edge Services in 5G Networks. *IEEE Access* 7, 67734–67744. doi:10.1109/access.2019.2918585

Yang, Y., Lv, Y. X., and Liu, Z. W. (2017). Research on a Software-Defined Power Communication Network Service-Aware Traffic Dispatch Mechanism. *Automation and Instrumentation* 8, 146. doi:10.14016/j.cnki.1001-9227.2017.08.146

Yao, D., Wen, M., Liang, X., Fu, Z., Zhang, K., and Yang, B. (2019). Energy Theft Detection with Energy Privacy Preservation in the Smart Grid. *IEEE Internet Things J.* 6 (5), 7659–7669. doi:10.1109/jiot.2019.2903312

Yu, W., Liang, F., He, X., Hatcher, W. G., Lu, C., Lin, J., et al. (2017). A Survey on the Edge Computing for the Internet of Things. *IEEE Access* 6, 6900

Zahoor, S., Javaid, S., Javaid, N., Ashraf, M., Ishmanov, F., and Afzal, M. (2018). Cloud–fog–based Smart Grid Model for Efficient Resource Management. *Sustainability* 10 (6), 1154–1160. doi:10.3390/su10062079

Zhang, J. L., Zhao, Y. C., Chen, B., Hu, F., and Zhu, K. (2018). Survey on Data Security and Privacy-Preserving for the Research of Edge Computing. *J. Commun.* 39 (3), 1

Zhang, K., Leng, S., Peng, X., Pan, L., Maharjan, S., and Zhang, Y. (2019a). Artificial Intelligence Inspired Transmission Scheduling in Cognitive Vehicular Communications and Networks. *IEEE Internet Things J.* 6 (2), 1987–1997. doi:10.1109/jiot.2018.2872013

Zhang, K., Zhu, Y., Leng, S., He, Y., Maharjan, S., and Zhang, Y. (2019b). Deep Learning Empowered Task Offloading for mobile Edge Computing in Urban Informatics. *IEEE Internet Things J.* 6 (5), 7635–7647. doi:10.1109/jiot.2019.2903191

Zhao, J. J., Yang, R. Y., Zhang, X. Yu. D. P., Zhang, X. T., and Yuan, X. D. (2021). Design and Analysis of the Special AI Video Identification Chip for the Power Industry in Edge Computing of Smart Coal-Fired Power Plants. *Energ. Sci. Technol.* 19 (05), 41

Zhong, Y., LongZhao, Y. J. W. Y., and Zhang, G. Y. (2021). *Power Cloud Data center Flow Control Technology Based on Software-Defined networkScience and Technology Innovation*, 15, 80.

Conflict of Interest: ML was employed by Shenzhen Power Supply Bureau Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Energy Research

METHODS
published: 25 April 2022
doi: 10.3389/fenrg.2022.891680

Short-Term Probability Density Function Forecasting of Industrial Loads Based on ConvLSTM-MDN

Yuan Y. Wang[1], Ting Y. Wang[1], Xiao Q. Chen[2], Xiang J. Zeng[1], Jing J. Huang[1] and Xia F. Tang[1]*

[1]Changsha University of Science and Technology, Changsha, China, [2]California Institute of Technology, Pasadena, CA, United States

Load forecasting for industrial customers is essential for reliable operation decisions in the electric power industry. However, most of the load forecasting literature has been focused on deterministic load forecasting (DLF) without considering information on the uncertainty of industrial load. This article proposes a probabilistic density load forecasting model comprising convolutional long short-term memory (ConvLSTM) and a mixture density network (MDN). First, a sliding window strategy is adopted to convert one-dimensional (1D) data into two-dimensional (2D) matrices to reconstruct input features. Then the ConvLSTM is utilized to capture the deep information of the input features. At last, the mixture density network capable of directly predicting probability density functions of loads is adopted. Experimental results on the load datasets of three different industries show the accuracy and reliability of the proposed method.

Keywords: load forecasting, probability density, convolutional long short-term memory, mixture density network, industrial customers

OPEN ACCESS

Edited by:
Linfeng Yang,
Guangxi University, China

Reviewed by:
Sumei Liu,
Beijing Forestry University, China
Hui Hou,
Wuhan University of Technology, China

****Correspondence:***
Xiao Q. Chen
xqchen@caltech.edu

Specialty section:
This article was submitted to Smart Grids, a section of the journal Frontiers in Energy Research

Received: *08 March 2022*
Accepted: *21 March 2022*
Published: *25 April 2022*

Citation:
Wang YY, Wang TY, Chen XQ, Zeng XJ, Huang JJ and Tang XF (2022) Short-Term Probability Density Function Forecasting of Industrial Loads Based on ConvLSTM-MDN. Front. Energy Res. 10:891680. doi: 10.3389/fenrg.2022.891680

1 INTRODUCTION

The improvement of the demand-side electrical energy management is of critical importance to reliable and economical operation of the modern power system (Wang et al., 2021a). Accurate short-term load forecasting (STLF) can help the department of demand-side management to understand and analyze electricity consumption behavior and further make intelligent control strategy to strengthen energy management. In many developing countries, electricity consumption by industrial customers is the major part of total electricity consumption on the demand side (Tan et al., 2020). For example, in China, about 67% of electrical energy is consumed by industrial customers (National Bureau of Statistics of the People's Republic of China, 2021). However, the complex electricity tariff rules (Wang et al., 2020) and the high uncertainty of industrial loads make it difficult for industrial customers to make a correct electricity strategy, which leads to excessively high electricity costs and non-essential losses. To solve the aforementioned problems, industrial customers can adjust production planning in advance to improve energy efficiency and economic benefits. Therefore, high-accuracy STLF for industrial customers is urgently needed.

As a typical time series forecasting problem, many STLF methods have been a hot topic in academia and industry (Cai et al., 2017; Hou et al., 2020a; Hou et al., 2020b; Cai et al., 2021; Hou et al., 2021), which can be roughly categorized into statistical methods and artificial intelligence methods (Kuster et al., 2017). Among them, the statistical methods (Zhao and Li, 2021; López et al., 2019) are difficult to handle load time series with high randomness and non-linearity (Wang et al., 2021b) and usually result in low forecasting accuracy. The artificial intelligence methods can be further divided into the shallow machine learning

(Wang et al., 2021c) and deep learning (Ruan et al., 2021). These methods have powerful non-linear processing capabilities, which address the drawback of statistical methods. However, shallow structure-based methods need additional feature extraction and selection due to their poor performance in feature mining, so they are not suitable to be implemented in different datasets (Afrasiabi et al., 2020). In addition, the depth limitation of shallow machine learning also restricts the forecasting accuracy.

Deep learning models can capture deep features from historical load data through multi-layer non-linear mapping and can handle various relevant factors. Jiao et al. (2018) used the long short-term memory (LSTM) network to predict the load of non-residential customers, which brings a significant improvement compared with several shallow machine learning models. A single model suffers the limitations of the algorithm and some accidental factors, resulting in poor generalization performance (Fallah et al., 2019). Hence, Farsi et al. (2021) adopted the combined model of CNN-LSTM, which can comprehensively utilize the information provided by each model to improve the forecasting accuracy.

However, commonly used two-dimensional (2D) CNNs are not suitable for one-dimensional (1D) time series data, while using 1D CNN to learn time series faces the problem of overfitting, unless increasing the number of CNN layers. Therefore, applying CNN to the time series forecasting problem is suboptimal (Essien and Giannetti, 2020). The convolutional LSTM (ConvLSTM) proposed by Shi et al. (2015) has both powerful feature extraction capabilities of CNN and excellent time sequence processing capability of LSTM, so it can not only capture features but also perform well in sequential learning. Essien and Giannetti (2020) established a deep ConvLSTM encoder–decoder architecture for multistep machine speed prediction. Experimental results show that the proposed method has higher test accuracy (root mean square error (RMSE) ranges from 64.23 to 64.93) than the deep LSTM and the CNN-LSTM encoder–decoder models. In addition, ConvLSTM has been successfully applied to time series forecasting problems such as wind power forecasting (Sun and Zhao, 2020) and solar irradiation forecasting (Hong et al., 2020). All the studies described before prove that ConvLSTM has a significant performance in forecasting time series data. Hence, applying ConvLSTM to STLF is expected to improve the probability forecasting accuracy of industrial customers.

The aforementioned approaches are implemented as point forecasts, which only provide the future point value without information about the associated uncertainty. To measure the uncertainty of load and accommodate the risk brought by the uncertainty of load, probabilistic load forecasting (PLF) gets more attention in industrial applications (Zhang et al., 2019a). The existing PLF methods can be divided into prediction intervals (PIs), quantile prediction, and probabilistic density function (PDF) forecasting according to the output form, and they provide the statistical information of the future load. Among all methods, PDF forecasting can fully reflect distribution information of future load data, which provides far more information than other forms of PLF (Xie et al., 2019), (Zhang et al., 2020). Therefore, PDF forecasting is an essential tool to quantify uncertainty in load forecasting.

On PDF forecasting, He et al. (2017) used kernel-based support vector quantile regression to generate complete probability distribution of future values and then predicted PDFs according to copula theory. He et al. (2019) developed the least absolute shrinkage and selection operator-quantile regression neural network (LASSO-QRNN) for electricity consumption forecasting. As mentioned before, many PDF forecasting methods focus on indirectly predicting PDF in current research, but the forecasting errors of indirect forecasting models grow with each iteration, resulting in low forecasting accuracy (Afrasiabi et al., 2021). It is necessary to research the method of directly forecasting PDF. In Zhang et al. (2020), an improved deep mixture density network (MDN) was built to predict wind power of multiple wind farms, and then a laconic and accurate PDF at each time step was produced. To enhance the learning ability of MDN, He et al. (2019) combined the deep learning approach and MDN to characterize PDF of wind speed. This method can directly construct PDFs by processing raw data and enhance forecasting accuracy and computational efficiency. Afrasiabi et al. (2020) also merged the deep learning model into MDN to directly predict PDFs of residential loads. In case studies, the accuracy rates of median prediction were 10.024 and 6.694% in terms of mean absolute percentage error (MAPE), respectively, which demonstrated the effectiveness of the deep mixture model.

A critical issue is that although PDF forecasting techniques based on MDN have been applied to wind power probabilistic forecasting and residential load probabilistic forecasting, none of the methodologies proposed so far are looking into industrial load forecasting. The amount of literature on PLF of industrial loads is quite limited. Berk et al. (2018) proposed an inhomogeneous Markov switching method to achieve PLF of industrial customers. Da Silva et al. (2019) combined the bottom-up approach with hierarchical linear models for PLF in the industrial sector. Due to the continuous development of the industry and the increasing variability of customers' activities (Wang et al., 2021a), PLF that can predict uncertain information is more suitable for industrial load forecasting. Based on the aforementioned analysis, we merged the deep learning model into MDN as a solution for industrial load forecasting.

With the aim of directly learning the severe uncertainty of industrial loads and providing accurate load forecasting results, we developed a new deep mixture model based on ConvLSTM and MDN. The model exploits the strengths of ConvLSTM in feature extraction and sequence learning to learn deep features of load data. ConvLSTM and MDN are combined by using a dense layer to directly predict PDF. The main contributions of this study are described as follows:

(1) This study introduces an emerging deep learning model into the field of industrial load forecasting, namely, ConvLSTM. Meanwhile, in order to make full use of the load and various types of data related to the load, a simple input construction method for ConvLSTM is proposed. ConvLSTM can extract key features of these data well to improve the performance of probabilistic forecast accuracy.
(2) The probability density function forecasting is new to the industrial load forecasting literature. We built a novel mixture model combining ConvLSTM and MDN. The model aims to acquire full statistical information about future industrial load consumption in the form of PDF.

The proposed method can predict industrial loads with strong non-linear relationship, high variability, and severe uncertainty.

(3) Comprehensive case studies are conducted on load datasets of different industrial customers and compared with multiple state-of-art models. Experimental tests results show that the proposed model has stronger robustness, better generalization performance, and higher forecasting accuracy. For instance, ConvLSTM improves the accuracy of LSTM by more than 20%.

The rest of the study is organized as follows: **Section 2** presents basic knowledge about CNN, LSTM, convolutional LSTM, and MDN. **Section 3** analyzes the relevant characteristics of the load, and the proposed ConvLSTM-MDN model and methodological approach are introduced in **Section 4**. Numerical simulations results are reported and discussed in **Section 5**. Finally, **Section 6** concludes the study.

2 TECHNICAL PRELIMINARIES

2.1 CNN

CNN is a deep neural network with convolution operation, which can extract features among input data with two advantages: local perception and weight sharing. Therefore, CNN has less number of parameters than ordinary neural networks. The typical CNN consists of convolutional layers, pooling layers, and fully connected layers. Convolutional layers employ a set of learnable kernels to perform the convolution operation on input data, in order to extract features or patterns from inputs. Pooling layers can shrink the parameter dimensions and control overfitting. Fully connected layers are put at the end of a sequence of the layers, which can summarize features extracted by previous layers to generate outputs.

2.2 LSTM

As a special variant of recurrent neural networks (RNNs), LSTM can effectively surmount the problems of gradient vanishing and gradient exploding when RNN learns long-term temporal correlations. Based on the architecture of RNNs, memory cell and three control gates are included in the architecture of the LSTM to control information flow. The memory cell can accumulate the state information and remain unchanged. Three control gates, namely, input gate, output gate, and forget gate, are used to record new information selectively and clear previous information selectively, thus solving the long-term dependence problem in sequence learning.

2.3 ConvLSTM

In order to satisfy the requirements of different tasks, various modified versions are developed from LSTM. On the basis of the fully connected LSTM (FC-LSTM) network, Shi *et al.* replaced the FC layer operators in the state-to-state and input-to-state transitions with convolution operators to obtain ConvLSTM models with the complementary strengths of LSTM and CNN models. Therefore, the network topology of ConvLSTM enables it to perform convolution operation on multidimensional data to capture the spatial and temporal features rather than just temporal features. **Figure 1** illustrates the inner structure of the ConvLSTM. Similarly, the ConvLSTM also uses the forget gate to decide which information is to be "remembered" or "forgotten." Different from LSTM, the input matrix x_t of ConvLSTM is fed as image (i.e., 2D or 3D matrix). In the ConvLSTM, the future cell state is determined by the input at the current time step, output at the previous time step, and cell state at the previous time step. The key formulas of the overall ConvLSTM connections are shown in **(1) to (5)**:

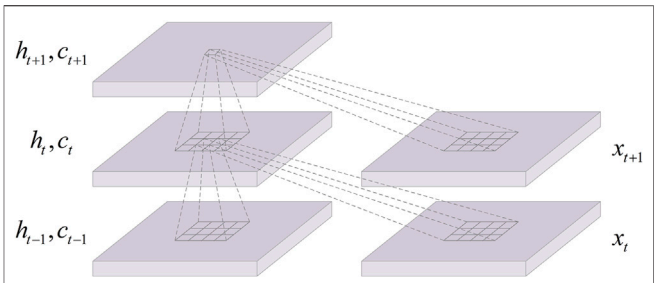


FIGURE 1 | Inner structure of ConvLSTM.

$$f_t = \sigma\left(w_{xf} * x_t + w_{hf} * h_{t-1} + w_{cf} \circ c_{t-1} + b_f\right), \quad (1)$$

$$i_t = \sigma\left(w_{xi} * x_t + w_{hi} * h_{t-1} + w_{ci} \circ c_{t-1} + b_i\right), \quad (2)$$

$$o_t = \sigma\left(w_{xo} * x_t + w_{ho} * h_{t-1} + w_{co} \circ c_t + b_o\right), \quad (3)$$

$$c_t = f_t \circ c_{t-1} + i_t \circ \tanh\left(w_{xc} * x_t + w_{hc} * h_{t-1} + b_c\right), \quad (4)$$

$$h_t = o_t \circ \tanh\left(c_t\right), \quad (5)$$

where t is the time step; f, i, o, and c represent forget gate, input gate, output gate, and cell state, respectively; the variables x, h, w, and b are input vector, output vector, weight matrix, and bias vector, respectively; σ denotes the sigmoid activation function; tanh denotes the hyperbolic tangent activation function; $\circ$ denotes the Hadamard product; and $*$ denotes the convolution operator.

2.4 MDN

The MDN can predict PDFs of the target variables, which was first introduced by Christopher M. Bishop in the 1990s. The structure of MDN is composed of a Gaussian mixture model and a feed-forward network. MDN uses Gaussian function as the basic component and superposes a sufficient number of Gaussian functions in a certain proportion to fit the final PDF. Gaussian function enables the MDN to flexibly and accurately represent arbitrary probability distributions (Zhang et al., 2019b). The output variables of model are used to construct final PDFs, which include mean, standard deviation, and proportion of Gaussian distribution. Theoretically, when the mixing coefficients and Gaussian parameters are correctly chosen, MDN can approximate any PDF (Bishop, 1994).

For any given value of a, the Gaussian mixture model provides a general form that approximates any conditional density function $p(y_t|a)$. The equation is as follows:

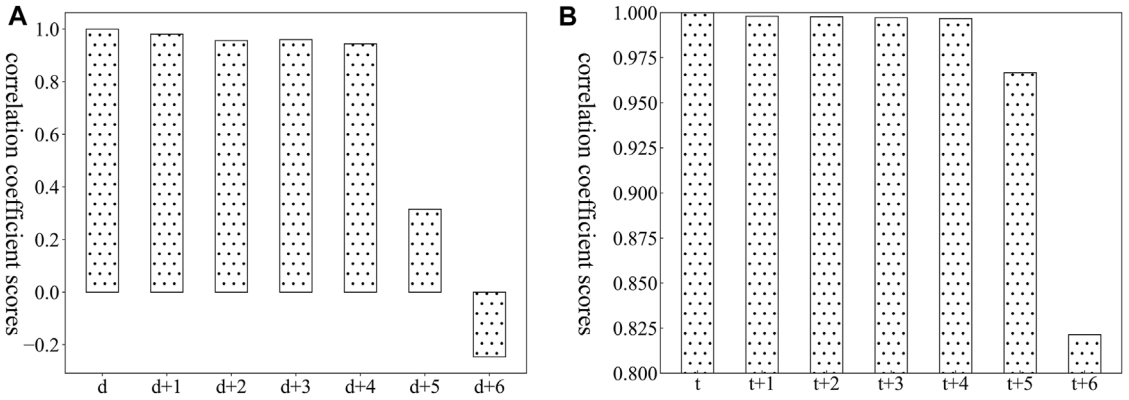


FIGURE 2 | Load correlation of adjacent time to *d* (day) and *t* (time). **(A)** shows the Pearson correlation analysis results of adjacent daily loads, and **(B)** shows the Pearson correlation analysis results of the load at adjacent times.

$$p(y_t|a) = \sum_{k=1}^{K} \pi_k(a,t)\varphi\left(y_t|\mu_k(a,t), \nu_k(a,t)\right), \quad \forall t \in \{1,\ldots,T\}, \forall k \in \{1,\ldots,K\}, \tag{6}$$

where K is the number of components in the mixture model; $\pi_k(a,t)$, $\mu_k(a,t)$, and $\nu_k(a,t)$ represent mixing coefficient, mean, and standard deviation of the *kth* component, respectively. It should be noted that the sum of mixing coefficients must be 1, which can be achieved by controlling the output through the softmax function. The softmax function is given by

$$\sum_{k=1}^{K} \pi_k(a,t) = 1, \quad 0 \le \pi_k(a) \le 1, \tag{7}$$

$$\alpha_k(a,t) = \frac{\exp(\pi_k(a,t))}{\sum_{l=1}^{K}\exp(\pi_k(a,t))}. \tag{8}$$

Similarly, to ensure that the variance is greater than or equal to 0, $\nu_k(a)$ can be reformulated as follows:

$$\beta_k(a,t) = \exp(\nu_k(a,t)). \tag{9}$$

In order to control the output value of the MDN model within reasonable bounds, the modified exponential linear unit (ELU) activation function can be used as follows:

$$g(a) = \mathrm{elu}(a) + 1. \tag{10}$$

The loss function of standard MDN is maximum likelihood method, which may lead the loss function to NaN value. The reason is that the function approaches infinitesimal as the input value approaches 0. To mitigate the possibility of NaN value, this study employs continuous ranked probability score (CRPS) as the loss function. CRPS is computed as the integral of the square, which avoids infinitesimal or infinite situations.

3 ANALYSIS OF INDUSTRIAL LOAD TEMPORAL RELEVANCY AND RELEVANT FACTORS

3.1 Industrial Load Temporal Relevancy

Industrial load is time series data, so an essential element of it is time. Temporal relevancy is important information that cannot be ignored. In this study, some industrial load data is randomly selected to perform Pearson correlation analysis on the load at adjacent times and the adjacent daily load at the same time. The analysis results are shown in **Figure 2**. It can be seen that the degree of correlation between loads in both cases tends to weaken with the increase of the time interval, that is, the load temporal relevancy gradually weakens. Therefore, it is important to select historical load data in a suitable time range as the input features of model.

3.2 Industrial Load Relevant Factors

This study considers two types of relevant factors: temperature and calendar information. First, the influence of temperature on load is analyzed by Pearson correlation coefficient. If the correlation coefficient score is greater than 0.6, temperature is selected as the input feature of the model. Taking the load data of an industrial customer in 2018 as an example, the correlation coefficient score is 0.827, which indicates a strong correlation. Meanwhile, **Figure 3** presents the load and temperature profiles. It is obvious that the load increases with decreasing or increasing temperature in winter and summer.

Calendar information includes working days, weekends and holidays. Industrial loads relate to production plan and activities of workers. Due to the work schedule, the load from Monday to Friday and the load on the weekend are significantly different. In addition, holidays are also an important factor. For example, in China, during important traditional festivals such as the Spring Festival, the National Day, and the Mid-Autumn Festival, employees of most industrial enterprises rest. Therefore, we cannot ignore the effect of calendar information on the load.

4 THE LOAD FORECASTING FRAMEWORK BASED ON CONVLSTM-MDN MODEL

This study proposes a deep model for probability density forecasting of industrial load, based on ConvLSTM-MDN. The framework of the proposed method is established by three steps, including data pro-processing, 1D time series to 2D matrices transforming and probability density forecasting of industrial

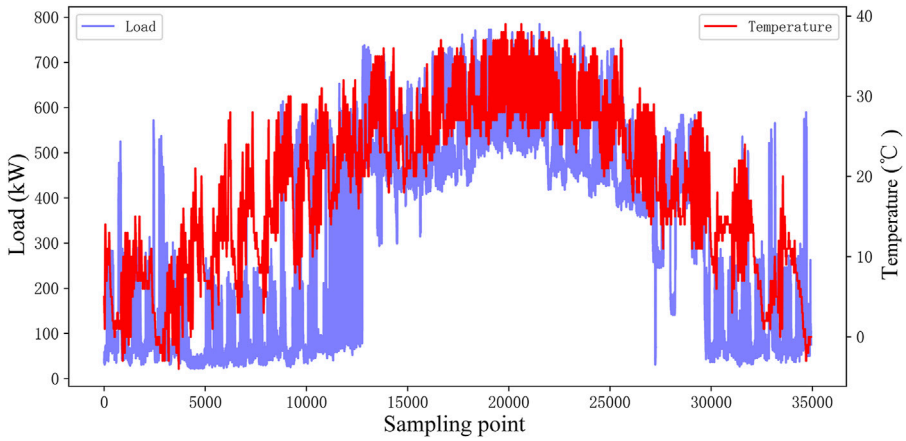


FIGURE 3 | Load and temperature profiles.

load. Each step in the framework is introduced in the following sections.

4.1 Data Pro-Processing

For load data of industrial customers, the main reasons for missing data include acquisition equipment failures and signal transmission interruption. In order to prevent missing data from destroying the continuity of load data, the following method is adopted: when the proportion of missing data on a certain day is low, linear interpolation is employed to process these data. Instead, all data for the day are deleted.

Then, in order to accelerate convergence speed of the model, we use the min-max normalization (Farsi et al., 2021) method to scale load data and temperature data to the range (0,1). Because the calendar information is discrete data, we adopt One-Hot encoding method to convert it into a form that can be processed by deep learning algorithms. This study marks Monday–Friday as 0 and weekends as 1. According to the actual holiday date of industrial customer, the holiday is marked as 1, and other days are marked as 0.

4.2 1D Time Series to 2D Matrices Transforming

Although the ConvLSTM algorithm takes into account the advantages of the CNN in feature extraction and the LSTM in sequential learning, the use of ConvLSTM in industrial load forecasting will face the problem of data dimension mismatch, that is, the structure of standard ConvLSTM is not suitable for directly processing load data which is 1D time series data. To tackle this problem, we adopted a method to convert the 1D data into 2D matrices that can be processed by ConvLSTM.

The load dataset is denoted as $D = \{l_1, l_2, \ldots, l_S\}$ with K instances per day; the temperature dataset, date dataset, and holiday dataset are denoted as $T = \{T_1, T_2, \ldots, T_S\}$, $W = \{W_1, W_2, \ldots, W_S\}$, and $H = \{H_1, H_2, \ldots, H_S\}$, respectively. We reconstruct these time series into a series of 2D matrices $[N \times M]$, where N is manually set, M is equal to $(K+3)$, and they are integers and greater than 2. The reconstruction method is as follows.

First, construct the first matrix graph: The l_1 to l_K data on the first day are the first row of the matrix. The l_{K+1} to l_{2K} data of the second day are the second row of the matrix. N days of data are selected to construct the 2D matrix $[N \times K]$. Then, temperature, date type, and holiday corresponding to the moment of the first column load are added to the matrix in order to obtain the first 2D matrix $[N \times M]$. This matrix can be used to predict $l_{N \times k+1}$ data. The first 2D matrix is as follows:

$$\begin{Bmatrix} l_1 & \cdots & l_K & T_1 & W_1 & H_1 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ l_{(N-1)\times K+1} & \cdots & l_{N\times K} & T_{(N-1)\times K+1} & W_{(N-1)\times K+1} & H_{(N-1)\times K+1} \end{Bmatrix}. \tag{11}$$

Second, other matrix graphs are constructed: a fixed-length sliding window method of 1D time series load dataset (with a length of $N \times K$ and one step size) is used to capture the other load matrix. That is, the l_2 to l_{K+1} data are the first row of the second matrix. The l_{K+2} to l_{2K+1} data are the second row of the second 2D matrix. N rows of data are selected to construct the second load matrix. The second matrix is also obtained by adding temperature, date type, and holidays to the load matrix in order. The second matrix can be used to predict $l_{N \times K+2}$ data.

Finally, following the fixed-length sliding window method described before, the last 2D matrix is obtained. A total of $(S-N \times K)$ 2D matrices can be obtained. **Figure 4** shows the diagram of 2D matrix conversion.

4.3 Probability Density Forecasting of Industrial Load

1) Hyper-Parameters Optimization

Hyper-parameters are important factors that directly influence the prediction accuracy of models. In this study, the prediction accuracy is largely related to ConvLSTM layer parameters

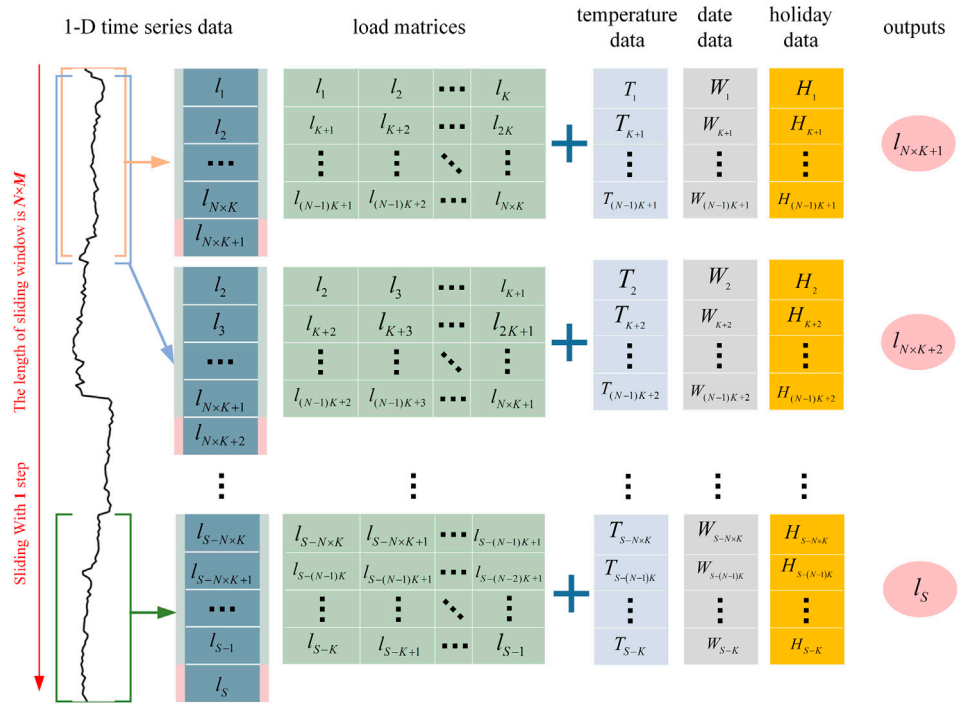


FIGURE 4 | Diagram of 1D time series to 2D matrix transforming.

TABLE 1 | Hyper-parameters of ConvLSTM layer and MDN.

Layer	Parameter(s)	*N* of 2D matrix	Value
ConvLSTM	Filters		15, 30, 45, 60, 75, and 90
	Kernel size	2	(2×2)
		3	(2×2), (3×3)
		⋮	⋮
		d	(2×2), (3×3), . . ., (*d*×*d*)
MDN	Number of components	—	3, 4, 5, 6, 7, 8, 9, and 10

(number of filters, kernel size) and MDN parameters (the number of components). Too many filters will increase model training time and result in overfitting. On the contrary, models may not get good accuracy results. In addition, a larger kernel size can capture better global features, but the amount of calculation will slow down training process. A smaller kernel size can improve learning speed, but may not capture features well. Similarly, appropriate number of components can better fit PDFs.

Therefore, we applied the grid search method for the hyper-parameter optimization. The grid search method loops through all selectable candidates to find the optimal parameters set. This method is simple to implement and has great versatility. It is described as follows: Given the possible values of three hyper-parameters, then find the optimal set of hyper-parameters that minimizes the validation loss. **Table 1** shows the search space of the three hyper-parameters in ConvLSTM layer and MDN. The other benchmark models included in this study use the similar approach to optimize the model hyper-parameters.

2) ConvLSTM-MDN Model

The converted 2D matrices are used as the input data of the model. The input set is a dimensional tensor with (Y, 1, N, M, 1) size, where Y is the number of samples. The model designed in this study is composed of ConvLSTM layer, flatten layer, dense layer, and MDN. The number of three layers is 1. The activation function of ConvLSTM layer is a rectified linear unit (ReLU). The flatten layer is a transition layer which can flatten the multidimensional array into a linear vector. The dense layer is used to extract the association between the previous features, and the activation function of dense layer is linear function. Afterward, the hidden layers mentioned earlier are merged into MDN to output the approximated parameters (m mean values, m standard deviation, and m mixing coefficients) in parallelized manner, where m is the number of components in the MDN model. The future PDFs are obtained according to the approximated parameters. Furthermore, the loss function of the model is CRPS.

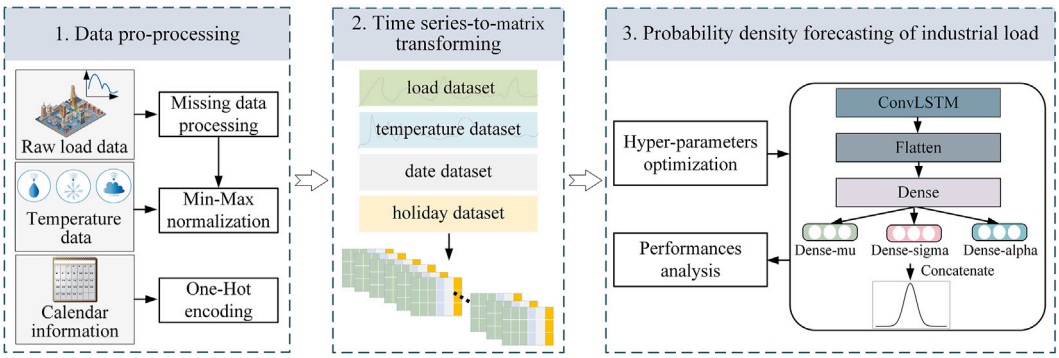


FIGURE 5 | Load forecasting framework based on the ConvLSTM-MDN model.

TABLE 2 | Hyper-parameters of the contrast models.

Model	Hyper-parameter(s)	Value
ARIMA-MDN	(p, d, q)	(2, 0, 1)
SVR-MDN	Kernel function	RBF
	Penalty coefficient	1
RF-MDN	Trees	80
FFNN-MDN	FFNN layer, units	1, 120
LSTM-MDN	LSTM layer, units	1, 120
2D-CNN-MDN	Convolutional layer	1
	Filters, kernel size	60, (2×3)
	Max-pooling layer	1
2D-CNN-LSTM-MDN	Convolutional layer	1
	Filters, kernel size	45, (2×3)
	Max-pooling layer	1
	LSTM layer, units	1, 60

TABLE 3 | Forecasting results of each *N*.

N of 2D matrix	Evaluation metrics		
	Median		CRPS (%)
	RMSE (kW)	MAPE (%)	
2	19.717	2.813	13.475
3	**19.052**	**2.605**	**12.774**
4	19.094	2.708	13.192
5	21.795	3.016	15.385
6	20.555	2.813	14.086
7	24.123	3.802	18.519

The bold values indicate the minimum value of each evaluation metrics, that is, the best forecasting effect.

3) Evaluation Metrics

To evaluate the deterministic forecasting performance of the proposed model, two commonly used evaluation metrics are adopted in this work, which are RMSE and MAPE. RMSE can measure the deviation between the forecasted and the actual value. But it is sensitive to data that fluctuates greatly in short time. MAPE measures the accuracy by calculating the relative error between the forecasted value and the actual value, which can solve the problem of RMSE. If the actual value is zero, MAPE cannot be calculated. The advantages of two evaluation metrics can be leveraged. The aforementioned evaluation metrics are defined as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}, \tag{12}$$

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right|. \tag{13}$$

Furthermore, we select CRPS as probabilistic evaluation metrics to evaluate the performance of predicted PDF. CRPS is widely used in the field of probabilistic forecasting, which can comprehensively assess the calibration and sharpness of the forecasted PDF. CRPS is expressed as follows:

$$CRPS = \frac{1}{n}\sum_{i=1}^{n}\int_{-\infty}^{+\infty}\left[\widehat{F}(x) - F(x)\right]^2 dx. \tag{14}$$

The general research framework based on the ConvLSTM-MDN model is visualized in **Figure 5**.

5 CASE STUDY

The proposed model is tested on three different types of industrial customers to assess the feasibility of the probabilistic forecasting method in load forecasting for industrial customers. Two industrial datasets are collected from a nonferrous metal smelting industry and a medical industry in Hunan Province, China, with a temporal resolution is 15-min interval. Another dataset is retrieved from the Irish Smart Metering Electricity Customer Behaviour Trials (CBTs) (Commission for Energy Regulation (CER), 2012). Temperature data are acquired from the National Oceanic and Atmospheric Administration (NOAA) website. After converting the aforementioned data into 2D matrices by using the fixed-length sliding window method mentioned in **Subsection 4.2**, the input dataset of the model is obtained. Then, 70% of the input dataset is dedicated to training, 10% for validating, and 20% for testing.

For the sake of comparison, eight different models are integrated into the MDN to construct PDFs with the same dataset, including linear regression (LR), autoregressive integrated moving average model (ARIMA), SVR, random forest (RF), feedforward neural

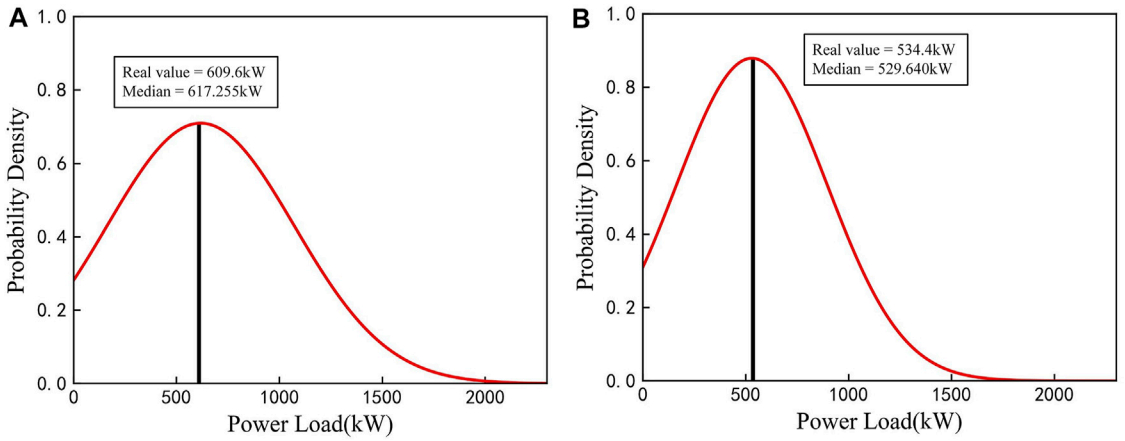


FIGURE 6 | ConvLSTM-MDN model predictive distribution and real values for a day at **(A)** 8:00 and **(B)** 16:00 in experiment 1.

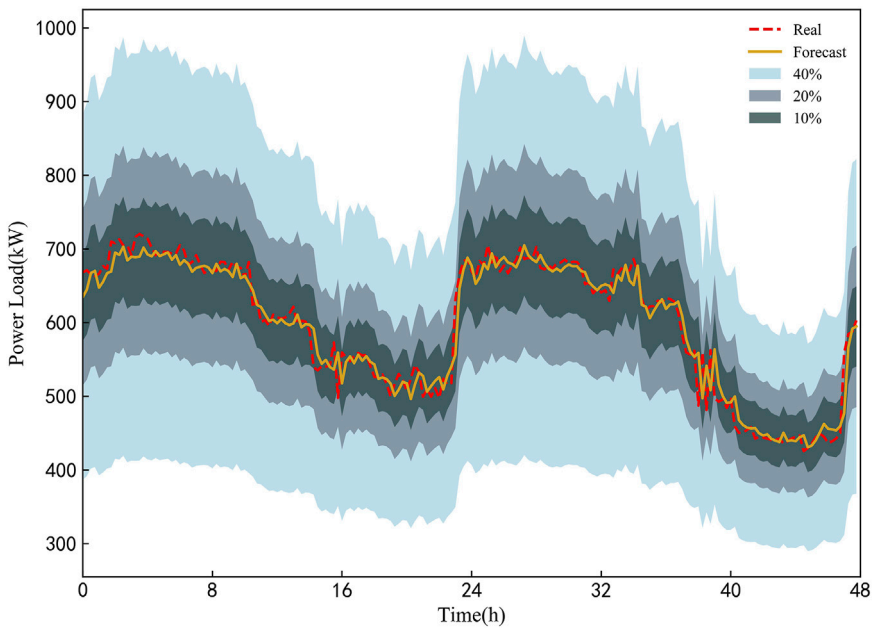


FIGURE 7 | PIs under different confidence levels obtained by the proposed ConvLSTM-MDN model in experiment 1.

network (FFNN), LSTM, 2D-CNN, and 2D-CNN-LSTM. The hyper-parameters of all models are given in **Table 2**. The first to sixth model in the contrast models directly use 1D time series data as the input data. The rest of model use 2D matrices as input data. The activation function of the neural network in all models is ReLU. Adam Optimizer is used to optimize the network parameters to minimize loss function. The number of components in the MDN network of all contrast models is 3.

5.1 Experiment 1: Load Forecasting of the Nonferrous Metal Smelting Industrial Customer

The nonferrous metal smelting industry data collected between 1 March 2018 and 31 August 2018 are selected for a short-term probability density function forecasting case study. After converting a 1D time series to 2D matrices, the input dataset is split into three parts for training (1 March 2018 to 6 July 2018), validation (7 July 2018 to 24 July 2018), and testing (25 July 2018 to 31 August 2018).

5.1.1 2D Matrix Size Analysis

According to the conversion method described in **Subsection 4.2**, the size of converted 2D matrix is $[N \times M]$. From **Section 3.1**, we know that temporal relevancy between loads weakens as the time interval increases. Since the M of the matrix is fixed, time length of historical load is determined by N. Too large N may result in longer processing time and running out of memory. On the contrary, too small N may lead to insufficient extraction of information. Therefore, we need to select an appropriate value of N. Considering the processing time, we

TABLE 4 | Load forecasting evaluation on the testing set.

Time	From 25-Jul-2018 to 31-Aug-2018			
Evaluation metrics	**Median**		**CRPS (%)**	**Forecasting time(s)**
	RMSE (kW)	**MAPE (%)**		
LR-MDN	53.616	10.028	40.391	0.050
ARIMA-MDN	52.296	9.690	40.376	0.459
SVR-MDN	42.826	7.681	32.897	0.523
RF-MDN	34.196	5.816	24.496	0.500
FFNN-MDN	33.184	5.640	23.125	0.574
LSTM-MDN	28.628	4.554	21.002	0.140
2D-CNN-MDN	28.625	3.989	19.622	0.180
2D-CNN-LSTM-MDN	23.707	3.600	17.418	0.401
Proposed model	**19.051**	**2.604**	**12.774**	0.606

The bold values indicate the minimum value of each evaluation metrics, that is, the best forecasting effect.

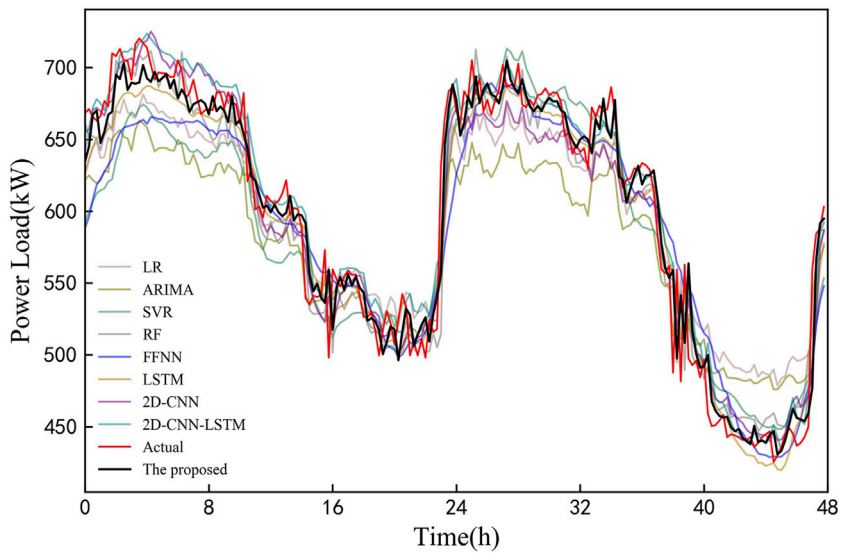


FIGURE 8 | Load forecasting profiles of all models in experiment 1.

TABLE 5 | Forecasting results of each *N*

***N* of 2D matrix**	**Evaluation metrics**		
	Median		**CRPS (%)**
	RMSE (kW)	**MAPE (%)**	
2	17.870	2.372	13.517
3	17.520	2.371	13.433
4	**17.404**	**2.343**	**13.283**
5	17.865	2.414	13.702
6	18.094	2.487	13.993
7	18.692	2.593	14.545

The bold values indicate the minimum value of each evaluation metrics, that is, the best forecasting effect.

set the value range of *N* to [2,7] in this study. By comparing the optimal result corresponding to each *N*, the *N* with minimum error is selected as the final *N*. The result corresponding to each *N* is shown in **Table 3**.

It is observed that the most accurate results are obtained when $N = 3$, that is, the RMSE, MAPE, and CRPS values are lower than the other values of *N*. When $N = 7$, the three error are the largest. The difference between the error of other *N* is little. Furthermore, the optimal hyper-parameters (filters, kernel size, and the number of components in the MDN) of the model corresponding to $N = 3$ are 30, (2×2), and 3.

5.1.2 Forecasting Result

After determining the size of 2D matrix and hyper-parameters of the model, training set and validation set are used to train the model to obtain the optimal model. Finally, the testing set is input into the optimized ConvLSTM-MDN model to forecast PDFs of industrial load. **Figure 6** shows the predicted PDFs for two different times of a day in the testing set and associated real values. **Figure 6A,B** show the PDFs for peak hours and off-peak hours, respectively. As shown in the figures, real values are very close to the peak of the PDF curve, especially real value in

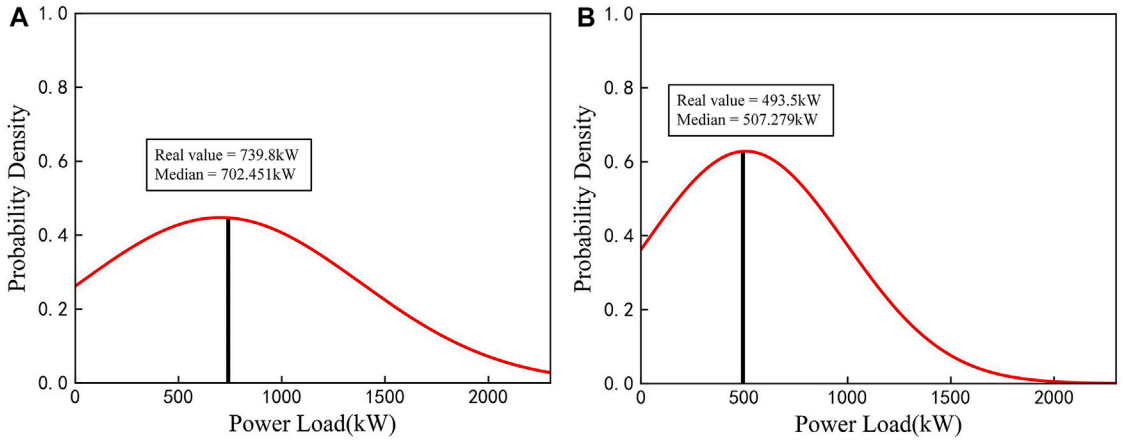


FIGURE 9 | ConvLSTM-MDN model predictive distribution and real values for a day at **(A)** 8:45 and **(B)** 23:00 in experiment 2.

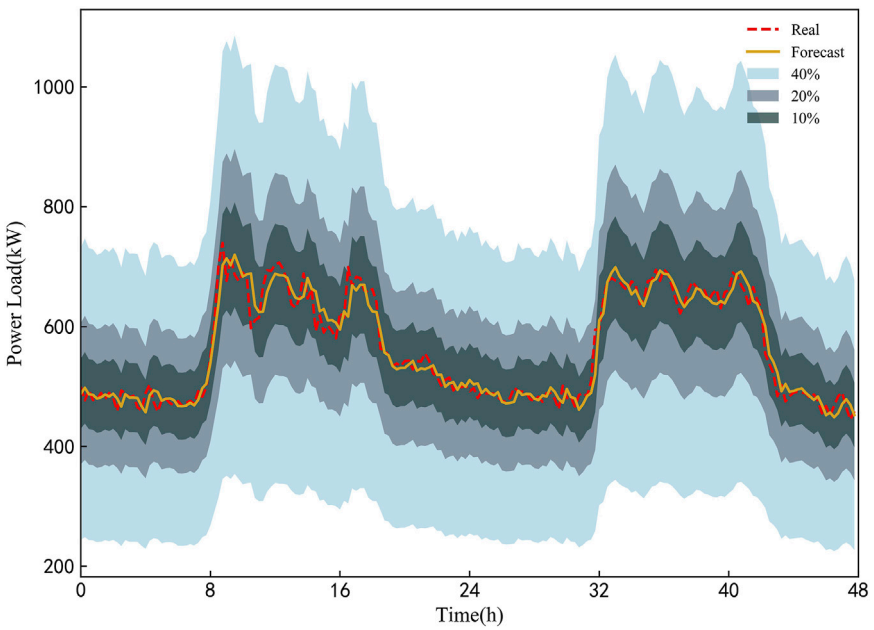


FIGURE 10 | PIs under different confidence levels obtained by the proposed ConvLSTM-MDN model in experiment 2.

Figure 6B almost coincides with the peak, which indicates that the sharpness of predicted PDF is clear.

Taking two sample days in testing set as an example, the PIs under different confidence level obtained by the proposed ConvLSTM-MDN framework are shown in **Figure 7**. As presented in **Figure 7**, the PIs under a higher confidence level can cover the PIs under a lower confidence level. In addition, the PIs under different confidence levels and real values of the load have similar fluctuation. Since the PIs under a low confidence level is narrower than the PIs under a high confidence level, the small number of real loads falls outside the PI with lowest confidence level. The PIs becomes narrower when the value of load rises or falls rapidly, and the PIs for peak hours become wider.

5.1.3 Comparisons Analysis

The forecasting result comparisons between the proposed model and contrast models are all provided in **Table 4**, with the best performance being highlighted. In the table, the forecasting effect of statistical method integrated into the MDN obviously worse than machine learning algorithms integrated into the MDN, and shallow machine learning algorithms worse than deep learning algorithms. For deep learning algorithms, the model combining 2D-CNN and LSTM can achieve better performance than LSTM and 2D-CNN alone for forecasting. The ConvLSTM, which leverages the strengths of CNN and LSTM, performs better than 2D-CNN-LSTM. For instance, the RMSE improvement rates is 19.64%, the MAPE improvement rates is 27.67%, and the CRPS improvement rates

TABLE 6 | Load forecasting evaluation on the testing set.

Time	From 28-Jul-2018 to 31-Aug-2018			
Evaluation metrics	**Median**		**CRPS (%)**	**Forecasting time(s)**
	RMSE (kW)	**MAPE (%)**		
LR-MDN	53.308	9.231	65.224	0.705
ARIMA-MDN	33.045	5.725	30.018	1.945
SVR-MDN	50.035	7.692	41.837	0.949
RF-MDN	51.780	8.285	46.491	0.466
FFNN-MDN	25.141	3.072	17.822	0.061
LSTM-MDN	22.499	2.956	16.804	0.374
2D-CNN-MDN	22.719	3.091	17.434	0.693
2D-CNN-LSTM-MDN	22.018	2.736	16.078	0.281
Proposed model	**17.404**	**2.343**	**13.283**	1.768

The bold values indicate the minimum value of each evaluation metrics, that is, the best forecasting effect.

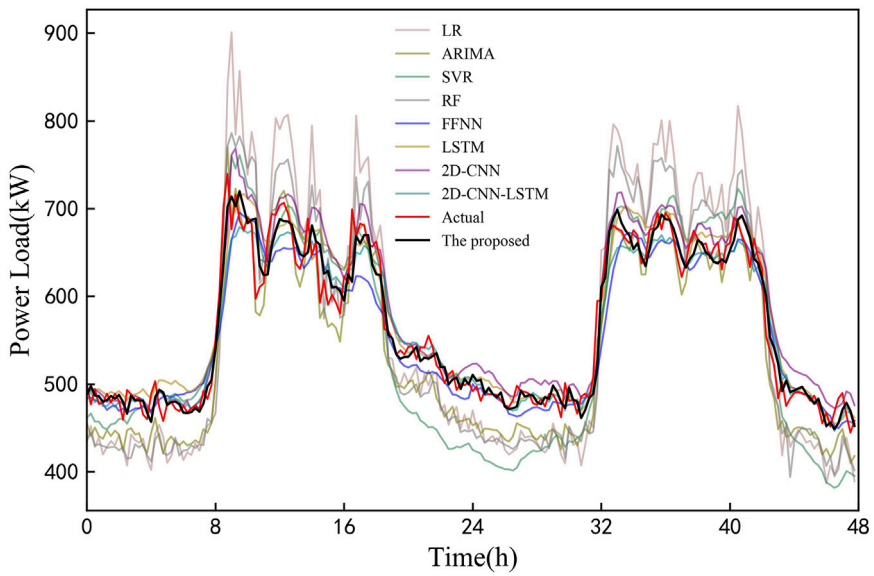


FIGURE 11 | Load forecasting profiles of all models in experiment 2.

TABLE 7 | Forecasting results of each *N*.

***N* of 2D matrix**	**Evaluation metrics**		
	Median		**CRPS (%)**
	RMSE (kW)	**MAPE (%)**	
2	0.407	15.986	0.194
3	0.406	14.972	0.191
4	**0.393**	**14.516**	**0.183**
5	0.415	18.867	0.205
6	0.417	16.916	0.202
7	0.435	17.910	0.215

The bold values indicate the minimum value of each evaluation metrics, that is, the best forecasting effect.

is 26.67%. It indicates that ConvLSTM can better capture features of industrial load. In addition, the application of ConvLSTM network to the load forecasting of industrial customers is feasible and effective. Compared with comparison models, the proposed model achieves the most accurate results, which shows the superiority of the model in PLF and DLF. Above all, it can be concluded that the ConvLSTM-MDN model is effective for solving the load forecasting problem of industrial customers.

In order to clearly display the prediction results of all models, the load of 2 days in the testing set is selected for further analysis, as shown in **Figure 8**. It is observed that the load of industrial customer in experiment 1 has strong volatility and high nonlinearity, which brings challenges to all models. According to curves in the figure, contrast models have large deviations from the real load, but the proposed model in this study generally fits and catches the trend of actual load.

5.2 Experiment 2: Load Forecasting of Medical Industrial Customer

In this experiment, the load data of the medical industry from 1 March 2018 to 31 August 2018 are used to run simulations.

TABLE 8 | Load forecasting evaluation on the testing set.

Time	From 21-Oct-2010 to 31-Dec-2018			
Evaluation metrics	**Median**		**CRPS (%)**	**Forecasting time(s)**
	RMSE (kW)	**MAPE (%)**		
LR-MDN	0.684	86.362	0.523	0.450
ARIMA-MDN	0.961	111.577	0.775	0.845
SVR-MDN	0.749	84.804	0.531	0.846
RF-MDN	0.584	81.013	0.464	0.654
FFNN-MDN	0.475	19.321	0.240	1.102
LSTM-MDN	0.496	18.106	0.249	0.744
2D-CNN-MDN	0.521	33.586	0.303	0.746
2D-CNN-LSTM-MDN	0.425	20.494	0.218	0.884
Proposed model	**0.393**	**14.516**	**0.183**	0.393

The bold values indicate the minimum value of each evaluation metrics, that is, the best forecasting effect.

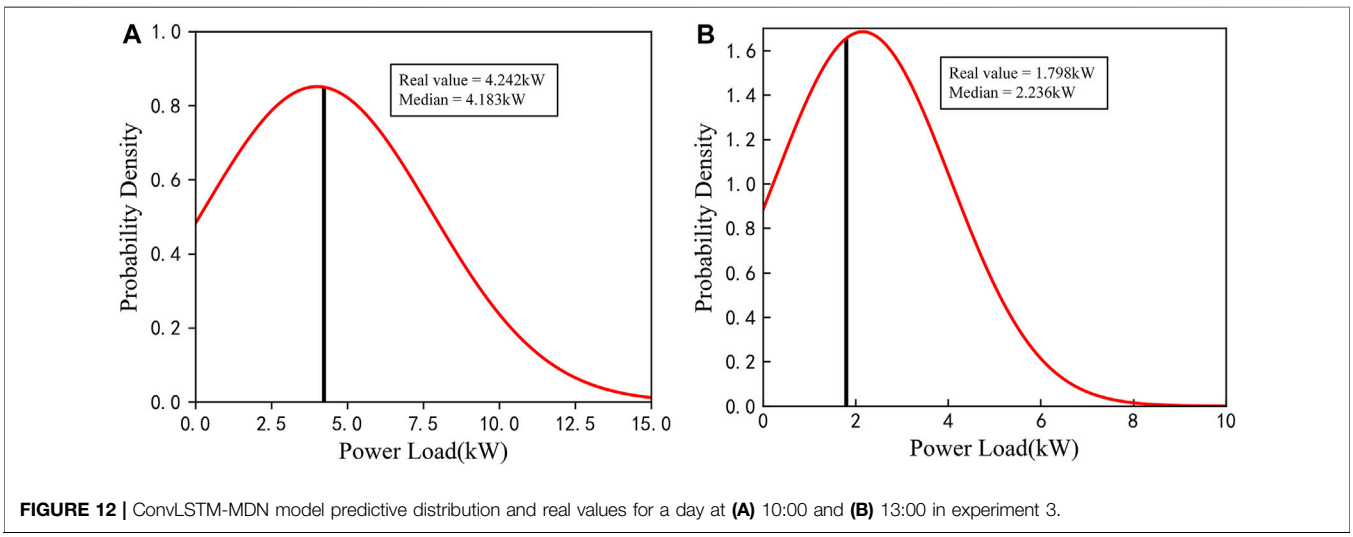


FIGURE 12 | ConvLSTM-MDN model predictive distribution and real values for a day at **(A)** 10:00 and **(B)** 13:00 in experiment 3.

Similarly, the approximate time range of testing set is from 25 July 2018 to 31 August 2018.

5.2.1 2D Matrix Size Analysis

According to the comparison in **Table 5**, the most accurate results are obtained when $N = 4$. Although the values of three evaluation metrics fluctuate with the change of N, the fluctuation range is small, that is, the difference between the maximum and minimum of RMSE, MAPE, and CRPS is 1.288, 0.25, and 1.315, respectively. It reflects the stability of the model under different input features. Furthermore, in the case of $N = 4$, filters, kernel size and the number of components are 75, (2 × 2), and 3.

5.2.2 Forecasting Result

Figure 9 shows the PDFs of the peak hour (**Figure 9A**) and off-peak hour (**Figure 9B**) of a day in the testing set and the associated real values. It can be observed from the figure that real values all falls near the peak of the PDF curve, especially real value in **Figure 9B** is the closest to the peak. It indicates that the high forecasting accuracy of the proposed model.

Figure 10 shows the PIs under different confidence level obtained by the proposed ConvLSTM-MDN model and real values in two sample days. It can be seen that a few real values fall outside the PI with lowest confidence level, and all real values fall within the PI with highest confidence level. The PIs under all confidence levels and real values of the load have similar fluctuation trend, which shows that the proposed model can capture the dynamic changes of the load.

5.2.3 Comparisons Analysis

Table 6 shows the evaluation metrics of all models on the testing set. The performance of deep learning algorithms integrated into the MDN significantly outperform the performance of shallow machine learning algorithms and statistical method integrated into the MDN in the term of all metrics. Among contrast models, the 2D-CNN-LSTM-MDN model has the best forecasting results. But compared with the 2D-CNN-LSTM-MDN model, the ConvLSTM-MDN model has the minimum errors, which indicate that the proposed model improves the forecasting accuracy. Although the forecasting time is the longest, it is acceptable in practical application with the popularization of cloud computing.

Figure 11 shows the comparison between the prediction results and actual values for 2 days in the testing set. It is observed that the proposed model can better fit the trend of

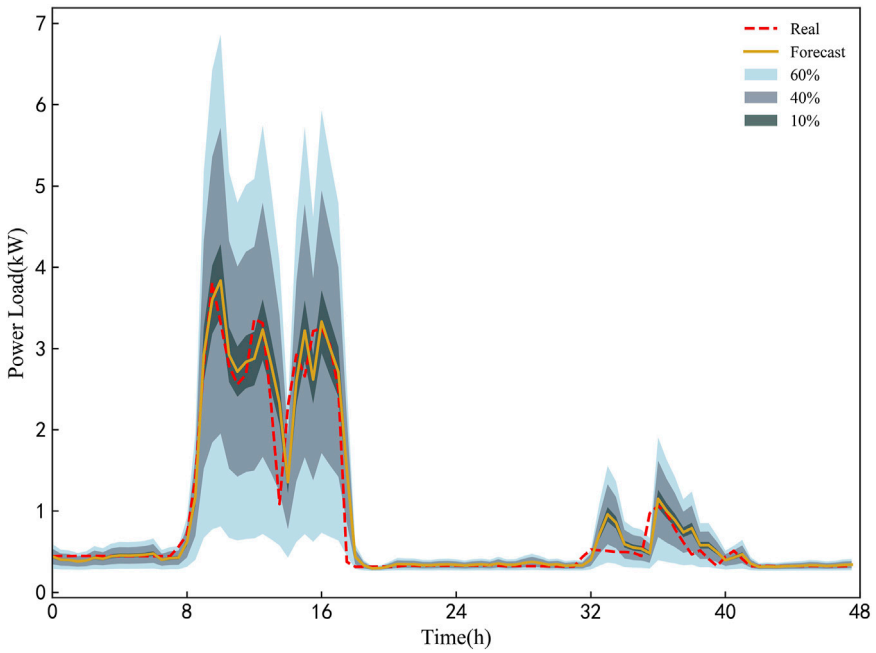


FIGURE 13 | PIs under different confidence levels obtained by the proposed ConvLSTM-MDN model in experiment 3.

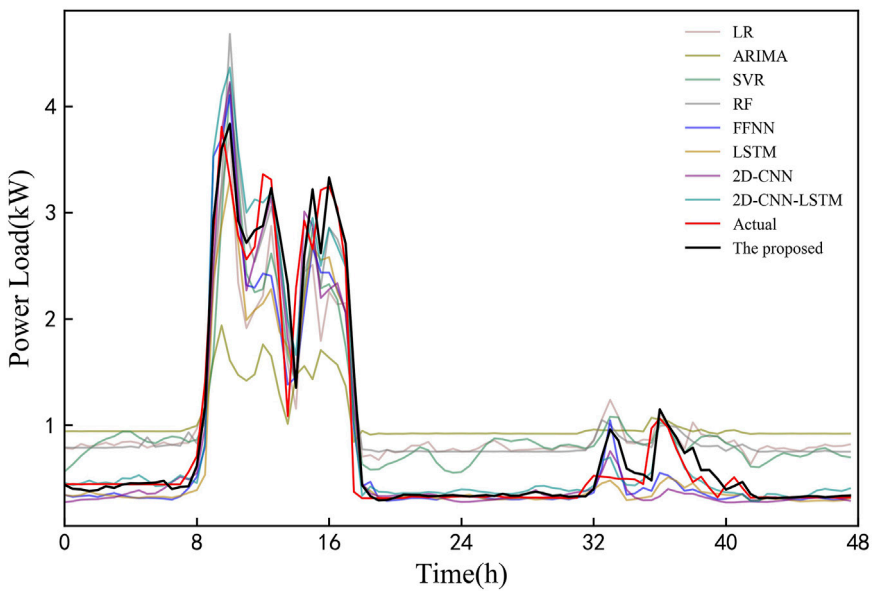


FIGURE 14 | Load forecasting profiles of all models in experiment 3.

actual load, and other models have large deviations from the real load, especially the LR-MDN model.

5.3 Experiment 3: Load Forecasting of the Irish Industrial Customer

In this experiment, public dataset small-to-medium industrial customer collected from Irish is from 1 January 2010 to 31 December 2010 with a 30-min interval. According to the splitting rules, the approximate time range of testing set is from 21 October 2010 to 31 December 2010.

5.3.1 2D Matrix Size Analysis

Table 7 shows the minimum error for each N. This experiment also achieves the best forecasting result when $N = 4$, which is the same as experiment 2. In the case of $N = 4$, filters, kernel size and

the number of components are 15, (2 × 2), and 3. There is a significant difference between Irish industrial load and other industrial loads, for example, the maximum values of the three industrial loads are 8.257 kWh, 971.2 kWh, and 785.4 kWh. Therefore, the value of the evaluation metrics in this experiment is completely different from the evaluation metrics in the previous two experiments.

5.3.2 Forecasting Result

Figure 12 shows the PDFs of the peak hour and off-peak hour of a day in the testing set and the associated real values. It can be observed from the figure that the real value in **Figure 12A** is the closest to the peak. **Figure 13** shows the PIs under different confidence level and real values in two sample days. In order to better display the forecasting result, we adjust the value of confidence level. As the value of confidence level decreases, real values that falls outside the PI increases. In addition, The PIs under all confidence levels can capture the dynamic changes of the load.

5.3.3 Comparisons Analysis

Table 8 shows the evaluation metrics of all models on the testing set. The results of shallow machine learning algorithms are far inferior to deep learning algorithms. Among all the models, the proposed model has the best forecasting performance. **Figure 14** shows the comparison between the prediction results and actual values for 2 days in the testing set. It can be seen that contrast models are quite different from the real load, and the proposed model can better fit the trend of actual load.

6 CONCLUSION

In this study, we propose a new probabilistic forecasting method, which can capture the uncertainty of a single industrial customer's load. By restructuring the load and various relevant factors into 2D matrices using a sliding-window approach, the forecasting model—ConvLSTM-MDN—was applied to the short-term probability forecasting problem of industrial loads. In order to verify the performance of the proposed method, this study builds the classical statistical methods, state-of-the-art deep- and shallow-based models for comparison, and conducts numerical simulations in three experiments. The following results were noted.

1) For three completely different industrial customers, the experimental results of 2D matrix size analysis show that the best forecasting results are obtained when N is 3 or 4. Therefore, when N is in the range of (3, 4), the historical load of a reasonable time interval can be obtained as input feature, which can not only ensure the forecasting accuracy but also reduce the training time.
2) The proposed model takes full advantage of the capabilities of ConvLSTM in feature extraction and sequence learning and the capabilities of MDN in describing uncertainty. The final results show that the model can effectively improve the forecasting accuracy of industrial load. For example, 15–60% improvement in accuracy compared to deep-based models.
3) The hybrid model used in this study is not complex, and a variety of external factors are considered, which is beneficial to be extended to various industrial customers. The forecasting results for three industrial customers show that the model generalizes well.

In our future work, we will study how to reduce computational demand and training time when the value of N increases. Moreover, the feature factors affecting industrial customer such as electricity price will be considered to improve forecasting accuracy.

DATA AVAILABILITY STATEMENT

The data analyzed in this study are subject to the following licenses/restrictions: Because the load data of industrial enterprises are a commercial secret, they cannot be disclose. Requests to access these datasets should be directed to yuanyuan.wang.1980@ieee.org.

AUTHOR CONTRIBUTIONS

YW: resources, supervision, project administration, and funding acquisition. TW: experiments, research methods, data processing, and write the original draft. XC: guide experiments. XZ: project administration, fund acquisition. JH: perform the writing-review on references. XT: funding acquisition.

FUNDING

This work is supported by the National Natural Science Foundation of China (No. 52177069), National Natural Science Foundation of China (No. 51777014), Hunan Provincial Key Research and Development Program (No. 2018GK 2057), Research projects funded by Department of Education of Hunan Province of China (18A124), and National Natural Science Foundation of Hunan Provincial (2020JJ5585).

REFERENCES

Afrasiabi, M., Mohammadi, M., Rastegar, M., and Afrasiabi, S. (2021). Advanced Deep Learning Approach for Probabilistic Wind Speed Forecasting. *IEEE Trans. Ind. Inf.* 17 (1), 720–727. doi:10.1109/tii.2020.3004436

Afrasiabi, M., Mohammadi, M., Rastegar, M., Stankovic, L., Afrasiabi, S., and Khazaei, M. (2020). Deep-Based Conditional Probability Density Function Forecasting of Residential Loads. *IEEE Trans. Smart Grid* 11 (4), 3646–3657. doi:10.1109/tsg.2020.2972513

Berk, K., Hoffmann, A., and Müller, A. (2018). Probabilistic Forecasting of Industrial Electricity Load with Regime Switching Behavior. *Int. J. Forecast.* 34 (2), 147–162. doi:10.1016/j.ijforecast.2017.09.006

Bishop, C. M. (1994). *Mixture Density Networks*. Citeseer, Princeton, NJ, USA: Rep. NCRG/94/004.

Cai, Y., Huang, T., Bompard, E., Cao, Y., and Li, Y. (2017). Self-Sustainable Community of Electricity Prosumers in the Emerging Distribution System. *IEEE Trans. Smart Grid* 8 (5), 2207–2216. doi:10.1109/tsg.2016.2518241

Cai, Y., Liu, Y., Tang, X., Tan, Y., and Cao, Y. (2021). Increasing Renewable Energy Consumption Coordination with the Monthly Interprovincial Transaction Market. *Front. Energ. Res.* 9, 719419. doi:10.3389/FENRG.2021.719419

Commission for Energy Regulation (Cer) (2012). CER Smart Metering Project Electricity Customer Behaviour Trial. [Online]. Available at: https://www.ucd.ie/issda.

Da Silva, F. L. C., Cyrino Oliveira, F. L., and Souza, R. C. (2019). A Bottom-Up Bayesian Extension for Long Term Electricity Consumption Forecasting. *Energy* 167, 198–210. doi:10.1016/j.energy.2018.10.201

Essien, A., and Giannetti, C. (2020). A Deep Learning Model for Smart Manufacturing Using Convolutional LSTM Neural Network Autoencoders. *IEEE Trans. Ind. Inf.* 16 (9), 6069–6078. doi:10.1109/tii.2020.2967556

Fallah, S. N., Ganjkhani, M., Shamshirband, S., and Chau, K. W. (2019). Computational Intelligence on Short-Term Load Forecasting. *A Methodological Overview"* 12 (3). doi:10.3390/en12030393

Farsi, B., Amayri, M., Bouguila, N., and Eicker, U. (2021). On Short-Term Load Forecasting Using Machine Learning Techniques and a Novel Parallel Deep LSTM-CNN Approach. *IEEE Access* 9, 31191–31212. doi:10.1109/access.2021.3060290

He, Y., Liu, R., Li, H., Wang, S., and Lu, X. (2017). Short-term Power Load Probability Density Forecasting Method Using Kernel-Based Support Vector Quantile Regression and Copula Theory. *Appl. Energ.* 185, 254–266. doi:10.1016/j.apenergy.2016.10.079

He, Y., Qin, Y., Wang, S., Wang, X., and Wang, C. (2019). Electricity Consumption Probability Density Forecasting Method Based on LASSO-Quantile Regression Neural Network. *Appl. Energ.* 233-234, 565–575. doi:10.1016/j.apenergy.2018.10.061

Hong, Y.-Y., Martinez, J. J. F., and Fajardo, A. C. (2020). Day-Ahead Solar Irradiation Forecasting Utilizing Gramian Angular Field and Convolutional Long Short-Term Memory. *IEEE Access* 8, 18741–18753. doi:10.1109/access.2020.2967900

Hou, H., Liu, P., and Xiao, Z. f. (2021). Capacity Configuration Optimization of Standalone Multi-Energy Hub Considering Electricity, Heat and Hydrogen Uncertainty. *IET Energ. Convers. Econ.* 2021, 12028. doi:10.1049/enc2.12028

Hou, H., Xu, T., and Wu, X. X. (2020). Optimal Capacity Configuration of the Wind-Photovoltaic-Storage Hybrid Power System Based on Gravity Energy Storage System. *Appl. Energ.* 271, 115052. doi:10.1016/j.apenergy.2020.115052

Hou, H., Xue, M. Y., and Xu, Y. (2020). Multi-objective Economic Dispatch of a Microgrid Considering Electric Vehicle and Transferable Load. *Appl. Energ.* 262, 114489. doi:10.1016/j.apenergy.2020.114489

Jiao, R., Zhang, T., Jiang, Y., and He, H. (2018). Short-Term Non-residential Load Forecasting Based on Multiple Sequences LSTM Recurrent Neural Network. *IEEE Access* 6, 59438–59448. doi:10.1109/access.2018.2873712

Kuster, C., Rezgui, Y., and Mourshed, M. (2017). Electrical Load Forecasting Models: A Critical Systematic Review. *Sustain. Cities Soc.* 35, 257–270. doi:10.1016/j.scs.2017.08.009

López, J. C., Rider, M. J., and Wu, Q. (2019). Parsimonious Short-Term Load Forecasting for Optimal Operation Planning of Electrical Distribution Systems. *IEEE Trans. Power Syst.* 34 (2), 1427–1437.

National Bureau of Statistics of the People's Republic of China (2021). *China Energy Statistical Yearbook 2020*. Beijing, China: China Statistics Press, 195–199.

Ruan, Y. J., Wang, G., Meng, H., and Qian, F. Y. (2021). A Hybrid Model for Power Consumption Forecasting Using VMD-Based the Long Short-Term Memory Neural Network. *Front. Energ. Res.* 9.

Shi, X., Chen, Z., and Wang, H. (2015). Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting. *Proc. 28th Int. Conf. Neural Inf. Process. Syst.* 2015, 802–810.

Sun, Z., and Zhao, M. (2020). Short-Term Wind Power Forecasting Based on VMD Decomposition, ConvLSTM Networks and Error Analysis. *IEEE Access* 8, 134422–134434. doi:10.1109/access.2020.3011060

Tan, M., Yuan, S., Li, S., Su, Y., Li, H., and He, F. H. (2020). Ultra-Short-Term Industrial Power Demand Forecasting Using LSTM Based Hybrid Ensemble Learning. *IEEE Trans. Power Syst.* 35 (4), 2937–2948. doi:10.1109/tpwrs.2019.2963109

Wang, R., Xia, X. Y., Li, Y. P., and Cao, W. M. (2021). Clifford Fuzzy Support Vector Machine for Regression and its Application in Electric Load Forecasting of Energy System. *Front. Energ. Res.* 9. doi:10.3389/fenrg.2021.793078

Wang, Y., Chen, J., Chen, X., Zeng, X., Kong, Y., Sun, S., et al. (2021). Short-Term Load Forecasting for Industrial Customers Based on TCN-LightGBM. *IEEE Trans. Power Syst.* 36 (3), 1984–1997. doi:10.1109/tpwrs.2020.3028133

Wang, Y., Kong, Y., Tang, X., Chen, X., Xu, Y., Chen, J., et al. (2020). Short-Term Industrial Load Forecasting Based on Ensemble Hidden Markov Model. *IEEE Access* 8, 160858–160870. doi:10.1109/access.2020.3020799

Wang, Y., Sun, S., Chen, X., Zeng, X., Kong, Y., Chen, J., et al. (2021). Short-term Load Forecasting of Industrial Customers Based on SVMD and XGBoost. *Int. J. Electr. Power Energ. Syst.* 129, 106830. doi:10.1016/j.ijepes.2021.106830

Xie, W., Zhang, P., Chen, R., and Zhou, Z. (2019). A Nonparametric Bayesian Framework for Short-Term Wind Power Probabilistic Forecast. *IEEE Trans. Power Syst.* 34 (1), 371–379. doi:10.1109/tpwrs.2018.2858265

Zhang, H., Liu, Y., Yan, J., Han, S., Li, L., and Long, Q. (2020). Improved Deep Mixture Density Network for Regional Wind Power Probabilistic Forecasting. *IEEE Trans. Power Syst.* 35 (4), 2549–2560. doi:10.1109/tpwrs.2020.2971607

Zhang, J., Yan, J., Infield, D., Liu, Y., and Lien, F.-s. (2019). Short-term Forecasting and Uncertainty Analysis of Wind Turbine Power Based on Long Short-Term Memory Network and Gaussian Mixture Model. *Appl. Energ.* 241, 229–244. doi:10.1016/j.apenergy.2019.03.044

Zhang, W., Quan, H., and Srinivasan, D. (2019). An Improved Quantile Regression Neural Network for Probabilistic Load Forecasting. *IEEE Trans. Smart Grid* 10 (4), 4425–4434. doi:10.1109/tsg.2018.2859749

Zhao, X., and Li, R. (2021). A Novel Decomposition and Combination Technique for Forecasting Monthly Electricity Consumption. *Front. Energ. Res.* 9.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Energy Research

ORIGINAL RESEARCH
published: 02 May 2022
doi: 10.3389/fenrg.2022.900824

Cluster Partition-Based Zonal Voltage Control for Distribution Networks With High Penetrated PVs

Xingyu Zhao, Chuanliang Xiao, Ke Peng, Jiajia Chen and Xinhui Zhang*

School of Electronic Engineering, Shandong University of Technology, Zibo, China

With the high penetrated photovoltaics (PVs) accessing distribution networks (DNs) in the future, the overvoltage of distribution network will be more serious, and the voltage control will be more complex. To solve this problem, a cluster partition-based zonal voltage control method for DN is proposed in this study. First, a cluster partition method using a comprehensive performance index is proposed: in terms of DN structure, a global density quality function index is introduced to measure the coupling degree of nodes in clusters. In terms of cluster function, the inconsistency coefficient index is presented to reflect the PV output characteristics, and a node membership function index is presented to limit the scale of the cluster. Then, the DN cluster partition is carried out under the fast Newman (FN) algorithm. Second, a second-order cone programming (SOCP) model of voltage control is established to maximize the PV consumption in each cluster. To achieve the coordinated optimization of clusters, an iterative optimization strategy among clusters is proposed. Finally, the effectiveness and rationality of the proposed method are verified in a 10 kV actual feeder system in Zhejiang Province, China.

Keywords: distributed network, photovoltaic generation, cluster partition, distributed optimization, voltage control

OPEN ACCESS

Edited by:
Yunfei Mu,
Tianjin University, China

Reviewed by:
Qianlong Zhu,
Anhui University, China
Zhijian Feng,
Qingdao University, China

****Correspondence:***
Chuanliang Xiao
xiaocl@sdut.edu.cn

Specialty section:
This article was submitted to Smart Grids, a section of the journal Frontiers in Energy Research

Received: *21 March 2022*
Accepted: *04 April 2022*
Published: *02 May 2022*

Citation:
Zhao X, Xiao C, Peng K, Chen J and Zhang X (2022) Cluster Partition-Based Zonal Voltage Control for Distribution Networks With High Penetrated PVs. Front. Energy Res. 10:900824. doi: 10.3389/fenrg.2022.900824

1 INTRODUCTION

In order to promote the realization of peaking and carbon neutrality goals, China will carry out the development of roof-distributed PVs for the whole county. However, the integration of large-scale PVs aggravates the overvoltage in the DN (Wang et al., 2021), and the voltage control model turns into a high-dimensional complex model (Liao et al., 2017; Li et al., 2018), which is difficult to solve. Therefore, the investigation of voltage control strategies for active distribution networks (ADNs) with a high penetration of PVs is necessary to improve the consumption of PVs. Zonal voltage control (Han et al., 2019) provides an effective solution for solving the complex voltage control. Under the network partition, the high-dimensional voltage control model can be transformed into several subproblems which are easy to be solved (Xiao, 2020). Zonal voltage control mainly exists in two aspects: network partition and subnetwork voltage regulation.

To implement the zonal voltage control, the network partition is first carried out. In terms of the network partition index, the DN structure and power balance are the main basis for a network partition. Hu et al. (2021) used the reactive voltage sensitivity matrix to express the electrical coupling degree between nodes, and the sensitive nodes are found out to complete the network partition. Dou et al. (2018) proposed a division index system of virtual cluster, including the spatial location, output characteristics, and response mode of PVs. Kou et al. (2019) proposed the network partition indexes which include the net load of nodes and active and reactive power regulation capacity of PVs. Chai et al. (2007) proposed the cluster performance index. The index considers the electrical distance

between nodes and regional voltage regulating ability, which can ensure each cluster solves the voltage violations within the clusters. In the context of PV accessing in the whole counties, the scale of PVs in DNs is continuous expanded, the influence of PVs on network partition cannot be ignored. Therefore, it is not enough to employ the electrical distance and power balance as the network partition index. If the consistency of PV output and nodal voltage in the cluster can be considered during the network partition, the output of the subnetwork will be more smooth and controllable. In terms of network partition algorithms, there are three main types: intelligent algorithms, clustering algorithms, and community discovery algorithms of complex networks. Bi et al. (2019) proposed an improved genetic algorithm, which can find the optimal network partition according to the state of tie switch. Although the amount of calculation is reduced, this algorithm has a long time scale, which is not suitable for real-time optimization. Yan et al. (2021) used the Tabu search algorithm to complete the network partition. But this algorithm cannot automatically generate the optimal number of clusters, which can lead to an inaccurate result. Wang et al. (2021) proposes a fast-unfolding clustering algorithm for cluster partitioning, which employs the complex network modularity function to avoid setting the threshold value. When the network structure changes, the algorithm is no longer applicable. In different scenarios, and the time scale and accuracy of the network partition are different. For the zonal voltage control of DNs, the voltage control belongs to a real-time optimization, which requires a short-time scale of network partition. Therefore, determining how to establish a network partition algorithm with fast computing speed and accurate performance is the key to achieving zonal voltage control.

After DNs are partitioned into several subnetworks, the zonal voltage control method can be applied. The zonal voltage control mainly focuses on the optimization within subnetworks and the coordination among subnetworks. Xiao et al. (2017) adopted the zonal voltage control strategy of "maximize reactive power regulating first, and then minimize active power curtailing." Although the control process is simplified, the active power–reactive power interaction between subnetworks is not considered. According to the overvoltage degree, Li et al. (2021) proposed a voltage regulation control strategy based on the energy storage cluster. Autonomous optimization is realized by utilizing the energy storage system in clusters, but the coordination between clusters is not considered. Chai et al. (2019) used the alternating direction method of multipliers (ADMM) to optimize the zonal voltage between upstream and downstream clusters, which only achieves the coordination between adjacent clusters. With the high penetrated PVs accessing distribution networks, it is necessary to further establish a reasonable and effective coordination strategy among clusters. Therefore, realizing the independent voltage control in each subnetwork and the coordinating control among the subnetworks deserves deep research.

In light of the aforementioned issues, this study proposes a cluster partition-based zonal voltage control method for DNs. The main contributions of this study are summarized as follows:

1) Based on the community partition theory, a cluster partition method combined with the cluster comprehensive performance index and FN algorithm is proposed. The electrical information of the network structure and the impact of PVs on the partitioning solutions are considered, and the size of each subnetwork can be restricted.
2) A SOCP-based voltage control model aiming at the maximum PV consumption in the cluster is established. The high-dimensional voltage control model of the whole DN is transformed into several subproblems within the clusters, which are easily solved.
3) An iterative optimization-based active power coordination strategy among clusters is proposed to realize the coordination among different clusters. The maximum PV consumption for all clusters can be guaranteed.

The remainder of this study is organized as follows. In **Section 2**, a comprehensive performance index-based cluster partition method is proposed. In **Section 3**, an independent optimization model in the cluster is proposed. In **Section 4**, the active power coordination and interaction strategy among clusters is proposed. The case study is described in **Section 5**, and a 10 kV actual feeder network is used for simulation verification, followed by discussion and conclusions in **Section 6**.

2 COMPREHENSIVE PERFORMANCE INDEX-BASED CLUSTER PARTITION METHOD

2.1 Comprehensive Performance Index of Cluster Partition

In order to solve the problem that the factors of the existing cluster partition index are not comprehensive, in this study, based on the DN structure and cluster function, the comprehensive performance index is proposed to complete the cluster partition. The comprehensive performance index is composed of the global density quality function index, the inconsistency coefficient index, and the node membership function index.

The global density quality function is selected to measure the coupling degree of nodes in clusters according to the DN, so as to satisfy the structural principle which means that the nodes with a strong coupling degree should be divided into the same clusters, and nodes with a weak coupling degree should be divided into different clusters. First, the global internal density quality function Q^i_{GD} and the global external density quality function Q^e_{GD} are defined, where Q^i_{GD} represents the coupling degree of nodes within the cluster and Q^e_{GD} represents the coupling degree of nodes between clusters.

$$\boldsymbol{Q}^{\boldsymbol{i}}_{\boldsymbol{GD}} = \frac{\sum_{k=1}^{K} \sum_{x \in V_k} \sum_{y \in V_k} \boldsymbol{A}_{\boldsymbol{xy}}}{\sum_i \sum_j \boldsymbol{A}_{\boldsymbol{ij}}}, \tag{1}$$

$$\boldsymbol{Q}^{\boldsymbol{e}}_{\boldsymbol{GD}} = \frac{\sum_{k=1}^{K} \sum_{x \in V_k} \sum_{y \in V - V_k} \boldsymbol{A}_{\boldsymbol{xy}}}{\sum_i \sum_j \boldsymbol{A}_{\boldsymbol{ij}}}, \tag{2}$$

where k is the number of clusters, V_k represents the k-th cluster, A_{xy} represents the weight of the edges of nodes x and y, $\sum_{k=1}^{K}\sum_{x\in V_k}\sum_{y\in V-V_k} A_{xy}$ represents the sum of edge weights of all nodes in the cluster K, $\sum_{k=1}^{K}\sum_{x\in V_k}\sum_{y\in V-V_k} A_{xy}$ represents the sum of edge weights of all nodes in different clusters, A_{ij} represents the weight of the edges of nodes i and j, and $\sum_i\sum_j A_{ij}$ represents the sum of edge weights of all nodes in the whole network.

In this study, the edge weight of the DN is determined by the active voltage sensitivity matrix $S_{PU,}$ which can be got from Xiao et al. (2017). In order to reflect the coupling degree between nodes, the average value of edge weight is used to represent the weight A_{ij}:

$$\boldsymbol{A_{ij}} = \frac{\boldsymbol{S_{PU}^{ij}} + \boldsymbol{S_{PU}^{ji}}}{\boldsymbol{2}}, \tag{3}$$

where the value of Q_{GD}^{i} and Q_{GD}^{e} belongs to (0,1). As Q_{GD}^{i} is close to 1, it indicates that the nodes in a cluster have a strong coupling degree. As Q_{GD}^{e} is close to 0, indicates that the nodes in different clusters have a weak coupling degree. The global density quality function index Q_{GD} is defined as follows:

$$\boldsymbol{Q_{GD}} = \frac{\boldsymbol{1}}{\boldsymbol{2}}\left(\boldsymbol{Q_{GD}^{i}} + \boldsymbol{1} - \boldsymbol{Q_{GD}^{e}}\right). \tag{4}$$

The value of Q_{GD} belongs to (0,1). As Q_{GD} is close to 1, it indicates that in the same clusters, the nodes have a strong coupling degree and in the different clusters, the nodes have a weak coupling degree, that is, the larger the value of Q_{GD}, the better the partition result.

Due to the strong fluctuation of PV output, it is difficult to ensure that the clustering results can satisfy the DN topology constraints. If the PV nodes with inconsistent output are partitioned into the same cluster, the PV output in each cluster will be smoother, and the regulation ability of clusters can be increased. Therefore, the inconsistency coefficient index X is proposed to describe the variation trend of PV output in the cluster at different times.

$$\boldsymbol{X} = \frac{\boldsymbol{1}}{\boldsymbol{\Delta T}}\sum_{t=2}^{T}\sum_{i\in V_k}\sum_{j\in V_k}\boldsymbol{f_t}\left(\boldsymbol{P^i}, \boldsymbol{P^j}\right)/\left(\boldsymbol{T}-\boldsymbol{1}\right), \tag{5}$$

where P^i and P^j are the output sequences of PV i and PV j in the network, T is the sampling number of PV output in the whole day, and ΔT is the sampling time interval, which is set to 15 min. The expression of $f_t(P^i, P^j)$ is given as follows:

$$\boldsymbol{f_t}\left(\boldsymbol{P^i}, \boldsymbol{P^j}\right) = \begin{cases} \boldsymbol{1}, & \left(\boldsymbol{P_t^i} - \boldsymbol{P_{t-1}^i}\right)\left(\boldsymbol{P_t^j} - \boldsymbol{P_{t-1}^j}\right) < \boldsymbol{0} \\ \boldsymbol{0}, & \left(\boldsymbol{P_t^i} - \boldsymbol{P_{t-1}^i}\right)\left(\boldsymbol{P_t^j} - \boldsymbol{P_{t-1}^j}\right) \geq \boldsymbol{0} \end{cases}, \tag{6}$$

where $f_t(P^i, P^j)$ is the inconsistency of the output curve trend between PV i and PV j during the t-th to the t-1 sampling point within a cluster. When the output curve trend of two PVs is inconsistent (one rising and one falling), $f_t(P^i, P^j)$ is taken as 1, and $f_t(P^i, P^j)$ is taken as 0 when they are consistent.

A suitable scale of cluster will affect the difficulty of subsequent zonal voltage control. The balance of nodal size in clusters will reduce the complexity of zonal voltage control. Therefore, in order to judge the nodal size of each cluster, a node membership index is established:

$$\boldsymbol{M} = \frac{\mu\left(\boldsymbol{x}, \boldsymbol{V}|\boldsymbol{x}|\right)}{\mu\left(\boldsymbol{x}, \boldsymbol{V}-\boldsymbol{V}|\boldsymbol{x}|\right)}, \tag{7}$$

$$\mu\left(\boldsymbol{x}, \boldsymbol{V}|\boldsymbol{x}|\right) = \frac{\boldsymbol{1}}{|\boldsymbol{V}[\boldsymbol{x}]|}\sum\nolimits_{\boldsymbol{y}\in\boldsymbol{V}|\boldsymbol{x}|}\boldsymbol{A_{xy}}, \tag{8}$$

$$\mu\left(\boldsymbol{x}, \boldsymbol{V}-\boldsymbol{V}|\boldsymbol{x}|\right) = \frac{\boldsymbol{1}}{|\boldsymbol{V}-\boldsymbol{V}[\boldsymbol{x}]|}\sum\nolimits_{\boldsymbol{y}\in\boldsymbol{V}-\boldsymbol{V}|\boldsymbol{x}|}\boldsymbol{A_{xy}}, \tag{9}$$

where $V[x]$ represents the cluster which node x belongs to, that is, $x \in V[x]$; $|V[x]|$ represents the sum of total edges between nodes in the cluster, where node x belongs to; $\mu(x, V[x])$ indicates the membership degree of the nodes connected to x in the same cluster $V[x]$; V-$V[x]$ represents the cluster which does not contain node x; $|V$-$V[x]|$ indicates the sum of edges between nodes within clusters, except $V[x]$; and $\mu(x, V$-$V[x])$ indicates the membership degree of node y in cluster V-$V[x]$.

Based on the aforementioned indexes, the comprehensive performance index is defined as follows:

$$\omega = \omega_1\boldsymbol{Q_{GD}} + \omega_2\boldsymbol{X} + \omega_3\boldsymbol{M}, \tag{10}$$

where ω_1, ω_2, and ω_3 are the weights of global density quality function index, inconsistency coefficient index, and node membership index, respectively, which meet $0\leq\omega_1$; ω_2; $\omega_3 \leq 1$ and $\omega_1+\omega_2+\omega_3 = 1$.

2.2 Cluster Partition Algorithm

Community structures are formed by the aggregation of network nodes, which generally utilize regional coupling of physical, chemical, or social interaction relations (Cohen, 2009). The community should satisfy that the nodes with a strong coupling degree belong to the same clusters, and nodes with a weak coupling degree belong to different clusters. A power network is also a complex network composed of points and edges, in which the point corresponds to the node in the network, and the edge corresponds to the connecting line between nodes (Liu, 2019). The community not only reflects the structural characteristics of the network but also reveals the internal relations of the network. Based on the community discovery algorithm, the community structure, which has a close connection within the community and a sparse connection between communities, can be analyzed. The FN algorithm is a complex network community partition algorithm based on a local search proposed by Newman et al. (2004). The steps of using the FN algorithm to divide clusters are as follows:

Step 1: Each node in the network is regarded as a separate cluster, and the initial comprehensive performance index value ω^0 is calculated according to **Eq. 10**.

Step 2: For node i, node j is selected randomly from the remaining nodes to merge into a node pair. The comprehensive performance index ω^1 is calculated. Then calculate the value $\Delta\omega = \omega^1$-ω^0.

Step 3: Step 2 is repeated for all node pairs. The node pair with the largest $\Delta\omega$ will be merged into a cluster and referred to as an individual node.

Step 4: Steps 1-3 are repeated until the comprehensive performance index value for all node pairs does not increase. The optimal cluster partition and the associated comprehensive performance index value can be obtained.

3 ZONAL VOLTAGE CONTROL MODEL IN CLUSTERS

3.1 Objective Function

For the overvoltage in the DN, a SOCP-based voltage control model is established with the objective function to maximize the PV consumption. Based on the cluster partition, combined with the regulation capacity of energy storage, the objective functions of clusters can be established as follows:

$$\boldsymbol{F} = \mathbf{max}\left\{ \sum_{i=1}^{n_K^{PV}} \boldsymbol{P}_{i,t}^{PV} - \boldsymbol{P}_{ESS,t} - \boldsymbol{P}_{loss,t} \right\}, \tag{11}$$

where n_K^{PV} is the number of installed PVs in cluster K, $P_{i,t}^{PV}$ is the output power of PV i at time t, $P_{ESS,t}$ is the energy storage power at time t, and $P_{loss,t}$ is the network loss at time t.

3.2 Constraints of the Proposed Model

1) Power flow constraints of DN

The traditional nonlinear power flow model can be transformed into a SOCP power flow model by second-order cone relaxation (Peng et al., 2021).

$$\begin{cases} \mathbf{P}_{ij,t}^2 + \boldsymbol{Q}_{ij,t}^2 = \boldsymbol{L}_{ij,t}\boldsymbol{U}_{i,t}^2 \\ \boldsymbol{U}_{i,t}^2 = \boldsymbol{u}_{i,t} \end{cases} = \left\| \begin{matrix} 2\boldsymbol{P}_{ij,t} \\ 2\boldsymbol{Q}_{ij,t} \\ \boldsymbol{L}_{ij,t} - \boldsymbol{u}_{i,t} \end{matrix} \right\|_2 \le \boldsymbol{L}_{ij,t} + \boldsymbol{u}_{i,t}, \forall (\boldsymbol{i},\boldsymbol{j}) \in \chi, \tag{12}$$

$$\boldsymbol{P}_{PV,j,t} - \boldsymbol{P}_{L,j,t} - \boldsymbol{P}_{ESS,j,t} = \sum_{l\in\Phi(j)} \boldsymbol{P}_{jl,t} - \sum_{i\in\Lambda(j)} (\boldsymbol{P}_{ij,t} - \boldsymbol{r}_{ij}\boldsymbol{L}_{ij,t}), \forall \boldsymbol{j} \in \Pi_K^P, \tag{13}$$

$$\boldsymbol{P}_{PV,j,t}^* - \boldsymbol{P}_{L,j,t} - \boldsymbol{P}_{ESS,j,t} = \sum_{l\in\Phi(j)} \boldsymbol{P}_{jl,t} - \sum_{i\in\Lambda(j)} (\boldsymbol{P}_{ij,t} - \boldsymbol{r}_{ij}\boldsymbol{L}_{ij,t}), \forall \boldsymbol{j} \in \Pi_K^P, \tag{14}$$

$$\boldsymbol{Q}_{PV,j,t} - \boldsymbol{Q}_{L,j,t} = \sum_{l\in\Phi(j)} \boldsymbol{Q}_{jl,t} - \sum_{i\in\Lambda(j)} \left(\boldsymbol{Q}_{ij,t} - \boldsymbol{x}_{ij}\boldsymbol{L}_{ij,t}\right), \forall \boldsymbol{j} \in \Pi, \tag{15}$$

$$\left(\boldsymbol{U}^{min}\right)^2 \le \boldsymbol{u}_{i,t} \le (\boldsymbol{U}^{max})^2, \forall \boldsymbol{j} \in \Pi, \tag{16}$$

$$\boldsymbol{L}_{ij,t} \le \left(\boldsymbol{I}_{ij}^{max}\right)_2, \forall (\boldsymbol{i},\boldsymbol{j}) \in \chi. \tag{17}$$

In the aforementioned constraints, **Eq. 12** represents the second-order cone relaxation transformation of branch power flow, where $L_{ij,t}$ represents squared current magnitude of line i-j at time t; $P_{ij,t}$ and $Q_{ij,t}$ represent the active power and reactive power of the line from node i to node j at time t, respectively; $U_{i,t}$ represents the voltage amplitude on node i at time t; $u_{i,t}$ represents the square of the voltage amplitude on node i at time t; and χ represents all line sets. **Eqs 13–15** represent the injection power balance equation of active and reactive nodes, where $\Phi(j)$ and $\Lambda(j)$ represent the parent branch set and child branch set of node j, respectively; $P_{PV,j,t}$ represents the adjusted active power generation of the j-th PV node at time t; $P_{PV,j,t}^*$ represents the output active power of the j-th PV at time t, which is the constant value; $Q_{PV,j,t}$ represents the reactive power regulation amount of the PV inverter of the j-th PV at time t, which is the constant value; Π represents a collection of all nodes; Π_K^P represents the collection of all nodes in cluster K; $P_{L,j,t}$ and $Q_{L,j,t}$ represent the active and reactive load demand of the j-th node at time t; $P_{ESS,j,t}$ represents the energy storage absorbed power at the j-th node at time t; and r_{ij} and x_{ij} represent the resistance and reactance of line i-j. **Eqs 16**, **17** represent the node voltage amplitude and branch current constraints, where I_{ij}^{max} represents the maximum current of line i-j.

2) Upper and lower limits of PV output:

$$\mathbf{0} \le \boldsymbol{P}_{PV,i,t} \le \boldsymbol{P}_{PV,i,t}^{max}, \tag{18}$$

where $P_{PV,i,t}^{max}$ is the upper limit value of active power output of PV i at time t.

3) Energy storage state constraints:

$$\boldsymbol{S}_{OC,min} \le \boldsymbol{S}_{OC,t} \le \boldsymbol{S}_{OC,max}, \tag{19}$$

$$-\boldsymbol{P}_{ESS,N} \le \boldsymbol{P}_{ESS,t} \le \boldsymbol{P}_{ESS,N}, \tag{20}$$

where $S_{OC,t}$ is the energy storage state of charge at time t; $S_{OC,max}$ and $S_{OC,min}$ are the upper and lower limits of the energy storage state of charge, respectively; $P_{ESS,t}$ is the energy storage power at time t; and $P_{ESS,N}$ is the rated power of energy storage.

4 ITERATIVE OPTIMIZATION-BASED ACTIVE POWER COORDINATION STRATEGY OF CLUSTERS

To realize the coordination among clusters, an iterative optimization-based active power coordination strategy among clusters is proposed, which takes advantage of the weak coupling between clusters. The specific process of the proposed strategy is shown in **Figure 1**.

Step 1: Based on the cluster partition, check whether the node voltage in each cluster exceeds the voltage limit. If the node voltage is within the normal range, it ends; if there exists overvoltage, proceed to the next step.

Step 2: The clusters with overvoltage are recorded as $C^V = \{C_1^V, C_2^V, \ldots, C_K^V, \ldots, C_N^V\}$. Then, each cluster in C^V optimizes the objective of **Eq. 11** to obtains the optimal output power of PV $P_{PV}^0 = \{P_{C_1^V,PV}^0, P_{C_2^V,PV}^0, \ldots, P_{C_K^V,PV}^0, \ldots, P_{C_N^V,PV}^0\}$, and the charging power of stored energy $P_{ESS}^0 = \{P_{C_1^V,ESS}^0, P_{C_2^V,ESS}^0, \ldots, P_{C_K^V,ESS}^0, \ldots, P_{C_N^V,ESS}^0\}$.

Step 3: Based on the intra-cluster optimal solution P_{PV}^0 and P_{ESS}^0, a power flow calculation is carried out for the DN after the

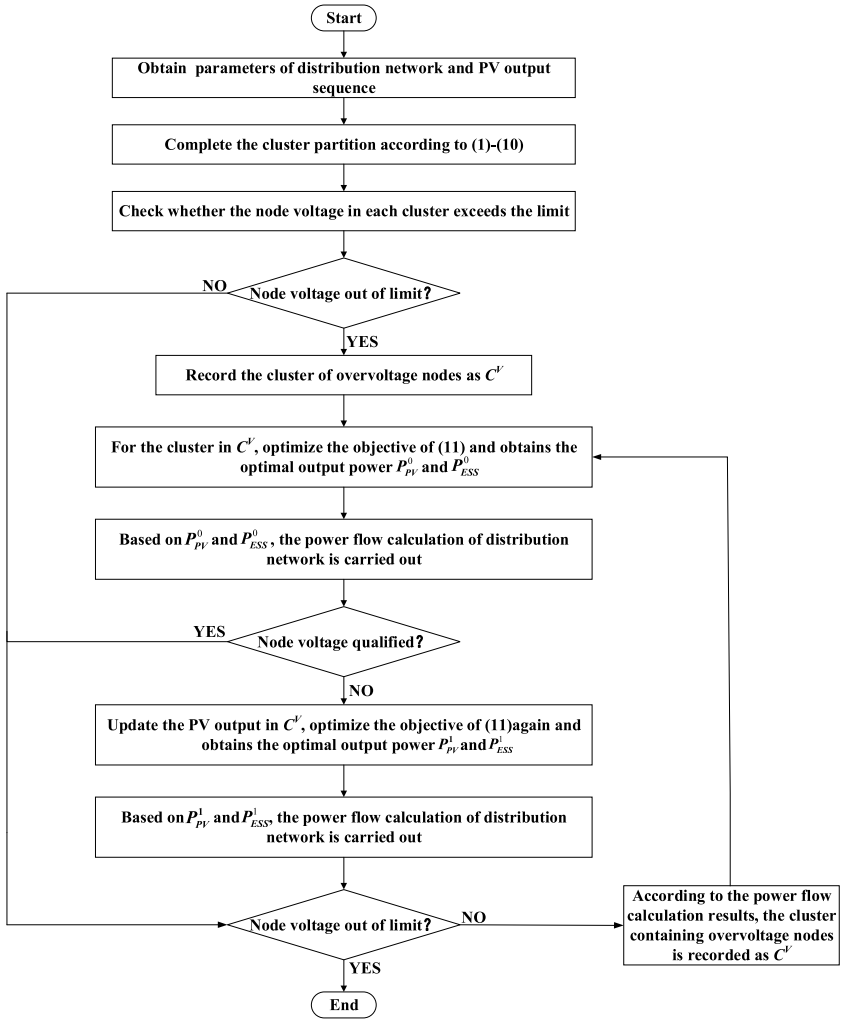


FIGURE 1 | Flowchart of the proposed voltage control.

active power cutting of the PV in cluster C^V. When the voltage of the whole network is qualified, it ends. Otherwise, go to the next step.

Step 4: Taking P^0_{PV} as the initial value, each cluster in C^V optimizes the objective of **Eq. 11** again to obtains the best output power of PV $P^1_{PV} = \{P^1_{C^V_1,PV}, P^1_{C^V_2,PV}, \ldots, P^1_{C^V_K,PV}, \ldots, P^1_{C^V_N,PV}\}$, and the charging power of stored energy $P^1_{ESS} = \{P^1_{C^V_1,ESS}, P^1_{C^V_2,ESS}, \ldots, P^1_{C^V_K,ESS}, \ldots, P^1_{C^V_N,ESS}\}$.

Step 5: Perform a power flow calculation based on the optimized results in step 4. If the voltage of the whole network is qualified, it ends; otherwise, proceed to the next step.

Step 6: Repeat steps 2-5 until all node voltages are within the normal range.

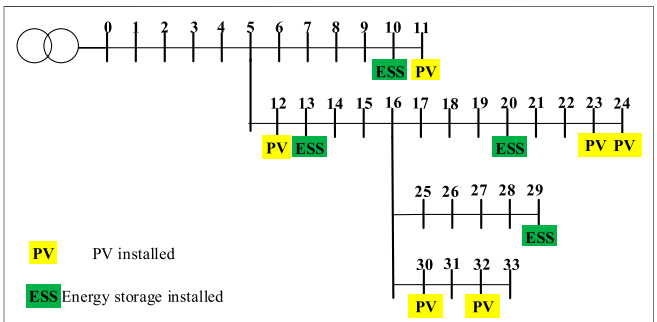


FIGURE 2 | Topology of 10 kV feeder under study.

TABLE 1 | Capacity of installed PVs.

Node location	11	12	23	24	30	32
Capacity (kW)	500	500	500	500	500	1,500

TABLE 2 | Capacity of installed energy storage.

Node location	10	13	20	29
Capacity (kW)	300	300	300	1,000

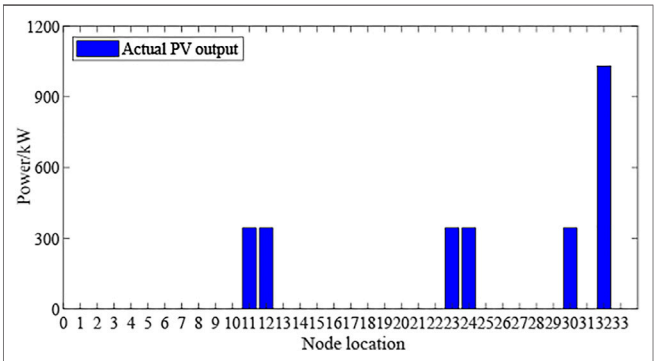


FIGURE 3 | Daily maximum outputs of PVs.

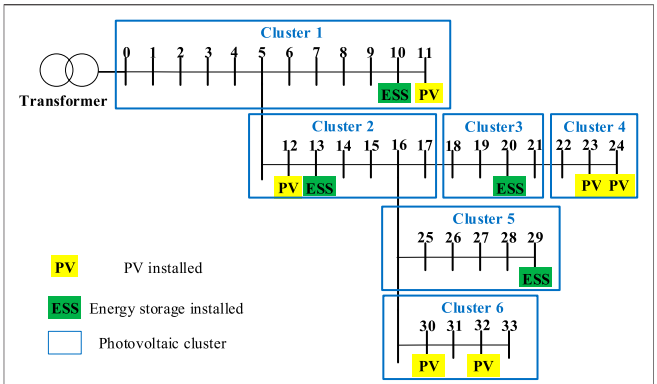


FIGURE 4 | Cluster partition results under the proposed method.

TABLE 4 | Results of cluster partition under different indexes.

Cluster	Node	
	ω	ρ
1	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11
2	12, 13, 14, 15, 16, 17	12, 13
3	18, 19, 20, 21	14, 15, 16, 17
4	22, 23, 24	18, 19, 20, 21
5	25, 26, 27, 28, 29	22, 23, 24
6	30, 31, 32, 33	25, 26
7		27, 28, 29
8		30, 31, 32, 33

5 CASE STUDY

5.1 Case Study System

In order to verify the effectiveness of the proposed method, an actual three phase balanced 10 kV radial active distribution network in Zhejiang Province, China, was employed. The feeder has 34 nodes in total. Among them, bus 0 is employed as the reference node, its voltage value is set to 1.05 p.u. The total load in the feeder is 7.3MVA, and a total of 4 MW PVs and 1.2 MW energy storage are added to the network. The topology of the feeder is shown in **Figure 2**.

In this network, the capacity of installed PVs and energy storages are shown in **Tables 1**, **2**, respectively.

First, the feeder is modeled by OpenDSS simulation platform, and then the cluster partition and the zonal voltage control are carried out under MATLAB. In order to verify the effectiveness of the proposed method, the day with the largest PV output in 2021 is taken as a typical scenario to be analyzed. The daily maximum PV output is shown in **Figure 3**.

5.2 Cluster Partition of the Feeder

Different partition results can be obtained under the comprehensive performance index with different weights. In the cluster partition, different weights can be set for each index according to the experimental conditions. In this study, three weight combinations (numbered C1, C2, and C3) are employed to illustrate the influence of a comprehensive performance index on cluster partition. **Table 3** shows the cluster partition results obtained by different weight combinations.

It can be seen from **Table 3** that as a certain weight creasing, the influence of the corresponding index will be enhanced in the cluster partition. For example, from the weight combination C1 to C2, the weight ω_1 becomes smaller; meanwhile, the global density function index becomes smaller. It can be known that the value of ω_1 in **Eq. 10** will be smaller, which means the influence of the global density function index is reduced. Therefore, the weights ω_1, ω_2, and ω_3 should be set according to the experimental

TABLE 3 | Results of cluster partition under different weight combinations.

Number	ω_1	ω_2	ω_3	Q_{GD}	X	M	ω	Number of clusters
C1	0.4	0.3	0.3	0.8639	0.6637	0.7233	0.7617	6
C2	0.3	0.4	0.3	0.8322	0.6866	0.7016	0.7308	7
C3	0.3	0.3	0.4	0.8148	0.6748	0.7039	0.7311	7

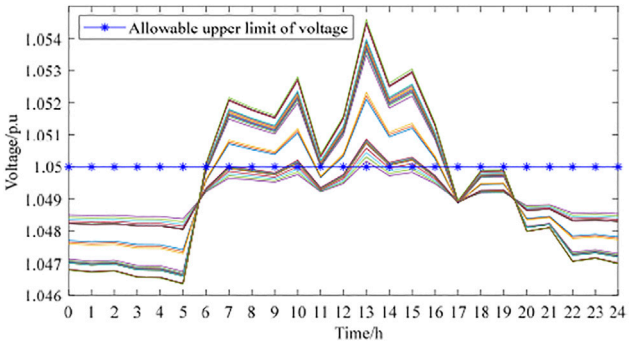


FIGURE 5 | Voltage profile of 10 kV feeder during a day.

TABLE 5 | Actual outputs of PVs under the proposed method.

PV	11	12	23	24	30	32
Outputs (kW)	343.6	343.6	343.6	343.6	343.6	1,030.9

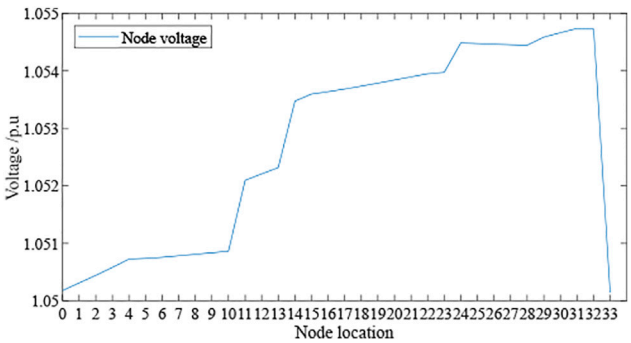


FIGURE 6 | Voltage profile of 10 kV feeder at 13:00.

conditions. According to the actual needs in this study, the cluster partition is analyzed by setting $\omega_1 = 0.4$, $\omega_2 = 0.3$, and $\omega_3 = 0.3$. The cluster partition results are shown in **Figure 4**.

It can be seen from **Figure 4** that the feeder is partitioned into 6 clusters using the proposed cluster partition method, and there are no isolated nodes exist. Meanwhile, each cluster has the PVs or energy storage, which means each cluster has an adjustable capacity for zonal voltage. To further verify the effectiveness of the proposed method, the modular degree, function ρ (Li and Yang, 2019) is employed to compare with the comprehensive performance index. **Table 4** shows the cluster partition results under the two different indexes.

From **Table 4**, it can be seen that the feeder can be partitioned into several clusters under the two indexes, and there are no isolated nodes exist. But it should be noticed that under the function ρ, two clusters have no PVs or energy storage, and the nodal size of the cluster is more unbalanced

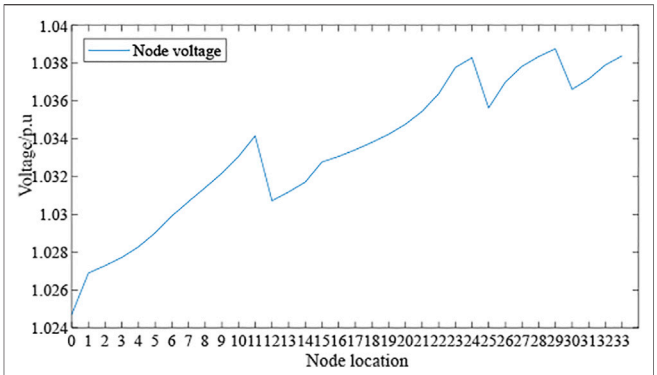


FIGURE 7 | Optimized nodal voltage under the proposed method.

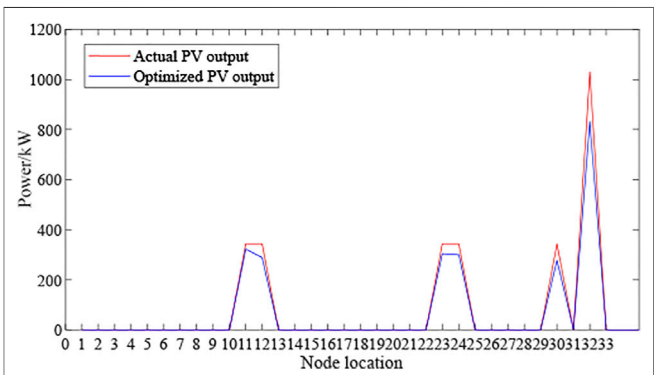


FIGURE 8 | PV outputs under different scenarios.

than the clusters obtained by comprehensive performance index, which means the proposed a comprehensive performance index can generate a more reasonable cluster partition for subsequent zonal voltage.

5.3 Cluster Partition-Based Zonal Voltage Control

Without any voltage regulation measures, the nodal voltage of the feeder is shown in **Figure 5**. It can be seen from **Figure 5** that the voltage of some nodes exceeds the limit at 13:00 noon. To illustrate that the proposed method can effectively adjust the node voltage to a safe operating range, the operating condition of the feeder at 13:00 noon is selected as a typical scenario to be analyzed. In typical scenarios, the actual output of each PVs is shown in **Table 5**.

In the typical scenario, the solar irradiance comes to the largest during a day, and some nodal voltages exceed the allowable upper limit, which is shown in **Figure 6**. Under the proposed method, the cluster containing overvoltage nodes is recorded as $C^V = \{C_1^V, C_2^V, C_3^V, C_4^V, C_5^V, C_6^V\}$. Under the proposed voltage control method, the profile of nodal voltage and PV outputs are shown in **Figures 7**, **8**, respectively.

It is shown in **Figure 7** that under the proposed voltage strategy, there is no overvoltage existing in the whole network. Meanwhile, **Figure 8** shows that the minimum curtailed PV output is 21.34 kW and the maximum is 198.94 kW, which achieves the goal of maximum PV consumption and avoids the waste of resources.

6 CONCLUSION

Aiming at the overvoltage problem caused by the high penetrated PVs, a cluster partition-based zonal voltage control method for the DN is proposed in this study. Considering the structure of DNs, output characteristics of distributed PV in clusters, and the cluster scale, a comprehensive performance index-based cluster partition method is proposed in this study. Based on the community discovery algorithm, the FN algorithm is employed to carry out the partition. The electrical information of the network structure and the impact of PVs on the partitioning solutions are considered, and the size of each subnetwork can be restricted. Based on the cluster partition, an SOCP-based voltage control model aiming at the maximum PV consumption in the cluster is established, and an iterative optimization-based active power coordination strategy among clusters is proposed. The maximum PV consumption for all clusters can be guaranteed, and the overvoltage problem caused by high penetrated PVs can be effectively solved.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

XZ: responsible for resources, software, formal analysis, program compilation, and writing—original draft. CX: responsible for writing—review and editing. KP: responsible for methodology, project administration, and funding acquisition. JC: responsible for obtaining the experimental data. XZ: responsible for investigation and resources.

FUNDING

This study was sponsored by the National Science Foundation for Young Scientists of China (51807026) and Science and Technology Project of Guangzhou Power Supply Bureau of Guangdong Power Grid Corporation (GZHKJXM20190061).

REFERENCES

Bi, R., Liu, X. F., Ding, M., Fang, H., Zhang, J. J., and Chen, F. (2019). Renewable Energy Generation Cluster Partition Method Aiming at Improving Accommodation Capacity. *Proc. CSEE* 39 (22), 6583–6592. CNKI:SUN:ZGDC.0.2019-22-010.

Chai, Y., Guo, L., Wang, C., Zhao, Z., Du, X., and Pan, J. (2007). Network Partition and Voltage Coordination Control for Distribution Networks with High Penetration of Distributed PV Units. *IEEE Trans. Power Syst.* 33, 3396–3407. doi:10.1109/TPWRS.2018.2813400

Chai, Y. Y., Liu, Y. X., Wang, C. S., Guo, L., Zhao, Z. Z., and Gao, S. (2019). Coordinated Voltage Control for Dis-Tributed PVs Clusters with Incomplete Measurements. *Proc. CSEE* 39 (08), 2202–2212+3. doi:10.13334/j.0258-8013.pcsee.182485

Cohen, J. (2009). Graph Twiddling in a MapReduce World. *Comput. Sci. Eng.* 11 (4), 29–41. doi:10.1109/mcse.2009.120

Ding, M., Wang, W. L., Wang, X. L., Song, Y. T., Chen, D. Z., and Sun, M. (2014). A Review on the Effect of Large-Scale Pv Generation on Power Systems. *Proc. CSEE* 34 (01), 1–14. doi:10.13334/j.0258-8013.pcsee.2014.01.001

Dou, X. B., Chang, L. M., Ni, C. H., Duan, X. M., Ge, P. D., and Wu, Z. J. (2018). Multi-level Dispatching and Control of Active Distribution Network for Virtual Cluster of Distributed Photovoltaic. *Automation Electric Power Syst.* 42 (03), 21–31. doi:10.7500/AEPS20170221007

Han, Y., Zhang, K., Li, H., Coelho, E. A. A., and Guerrero, J. M. (2018). MAS-based Distributed Coordinated Control and Optimization in Microgrid and Microgrid Clusters: a Comprehensive Overview. *IEEE Trans. Power Electron.* 33 (3), 6488–6508. doi:10.1109/TPEL.2017.2761438

Hu, X. K., Yin, R., and Shi, M. (2021). Distributed Photovoltaic Cluster Partition and Reactive Power Optimization Strategy Based on Improved Particle Swarm Optimization Algorithm. *Power Capacitor & Reactive Power Compensation* 42 (04), 14–21. doi:10.14044/j.1674-1757.pcrpc.2021.04.003

Kou, L. F., Xu, Y. H., Hou, X. G., and Zhang, F. (2019). Cluster Partition Method for Large Scale Distributed Photovoltaic in Distribution Network. *Renew. Energ. Resour.* 37 (04), 525–530. doi:10.3969/j.issn.1671-5292.2019.04.009

Li, S., Wei, Z. N., Sun, G. Q., and Wang, Z. Q. (2018). Stability Research of Transient Voltage for Multi-Machine Power Systems Integrated Lager-Scale PV Power Plant. *Acta Energiae Solaris Sinica* 39 (12), 3356–3362. CNKI:SUN:TYLX.0.2018-12-010.

Li, X. J., and Yang, J. H. (2019). Research on Realization and Comparison of Community Division Algorithms in Network. *Comput. Digital Eng.* 47 (11), 2861–2865. doi:10.3969/j.issn.1672-9722.2019.11.043

Liao, Q. P., Lyu, L., Liu, Y. B., Zhang, Y., and Xiong, J. (2017). Reconfiguration Based Model and Algorithm of Voltage Regulating for Distribution Network with Renewable Energy. *Automation Electric Power Syst.* 41 (18), 32–39. doi:10.7500/AEPS20170101003

Liu, X. F. (2019). *Study on Cluster Partition Method of High Permeability Distributed Renewable Energy Power Generation*. Master's thesis (Hefei, China: Hefei University of Technology).

Peng, Y., Xiong, W., Yuan, X. F., Zhou, X. S., Shuai, S. X., and Zhao, Z. (2021). Research on Optimal Power Flow of Active Distribution Network Based on Mixed Integer Second-Order Cone Programming. *Electr. Meas. Instrumentation* 2021, 1–7. Available at: https://kns.cnki.net/kcms/detail/23.1202.TH.20201124.1033.015.html (Accessed 12 17, 2021).

Wang, L., Zhang, F., Kou, L. F., Xu, Y. H., and Hou, X. G. (2021). Large-scale Distributed PV Cluster Division Based on Fast Unfolding Clustering Algorithm. *Acta Energiae Solaris Sinica* 42 (10), 29–34. doi:10.19912/j.0254-0096.tynxb.2018-0896

Xiao, C. L. (2020). *Cluster Partition Based Voltage Control and Optimal Scheduling for Active Distribution Networks*. Doctor's thesis (Hefei, China: Hefei University of Technology).

Xiao, C. L., Zhao, B., Zhou, J. H., Li, P., and Ding, M. (2017). Network Partition Based Cluster Voltage Control of High-Penetration Distribution Photovoltaic Systems in Distribution Net-Works. *Automation Electric Power Syst.* 41 (21), 147–155. doi:10.7500/AEPS20170101002

Yan, H. D., Cai, L. H., Sha, J., Lyu, Z. P., and Xu, Z. (2021). Distribution Network Cluster Division and Voltage Control with High Proportion of New Energy Access. *Power Demand Side Manage.* 23 (04), 8–13. doi:10.3969/j.issn.1009-1831.2021.04.003

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Energy Research

ORIGINAL RESEARCH
published: 26 May 2022
doi: 10.3389/fenrg.2022.892066

Secure Transmission and Intelligent Analysis of Demand-Side Data in Smart Grids: A 5G NB-IoT Framework

*Yongpeng Shen[1], Ting He[1], Qian Wang[1], Junmin Zhang[2] and Yanfeng Wang[1]**

[1]College of Electrical and Information Engineering, Zhengzhou University of Light Industry, Zhengzhou, China, [2]Pinggao Group Intelligent Power Technology Co., Ltd., Pingdingshan, China

In order to tap the advantages of Narrow Band-Internet of Things (NB-IoT) on the demand side of the smart grid, improve the security of information transmission, and make use of demand-side data, we focus on the secure transmission, trusted services, and intelligent analysis of "5G + Smart Grid," and we construct a comprehensive solution which consists of encrypted data terminals, management systems, and intelligent analysis methods. A 5G NB-IoT framework proposed in this study can serve grid planning and demand response, and it can further promote the deep integration of 5G and the smart grid. Therefore, this research will contribute to the implementation of a new generation of information technology on the smart grid, build a deep integration application scenario of the "5G + Smart Grid," improve the intelligence of the grids, and further promote the "dual carbon" goal of the power system.

Keywords: NB-IoT, the smart grid, security transmission, demand-side data, information technology

OPEN ACCESS

Edited by:
Yongxi Zhang,
Changsha University of Science and Technology, China

Reviewed by:
Yu Wang,
Guangdong University of Technology, China
Yitao Liu,
Shenzhen University, China

****Correspondence:***
Yanfeng Wang
wangyanfeng@zzuli.edu.cn

Specialty section:
This article was submitted to Smart Grids, a section of the journal Frontiers in Energy Research

Received: *08 March 2022*
Accepted: *14 April 2022*
Published: *26 May 2022*

Citation:
Shen Y, He T, Wang Q, Zhang J and Wang Y (2022) Secure Transmission and Intelligent Analysis of Demand-Side Data in Smart Grids: A 5G NB-IoT Framework. Front. Energy Res. 10:892066. doi: 10.3389/fenrg.2022.892066

1 INTRODUCTION

The smart grid is an electricity network that can cost-efficiently integrate the behavior and actions of all users, including generators, consumers, and those that both generate and consume, in order to ensure the power system is economically efficient and sustainable with low losses and high levels of quality, security of supply, and safety (TEN-E, 2017). The smart grid uses the IoT technology to add intelligence and monitoring to every node. The applications of the smart grid can balance the flow of power more efficiently. They can detect surges, outages, and energy waste. They can also deal with peak loads or fluctuations immediately and automatically.

With large-scale access to renewable energy sources such as wind, solar, and electric vehicles, and distributed access to demand-side energy storage, the power system is changing dramatically. The power system is becoming more diverse in terms of forms of energy supply, massive amounts of data, and ways of interaction and control.

The demand side is the "nerve ending" of the smart grid, which has the characteristics of a large number of terminals and extensive connection of equipment. A safe, reliable, and intelligent demand side is the cornerstone of the power system. Two-way communication among generators, transmitters, and customers is the key to the smart grid. This mutual intelligent system offers solid benefits, including energy management, reliability and resilience, and the integration of intermittent renewable energy generation and storage. It also accommodates distributed power generation and microgrids, enhances the value of electric vehicles, and gives customers greater choices of how and when to use electricity (IIOT Power, 2019; Xu et al., 2020a; Xu et al., 2020b; Yuan et al., 2021).

The downside of the smart grid is how to interact with each other between those interconnections—driven by IoT technologies, data flow, and information management. In terms

of its nature, do the disadvantages offer cybersecurity threats, opportunities for malevolent forces to intrude, disrupt, or destroy? (IIOT Power, 2019).

Most of the research studies on demand-side information security by scholars belong to the scope of a home area network (HAN) and a neighborhood area network (NAN) in the advanced metering infrastructure (AMI) framework, which can be divided into two categories: one does not use any encryption means, but other methods are used to ensure the security of demand-side information, such as connecting batteries and other equipment on the user side, or adding noise to the demand-side information artificially (Yuan et al., 2020; Tan et al., 2013; Lang et al., 2022; Chen et al., 2013). The aforementioned methods ensure the safety of the user's personal information but reduce the usability of the information. Lu et al (2012) proposed a privacy protection and data aggregation scheme based on homomorphic encryption, which can prevent the leakage of user privacy during smart grid communication. Abdallah and Shen (2015) proposed a lattice-based number theory research unit (NTRU) public key cryptosystem with low computing resource consumption, which enabled the use of a more secure public key cryptosystem in the demand-side network while still achieving low computing resource consumption. Nicanfar et al (2012) proposed an identity-based public key encryption system to construct an identity authentication strategy suitable for a HAN. Jokar et al (2011) proposed an intrusion detection scheme at the physical layer and 802.15.4 Media Access Control (MAC) layer to determine whether the HAN has been invaded by detecting signal strength, data size, format, and flow direction. Sanduleac and Ciornei (2021) proposed a general framework for extracting technical knowledge from high reporting rate smart meters (HRRSM) data to strengthen distribution system operator (DSO) monitoring tools to protect the privacy of users.

Through the analysis of the aforementioned secure transmission methods, the following conclusions can be drawn.

1) At present, the demand side mostly adopts the neighborhood area network and local area network framework in the AMI framework, and the advantages of the 5G Massive Machine-Type Communication Wide Area Network (mMTC WAN) have not yet emerged.
2) The current research studies only focus on the efficiency and structure of transmission but have not paid enough attention to the issue of information security transmission in the context of a two-way interaction.
3) How to use intelligent approaches to increase the value of demand-side data requires urgent attention.

In response to the aforementioned problems, we focus on the secure transmission, trusted services, and intelligent analysis of the "5G + Smart Grid" and construct an all-round solution consisting of encrypted data terminals, management systems, and intelligent analysis methods. The proposed "5G + Smart Grid" framework can service grid planning, demand response, and promote the deep integration of 5G and the smart grid.

The rest of this article is organized as follows. In **section 2**, the features of the demand side of the smart grid are analyzed. **Section 3** describes the overall technical framework of the proposed 5G NB-IoT secure transmission and intelligent analysis of demand-side data in the smart grid. **Sections 4–6** describe the secure transmission system, trusted service management system, and data intelligent analysis system in detail, respectively. Finally, the conclusion is stated in **section 7**.

2 ANALYSIS OF THE DEMAND SIDE IN THE SMART GRID

We consider the demand side in the smart grid that contains a distributing substation, distributed power source, distributed energy storage, industrial electricity supply, and residential electricity supply, as shown in **Figure 1**. Sometimes, the concepts of the distributed power source and distributed energy storage are confused. For example, electric vehicles are both distributed power sources and distributed energy storage. Photovoltaic power stations as distributed power sources may also be equipped with energy storage, to become distributed energy storage.

Generally, each distributed power source, distributed energy storage, industrial electricity customer, and residential electricity customer is connected to the distribution network through a smart meter (SM). In industrial application scenarios, various pieces of electrical equipment are connected to the mains through smart sockets (SS), such as lighting equipment, charging equipment, hoister, blowers, heaters, refrigerators, air conditioners, water heaters, cookers, and washing machines. The Intelligent Distribution Trans-former supervisory Terminal Unit (iTTU) is responsible for monitoring the working status of transformers and the entire distribution network.

In the next-generation power distribution network, each iTTU, SM, and SS is responsible for monitoring the voltage, current, active power, reactive power, and other information of relevant nodes. They must have communication functions to transfer the grid data to the data center in real time for further analysis. The basic features of the next-generation demand side in the smart grid can be summarized as follows.

1) Large number of nodes.

As mentioned previously, on the demand side of the smart grid, each node will have at least one SM. Each electrical apparatus will be connected to the mains through SS. It will result in a user having dozens of data nodes, while a power distribution network will have thousands of data nodes.

2) Many application scenarios.

The application scenarios on the user side include a distributing substation, distributed power source, distributed energy storage, industrial electricity supply, residential electricity supply, smart power distribution, and others. At the same time, it covers multiple stakeholders, such as equipment manufacturers, telecom operators, and users. Once a data leakage

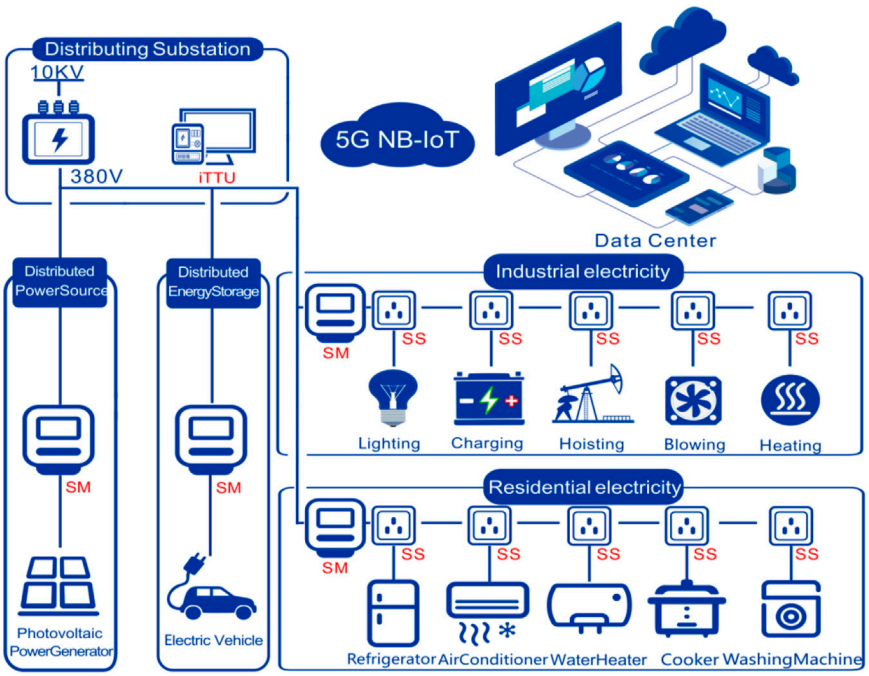


FIGURE 1 | Demand side in the smart grid.

problem occurs, it is difficult to clarify the security responsibilities.

3) Fast iteration speed and emerging security upgrade risks.

At present, the communication architecture of the smart grid has not yet been unified, the business and profit models have not yet been clarified, and the communication protocol between the various entities has not yet formed a unified standard; meanwhile, with the rapid development of the smart grid, new things and new models are constantly emerging. Therefore, smart grid data terminals must have high-security upgrade management functions and convenient software iteration speed to adapt to the new models and new requirements.

In view of the aforementioned characteristics of the demand side of the smart grid, combined with its basic characteristics, the smart grid data transmission system should have the following attributes.

1) Low power consumption and massive connection characteristics to facilitate massive deployment in the area and long-term reliable operation in battery-powered scenarios.
2) Complete the perception layer security functions such as physical data security, trusted service management, security upgrade management, and key life cycle management.
3) A complete network layer security solution adopts a unified communication system to facilitate unified deployment on a global scale and has a complete IoT card security management chain.
4) Access stratum (AS) and non-access stratum (NAS) network layer security features.

3 OVERALL TECHNICAL FRAMEWORK

NB-IoT is a new and streamlined IoT terminal communication technology proposed by 3 GPP R13 on the basis of long-term evolution (LTE) technology. It is a narrowband wireless cellular communication technology specially designed for the IoT to achieve the requirements of large connection, wide coverage, low power consumption, etc. From the perspective of basic features and network layer security features, NB-IoT highly matches the needs of the demand side in the smart grid. The basic characteristics of NB-IoT can be summarized as follows.

1) Deep coverage.

NB-IoT provides better deep coverage than other competing technologies. It has a high transmit power spectral density (PSD). In the downlink standalone mode, PSD is 43 dBm/180 kHz, which is 17 dB higher than that of LTE; in the in-band mode, PSD is 35 dBm/180 kHz, which is 9 dB higher than LTE. A multiple retransmission technology is employed to increase hybrid automatic repeat request (HARQ) gain and exchange coverage gain at a lower rate, and the maximum downlink retransmission and uplink retransmission are 2,048 and 128 times, respectively. In general, the coverage radius of NB-IoT is about 4 times that of GSM/LTE (Ratasuk et al., 2016; Martinez et al., 2019).

2) Ultralow power consumption.

NB-IoT has a power saving mode (PSM) and an extended discontinuous reception (eDRX) mode, which greatly reduce the power consumption of the module. In the PSM, the terminal is still registered on the network but cannot obtain the signal, so the terminal maintains a long time of deep sleep to save power. In the eDRX mode, the idle mode discontinuous reception cycle ranges from seconds to hours, and the connection mode discontinuous reception cycle supports 5.12 and 10.24 s, which greatly improves the downlink reachability during a low-power operation. Ultralow power consumption can ensure that a battery-powered NB-IoT terminal has a long service life of up to 10 years (depending on the specific application) (Wang et al., 2017; Kanj et al., 2020).

3) Ultralow cost.

By simplifying the protocol stack, the radio frequency circuit, and the complexity of baseband processing, NB-IoT does not require a duplexer, so out-of-band and blocking indicators are reduced. For example, it reduces baseband complexity and radio frequency circuits by 10% and 65%, respectively (Hoglund et al., 2020; Ballerini et al., 2020; Li et al., 2017). At present, the cost of NB-IoT modules has fallen below 15 yuan.

4) Massive connections.

NB-IoT adopts the narrowband technology and upper/lower equivalent power enhancement technology, thereby greatly increasing the channel capacity. NB-IoT improves spectral density by reducing the signaling overhead of the air interface. Through the optimization of the base station and the core network, NB-IoT realizes independent admission congestion control, downlink data buffering, and terminal context storage, which can achieve 50,000 connections/cell (Zayas et al., 2017; Sultania et al., 2020).

In terms of network layer security, NB-IoT has the following characteristics:

1) For terminals that support both the control plane optimized transmission scheme and the user plane optimized transmission scheme, NB-IoT adopts two layers of security mechanisms: AS and NAS. The former ensures Radio Resource Control (RRC) security and user plane security in the access network. The latter ensures NAS security in Evolved Packet Core (EPC).
2) By defining access safe management entity (ASME), NB-IoT realizes that the access network receives the highest level key from a home subscriber server (HSS). In addition, NB-IoT constructs a four-layer key structure, including terminal and HSS shared keys, terminal and ASME shared keys, terminal and mobility management entity (MME) shared keys, and terminal and base station shared keys.
3) The security activation of the access layer and the non-access layer is completed through security mode control (SMC), and the security of the access layer can be reactivated through the RRC connection re-establishment process and the RRC connection recovery process.
4) By using the integrity protection key and integrity check code (consisting of count value, bearer identification, upstream and downstream direction indication, and data content), a data integrity protection mechanism is constructed to ensure data integrity (Lu et al., 2020).

The basic characteristics and network layer security characteristics of NB-IoT show that it has the ability for large-scale deployment in the demand side of the smart grid. However, the basic characteristics of the energy IoT, such as the large number of nodes, many application scenarios, fast iteration speed, and constant risks of security upgrade, also put forward strict security requirements for NB-IoT module hardware, module software, and IoT card management.

Focusing on the secure transmission and intelligent analysis of massive data on the demand side of the smart grid and based on application scenarios such as a distributing substation, distributed power source, distributed energy storage, industrial electricity supply, and residential electricity supply, this study constructs a technical framework of "5G NB-IoT + Smart Grids," as shown in **Figure 2**. The framework includes data secure transmission, trusted service management of data transmission, and the intelligent analysis of demand-side data based on machine learning. The technical details of the three aspects will be described in detail in the following sections.

4 SECURE TRANSMISSION SYSTEM

The smart grid demand-side data secure transmission system consists of two parts, the NB-IoT encrypted data terminal and the data transmission management system. As shown in **Figure 3**, the NB-IoT smart grid encrypted data terminal hardware takes the NB-IoT security module as the core, supplemented by peripheral circuits such as data acquisition and power management. The NB-IoT smart grid data transmission system is composed of a smart grid data terminal on the demand side, an NB-IoT base station of a telecom operator, a smart grid data server and application server, as well as a visualization platform, a Web client, and a mobile app, as shown in **Figure 4**.

The core of the smart grid demand-side data secure transmission system lies in the NB-IoT security module, which consists of the highly integrated SoC of NB-IoT R16/R17, SE-SIM (ESAM), and SE-SIM (eSIM). The software of the NB-IoT security module consists of AES/3DES, SDK software, RTOS software, eSIM application software, and eSIM OS. Among these, Advanced Encryption Standard (AES) and triple data encryption standard (3DES) are two standards in present data encryption. AES is a new encryption using an alternative replacement network, while 3DES is only an adaptation of the old DES encryption relying on a balanced Feistel network.

5 TRUSTED SERVICE MANAGEMENT PLATFORM

The main function of the trusted service management platform is to provide cloud management services for data secure

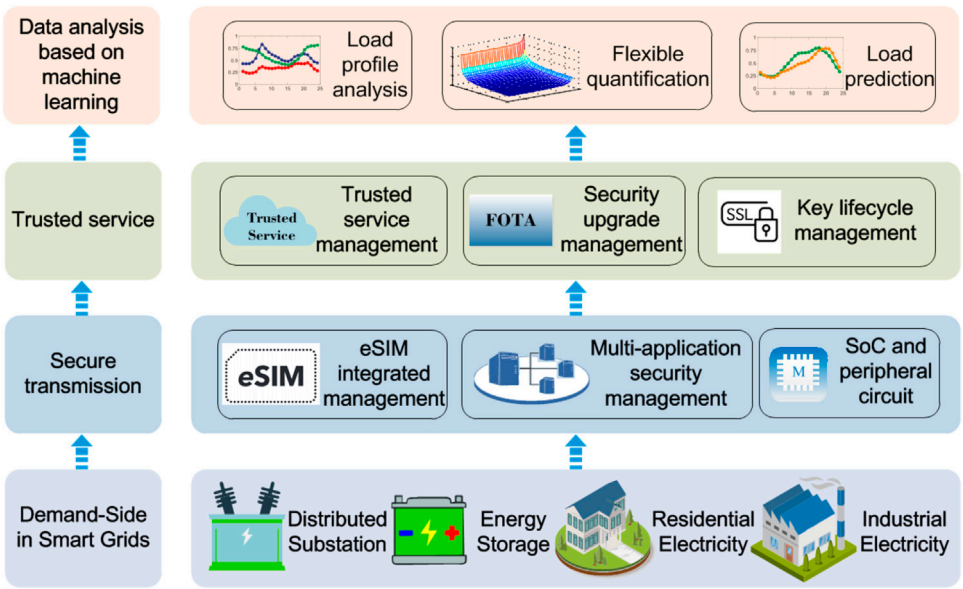


FIGURE 2 | Overall technical framework of secure transmission and intelligent analysis of demand-side data in smart grids based on NB-IoT.

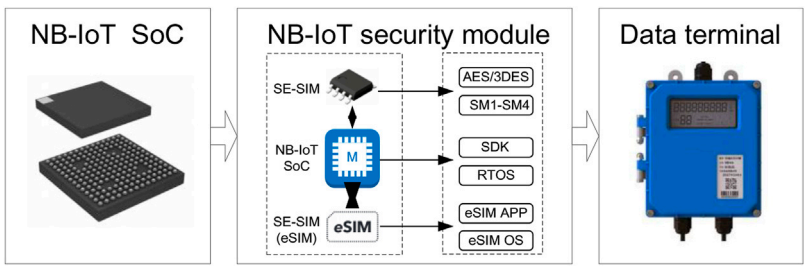


FIGURE 3 | Hardware structure of the NB-IoT encrypted data terminal.

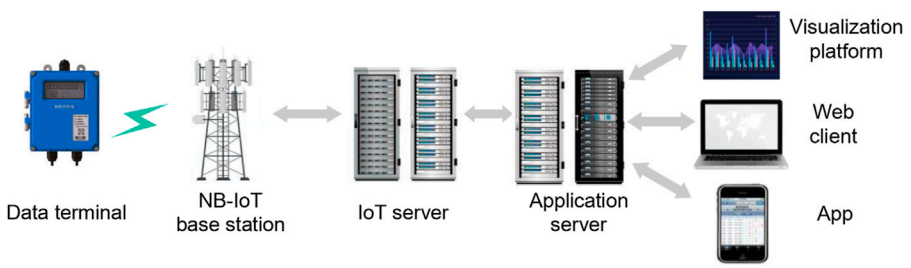


FIGURE 4 | Structure of a secure transmission system.

transmission systems in the demand side of the smart grid. Its core technologies include trusted service management, security upgrade management, and key lifecycle management, as shown in **Figure 5**.

1) Trusted service management.

Trusted service management is the top level of the whole platform. It is mainly responsible for the function management of the entire platform, including security element management, security domain management, application provider management, application information management, application lifecycle management, security element (SE) life

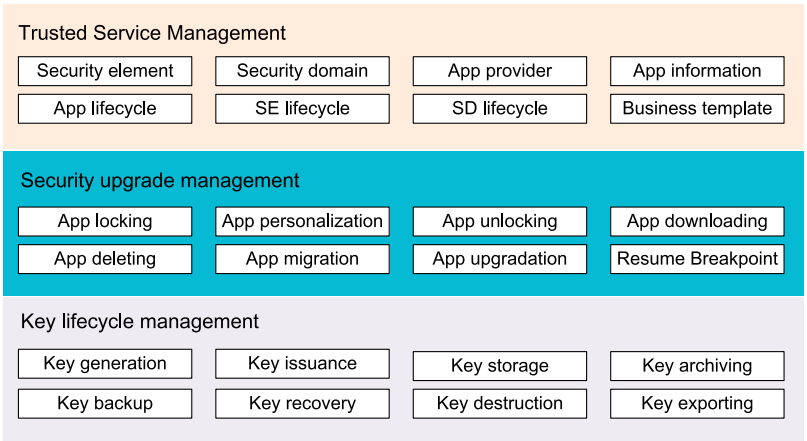


FIGURE 5 | Structure of a trusted service management.

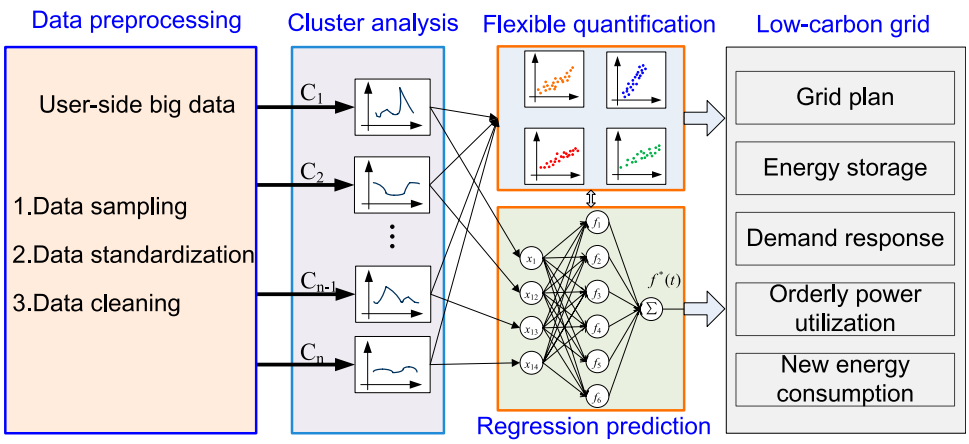


FIGURE 6 | Framework work of data intelligent analysis based on machine learning.

management, security domain (SD) lifecycle management, and business template management.

2) Security upgrade management.

If the software code of the data terminal needs to be upgraded due to functional changes, bug corrections, version updates, and other reasons, the safety of the upgrade process must be ensured. Security upgrade management mainly includes application locking, application personalization, application unlocking, application downloading, application deleting, application migration, application upgradation, and resume breakpoint.

3) Key lifecycle management.

The lifecycle management of keys is a necessary means to ensure data security. The key lifecycle management system includes key generation, key issuance, key storage, key archiving, key backup, key recovery, key destruction, and key exporting.

6 DATA INTELLIGENT ANALYSIS

The overall structure of the demand-side data intelligent analysis system is shown in **Figure 6**. First, pre-processing operations such as sampling, data standardization, and data cleaning are performed on massive data. Then the user load pattern is determined through cluster analysis. On this basis, demand-side flexibility is quantitatively analyzed, and regression prediction is performed on the grid load. The analysis results will serve for grid planning, demand response, new energy consumption, etc.

1) Cluster analysis of daily load.

According to different electricity consumption habits, the load on the demand side presents diverse characteristics. The daily load curve at a specific time interval can be used as the clustering feature. Taking 24 typical load characteristics in different seasons as clustering objectives, the user load characteristics can be curved through cluster analysis.

Assuming that the daily load curve of user i at hour intervals is $\mathbf{R}_i$, a 24-dimensional vector is obtained after normalization, and the Euclidean distance between the daily load data of user i and the typical load characteristic j is

$$d(\mathbf{R}_i, \mathbf{R}_j) = \sqrt{\sum_{t=1}^{24} \left(R_{it} - R_{jt}\right)^2}, \tag{1}$$

where R_{it} and R_{jt} are the loads at t hours of user i and of the typical load characteristic j, respectively. The clustering objective function is the sum of the squares of minimum errors of k clusters, that is,

$$E = \sum_{i=1}^{k} \sum_{R \epsilon C_i} d^2 (\mathbf{R}, \mu(C_i)), \tag{2}$$

where, $\mu(C_i)$ is the cluster center of C_i.

2) Demand-side flexible evaluation.

According to the clustering results of daily load characteristics, through the evaluation of the user's removable and adjustable load, the demand-side flexible quantitative evaluation can be realized (Yuan et al., 2020; Yuan et al., 2021). The daily peak contribution of user i on the jth month is defined as

$$F_{i,j} = \frac{1}{n} \sum_{d=1}^{n} \frac{p_{i,j}^{d}(t_d)}{P_j^{d}(t_d)}, \tag{3}$$

where $p_{i,j}^{d}(t_d)$ is the power demand of user i at time t_d on the jth day, t_d is the time of system peak demand on the dth day of mth month, and $P_j^{d}(t_d)$ is the system peak demand at t_d (Yuan et al., 2020).

3) Regression forecasting of daily load.

Based on the demand-side flexible measurement evaluation results, combined with multi-user daily load characteristics, the neural network regression prediction model of the time series of the daily load can be constructed. The inputs of this model are season, time, demand-side flexibility, user-side demand management, and energy storage ratio.

7 CONCLUSION

Essentially, the smart grid is an IOT-enabled application that allows utilities and their customers to exchange electricity and information and thereby improves energy efficiency. On the demand side, the value of IoT will be more prominent, due to the features such as a large number of nodes, many application scenarios, fast iteration speed, and emerging security upgrade risks. The rapid development of 5G communication technology will provide new power for the development of the smart grid. As the main technology of 5G mMTC, NB-IoT highly matches the needs of the demand side in the smart grid. Currently, the advantages of NB-IoT LPWAN in the demand side of the smart grid have not yet emerged, the issue of information security transmission has not been paid enough attention to, the value of demand-side data has not been fully tapped. Based on the aforementioned reasons, we constructed an all-round solution which consists of encrypted data terminals, management systems, and intelligent analysis methods. The main contributions of this study are as follows:

1) We analyzed the basic features of the next-generation demand side in the smart grid, the requirements of the data transmission system, and the basic features and network layer security features of NB-IoT. We also revealed that NB-IoT has the ability for large-scale deployment in the demand side of the smart grid.
2) We constructed the "5G NB-IoT + Smart Grids" technical framework from aspects of data secure transmission, trusted service management of data transmission, and the intelligent analysis of demand-side data based on machine learning and provided the technical details of the aforementioned three aspects in detail.

This research will contribute to the implementation of new generation information technologies on the smart grid, build a "5G + smart grid" in-depth integration application scenario, improve the intelligence of the grids, and further promote the "dual carbon" goal in power systems.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

Conceptualization, methodology, and writing—original draft: YS; writing, data curation, formal analysis, and visualization: TH; writing—review and editing: QW; resources: JZ; and resources and funding acquisition: YW.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the support of the National Natural Science Foundation of China (Grant Nos 61803345, 51807013, and 52177068); Science and Technology Development Project of Henan Province (222102240005); and Science and Technology Innovation Team in Universities of Henan Province (22IRTSTHN017).

REFERENCES

Abdallah, A., and Shen, X. (2015). Lightweight Security and Privacy Preserving Scheme for Smart Grid Customer-Side Networks[J]. *IEEE Trans. Smart Grid* 99, 1.doi:10.1109/tsg.2015.2463742

Ballerini, M., Polonelli, T., Brunelli, D., Magno, M., and Benini, L. (2020). NB-IoT versus LoRaWAN: An Experimental Evaluation for Industrial Applications. *IEEE Trans. Ind. Inf.* 16 (12), 7802–7811. doi:10.1109/tii.2020.2987423

Chen, Z., and Wu, L. (2013). Residential Appliance DR Energy Management with Electric Privacy Protection by Online Stochastic Optimization. *IEEE Trans. Smart Grid* 4 (4), 1861–1869. doi:10.1109/tsg.2013.2256803

Hoglund, A., Medina-Acosta, G. A., Veedu, S. N. K., Liberg, O., Tirronen, T., Yavuz, E. A., et al. (2020). 3GPP Release-16 Preconfigured Uplink Resources for LTE-M and NB-IoT. *IEEE Comm. Stand. Mag.* 4 (2), 50–56. doi:10.1109/mcomstd.001.2000003

IIOT Power (2019). *The Dark Side of the Smart Grid"*. Available from: https://www.powermag.com/the-dark-side-of-the-smart-grid/.

Jokar, P., Nicanfar, H., and Leung, V. C. M. Specification-based Intrusion Detection for Home Area Networks in Smart Grids[C]. Proceeding of the IEEE Third International Conference on Smart Grid Communications, Oct. 2011, Brussels, Belgium, IEEE, 2011:208–213.

Kanj, M., Savaux, V., and Le Guen, M. (2020). A Tutorial on NB-IoT Physical Layer Design[J]. *IEEE Commun. Surv. Tutor.* 22:2408-2446. doi:10.1109/comst.2020.3022751

Lang, A., Wang, Y., Feng, C., Stai, E., and Hug, G. (2022). Data Aggregation Point Placement for Smart Meters in the Smart Grid. *IEEE Trans. Smart Grid* 13, 541–554. doi:10.1109/TSG.2021.3119904

Li, Y., Cheng, X., Cao, Y., Wang, D., and Yang, L. (2017). Smart Choice for the Smart Grid: Narrowband Internet of Things (NB-IoT)[J]. *IEEE Internet Things J.* 5 (3), 1505–1515. doi:10.1109/jiot.2017.2781251

Lu, R., Liang, X., and Li, X. (2012). EPPA: An Efficient and Privacy-Preserving Aggregation Scheme for Secure Smart Grid Communications. *IEEE Trans. Parallel Distrib. Syst.* 23 (9), 1621–1631. doi:10.1109/tpds.2012.86

Lu, T., Fang, H., Yuan, Y., Dai, B., and Sha, X. (2020). *The Evolution of NB-IoT Protocols—The Road to 5G IoT from R13 to R16[M]*. Beijing: People Post Press, 07.

Martinez, B., Adelantado, F., Bartoli, A., and Vilajosana, X. (2019). Exploring the Performance Boundaries of NB-IoT. *IEEE Internet Things J.* 6 (3), 5702–5712. doi:10.1109/jiot.2019.2904799

Nicanfar, H., Jokar, P., and Leung, V. C. M. (2012). Efficient Authentication and Key Management for the Home Area Network[C]. Proceeding of the IEEE International Conference on Communications, June 2012, Ottawa, ON, Canada, IEEE, 878–882.

Ratasuk, R., Vejlgaard, B., Mangalvedhe, N., and Ghosh, A. (2016). NB-IoT System for M2M Communication[C]. Proceeding of the 2016 IEEE wireless communications and networking conference, April 2016, Doha, Qatar, IEEE, 1–5.

Sanduleac, M., and Ciornei, I. (2021). High Reporting Rate Smart Metering Data for Enhanced Grid Monitoring and Services for Energy Communities[J]. *IEEE Trans. Industrial Inf.* 18:4039-4048. Accepted. doi:10.1109/tii.2021.3095101

Sultania, A. K., Mahfoudhi, F., and Famaey, J. (2020). Real-Time Demand Response Using NB-IoT. *IEEE Internet Things J.* 7 (12), 11863–11872. doi:10.1109/jiot.2020.3004390

Tan, O., Gunduz, D., and Poor, H. V. (2013). Increasing Smart Meter Privacy through Energy Harvesting and Storage Devices. *IEEE J. Sel. Areas Commun.* 3l (7), 133 1–1341. doi:10.1109/jsac.2013.130715

Trans-European Networks for Energy (2017). *Smart Grid Regional Group"*. Available from: https://ec.europa.eu/energy/topics/infrastructure/projects-common-interest/regional-groups-and-their-role/smart-grid-regional-group_en.

Wang, Y. P. E., Lin, X., Adhikary, A., Grovlen, A., Sui, Y., Blankenship, Y., et al. (2017). A Primer on 3GPP Narrowband Internet of Things[J]. *IEEE Commun. Mag.* 55 (3), 17–123. doi:10.1109/mcom.2017.1600510cm

Xu, X., Xu, Y., Wang, M. H., Li, J., Xu, Z., Chai, S., et al. (2020a). Data-driven Game-Based Pricing for Sharing Rooftop Photovoltaic Generation and Energy Storage in the Residential Building Cluster under Uncertainties[J]. *IEEE Trans. Industrial Inf.* 17 (7), 4480–4491.doi:10.1109/tii.2020.3016336

Xu, X., Jia, Y., Xu, Y., Xu, Z., Chai, S., and Lai, C. S. (2020b). A Multi-Agent Reinforcement Learning-Based Data-Driven Method for Home Energy Management. *IEEE Trans. Smart Grid* 11 (4), 3201–3211. doi:10.1109/tsg.2020.2971427

Yuan, Y., Dehghanpour, K., Bu, F., and Wang, Z. (2020). A Data-Driven Customer Segmentation Strategy Based on Contribution to System Peak Demand. *IEEE Trans. Power Syst.* 35 (5), 4026–4035. doi:10.1109/tpwrs.2020.2979943

Yuan, Y., and Wang, Z. (2021). Mining Smart Meter Data to Enhance Distribution Grid Observability for Behind-The-Meter Load Control: Significantly Improving System Situational Awareness and Providing Valuable Insights. *IEEE Electrific. Mag.* 9 (3), 92–103. doi:10.1109/mele.2021.3093636

Zayas, A. D., and Merino, P. (2017).The 3GPP NB-IoT System Architecture for the Internet of Things[C]. Proceeding of the 2017 IEEE International Conference on Communications Workshops (ICC Workshops), May 2017, Paris, France. IEEE, 277–282.

Conflict of Interest: Author JZ is employed by Pinggao Group Intelligent Power Technology Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Energy Research

ORIGINAL RESEARCH
published: 08 June 2022
doi: 10.3389/fenrg.2022.892541

Edge Computing Based Electricity-Theft Detection of Low-Voltage Users

Yingjun Zheng[1], *Feng Chen*[2], *Hongming Yang*[1] *and Sheng Su*[1]*

[1]School of Electrical and Information Engineering, Changsha University of Science & Technology, Changsha, China, [2]Changsha Electric Power Corporations, State Grid of China, Changsha, China

Electricity theft of low voltage (LV) users could result not only in the escalation of power loss but also in dangerous electric shock. Since LV users are served by distribution transformers, electricity theft of an LV user will cause line loss escalation of the associated distribution serving zone (DTSZ). Therefore, it seems promising to identify anomaly users of electricity theft with a Granger causality test to find out the user causing an escalation of line loss in DTSZ with time series of users' usage and line loss. However, meters of LV users in severe environments occasionally suffer from communication failure to upload metering data to the head end of advanced metering infrastructure (AMI), which could distort the daily electricity usage of the associate user. Consequently, it could cause false alarms unavoidably once we detect electricity theft with these distorted data. Since the distribution transformer unit (DTU) collects metering data of LV users within associate DTSZ without distortion, an edge computing–based electricity theft detection approach is proposed in this article. The correlation between line loss of a DTSZ and electricity usage of anomaly users of electricity theft is first analyzed. Thereafter, the Granger causality test is used to identify anomaly users with authentic usage data with edge computing in DTU. Finally, the abnormal data and the data repaired by different missing data filling algorithms are used on the main station to detect electricity theft. Numerical simulation suggests that although missing data completion could recover information in missing data partially, it could result in notable false positive alarms in electricity theft, while the proposed method based on edge computing can completely eliminate the data distortion caused by communication failure.

Keywords: electricity theft, communication failure, edge computing, missing data completion, distribution transformer terminal, attribution analysis

OPEN ACCESS

Edited by:
Yunfei Mu,
Tianjin University, China

Reviewed by:
Lin Zhu,
The University of Tennessee, Knoxville, United States
Gaoqi Liang,
Nanyang Technological University, Singapore
Jian Liu,
Wuhan University, China

***Correspondence:**
Sheng Su
eessheng@163.com

Specialty section:
This article was submitted to Smart Grids, a section of the journal Frontiers in Energy Research

Received: *09 March 2022*
Accepted: *25 April 2022*
Published: *08 June 2022*

Citation:
Zheng Y, Chen F, Yang H and Su S (2022) Edge Computing Based Electricity-Theft Detection of Low-Voltage Users. Front. Energy Res. 10:892541. doi: 10.3389/fenrg.2022.892541

1 INTRODUCTION

Electricity theft of low voltage (LV) users could cause substantial revenue loss to power utilities. Moreover, anomaly wire hooks result in numerous electric shocks to users. Therefore, it is highly preferred to identify anomaly LV users (Wang Y. et al., 2019; Zhang et al., 2019; Partha et al., 2020). Since there are millions of LV residential users with diversified usage patterns, it is rather difficult to identify anomaly ones of electricity theft, and electricity detection of LV users remains a difficulty in industrial applications (Li et al., 2018).

The extensive application of smart meters could provide substantial meter data on electricity usage, which can lay a solid foundation for data-driven electric theft detection. Zheng et al. (2019)

detected electricity theft with maximum information coefficient and density peak fast clustering algorithm in combination. Zhuang et al. (2016) and Sun et al. (2018) detected electricity theft with fluctuation of multi-day usage, fluctuation of the SD of usage, and trend of usage with an improved outlier identification algorithm. Since extracted features play a key role in the precision of anomaly detection, a stacked de-correlation auto-encoder is employed (Hu et al., 2019) to extract highly condensed independent features. Thereafter, a support vector machine is used to identify anomaly users (Hu et al., 2019). Since power utilities have limited market crews for onsite inspection, a false positive rate (FPR) is key to evaluating the performance of electricity theft (Jin et al., 2022). To prevent false positives alarm, marketing crews implement onsite inspection of DTSZ with a high loss ratio above 8%. Since there is an underlying correlation between anomaly users of electricity theft and line loss of associated feeder, a Granger causality analysis–based approach is proposed by Jin et al. (2020) to detect users who cause fluctuations in line loss. Since high line loss in a DTSZ is usually caused by electricity theft, once we detect electricity theft in DTSZ with a high loss rate, it could achieve a low false alarm rate (Tang et al., 2020).

It should be pointed out that the meters of LV users communicating *via* power line communication (PLC) could suffer failure occasionally. Once a smart meter fails to upload its usage data to the head end of AMI, it will upload it in the following days. The head end calculates the daily line loss of a DTSZ according to the daily served electricity and accumulation of all users' daily usage. Line loss of the DTSZ escalates on the days when the meter fails and declines to even below zero in the following days when meters upload usage data of communication failure and that very day. Existing approaches to electricity detection identify anomaly users with accurate metering data, while false data cause misleading results inevitably.

To overcome the problem of detecting electricity theft with metering data in the head end of AMI, an edge computing–based electricity detection approach is proposed in this article. The rest of the article is organized as follows. Existing approaches to missing data completion are investigated in **Section 2**. Correlation of anomaly users' usage and line loss of associate DTSZ in investigated and edge computing–based approach is developed in **Section 3**. Numerical simulation of real-world metering data is analyzed in **Section 4** to demonstrate the superiority of the proposed method to that of data restored with various missing data completion algorithms. **Section 5** concludes the article.

2 ELECTRICITY MISSING DATA COMPLETION METHOD

Metering data could suffer interference and failure in the process of data acquisition, conversion, and communication, and missing data and false data are common for industrial applications of power utilities. Traditionally, power systems are measured with redundancy. Therefore, some missing data or false data can be identified and corrected with state estimation. There are similar missing data and false data in AMI. However, since these data are not closely coupled with each other, they can be corrected and filled with state estimation (Yang Y. et al., 2020). Traditionally, missing data and false data of AMI are filled or corrected with the mean of previous and following data, interpolation mode, closest distance data, regression model, and maximum expectation based algorithm (Sundararajan et al., 2019). However, most of them implement data completion with statistic-based and mechanism-based models and neglect underlying features of a single time series and correlation among various time series. Data completion with these approaches is not as satisfying as expected (Chen et al., 2019; Yang et al., 2019).

Since missing data is rather common in various fields, numerous researchers have researched missing data completion and achieved notable progress in recent years (Siamkaz et al., 2018; Song et al., 2019). Based on the inertial effect of the measured data, Ruan deduced coarse values of pre-attack measurements. Then, based on the deduced coarse values and suggested state bounds, an optimization model is proposed to recover the measurements (Ruan et al., 2022). The matrix filling method used in the Netflix recommendation system is established on the premise that the data matrix has low rank and sparsity. It could reconstruct the original matrix precisely in the case of partial loss of original data. The low-rank matrix completion theory is based on the low rank of the data to recover the missing data. It takes matrix rank minimization as the objective function. The classical mathematical model of data recovery is expressed as follows.

$$\begin{cases} \min\limits_{K} \ \|K\|_* \\ \text{s.t. } P_\Omega(M) = P_\Omega(K) \end{cases} \tag{1}$$

$$P_\Omega(M) = \begin{cases} M_{i,j} & (i,j) \in \Omega \\ 0 & (i,j) \notin \Omega \end{cases} \tag{2}$$

where $\|\cdot\|_*$ denote the matrix kernel norm; K denotes the restored low-rank matrix; M denotes the matrix to be repaired with only some elements observed; Ω denotes the set of positions of non-empty elements in M. If $M_{i,j}$ a member of matrix M is observed, then $(i,j) \in \Omega; P_\Omega$ is the operator. Since there are Gaussian noise, spikes, and other formal noise for the most real-world system, the data recovery model can be depicted in **Eq. 3** as follows.

$$\begin{cases} \min\limits_{K,E,G} \ \left(\|K\|_* + \rho\|E\|_1 + \delta\|G\|_F^2\right) \\ \text{s.t. } P_\Omega(M) = K + E + G + N \end{cases} \tag{3}$$

where E denotes peak outlier matrix, G denotes Gaussian noise matrix, ρ and δ denote weight coefficients correspondingly, and N denotes auxiliary matrix. The augmented Lagrange function in **Eq. 3** can be transformed into an unconstrained optimization problem and solved with the alternating direction method of multipliers (ADMM) according to literature (Yang T. et al., 2020).

Tensor completion is a high dimensional matrix completion. Since the electricity usage of multiple users on different days may have an underlying multi-dimensional internal correlation, tensor completion could be utilized to recover missing or false with high precision (Zhao et al., 2020). The fundamental principle

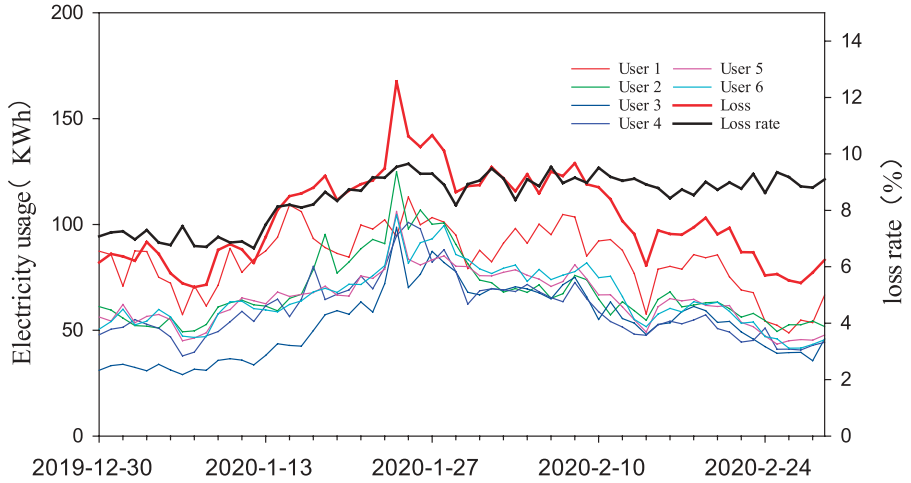


FIGURE 1 | Electricity usage of users and loss of DTSZ with a high loss rate.

of tensor completion is similar to matrix completion and could be referred to (Zhao et al., 2020).

Unlike matrix completion and tensor completion to recover missing data with low-rank data, generative adversarial networks (GAN) is a data-driven approach which extracts features from large amounts of unlabeled data through GAN's adversarial game. A discriminant model that can accurately identify the authentic and false/missing data and a generic model that can capture the potential features and spatial and temporal features of the data are obtained. Thereafter, the dual semantic perception constraint is utilized to retrain the model to find the candidate data that has the greatest similarity to the data to be reconstructed with missing values (Wang S. et al., 2019). It should be pointed out that the GAN based approach works on a large number of data, and it is not appropriate for the date completion of limited LV users in a DTSZ. Matrix completion and tensor completion are utilized to recover false/missing data in this article.

3 DETECTION OF POWER THEFT IN LOW VOLTAGE PLATFORM BASED ON EDGE COMPUTING

3.1 Correlation Analysis of User Power Quantity and Line Loss in Low Voltage Station Area

Non-technical loss (NTL) in DTSZ is mainly caused by electricity theft, and the NTL caused by anomaly users is usually correlated to its metering data and associated NTL. Therefore, there is an underlying correlation between the metering data of anomaly users and the NTL of DTSZ. The correlation could be identified with Granger causality analysis to find out the anomaly users caused an escalation of loss of DTSZ.

Metering data of a real-world DTSZ is employed to analyze the correlation between metering data and loss of DTSZ. Loss of the DTSZ in 62 consecutive days is shown in **Figure 1**. There are six industrial and commercial users and 33 low-voltage residential users in the DTSZ. The served daily mean electricity is about 1200 kWh, while the daily mean loss is about 100 kWh in January and February 2020. The loss rate came up to 9.6%, and it is highly suspected that there is an anomaly user of electricity theft. Since electricity theft of industrial and commercial users is contributing much more than average residential users, electricity usage of six industrial and commercial users within the DTSZ from 30 December 2019 to February 2020 is analyzed as follows.

The red line denotes the daily loss of DTSZ (G), and the black line denotes the daily loss rate of DTSZ. The other six lines denote the electricity consumption of six industrial and commercial users, which are depicted as H_1~H_6 in the following section. It can be observed that the loss profile of DTSZ has a similar trend as that of users' usage. Most of them escalate in the beginning and then decline in the end.

3.2 Data Communication and Anomaly Analysis in LV DTSZ

DTU is generally installed on the secondary side of the distribution cabinet (Liu et al., 2020). It communicates *via* protocol RS485 or PLC within the DTSZ and communicates with the head end of AMI with wireless communication or optical fiber. DTU collects metering data of the distribution transformer and associates LV users within DTSZ and uploads it to the head end of AMI (Huang et al., 2021; Zhong et al., 2021). Communication architecture within a typical DTSZ is shown in **Figure 2**.

In real-world AMI, meters of users operate in severe environments and it could suffer communication occasionally. Once it suffers communication failure, it could upload daily metering data in the following days, which could cause zero electricity usage on the previous day and electricity

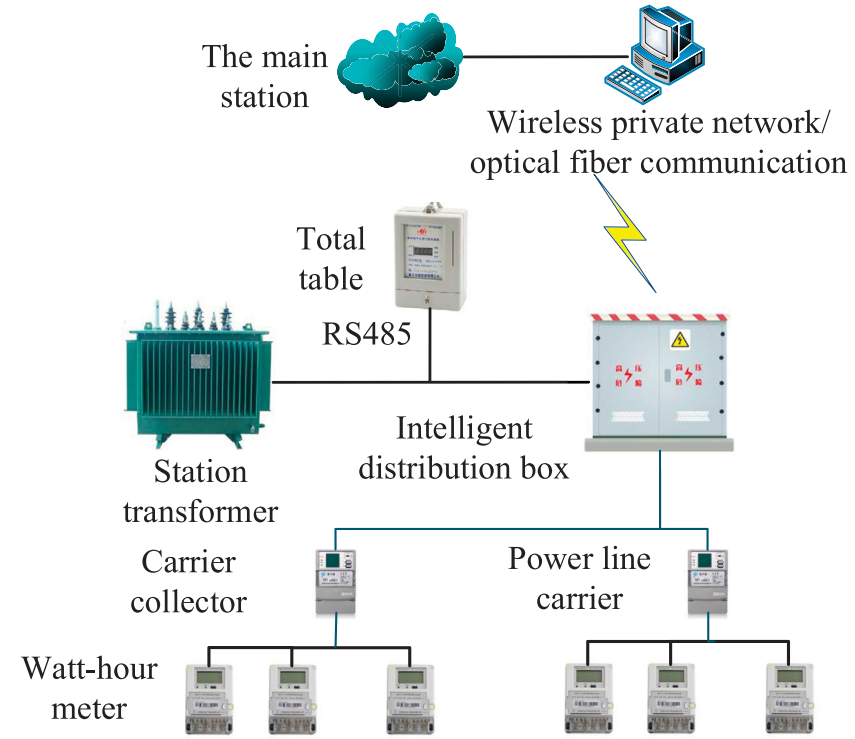


FIGURE 2 | Basic communication diagram of the distribution area.

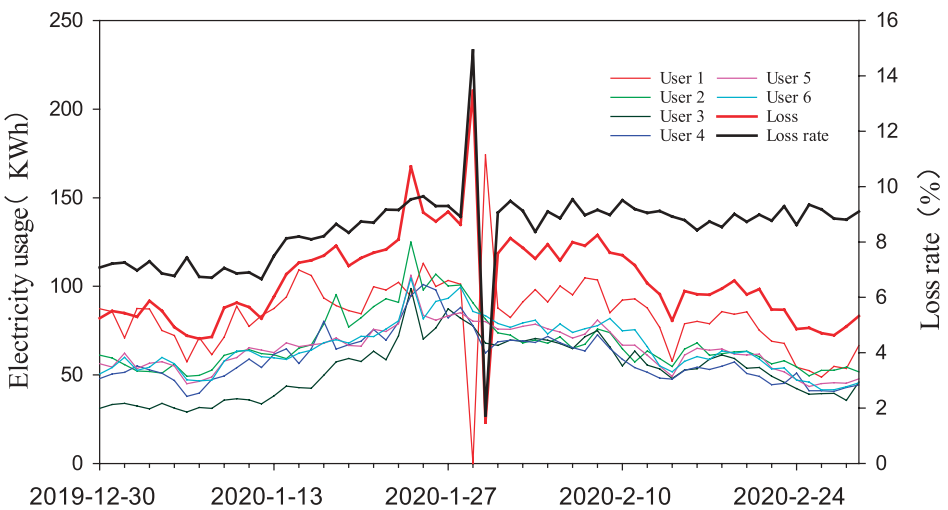


FIGURE 3 | Electricity usage of users and loss of DTSZ with communication failure.

usage of 2 days on the following day in the head end of AMI. In order to demonstrate the impact of communication failure on data quality of the head end of AMI, user one of DTSZ was selected to report 0 electricity usage on 29 January 2020, and the usage was accumulated and uploaded on the following day. The usage of users and loss data of DTSZ is plotted as shown in **Figure 3**. It can be observed that the loss and loss rate of DTSZ escalated notably on 29 January 2020 due to communication failure. While it decreases notably on the following day since usage in two consecutive days is accumulated. As a consequence, communication failure could distort the underlying correlation between anomaly users' electricity usage and loss of DTSZ and result in misleading electricity theft identification as a consequence.

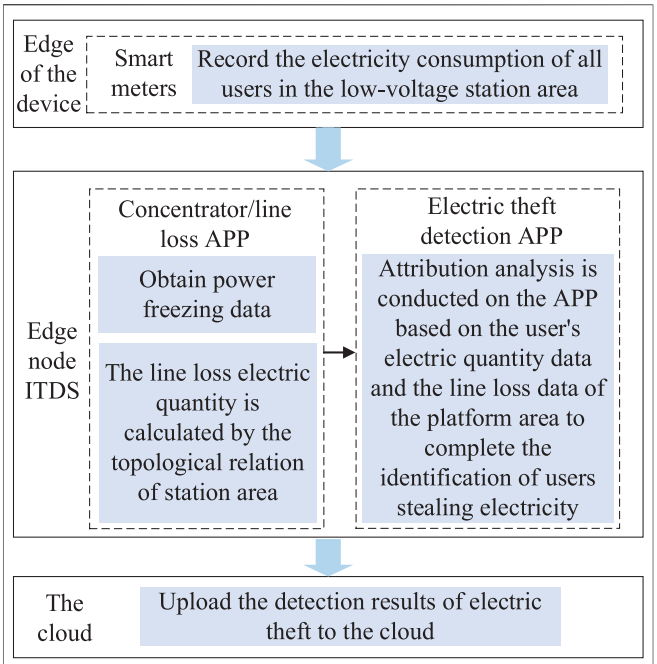


FIGURE 4 | Electricity theft detection framework with edge computing.

3.3 Edge Computing Based Electricity Theft Detection in Distribution Serving Zone

Co-integration test and Granger causality analysis are commonly used in economics to analyze the correlation among time series. In general, a co-integration test is used to test whether there is a long-term equilibrium among time series. Thereafter, Granger causality analysis is used to determine whether a variable impacts another variable (Zhu et al., 2017; Fan et al., 2019; Tian et al., 2019). Since electricity usage of an anomaly user of electricity theft is correlated to its usage in most theft modes, the fluctuation of loss pf DTSZ caused by anomaly users has a similar profile as time series of economic variables disturbed by other factors. Therefore, the equilibrium relationship and causal relationship between the loss of DTSZ and users' metering data can be analyzed to detect anomaly users.

Traditionally, field terminals, such as DTU, have limited computing and storage resources. Complex functions such as identification of anomaly users of electricity theft can only be implemented in the head end of AMI. Communication failure induced data missing could impact its performance notably (Shi et al., 2016; Covi et al., 2021). In recent years, more and more meters are being deployed in distribution systems with the rapid development of Internet of Things (IoT) technology (Deng et al., 2021). Since computing and storage resources of concurrent IoT terminals escalate notably, it is technically feasible to implement some of the complex functions in IoT terminals with edge computing (Li et al., 2020; Wang et al., 2020; Liu et al., 2022).

DTU could be utilized as a platform for edge computing in DTSZ. Since it collects substantial data within the DTSZ, it can provide loss analysis, power quality monitoring, and topology analysis with edge computing. Since different vendors implement different business functions in the diversified OS environment, Docker technology is employed to provide an appropriate container for the APP of various vendors on the same DTU platform (Gong et al., 2018). Docker-based DTU is composed of a system layer and an application (APP) layer. The APP layer is divided into acquisition APP and business APP, which can interact with each other through the message bus. The former collects real-time operation data and load data; The latter accesses the data center through the device bus, extracts the required data for calculation and analysis, and implements edge computing of business functions (Nie et al., 2020).

Electricity theft detection can be implemented in the DTU with edge computing. The fundamental of the approach is that the concentrator APP of DTU collects users' metering data within the DTSZ. Loss analysis APP collects serving electricity and calculates the loss of DTSZ. Thereafter, the electricity theft detection APP identifies anomaly users with Granger causality analysis with loss data and metering data of associated users in the DTSZ. The framework of the implementation process is shown in **Figure 4**. The method proposed in this paper transfers the detection of electricity theft from the head end of AMI to the DTU in the edge, which can eliminate communication associated data missing to identify anomaly users with lower fails positive rate.

Granger causality analysis is first used to identify anomaly users in DTSZ depicted in **Section 1**. Thereafter, the distorted data and distorted data recovered with various data completion approaches are analyzed in this section.

4 NUMERICAL SIMULATION

4.1 Edge Computing–Based Electricity Theft Detection

ADF unit root tests were performed for G and H_1-H_6, and their differential sequences were on the edge side. With 5% confidence as the standard, all the sequences were first-order unitary sequences, and the test results are shown in **Table 1**.

The Engle–Granger co-integration test was applied to G and H_1-H_6, respectively. Thereafter, the stationarity test of residual series was implemented with the ADF test. The stationarity test results of the residual series are listed in **Table 2**, and the shaded area in the table indicates that the result is less than the threshold of −3.4363.

$$\begin{cases} G = 22.9015 + 1.3434\mathrm{H}_4 \\ T_1 = 4.5562, T_2 = 16.4799 \\ R^2 = 0.8190, \bar{R}^2 = 0.8160 \\ F = 271.5902, D = 1.4099 \end{cases} \tag{4}$$

$$\begin{cases} G = -0.6793 + 1.6065\mathrm{H}_5 \\ T_1 = -0.1428, T_2 = 22.2943 \\ R^2 = 0.8923, \bar{R}^2 = 0.8905 \\ F = 497.0337, D = 0.9883 \end{cases} \tag{5}$$

TABLE 1 | Results of stationary test for DTSZ with high loss rate.

Time series	ADF	5% Confidence	Stationary	Time series	ADF	5% Confidence	Stationary
G	−1.7940	0.3801	N	$\triangle_G$	−8.9978	0.0000	Y
H1	−2.4498	0.1328	N	$\triangle_{H1}$	−9.7337	0.0000	Y
H2	−0.4467	0.5171	N	$\triangle_{H2}$	−10.7129	0.0000	Y
H3	−1.9382	0.3130	N	$\triangle_{H3}$	−7.9962	0.0000	Y
H4	−0.4595	0.5121	N	$\triangle_{H4}$	−8.0350	0.0000	Y
H5	−1.8255	0.3650	N	$\triangle_{H5}$	−9.3035	0.0000	Y
H6	−0.4429	0.5187	N	$\triangle_{H6}$	−9.7269	0.0000	Y

TABLE 2 | Residual sequence smoothness test results of user and loss.

User	Inspection results	User	Inspection results
H_1	−3.1573	H_4	−5.7480
H_2	−3.3738	H_5	−4.5032
H_3	−2.9929	H_6	−4.5394

$$\begin{cases} G = 13.0982 + 1.3801\mathrm{H}_6 \\ T_1 = 2.7918, T_2 = 19.7185 \\ R^2 = 0.8663, \bar{R}^2 = 0.8641 \\ F = 388.8198, D = 1.0359 \end{cases} \tag{6}$$

When the residual sequence is stationary, the regression equation between the corresponding variables is shown as above. In **Eqs 4–6**, T_1 and T_2 are the t-test values of corresponding parameters, R^2 denotes the determinability coefficient, $\bar{R}^2$ denotes the adjusted determinability coefficient, F denotes the model test value, and D denotes the Dubin Watson statistic. If F test value and T test value are significant, the regression effect of the equation is better.

After constructing the least squares regression model for H_1–H_6 and G, the co-integration test results show that the test value in the stationarity test of residual sequences of H_1–H_3 and G is greater than the threshold while that in the stationarity test of residual sequences of H_4–H_6 and G is less than the threshold, which indicates H_4–H_6 has a co-integration relationship with G. The subsequent Granger causality analysis can be continued, and the threshold refers to the critical value of McKinnon's co-integration test (Pan, 2017). To further clarify the dynamic relationship between H_4~H_6 and G, an error correction model among H_4~H_6 and G is established, and the results are listed in **Eqs 7–9**. It can be observed that when H_4–H_6 fluctuates (increases) by 1% in the short term, H_4–H_6 will increase by 0.7040%, 1.2302%, and 1.1019%, respectively. According to the coefficient of error correction term, when the short-term fluctuation of H_4–H_6 and G deviates from the long-term equilibrium relationship among them, the non-equilibrium state among H_4–H_6 and G will be corrected to the equilibrium state with the adjustment force of −0.5700, −0.4457, and −0.4409, respectively.

$$\begin{cases} \Delta G = 0.7040\Delta H_4 - 0.5700e_{t-1} \\ T_1 = 4.1511, T_2 = -4.8494 \\ R^2 = 0.3567, \bar{R}^2 = 0.3228 \\ D = 2.1373 \end{cases} \tag{7}$$

TABLE 3 | Granger causality test results of edge computing.

Assuming	Significance
G is not a Granger reason for H_4	0.0137
H_4 is not a Granger reason for *G*	0.0465
G is not a Granger reason for H_5	0.5068
H_5 is not a Granger reason for *G*	0.4232
G is not a Granger reason for H_6	0.0275
H_6 is not a Granger reason for *G*	0.7291

where $e_{t-1} = G_{(t-1)} - 20.0735 - 1.3920H_{4(t-1)}$, ΔG is the first-order difference of G, ΔH_4 is the first-order difference of H_4, $G_{(t-1)}$ is the first-order lag sequence of G, and $H_{4(t-1)}$ is the first-order lag sequence of H_4.

$$\begin{cases} \Delta G = 1.2302\Delta H_5 - 0.4457e_{t-1} \\ T_1 = 10.0978, T_2 = -4.1534 \\ R^2 = 0.6645, \bar{R}^2 = 0.6469 \\ D = 2.0424 \end{cases}, \tag{8}$$

where $e_{t-1} = G_{(t-1)} + 2.2644 - 1.6356H_{5(t-1)}$, ΔH_5 is the first-order difference of H_5, and $H_{5(t-1)}$ is the first-order lag sequence of H_5.

$$\begin{cases} \Delta G = 1.1019\Delta H_6 - 0.4409e_{t-1} \\ T_1 = 7.3627, T_2 = -3.6399 \\ R^2 = 0.5144, \bar{R}^2 = 0.4888 \\ D = 2.1190 \end{cases} \tag{9}$$

where $e_{t-1} = G_{(t-1)} - 17.0197 - 1.3225H_{6(t-1)}$, ΔH_6 is the first-order difference of H_6, and $H_{6(t-1)}$ is the first-order lag sequence of H_6.

The co-integration relationship among H_4~H_6 and G and the error correction model were analyzed. It can be observed that the error correction coefficients in the three error correction models of H_4~H_6 and G all conform to the reverse adjustment mechanism. Characteristics of long-term stability and the dynamic relationship between H_4~H_6 and G with co-integration relationship are further clarified. Since there is a co-integration relationship between H_4~H_6 and G, the causal relationship between their influences can be further analyzed with Granger causality analysis. Granger causality analysis results of edge computing are shown in **Table 3**. It can be observed that the significance of "H_4 is not the Granger cause of G" is less than the critical level of 5%, which indicates the null hypothesis is rejected. Therefore, H_4 is the cause of the change of G, and H_4 can be

TABLE 4 | Daily electricity usage of each anomaly user and loss of DTSZ.

Issue	Date	User 1	User 2	User 3	User 4	User 5	User 6
Daily metering usage	2020/1/29	0.00	0.00	0.00	0.00	0.00	0.00
	2020/1/30	174.15	172.04	145.58	140.06	160.32	169.20
Loss of DTSZ	2020/1/29	210.34	205.93	192.97	193.00	195.55	201.10
	2020/1/30	23.09	27.50	40.46	40.43	37.88	32.33

TABLE 5 | Result of Granger attribution test at the head end of AMI.

	Missing user	Stationary (loss and electricity usage)	Co-integration relationship (loss and electricity usage)	Significance test results	Audit results
Edge computing	Null	First order differential stationary	H_4, H_5, H_6	H_4	Correct
Detecting in the head end of AMI	User 1	First order differential stationary	H_2, H_3, H_4, H_5, H_6	H_2, H_3, H_4, H_5, H_6	Miscalculation
	User 2	First order differential stationary	H_1, H_3, H_4, H_5, H_6	H_1, H_3, H_4, H_5, H_6	Miscalculation
	User 3	First order differential stationary	H_1, H_2, H_4, H_5, H_6	H_1, H_2, H_4, H_5, H_6	Miscalculation
	User 4	First order differential stationary	H_1, H_2, H_3, H_5, H_6	H_1, H_2, H_3, H_5, H_6	Miscalculation
	User 5	First order differential stationary	H_1, H_2, H_3, H_4, H_6	H_1, H_2, H_3, H_4, H_6	Miscalculation
	User 6	First order differential stationary	H_1, H_2, H_3, H_4, H_5	H_1, H_2, H_3, H_4, H_5	Miscalculation

regarded as the anomaly user of electricity theft in the DTSZ. This has been verified by on-site inspection.

4.2 Master Station Detection Comparative Experiment

When LV users suffer communication failure and upload metering data on the following day, the data at the head end of AMI is distorted with missing data and could cause misleading electricity theft detection results. In order to verify the superiority of the proposed approach, six industrial and commercial users in DTSZ were set to upload zero usage on the 30th day (28 January 2020) and upload usage on two consecutive days on the following day (29 January 2020). Therefore, the loss of DTSZ increased on the first day and decreased on the second day. The metered usage of each user and loss of DTSZ are listed as shown in **Table 4**. Granger causality test is used to test whether there is a correlation between each user and the loss of DTSZ based on distorted false data. The ultimate results of Granger causality analysis at the head end of AMI are shown in **Table 5**.

It can be concluded from **Table 5** and Schedules as follows.

- When users suffer communication failure and fail to upload metering data, there is no co-integration relationship between the user's electricity usage and loss of DTSZ. Therefore, we cannot analyze it with Granger causality analysis. The co-integration relationship between electricity usage of other users and loss of DTSZ remains, and they can be analyzed with Granger causality analysis.
- Except for the user who suffers communication failure, all other users could be identified as anomaly users of electricity theft. Since user 4 has been confirmed to be the anomaly user of electricity theft by onsite inspection, once we identify the anomaly user of electricity theft with distorted data with data missing, the false positive rate escalates to 80%, which is not acceptable for industrial applications.

Missing data completion got broad research in recent years. Yang Y. et al. (2020) proposed a low-rank matrix theory based on matrix completion of power quality data. It designs a multi-norm joint low-rank optimization model and solves it with an alternating direction multiplier method. Zhao et al. (2020) proposed a tension completion based approach to recover missing data of multiple-user, and a low-rank tensor completion model was employed to recover missing data in DTSZ. It analyzes the characteristics of the LV data in DTSZ and constructs the standard missing tensor.

In order to find outperformance of matrix completion and Tensor completion, missing data of each user is recovered with these two approaches. The authentic usage and recovered usage with the two completion approaches are listed in **Table 6**. The authentic loss on 29 January and 30 January is 115.30 and 118.13 kWh, respectively. According to the user usage data recovered with correction, the loss of the DTSZ calculated with recovered data is listed in **Table 7**.

It can be observed from **Tables 6**, **7** that although it is widely supposed that matrix completion and tensor completion can recover missing data ideally, its premise is that time series are of low rank. Missing data cannot be recovered precisely once there is no strong correlation between users' usage data in the DTSZ. The data recovered with these two approaches are used to test whether the Granger causality test can accurately identify anomaly users of electricity theft. The Granger causality analysis results are listed in **Table 8**.

It can be concluded from **Table 8** as follows.

- According to the data analysis with matrix completion, Granger causality analysis cannot identify any anomaly user of electricity theft once users H_1–H_3 or H_6 suffer communication

TABLE 6 | Recovered electricity usage of each anomaly user.

Packing method	Date	Daily usage/(KWh)					
		H_1	H_2	H_3	H_4	H_5	H_6
Authentic usage	2020/1/29	95.04	90.63	77.66	77.70	80.25	85.80
	2020/1/30	79.11	81.41	67.91	62.36	80.07	83.40
Matrix completion	2020/1/29	102.52	102.37	95.85	97.57	96.22	98.56
	2020/1/30	97.45	95.63	90.46	90.75	92.81	91.38
Tensor completion	2020/1/29	107.39	108.31	79.33	79.14	85.41	86.79
	2020/1/30	89.22	95.16	82.31	71.01	83.93	90.97

TABLE 7 | Calculated loss for each recovered anomaly user.

Data completion	Date	Loss of power/(KWh)					
		H_1	H_2	H_3	H_4	H_5	H_6
Matrix completion	Recovered data on 29 January	104.37	103.15	100.64	97.12	99.27	106.31
	Error of 29 January	–9.48%	–10.54%	–12.71%	–15.77%	–13.90%	–7.80%
	Recovered data on 30 January	98.20	96.95	94.17	90.79	93.03	98.49
	Error of 30 January	–16.87%	–17.93%	–20.28%	–23.14%	–21.25%	–16.63%
Tensor completion	Recovered data on 29 January	102.65	106.38	105.85	102.97	103.47	107.23
	Error of 29 January	–10.97%	–7.74%	–8.20%	–10.69%	–10.26%	–7.00%
	Recovered data on 30 January	113.28	114.14	112.35	111.03	111.17	112.73
	Error of 30 January	–4.11%	–3.38%	–4.89%	–6.01%	–5.89%	–4.57%

TABLE 8 | Results of Granger causality analysis of anomaly user with data recovery.

Edge of the results	User with missing data	Stationary (loss and electricity usage))	Co-integration relationship (loss and electricity usage)	Significance test results	Audit results
	Without data missing	First-order differential stationary	H_4, H_5, H_6	H_4	T
Matrix fill	User 1	First order differential stationary	H_1, H_2, H_4, H_5, H_6	NULL	F
	User 2	First order differential stationary	H_1, H_2, H_4, H_5, H_6	NULL	F
	User 3	First order differential stationary	H_1, H_2, H_3, H_4, H_5, H_6	NULL	F
	User 4	First order differential stationary	H_1, H_2, H_4, H_5, H_6	H_4	F
	User 5	First order differential stationary	H_1, H_2, H_4, H_5, H_6	H_5	False-positive
	User 6	First order differential stationary	H_1, H_2, H_4, H_5, H_6	NULL	Failure
Tensor completion	User 1	First order differential stationary	H_1, H_2, H_4, H_5, H_6	NULL	F
	User 2	First order differential stationary	H_1, H_2, H_4, H_5, H_6	H_4	T
	User 3	First order differential stationary	H_1, H_2, H_3, H_4, H_5, H_6	NULL	F
	User 4	First order differential stationary	H_1, H_2, H_4, H_5, H_6	H_4	T
	User 5	First order differential stationary	H_1, H_2, H_4, H_5, H_6	H_4	T
	User 6	First order differential stationary	H_1, H_2, H_4, H_5, H_6	H_6	False-positive

failure and recover with matrix completion. User H_4 is correctly identified as an anomaly user of electricity theft once user H_4 suffers communication failure and recovers with matrix completion. User H_5 is incorrectly identified as an anomaly user once user H_5 suffers communication failure and recovers with matrix completion. The accuracy rate of electricity theft detection with data in the head end of AMI declined to 16.6%.

- According to the data analysis after tensor completion, Granger causality analysis cannot determine any anomaly user of electricity theft once user H_1 or H_3 suffers communication failure and recovers with tensor completion. User H_4 can be judged as an anomaly user once users H_2, H_4, or H_5 suffer communication failure and recover with tensor completion. User H_6 can be misjudged as an anomaly user once user H_6 suffers communication failure and recovers with tensor completion. The accuracy rate of electricity theft detection with data in the head end of AMI declined to 50%.

- The goal of either matrix completion or tensor completion is to get the minimum norm of low-rank matrix/tensor. They recover data within a certain error range with higher linear correlation for the low rank of the data. It can be observed from **Tables 5**, **6** that there are notable errors in the recovered data of both algorithms. According to the Granger causality analysis by Jin et al. (2020), the one-to-one correspondence

between electricity usage and loss of the detected user at the same time has a great influence on the final result, while both completion algorithms change correlation to a certain range, which results in the failure of Granger causality analysis–based approach.

In conclusion, both matrix completion and tensor completion based approaches cannot recover data missing ideally and could negatively impact the precision of Granger causality–based electricity theft detection. The proposed approach to implement electricity theft in DTU of DTSZ with edge computing could eliminate the impact of communication failure–induced difficulty and facilitate precise electricity theft detection.

5 CONCLUSION

The article analyzes communication failure's impact on data missing in the head end of AMI and points out that the false metering data could negatively impact electricity theft detection of LV users in DTSZ. Edge computing–based approach is proposed to detect electricity theft of DTSZ in DTU with edge computing, which can identify anomaly users with authentic metering data in the edge and mitigate the difficulty of data recovery of missing/false data caused by communication failure. The real world metering data of a DTSZ is employed to produce distorted data caused by communication failure. Thereafter, produced data is recovered with matrix completion and tensor completion. Numerical simulation of these data shows that the Granger causality analysis–based approach could identify anomaly users of electricity precisely with authentic data in the edge. However, all users are identified as anomaly users of electricity theft once the false data in the head end system is utilized. Once false/missing data are recovered with matrix completion or tensor completion, the accuracy of the Granger causality analysis–based approach declines to 16.7% or 50%.

It should be pointed out that there is an anomaly LV user bypass meter, and its electricity usage is zero around the clock. Since there are numerous vacant apartments without electricity usage, it is rather difficult to detect these anomaly users since its meter data do not provide any useful information. We cannot detect these users precisely, even with authentic metering data. The way to identify these users requires further investigation.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

YZ implemented numerical simulation and wrote the manuscript. FC helped collect low voltage users' metering data. HY revised the manuscript, and SS proposed the edge computing-based approach.

FUNDING

This research is supported by the National Natural Science Foundation of China (no. 51777015).

REFERENCES

Chen, X., Mcelroy, M. B., Wu, Q., Shu, Y., and Xue, Y. (2019). Transition towards Higher Penetration of Renewables: an Overview of Interlinked Technical, Environmental and Socio-Economic Challenges. *J. Mod. Power Syst. Clean. Energy* 7, 1–8. doi:10.1007/s40565-018-0438-9

Covi, E., Donati, E., Liang, X., Kappel, D., Heidari, H., Payvand, M., et al. (2021). Adaptive Extreme Edge Computing for Wearable Devices. *Front. Neurosci.* 15, 611300. doi:10.3389/fnins.2021.611300

Deng, Y., Zhang, Y., Luo, F., and Mu, Y. (2021). Operational Planning of Centralized Charging Stations Utilizing Second-Life Battery Energy Storage Systems. *IEEE Trans. Sustain. Energy* 12, 387–399. doi:10.1109/tste.2020.3001015

Fan, H., Wang, Y., Li, Y., and Zhao, C. (2019). Sparse Fault Degradation Oriented Fisher Discriminant Analysis Based Fault Trace. *Control Eng. China* 7, 1239–1244.

Gong, G., Luo, A., Chen, Z., Luan, A., An, X., Wang, X., et al. (2018). Information Physical System of Active Distribution Network Based on Edge Computing. *Power grid Technol.* 10, 3128–3135.

Hu, T, GUO, Q, and SUN, H (2019). Nontechnical loss detection based on stacked uncorrelated autoencoder and support vector machine. *Automation of Electric Power Systems* 1, 119–125.

Huang, Y., Yu, H., Yin, J., Meng, G., and Cheng, Y. (2021). Electric Power Network Data Transmission Schemepresent Situation and the Outlook Based on 5G Technology. *Trans. China Electrotech. Soc.* 17, 3581–3593.

Jin, C., Su, S., Cao, Y., Xue, Y., Yang, Y., and Liu, X. (2020). Detection of Power Theft in High-Loss Areas Based on Granger Attribution Analysis. *Automation Electr. Power Syst.* 23, 82–89.

Jin, C., Su, S., Xue, Y., Yang, Y., Liu, X., and Cao, Y. (2022). Review on Data-Driven Based Electricity Theft Detection Method and Research Prospect for Low False Positive Rate. *Automation Electr. Power Syst.* 1, 3–14.

Li, B., Yan, K., Luo, F., and Lu, Y. (2018). Comprehensive Application of Optimal Benchmarking in Line Loss Lean Management of City-Level Power Grid Enterprises. *Automation Electr. Power Syst.* 23, 184–191.doi:10.7500/AEPS20180129003

Li, Q., Zhang, Y., Chen, J., Yi, Y., and He, F. (2020). Development Forms and Challenges of Ubiquitous Power Internet of Things. *Automation Electr. Power Syst.* 1, 13–22.

Liu, D., Zeng, X., and Wang, Y. (2020). Security Situation Assessment of Smart Distribution Transformer Terminal Based on Information Entropy. *South. Power Grid Technol.* 1, 18–23.

Liu, D., Liang, H., Zeng, X., Zhang, Q., Zhang, Z., and Li, M. (2022). Edge Computing Application, Architecture, and Challenges in Ubiquitous Power Internet of Things. *Front. Energy Res.* 10, 850252. doi:10.3389/fenrg.2022.850252

Nie, Z., Zhang, J., and Fu, H. (2020). Key Technologies and Application Scenarios Designed for Edge Modernization and Containerization of Distribution Transformer Terminals. *Automation Electr. Power Syst.* 3, 154–161.

Pan, Z. (2017). *Time Series Analysis-Macroeconomic Data Analysis Model*. Beijing: Economic Science Press, 121–145.

Partha, P. B., Hongyun, C., Zhou, B., Chen, B., Mashima, D., Zheng, V. W., et al. (2020). Electricity Theft Pinpointing through Correlation Analysis of Master and Individual Meter Readings. *IEEE Trans. Smart Grid* 4, 3031–3042.doi:10.1109/TSG.2019.2961136

Ruan, J., Liang, G., Zhao, J., Qiu, J., and Dong, Z. Y. (2022). An Inertia-Based Data Recovery Scheme for False Data Injection Attack. *IEEE Trans. Ind. Inf.* 1, 1. doi:10.1109/TII.2022.3146859

Shi, W., Cao, J., Zhang, Q., Li, Y., and Xu, L. (2016). Edge Computing: Vision and Challenges. *IEEE Internet Things J.* 3, 637–646. doi:10.1109/jiot.2016.2579198

Siamakz, D., Dougherty, E. R., and Xiao, N. Q. (2018). Optimal Bayesian Classification with Missing Values. *IEEE Trans. Signal Process.* 16, 4182–4192.

Song, X., Guo, Y., Li, N., and Qian, P. (2019). A Novel Approach for Missing Data Prediction in Coevolving Time Series. *Computing* 101, 1565–1584. doi:10.1007/s00607-018-0668-8

Sun, Y., Li, S., Cui, C., Li, B., Chen, S., and Cui, G. (2018). Detection Method of Power Data Outliers Based on Gaussian Kernel Function. *Power Syst. Technol.* 5, 1595–1604.

Sundararajan, A., Khan, T., Moghadasi, A., and Sarwat, A. I. (2019). Survey on Synchrophasor Data Quality and Cybersecurity Challenges, and Evaluation of Their Interdependencies. *J. Mod. Power Syst. Clean. Energy* 7, 449–467. doi:10.1007/s40565-018-0473-6

Tang, D., Liu, Y., Xiong, Z., Ma, T., and Su, T. (2020). Early Warning Method of Electricity Anti-theft in Distribution Station Area Based on the Spatiotemporal Correlation Matrix. *Automation Electr. Power Syst.* 19, 168–176.

Tian, S., Gong, T., Huang, X., and Yu, W. (2019). Based on the Power of Big Data Area E - GDP Value Prediction. *Electr. power Autom. Equip.* 11, 198–204.

Wang, H., Li, S., Yu, H., and Zhang, J. (2020). Power Quality Data Compression Storage Method for Distribution Network Based on Distributed Compressed Sensing and Edge Computing. *Trans. China Electrotech. Soc.* 21, 4553–4564.

Wang, S., Chen, H., Pan, Z., and Wang, J. (2019b). A Reconstruction Method for Missing Data in Power System Measurement Using an Improved Generative Adversarial Network. *Proc. CSEE.* 1, 56–64. doi:10.13334/j.0258-8013.pcsee.181282

Wang, Y., Chen, Q., Hong, T., and Kang, C. (2019a). Review of Smart Meter Data Analytics: Applications, Methodologies, and Challenges. *IEEE Trans. Smart Grid* 10, 3125–3148. doi:10.1109/tsg.2018.2818167

Yang, T., Li, Y., He, Z., Han, X., Pen, H., and Lu, Y. (2020b). Power Quality Data Repair Algorithm for Ubiquitous Power Internet of Things Based on Matrix Filling. *Automation Electr. power Syst.* 2, 13–21.

Yang, T., Zhao, L., and Wang, C. (2019). Review on Application of Artificial Intelligence in Power Systems and Integrated Energy System. *Automation Electr. Power Syst.* 1, 2–14.

Yang, Y., Qi, L., Wang, H., Su, L., and Xu, Y. (2020a). Measurement Data Missing Reconstruction in Distribution Network Based on Generative Antagonism and Dual Semantic Perception. *Automation Electr. power Syst.* 18, 46–54.

Zhang, C., Xiao, X., and Zheng, Z. (2019). User-side Power-Stealing Behavior Detection Based on Real-Valued Deep Confidence Network. *Power Syst. Technol.* 3, 1083–1091.

Zhao, H., Shou, P., and Ma, L. (2020). Tensor Completion Method for Missing Data in Low-Voltage Platform Region. *Proc. CSEE.* 22, 7328–7337.

Zheng, K., Chen, Q., Wang, Y., Kang, C., and Xia, Q. (2019). A Novel Combined Data-Driven Approach for Electricity Theft Detection. *IEEE Trans. Ind. Inf.* 15, 1809–1819. doi:10.1109/tii.2018.2873814

Zhong, J., Xiong, X., He, Y., Shan, R., Jian, H., and Li, S. (2021). A Plug-And-Play and Topology Recognition Method for Platform Intelligent Terminal. *Automation Electr. Power Syst.* 10, 166–173.

Zhu, J. Y., Sun, C., and Li, V. O. K. (2017). An Extended Spatio-Temporal Granger Causality Model for Air Quality Estimation with Heterogeneous Urban Big Data. *IEEE Trans. Big Data* 3, 307–319. doi:10.1109/tbdata.2017.2651898

Zhuang, C., Zhang, B., Hu, J., Li, Q., and Zeng, Q. (2016). Detection of Abnormal Power Consumption Patterns of Power Users Based on Unsupervised Learning. *Proc. CSEE.* 2, 379–387.

Conflict of Interest: FC is employed by Changsha Electric Power Corporation.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Energy Research

ORIGINAL RESEARCH
published: 16 June 2022
doi: 10.3389/fenrg.2022.928919

An Ontology-Based Dynamic Attack Graph Generation Approach for the Internet of Vehicles

Shuning Hou[1], *Xiuzhen Chen*[2]*, *Jin Ma*[1], *Zhihong Zhou*[1] *and Haiyang Yu*[1]

[1]*Institute of Cyber Science and Technology, Shanghai Jiao Tong University, Shanghai, China,* [2]*Shanghai Municipal Key Lab of Integrated Management Technology for Information Security, Shanghai, China*

With the development of automobile intelligence, the security of the Internet of Vehicles has become a key factor that affects the development of intelligent vehicles. However, existing security risk analysis methods for the IoV either focus only on certain levels, such as the component level, or perform only a static analysis. This paper proposes a dynamic attack graph generation method for the IoV to identify and visually display the security risks caused by the associated vulnerabilities in an IoV system. First, using the actual architecture of the IoV, this paper shows how to model the security elements and their relationships in the IoV system and proposes a network security ontology model for this system. Second, it shows how to construct a reasoning rule base according to the causal relationship between the vulnerabilities using the Semantic Web Rule Language Finally, in view of the rapid change in the network topology of the IoV, a dynamic attack graph generation algorithm based on an ontology reasoning engine is proposed, which can effectively reduce the overhead caused by the changes in the attack graph. The effectiveness of the algorithm is demonstrated through an actual security event scenario and a constructed scenario. The experimental results show that the algorithm can dynamically and accurately display the network attack graph of the IoV. The proposed method is helpful in globally analyzing the threat caused by the combined exploitation of the vulnerabilities in an IoV system and risk management.

Keywords: internet of vehicles, dynamic attack graph, ontology, vulnerability, security risk

OPEN ACCESS

Edited by:
Linfeng Yang,
Guangxi University, China

Reviewed by:
Chen Zhang,
University of Shanghai for Science and Technology, China
Wei Lan,
Guangxi University, China

***Correspondence:**
Xiuzhen Chen
chenxz@sjtu.edu.cn

Specialty section:
This article was submitted to Smart Grids, a section of the journal Frontiers in Energy Research

Received: *26 April 2022*
Accepted: *16 May 2022*
Published: *16 June 2022*

Citation:
Hou S, Chen X, Ma J, Zhou Z and Yu H (2022) An Ontology-Based Dynamic Attack Graph Generation Approach for the Internet of Vehicles. Front. Energy Res. 10:928919. doi: 10.3389/fenrg.2022.928919

1 INTRODUCTION

Owing to the development of modern automobiles, people's daily travel is becoming increasingly more convenient and comfortable. In addition to efficiency and convenience, the rapid development of the technology of the Internet of Vehicles (IoV) has brought about a series of potential security threats, such as cloud supply chain issues, private data security, protocol cracking, and illegal data injection. For attackers, the attack surface of the IoV has become more extensive, and it is no longer limited to near-range physical attacks.

The 2020 Global Automotive Network Security Report (Upstream Security, 2020) released by Upstream Security reported an investigation of the security incidents happening in the automotive field since 2010 and concluded that the most common attack vectors use different entry points such as servers, mobile applications, and on-board diagnostic (OBD) ports. In addition, IOActive (Thuen, 2016), an Israeli automotive network information security company, scored the collected vehicle-related vulnerabilities on a one to five scale and found that approximately 72% of the vehicle-related

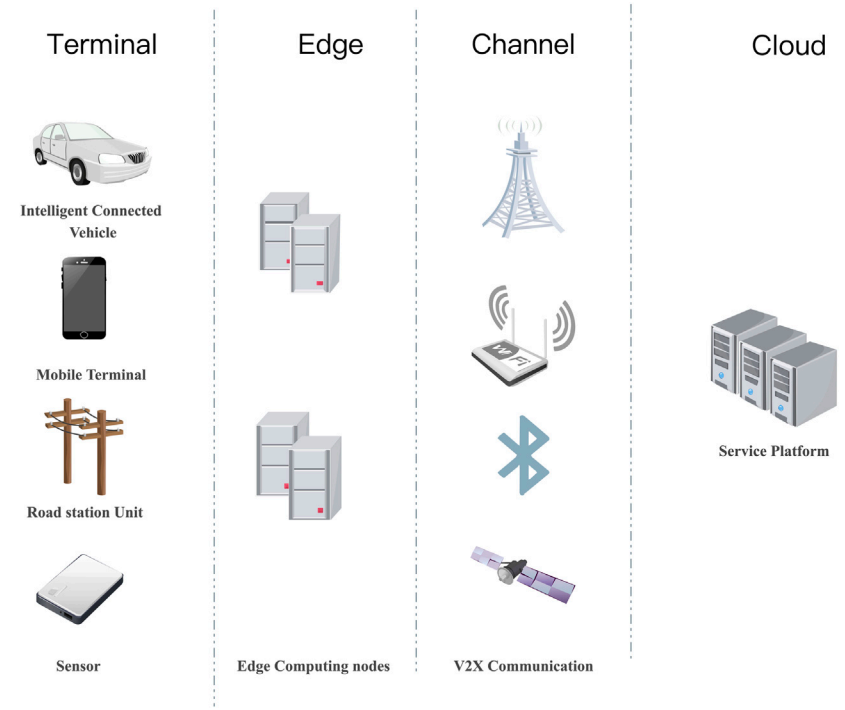


FIGURE 1 | The "cloud-channel-edge-terminal" architecture of the IoV.

vulnerabilities belonged to the medium- and low-risk categories. IOActive also pointed out that this did not necessarily mean that there was no significant risk. These single vulnerabilities may not be harmful; however, when multiple vulnerabilities are exploited in combination, the attack consequences can be immeasurable.

Taking the Tesla attack chain disclosed by Tencent Keen Lab as an example (Keen, Security, 2022), the researchers first considered the wireless network as the starting point of the attack. The vulnerability of the vehicle browser was used to execute arbitrary code in the browser. Furthermore, the kernel privilege escalation vulnerability was exploited to obtain the root privilege to extract information about the vehicle system. The researchers then bypassed the integrity check mechanism through the electronic control unit (ECU), modified and refreshed the firmware, and finally broke through the gateway to realize the transmission of any controller area network (CAN) message on the CAN bus. This is a complex attack path that exploits multiple vulnerabilities on various attack surfaces. The defense of an IoV system against an attack is insufficient if it only analyzes the vulnerabilities from a single level. It is necessary to comprehensively consider and analyze various attack surfaces to discover potential attack paths leading to the target under the "cloud-channel-edge-terminal" architecture of the IoV, as shown in **Figure 1**. Therefore, this study considered attack graph technology as an effective way to analyze the potential attack paths in an IoV system.

Attack graph technology (Lallie et al., 2020) can show potential attack paths and attack consequences through vertex and directed edge structures from the perspective of attackers in combination with specific network structure information. Most of the studies on attack graph are oriented toward ordinary enterprise network information systems. There are few studies on the attack graph of the IoV, a special information system, and there is also a lack of unified and standardized expression of IoV security knowledge. This paper proposes a dynamic attack graph generation approach based on ontology (McGuinness and Van Harmelen, 2004) for the IoV. A security ontology of the IoV is established, formulates a unified and standardized expression for security knowledge, which including the network topology, vulnerabilities and their relationships, and other security elements in the IoV, the proposed method can mine all potential attack paths in advance using the ontology inference engine HermiT (Glimm et al., 2014). The main contributions of this study are as follows:

1. This paper proposes an IoV network security ontology. The ontology models various security elements and their relationships in the IoV system, and then formulates a unified and standardized expression for IoV security knowledge.
2. This paper builds a knowledge base of the reasoning rules of the IoV. This paper analyzes different types of security vulnerabilities and corresponding attack methods, such as long-distance wireless attacks, short-range wireless attacks, and physical contact attacks, and describes specific attacks by Semantic Web Rule Language (SWRL) (Horrocks et al., 2004) rules.

3. This paper proposes a dynamic attack graph generation algorithm. The algorithm can update incrementally according to the change in network topology, which is more suitable for the IoV characterized by rapid changes in network topology. This can effectively demonstrate the vulnerability of the global IoV network and help with risk management and has lower computational complexity while updating attack graph.

The rest of this paper is organized as follows. **Section 2** discusses the existing related works. **Section 3** presents the components of the proposed model. **Section 4** shows how to construct different IoV network attack scenarios to verify the effectiveness of the algorithm. Finally, **Section 5** concludes the paper.

2 RELATED WORKS

The security threats of the IoV will not only cause economic losses to individuals and enterprises but also endanger personal safety and even national public security in severe cases. Therefore, several studies have been conducted to improve the security performance of the IoV. Existing studies on the security vulnerability analysis of the IoV are mainly divided into three levels: platform, network, and component levels (Li, 2019). They not only discover hidden vulnerabilities through penetration testing but also evaluate the risk globally using attack graphs and other methods such as attack trees and matrices. Detailed information on related works at each level is given as follows.

2.1 The Network Level

Studies on security at the network level mainly focus on the network communication security of the IoV, including identity authentication and privacy disclosure. Researchers have proposed group signature schemes (Shao et al., 2016), batch authentication schemes (Sutrala et al., 2020), and lightweight anonymous authentication schemes (Sadri and Rajabzadeh Asaar, 2020) to solve these problems.

2.2 The Platform Level

Studies on security at the platform level mainly aim at the security of the vehicle CAN bus and in-vehicle sensor network, which distinguishes normal behaviors from attack behaviors by extracting the vehicle characteristics (electrical, physical, and data packet characteristics) or by using a deep learning algorithm in the normal state (Choi et al., 2018; Song et al., 2020).

2.3 The Component Level

Studies on security at the component level focus on various vulnerabilities in the IoV system, analyze vulnerabilities and security risks based on the penetration testing results, and propose security protection suggestions or measures for the IoV network. Through the results of laboratory simulation experiments and actual road tests, Koscher et al. (Koscher et al., 2010) proved that attackers who gain control of key ECUs can evade the internal security features of intelligent networked vehicles and conduct malicious operations on them. In 2015, (Miller and Valasek, 2015), successfully hacked an in-vehicle communication system remotely by exploiting the vulnerability of the Uconnect system port in the Jeep vehicle. The final experimental results proved that attackers could exploit this vulnerability to control car braking and steering. In addition, they (Miller and Valasek, 2014) demonstrated the long-range attack surface of intelligent connected cars and analyzed the security of different manufacturers' car networks.

In general, the above studies on the security risks of the IoV mainly focus on the detection and mining of single vulnerabilities through simulation experiments, road experiments, and the use of fuzzing and feature extraction technologies, without considering the possibility of escalating the risks from the combined vulnerabilities at the same level or multiple levels.

In recent years, some researchers have looked into the application of graph technology in the field of IoV security. In the SAE J3601 "Cybersecurity Guidebook for Cyber-Physical Vehicle Systems" (SAE J3061 Vehicle Cybersecurity Systems Engineering Committee, 2016) launched by the American Society of Automotive Engineers in 2016, the attack tree technology is proposed to model the risk of the vehicle system network and quantify the system risk from the perspective of a threat. Kong et al. (Kong et al., 2018) proposed a security risk assessment framework for smart cars, to identify and assess security risk. The framework is based on the Guidelines for the Management of IT Security, using attack tree to analyze and categorize the assets, threats, and vulnerabilities. However, due to characteristics of attack tree, an attack tree can only analyze one attacker purpose. when faced with multi-target attackers, attack tree analysis is more complex than attack graph.

Salfer and Eckert, (2018) proposed an automatic generation model, called security analyzer for exploitability risks (SAlfER), which can semi-automatically quantify the risk of a given attacker by his/her exploitation steps with both budget and cost with cycles. In addition, they proposed an algorithm for a random attack graph for the security evaluation of a vehicle network, which creates a path for each starting node, continuously expands, clones, completes the path, and writes the path with similar dependencies in the attack graph to generate the result. The model mainly aims at the design stage before the vehicle goes into production, refers to the vehicle development documents, comprehensively considers various possibilities, calculates the security risk, and makes key business decisions in terms of safety and sustainability. Ibrahim et al. (Ibrahim et al., 2020) used the Architecture Analysis and Design Language (AADL) to assess the risk of vehicle system security cases. The vehicle system design, connection, weakness, resources, potential attack examples, and their pre-conditions and post-conditions were modeled using the AADL. The generated final attack diagram is displayed graphically to help the system administrator select the best countermeasures. The above two references demonstrate the feasibility of applying attack graph technology to security studies on the IoV. However, they either focused only on in-vehicle systems or performed only a static analysis using vehicle development documents.

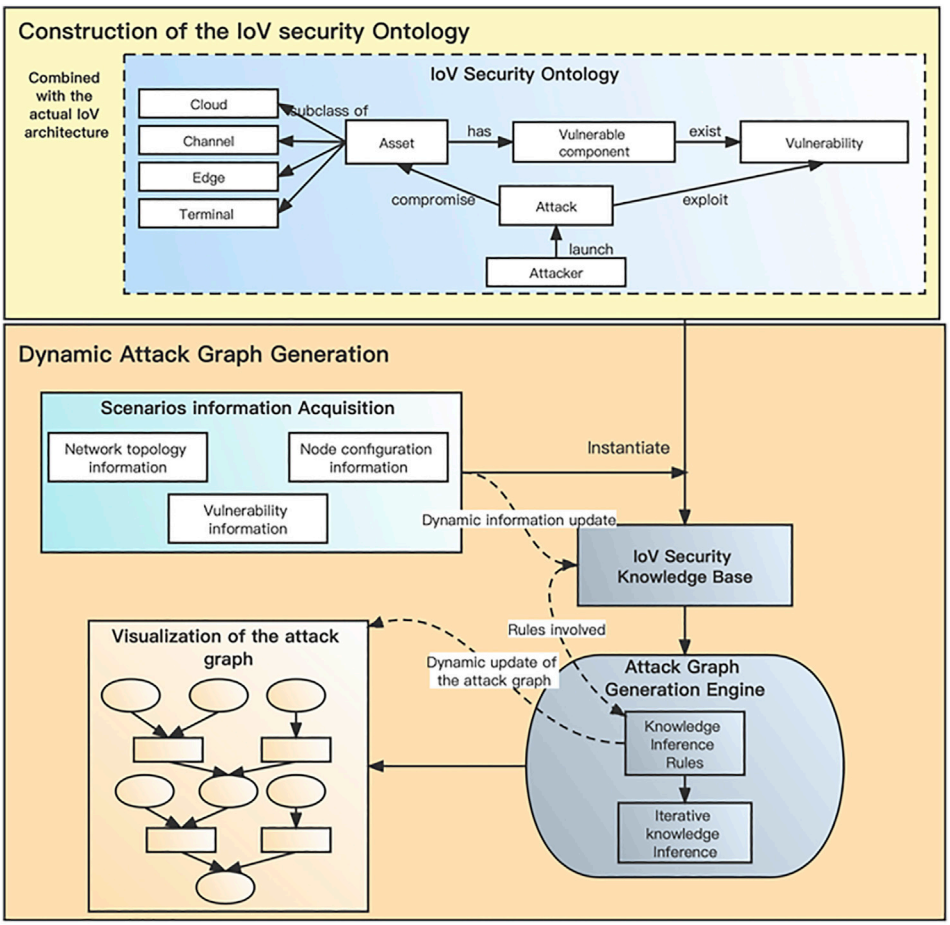


FIGURE 2 | IoV attack graph generation model.

In this paper, ontology and attack graphs were combined and applied to the field of IoV security to formulate a unified and standardized expression of IoV security knowledge. In view of the current situation of multistep attacks using the combined vulnerabilities in the IoV, this study integrates the overall architecture of the cloud-channel-edge-terminal to generate a comprehensive attack graph, which is conducive to the comprehensive risk analysis of IoV systems and has far-reaching significance for the construction of active defense systems for the IoV.

3 ONTOLOGY-BASED ATTACK GRAPH GENERATION ALGORITHM

Owing to the lack of a unified and standardized expression for security knowledge in existing studies on attack graphs of the IoV, this paper proposes a security ontology for the IoV system and formalizes a normative definition of the security elements in this system that can better describe its network architecture and vehicle-related security vulnerabilities. The IoV attack graph generation model proposed in this paper is shown in **Figure 2**. It includes two modules: construction of an IoV security ontology and generation of a dynamic attack graph. The former models the security elements and their relationships with the IoV system. The latter instantiates the entities to construct a knowledge base, inputs them into the inference engine HermiT, and finally generates a complete attack graph through the graph generation engine. When the scenario information changes, the inference engine locates the corresponding rules and nodes to update the attack graph in a timely manner.

3.1 Construction of the IoV Security Ontology

Ontology (Guarino et al., 2009) is a method and theory used to describe the essence of things, which refers to the formal specification of shared concepts in the same field. Research on building a network security ontology model (Iannacone et al., 2015) can better describe the intelligence and information related to attacks. This paper proposes an IoV security ontology by analyzing the cloud-channel-edge-terminal architecture of the IoV and abstracting the related security elements and their relationships, including the assets, vulnerabilities, and attacks.

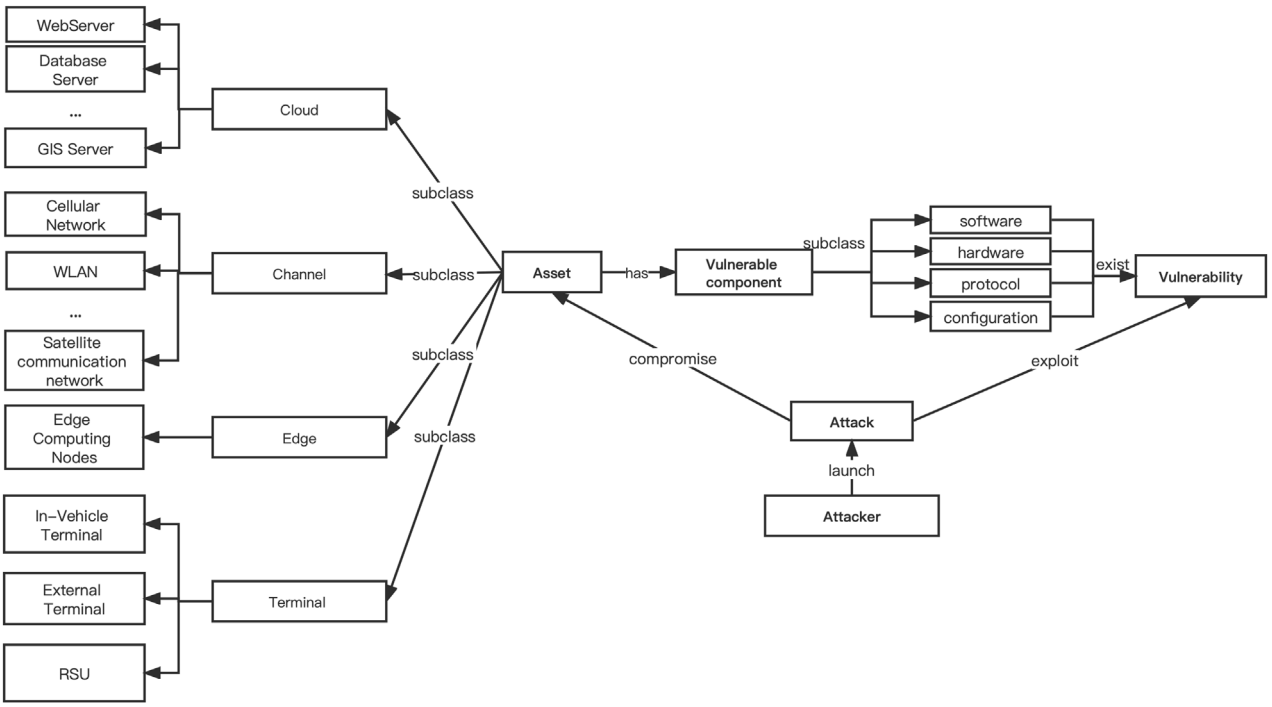


FIGURE 3 | IoV security ontology.

3.1.1 Definition of the IoV Security Ontology

This study combines the actual architecture of the IoV with existing security ontologies to build a specific security ontology for the IoV, as shown in **Figure 3**, which mainly includes five types of entities: assets, vulnerable components, vulnerabilities, attacks, and attackers. The entity classes are defined as follows:

Definition 1: The asset class. This class includes the equipment and sensitive data at all levels of the IoV system, i.e., a four-layer architecture of the cloud-channel-edge-terminal. The cloud layer corresponds to the application layer of the IoV, and the main equipment consists of various data collection and processing and cloud service support servers, including database, web, and travel navigation data processing servers. The channel layer corresponds to the network layer, including the cellular network 2G/3G/4G, WLAN, and the satellite communication network. The edge layer includes many edge computing nodes, which generally refer to the roadside equipment. The terminal layer corresponds to the perception layer, which is responsible for collecting vehicle location information and traffic information around the vehicles and for perceiving the environment and its state during driving. It primarily includes RFID readers and various communication terminals, such as vehicle terminals, mobile applications, and sensors. In addition, because ICVs have more than one attack surface, the asset entities at the terminal layer are further subdivided into in-vehicle terminals, external terminals, etc.

Definition 2: The vulnerable component class. This class shows the location of the vulnerability in the attack target assets. It is divided into subclasses such as software, hardware, programs, and services. These are all common types of components in the network assets of the IoV.

Definition 3: The vulnerability class. This class shows the technical drawbacks that can be exploited by attackers to launch attacks. Vulnerabilities are divided into software, hardware, and protocol vulnerabilities, according to the location where the vulnerability is hosted.

Definition 4: The attack class. The main subclasses include long-range wireless, short-range wireless, and physical access attacks. Among them, the long-distance wireless attacks are mainly carried out through Wi-Fi, cellular networks (3/4/5G), cloud platforms, GPS, etc. The short-range wireless attacks are mainly carried out through Bluetooth, keyless entrance systems, dedicated short-range communications, etc. The physical access attacks are performed through actual physical contact by the attackers.

Definition 5: The attacker class. According to the identity of the attacker, the attacker class is divided into internal and external attackers.

3.1.2 Relationship Model Between Entities

The relationship between entity classes is reflected by the entity properties, which are divided into datatype and object properties (McGuinness and Van Harmelen, 2004). Datatype properties are mainly properties of a single individual class, whereas object properties are properties that define the relationships between different individual classes.

- Datatype properties

The datatype properties of the asset class entities include the asset name, asset level, asset importance level, and asset motion status, which are only limited to vehicle entities. The value of the motion status is either still or moving.

The datatype properties of the vulnerable component class entities include the component name, component version, component function, and component-related information importance level.

The datatype properties of the vulnerability entities include the vulnerability ID, vulnerability description information, functions affected by the vulnerabilities, Common Vulnerability Scoring System score, patch information, and utilization probability.

The datatype properties of the attack entities include their Common Attack Pattern Enumeration and Classification ID, attack preconditions, and attack postconditions, i.e., attack benefits. Attack preconditions are prerequisite permissions required to implement this type of attack and are mainly composed of two parts: specific permissions and location. That is, the preconditions describe the specific type of software/hardware the attack needs to obtain the corresponding permissions. The attack postconditions are similar to the attack preconditions. The postconditions describe the expected benefits after the attack is launched and consist of specific permissions and locations.

The datatype properties of the attacker class include the attacker's location (distinguishing between long-distance and close-range attackers), capabilities, and permissions.

- Object properties

According to the relationship between the entity classes above, the object property defined in the IoV security ontology is as follows:

access property: This is a symmetric property, which indicates that both instances can access each other.

$$access(Asset, Asset)$$

compromise property: This property indicates that an attacker can successfully compromise an asset instance.

$$compromise(Attack, Asset)$$

hasComponent property: This property indicates that an asset instance has a containment relationship with a component instance.

$$hasComponent(Asset, Component)$$

exist property: This property represents the existence of a vulnerability in a component instance.

$$exist(Component, Vulnerability)$$

exploit property: This property indicates that an attacker needs to exploit the vulnerability of an instance to an attack.

$$exploit(Attack, Vulnerability)$$

launch property: This property indicates that an attacker needs to use a certain type of attack to launch an attack behavior.

$$launch(Attacker, Attack)$$

3.2 Attack Graph Generation Method for the IoV

Attack graph generation is divided into two stages: the initial attack graph generation and the dynamic attack graph update. The generation of attack graphs is based on the IoV security knowledge base, which is constructed by the SWRL rules of the security elements in IoV attack scenarios. Through inference engine HermiT, initial attack graph generates. When the network topology, vulnerability information, and other data change dynamically, the inference engine locates the corresponding rules and nodes to update the attack graph in a timely manner.

3.2.1 Inference Rules Knowledge Base

The SWRL (Horrocks et al., 2004) is a language that presents rules in a semantic manner and can be used in attack graph generation algorithms to infer the process of exploiting vulnerabilities to invade assets. An SWRL rule includes the body (inference precondition) and the head (inference result). The body part points out all the preconditions required for inference, including specific instances and the relationships between instances. And the head provides the inference results that can be obtained under the rule.?in rules represents this station is a variable. Entities can substate them.

The reasoning rules in this paper are divided into three types according to their different functions:

1. Vulnerability Existence Inference Rules

Vulnerability existence inference rules can infer whether an asset has a vulnerability based on its category and current version. An example of a vulnerability existence inference rule is provided below.

$$\begin{aligned} &Component(?comp) \hat{} hasversion(?comp, ?x) \\ &\hat{} Vulnerability(?vul) \\ &\hat{} hasupdateversion(?vul, ?y) \\ &\hat{} swrlb{:}\ lessThan(?x, ?y) \\ &\rightarrow exist(?comp, ?vul) \end{aligned}$$

2. Vulnerability Exploitability Inference Rules

Vulnerability exploitability inference rules can construct a single-step attack path and determine whether the attacker's resources and attack capabilities in the current state can attack the specified assets. An example of a vulnerability exploitability inference rule is provided below.

$$Asset(?asset)\hat{}Component(?comp)$$
$$\hat{}Vulnerability(?vul)\hat{}Attacker(?attacker)$$
$$\hat{}Attack(?attack)$$
$$\hat{}hasComponent(?asset,?comp)$$
$$\hat{}exist(?comp,?vul)$$
$$\hat{}exploit(?attack,?vul)$$
$$\hat{}launch(?attacker,?attack)$$
$$\hat{}assess(?asset,?attacker)$$
$$\rightarrow compromise(?attacker,?asset)$$
$$\hat{}attackbenefit(?attacker,?asset)$$

This rule means that when an asset instance has a vulnerability group and there is a vulnerable instance in the vulnerable component, there is a certain attack method that can exploit the vulnerability. If the attacker knows how to use this attack method and can successfully access the asset, it can be inferred that the asset can be compromised by the attacker and attacker gains the attack benefits.

Take attacker launches the command injection attack by exploit the vulnerability V1 in CAN bus to vehicle as an example, the corresponding rules rule are as follows:

$$Asset(vehicle)\hat{}Component(CANbus)$$
$$\hat{}Vulnerability(V1)\hat{}Attacker(attacker)$$
$$\hat{}Attack(commandInjection)$$
$$\hat{}hasComponent(vehicle,CANbus)$$
$$\hat{}exist(CANbus,V1)$$
$$\hat{}exploit(commandInjection,V1)$$
$$\hat{}launch(attacker,commandInjection)$$
$$\hat{}assess(vehicle,attacker)$$
$$\rightarrow compromise(attacker,CANbus)$$
$$\hat{}commandInjection(attacker,\ CANbus)$$

The rule means if the permissions obtained by the attacker meet all the preconditions above, the attacker can launch the command injection attack successfully and compromise the CAN bus on this vehicle.

3. Network Connectivity Inference Rules

The network connectivity inference rules can infer the attacker's access privileges according to the reachability between the resources and the assets compromised by the attacker. An example of a network connectivity inference rule is presented below.

$$compromise(?asset1,?attacker)$$
$$\hat{}assess(?asset1,?asset2)$$
$$\rightarrow assess(?attacker,?asset2)$$

This rule means that when the asset of instance 1 is compromised by an attacker instance and instance 1 can assess the asset of instance 2, it can be inferred that the attacker can access the asset of instance 2.

3.2.2 Initial Attack Graph Generation Algorithm

In the initial state, all the security elements in IoV attack scenarios and their relationships are instantiated. The inference engine then mines the causal relationships among the scattered vulnerabilities according to the inference rules, determines the potential attack paths of the attacker, and finally synthesizes all possible attack paths to generate a complete attack graph for the output.

Definition 6: Assume an attack graph $G=<C, V, E>$, where C represents the condition set (including all initial conditions, preconditions, and postconditions), V represents the set of vulnerabilities, and E represents the edge set.

The initial attack graph generation algorithm is mainly based on the breadth-first traversal (BFS) algorithm. The algorithm takes the security ontology instance of the target network, the attack scenario, and the attacker's target (optional) as inputs and generates an attack graph G in the form of a tuple. The algorithm adopts the method of forward chaining, starting from the initial conditions, based on the BFS to obtain additional properties by continuously searching for vulnerable hosts in the network. To generate an attack graph, line two to five in theAlgorithm 1 first obtains the initial object properties of the attacker CA and assets A through function Properties (). The second step (line 7–8) is to call inference engine by function onto. reasoner (), and obtain inferential facts, then locate the specific inference rule set F. The third step is to find the corresponding vulnerability nodes and pre-post-condition nodes for each rule of F and to construct the node set and edge set of the attack graph (line 12–25). The resources owned by the attacker are then updated until the inference ruleset is traversed (line 26). Then, the second step is repeated until the attacker's goal is reached and the algorithm terminates (loop starts on line 6). If the attacker does not specify a specific goal, it is assumed that the attacker's goal is to obtain all the resources that can be compromised in the system as much as possible, and the algorithm terminates when no new inference facts are generated.

3.2.3 Dynamic Attack Graph Update Algorithm

Owing to the rapid movement of vehicle nodes, the network topology update frequency is higher than that of traditional enterprise information systems. Therefore, this paper summarizes three situations that can cause changes in the IoV attack graph. These three situations are described below.

Network topology changes: Vehicles leaving/entering the communication area will cause network topology changes. This type of situation causes a change in the vehicle terminal node Nv. When a vehicle leaves or enters the communication area, the node should be deleted or added.

Network connectivity changes: Changes in firewall rules or when a moving vehicle communicates with different roadside base stations/cloud servers can cause changes in network connectivity. Changes in network connectivity are mapped to changes in attack scenarios, which are changes in object properties between asset nodes, and the properties of related nodes should be added or deleted.

Node information changes: The change in the motion state of a vehicle and in the node-related vulnerability information belongs to the change in the node information. When a vehicle is stationary, physical contact is possible; therefore, physical contact attacks are possible. When the motion state of a

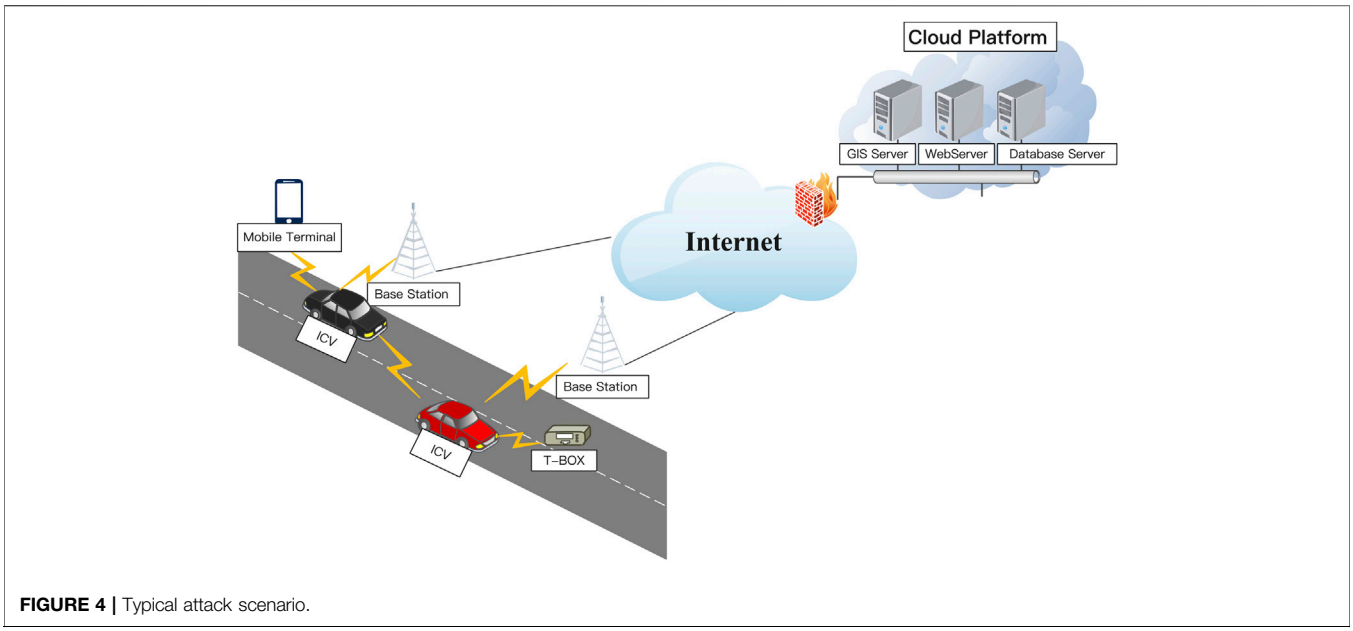


FIGURE 4 | Typical attack scenario.

vehicle changes, the connectivity relationship between the vulnerable nodes is related to physical contact changes, and the related security properties of the node should be added or deleted.

Algorithm 1 Attack graph generation algorithm

Input: $IoVOnto$, $attackScene$,$Goal$
output: G

$Instantiate(IoVOnto, attackScene)$
$A, P, CA \leftarrow \emptyset$
$A \leftarrow onto.asset()$
$P \leftarrow onto.Properties(A)$
$CA \leftarrow onto.Properties(attacker)$
while $Goal$ not in CA **do**
  $onto.reasoner()$
  $F \leftarrow reasoner.inferredFacts()$
  if $F == \emptyset$ **then**
    break
  else
    $FP \leftarrow onto.Properties(F.attack)$
    $P \leftarrow onto.Properties(F)$
    for $fp \in FP$ **do**
      $r \leftarrow getRuleOfAttack(fp)$
      $v_i \leftarrow r.vul()$
      $s_{pre} \leftarrow body(r)$
      $s_{post} \leftarrow head(r)$
      if $s_{pre} \in P$ **then**
        $c_i \leftarrow s_{pre}$
        $C \leftarrow C \cup c_i \cup c_{i+1}$
        $V \leftarrow V \cup v_i$
        $E \leftarrow E \cup (c_i, v_i) \cup (v_i, c_{i+1})$
      end if
    end for
    $CA \leftarrow onto.Properties(attacker)$
  end if
end while
checkLoop(C,V,E)
$G \leftarrow < C, V, E >$

Algorithm 2 Algorithm of Adding Security Properties

Input: $IoVOnto$,$changes$,$Goal$,G
output: G'

$CA \leftarrow onto.Properties(attacker)$
$p_{add} \leftarrow \emptyset$
$n_{add} \leftarrow \emptyset$
for $change \in changes$ **do**
  if $isvul(change)$ **then**
    $V \leftarrow V \cup change$
    $IoVOnto.add(change)$
    $p_{add} \leftarrow p_{add} \cup onto.Properties(change)$
  else
    $p_{add} \leftarrow p_{add} \cup change$
  end if
end for
$IoVOnto.add(p_{add})$
line $6-30$ *in algorithm*1

Algorithm of Adding Security Properties: When a node or security properties are added, first, the corresponding node should be instantiated to build the datatype properties associated with Nv and the object properties with the newly added vulnerability node Nv (line 1–13 in Algorithm 2). Then, the algorithm calls the inference engine to determine the new inference facts, finds the reasoning rules related to the reasoning fact, adds nodes and edges to the attack graph, and adds the attack consequences to the scene for further reasoning until no new reasoning facts are generated, which is the same diagram in Algorithm 1.

Algorithm of Deleting Security Properties: In the case of deleting a node or security properties, the algorithm first obtains all the properties of Nv, finds all sets of the vulnerable nodes V that Nv is related to and the condition nodes related to Nv, deletes all vulnerable nodes in V and the edges related to the vulnerable

nodes in the attack graph, and deletes the related condition nodes and the edges connected to them at the same time.

Algorithm 3 Algorithm of Deleting Security Properties

Input: $IoVOnto$,$changes$,$Goal$,G
output: $G^{'}$
$CA \leftarrow onto.Properties(attacker)$
$p_{delete} \leftarrow \emptyset$
$n_{delete} \leftarrow \emptyset$
for $change \in changes$ **do**
  if $isvul(change)$ **then**
    $V \leftarrow V \backslash change$
    $p_{delete} \leftarrow p_{delete} \cup onto.Properties(change)$
  else
    $p_{delete} \leftarrow p_{delete} \cup change$
  end if
end for
$delete \leftarrow p_{delete}$
$R \leftarrow getSWRLRule(IoVOnto)$
while $delete! = \emptyset$ **do**
  for $n \in delete$ **do**
    for $r \in R$ **do**
      if $n \in body(r)$ **then**
        $c_i \leftarrow body(r)$
        $v_i \leftarrow r.vul()$
        $C \leftarrow C \backslash c_i$
        $E \leftarrow E \backslash (v_i, c_i)$
        $c_{i+1} \leftarrow head(r)$
        $C \leftarrow C \backslash c_{i+1}$
        $E \leftarrow E \backslash (v_i, c_{i+1})$
        $temp \leftarrow temp \cup head(r)$
      end if
    end for
  end for
  $delete \leftarrow temp$
end while
$G^{'} \leftarrow < C, V, E >$

4 EXPERIMENT RESULTS AND ANALYSIS

4.1 Experiment on a Typical Attack Scenario

4.1.1 Construction of a Typical Attack Scenario

According to the White Paper on Security Penetration of Intelligent and Connected Vehicles (China Software Tesing Center, 2020), the types of attacks in typical attack scenarios can be divided into three categories according to different attack surfaces: long-distance wireless attacks, short-range wireless attacks, and physical contact attacks. Therefore, when constructing the attack scenario, considering these three attack methods, we selected ten vulnerabilities related to the IoV system and constructed the attack scenario shown in **Figure 4**. Among them, V1 and V2 are vulnerabilities in cloud servers, which are mainly related to long-distance wireless attacks, and the consequences include sensitive information collection and privilege escalation. V3 is a low-version browser vulnerability on in-Vehicle infotainment (IVI) system. V4 is an unverified vulnerability in a Wi-Fi connection, which is related to short-range wireless attacks. A short-range attacker can gain access to a vehicle from this vulnerability. V10 is a physical contact vulnerability exposed outside of the vehicle. An attacker can access the internal bus of a vehicle through physical contact with the OBD interface and implement command injection attacks on the vehicle. The vulnerability information is presented in **Table 1**.

4.1.2 Attack Graph Generation and Analysis

The attack graph generated by the experiment at T_0 is shown in **Figure 5**. There are four attack paths in total.

Path 1:Through the cloud vulnerability V1, attacker can launch attack $Info\_Collection(attacker, Vehicle1)$, and collect private information of vehicle 1. Based on that, attacker conducts social engineering attack $exploit(SocialEngineeringAttack, CVE-2015-5065)$, and successfully gains the account information of vehicle 1 $hasAccount(attacker, Vehicle1)$. The attacker obtains the invading interface and local user privilege of invading vehicle 1. Finally, by exploiting vulnerability V7 on ECU Gateway, attackers can finally achieve arbitrary code execution attacks $execArbitraCode(attacker, Vehicle1)$. Vehicle 1 will receive arbitrary codes that cause denial of service attacks or other malicious operations possibly.

Path 2 is a short-range wireless attack. Owing to the lower version of the installed car browser, a permission vulnerability V3 in IVI was exploited by attacker $gainUserPrivi(attacker, Vehicle1.IVI)$. The attacker obtains the local user privilege of vehicle 1. Then attacker discovers a shell vulnerability on the system, escalates the shell privilege to obtain the root permission of the IVI system

TABLE 1 | Vulnerability information.

Vulnerability	CVE ID	Vulnerability Component	Attack Consequence
V1	CVE-2015-1761	SQLServer	Information Collection
V2	CVE-2013-5065	WebServer	Privilege Escalation
V3	CVE-2009-1725	In-vehicle Browser	Permission Acquisition
V4	CVE-2018-11476	Wi-Fi	Unauthenticated Access
V5	CVE-2018-11477	Wi-Fi	Information Collection
V6	CVE-2018-11478	OBD Dongle	Command Injection
V7	CVE-2013-6282	ARM Linux	Privilege Escalation
V8	CVE-2016-9337	ECU Gateway	Arbitrary Code Execution
V9	CVE-2018-9311	T-Box	Initial Access
V10	CVE-2018-9322	OBD Interface	Verify Bypass

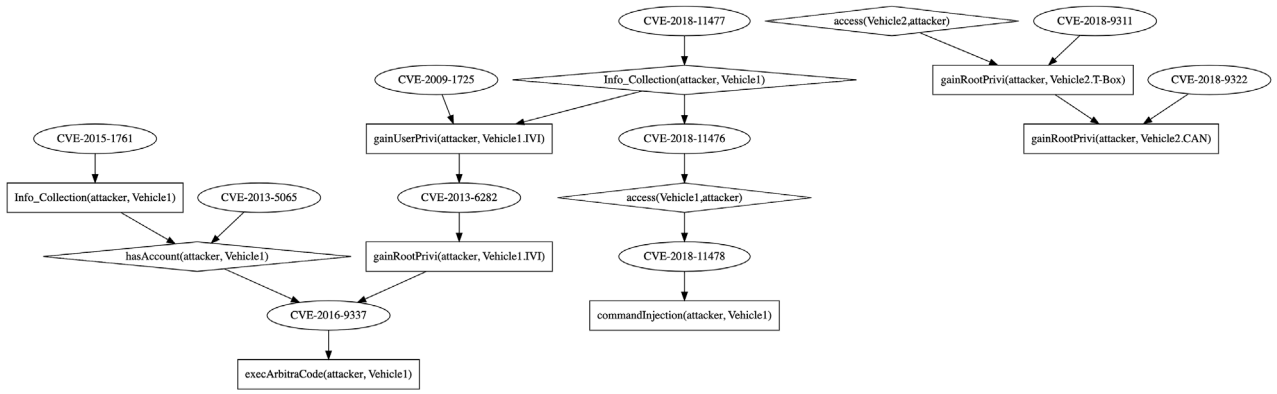


FIGURE 5 | Attack graph at time T_0.

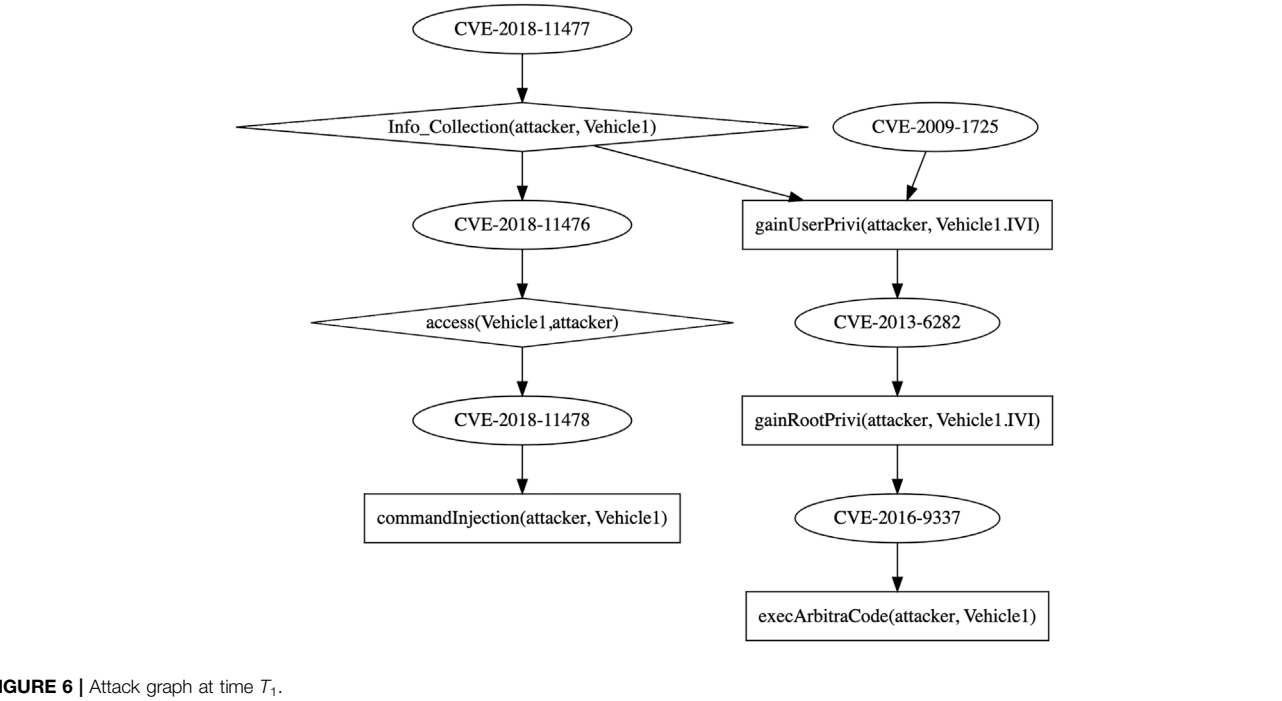


FIGURE 6 | Attack graph at time T_1.

gainRootPrivi(*attacker*, *Vehicle*1.*IVI*), and finally passes the ECU gateway vulnerability. Command injection attacks can thus be exploited *execArbitraCode*(*attacker*, *Vehicle*1) .Vehicle 1 will receive arbitrary codes that cause denial of service attacks or other malicious operations possibly.

Path 3 is a physical contact attack. Here, the attacker successfully collects sensitive vehicle information $Info\_Collection(attacker, Vehicle1)$ and accesses vehicle 1 through an unauthenticated Wi-Fi connection $access(Vehicle1, attacker)$, and then obtains command injection permissions through a vulnerability in the OBD dongle $commandInjection(attacker, Vehicle1)$. Vehicle 1 will receive and execute malicious commands possibly.

Path 4 is a combination of short-range wireless and physical contact attacks. Here, the attacker uses a pseudo base station and a signal amplifier to conduct a man-in-the-middle attack between vehicle 2 and the telematics service provider (TSP) and then uses the telematics communication box (T-Box) vulnerability V9 to implant a backdoor $gainUserPrivi(attacker, Vehicle2.T-Box)$. At the same time, the attacker obtains the access rights of the vehicle 2 bus by using the vulnerability V10 bypassed by the OBD interface verification and finally obtains the control rights

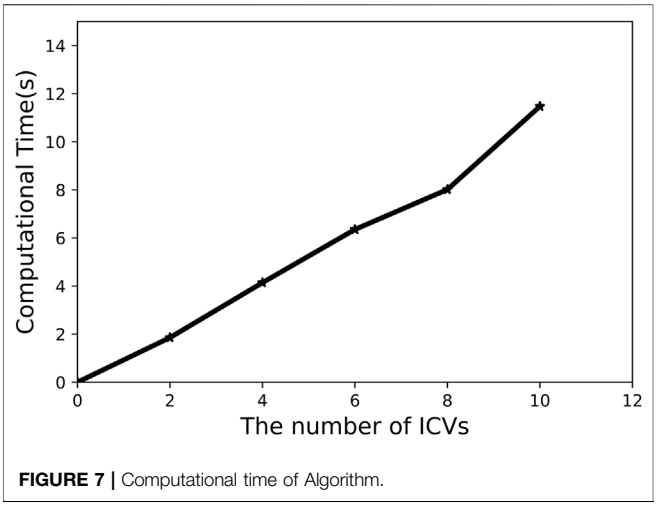


FIGURE 7 | Computational time of Algorithm.

of the T-Box and the vehicle CAN bus $gainRootPrivi(attacker, Vehicle2.CAN)$. The attacker gained top administrative privileges on the vehicle 2.

Consider the following situation: the cloud webserver has updated at T_1 and the vulnerability V2 is remedied. At the same time, the motion state of vehicle 2 changes from static to moving, and the V9 vulnerability exploitation condition is not established. In this case, the V2 and V9 vulnerability nodes should be deleted. Moreover, inference rules whose preconditions involve changing the object properties should be determined, as well as the relevant condition nodes, and then the nodes and the connected edges should be deleted. The dynamic attack graph generated after the node information changes at T_1 is shown in **Figure 6**.

4.2 Algorithm Complexity Analysis

To demonstrate the scalability of algorithm, a series of experiments are conduct. The network configuration information of cloud platform is the same as typical attack scenario. And we simulated a standard vulnerable ICV configuration which has four exploitable vulnerabilities. We gradually increase the number of standard ICVs and obtain the computational time of the algorithm. **Figure 7** shows average computational time in each attack scenario. The experimental results show that with the increase of the number of vehicles, the computational time of attack graph generation does not increase exponentially. After the initial attack graph generation, the attack graph only updates the local attack graph rather than the overall attack graph generation.

4.3 Algorithm Complexity Analysis

To verify the scalability of the algorithm, this paper evaluates the efficiency of the algorithm by analyzing its time complexity. Given the number of nodes in graph generated N and the number of edges E, when the network information of the IoV changes, the algorithm only updates the corresponding instance relationships and properties in the ontology and makes local changes to the attack graph. The time complexity of this local change is $O(\Delta n+\Delta e)$. The traditional attack graph generation tool MulVAL (Ou, Govindavajhala and Appel, 2005) needs to reconstruct the overall attack graph, and the time complexity of this algorithm is $O(n + e)$. Obviously, $O(\Delta n+\Delta e) < O(n + e)$. Compared with regenerating the complete attack graph, the dynamic attack graph generation algorithm based on the local update proposed in this paper reduces the computational overhead, improves the timeliness of the attack graph, and can better adapt to the rapid change in the IoV topology.

5 CONCLUSION

This paper proposes an IoV network security ontology model in combination with the "cloud-channel-edge-terminal" and constructs an IoV reasoning rule knowledge base by SWRL rules. Both of them help with applying attack graph technology to the field of network security of the IoV and can describe attack scenarios in IoV system well. On this basis, this paper proposes a dynamic attack graph generation algorithm that can be updated incrementally according to the changes in the network topology, which is more suitable for IoV networks characterized by rapid changes in network topology. The algorithm can show the vulnerability of the global IoV network effectively and help carry out better risk management and has lower algorithm complexity while updating attack graph. Finally, the experiment results demonstrated the effectiveness and feasibility of the proposed algorithm.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

HS:Conceptualization, Methodology,Software, Writing—Original Draf CX:Writing—Review and Editing, Supervision MJ:Supervision, Resources ZZ:Supervision, Resources YH:Data Curation.

FUNDING

This study was supported by the Joint Funds of the National Natural Science Foundation of China (no. U2003206).

REFERENCES

China Software Tesing Center (2020). Intelligent Internet Connected Vehicle Safety Penetration White Paper. Available at: https://www.cstc.org.cn/info/1202/231181.htm.

Choi, W., Joo, K., Jo, H. J., Park, M. C., and Lee, D. H. (2018). VoltageIDS: Low-Level Communication Characteristics for Automotive Intrusion Detection System. *IEEE Trans.Inform.Forensic Secur.* 13 (8), 2114–2129. doi:10.1109/TIFS.2018.2812149

Glimm, B., Horrocks, I., Motik, B., Stoilos, G., and Wang, Z. (2014). HermiT: An OWL 2 Reasoner. *J. Autom. Reason.* 53 (3), 245–269. doi:10.1007/s10817-014-9305-1

Guarino, N., Oberle, D., and Staab, S. (2009). "What Is an Ontology?," in *Handbook on Ontologies* (Springer), 1–17. doi:10.1007/978-3-540-92673-3_0

Horrocks, I.,(2004). SWRL: A Semantic Web Rule Language Combining OWL and RuleML. *W3C Memb. Submiss.* 21 (79), 1–31.

Iannacone, M. (2015). "Developing an Ontology for Cyber Security Knowledge Graphs," in Proceedings of the 10th Annual Cyber and Information Security Research Conference, April 2015, 1–4. doi:10.1145/2746266.2746278

Ibrahim, M., Al-Hindawi, Q., Elhafiz, R., Alsheikh, A., and Alquq, O. (2020). Attack Graph Implementation and Visualization for Cyber Physical Systems. *Processes* 8 (1), 12. doi:10.3390/pr8010012

Keen, Security (2022). *Lab Blog* (No Date) *Keen Security Lab Blog*. Available at: http://keenlab.tencent.com/index.html (Accessed May 13, 2022).

Kong, H.-K., Hong, M. K., and Kim, T.-S. (2018). Security Risk Assessment Framework for Smart Car Using the Attack Tree Analysis. *J. Ambient. Intell. Hum. Comput.* 9 (3), 531–551. doi:10.1007/s12652-016-0442-8

Koscher, K. (2010). "Experimental Security Analysis of a Modern Automobile," in Proceedings of the IEEE symposium on security and privacy, Oakland, CA, USA, May 2010 (IEEE), 447–462. doi:10.1109/sp.2010.34

Lallie, H. S., Debattista, K., and Bal, J. (2020). A Review of Attack Graph and Attack Tree Visual Syntax in Cyber Security. *Comput. Sci. Rev.* 35, 100219. doi:10.1016/j.cosrev.2019.100219

Li, X. (2019). Survey of Internet of Vehicles Security. *J. Cyber Secur.* 4 (03), 17–33. doi:10.19363/J.cnki.cn10-1380/tn.2019.05.02

McGuinness, D. L., and Van Harmelen, F. (2004). OWL Web Ontology Language Overview. *W3C Recomm.* 10 (10), 2004.

Miller, C., and Valasek, C. (2014). *A Survey of Remote Automotive Attack Surfaces*. South Las Vegas, Nevada: black hat USA, 94.

Miller, C., and Valasek, C. (2015).Remote Exploitation of an Unaltered Passenger Vehicle. Black Hat USA, S 91.

Ou, X., Govindavajhala, S., and Appel, A. W. (2005). "MulVAL: A Logic-Based Network Security Analyzer," in *USENIX Security Symposium* (Baltimore, MD, 113–128.

Sadri, M. J., and Rajabzadeh Asaar, M. (2020). A Lightweight Anonymous Two-Factor Authentication Protocol for Wireless Sensor Networks in Internet of Vehicles. *Int. J. Commun. Syst.* 33 (14), e4511. doi:10.1002/dac.4511

SAE J3061 Vehicle Cybersecurity Systems Engineering Committee (2016). *Cybersecurity Guidebook for Cyber-Physical Vehicle Systems*. Warrendale, Pennsylvania: SAE International.

Salfer, M., and Eckert, C. (2018). "Attack Graph-Based Assessment of Exploitability Risks in Automotive On-Board Networks," in Proceedings of the 13th International Conference on Availability, Reliability and Security. ARES 2018: International Conference on Availability (Hamburg Germany: Reliability and SecurityACM), 1–10. doi:10.1145/3230833.3230851

Shao, J., Lin, X., Lu, R., and Zuo, C. (2016). A Threshold Anonymous Authentication Protocol for VANETs. *IEEE Trans. Veh. Technol.* 65 (3), 1711–1720. doi:10.1109/TVT.2015.2405853

Song, H. M., Woo, J., and Kim, H. K. (2020). In-vehicle Network Intrusion Detection Using Deep Convolutional Neural Network. *Veh. Commun.* 21, 100198. doi:10.1016/j.vehcom.2019.100198

Sutrala, A. K., Bagga, P., Das, A. K., Kumar, N., Rodrigues, J. J. P. C., and Lorenz, P. (2020). On the Design of Conditional Privacy Preserving Batch Verification-Based Authentication Scheme for Internet of Vehicles Deployment. *IEEE Trans. Veh. Technol.* 69 (5), 5535–5548. doi:10.1109/TVT.2020.2981934

Thuen, C. (2016). Commonalities in Vehicle Vulnerabilities. Retrieved from Infosecurity.

Upstream Security (2020). Global Automotive Cybersecurity Report | Upstream. Available at: https://upstream.auto/upstream-security-global-automotive-cybersecurity-report-2020/(Accessed May 13, 2022).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Energy Research

ORIGINAL RESEARCH
published: 11 July 2022
doi: 10.3389/fenrg.2022.953082

Optimization and Control for Solid Oxide Fuel Cell System Hybrid DC Microgrids From the Perspective of High Efficiency, Thermal Safety, and Transient Response

Lin Zhang[1†], *Wenhui Tang*[1], *Feng Wang*[1], *Chao Xie*[1], *Weibin Zhou*[1] *and Hongtu Xie*[2†*]

[1]*Department of Early Warning Technology, Air Force Early Warning Academy, Wuhan, China,* [2]*School of Electronics and Communication Engineering, Sun Yat-sen University, Guangzhou, China*

OPEN ACCESS

Edited by:
Yongxi Zhang,
Changsha University of Science and Technology, China

Reviewed by:
Lepintg Chen,
National University of Defense Technology, China
Haishan Tian,
Hunan Normal University, China
Dong Feng,
National University of Defense Technology, China

****Correspondence:***
Hongtu Xie
xiehongtu@mail.sysu.edu.cn

†These authors share first authorship

Specialty section:
This article was submitted to Smart Grids, a section of the journal Frontiers in Energy Research

Received: *25 May 2022*
Accepted: *08 June 2022*
Published: *11 July 2022*

Citation:
Zhang L, Tang W, Wang F, Xie C, Zhou W and Xie H (2022) Optimization and Control for Solid Oxide Fuel Cell System Hybrid DC Microgrids From the Perspective of High Efficiency, Thermal Safety, and Transient Response. Front. Energy Res. 10:953082. doi: 10.3389/fenrg.2022.953082

Managing the power transients with both high efficiency and thermal management constraints is a difficult task in the solid oxide fuel cell (SOFC) hybrid direct-current (DC) microgrids. This article proposed a SOFC-based DC microgrid, which consists of the SOFC stand-alone system, DC microgrid network, and DC/DC boost converter, along with the associated DC/DC load. The control and optimization strategy of the SOFC-based DC microgrids has been designed, including the thermal and electrical characteristics analysis. First, the voltage and current regulator based on the proportional–integral (PI) is designed, which can maintain the voltage stability of the proposed SOFC system. In addition, the optimal regulator based on the optimal operating points (OOPs) is designed, which can realize high efficiency and steady-state thermal safety of the proposed SOFC system. Finally, the main performance, including the SOFC stack electrical characteristics verification, SOFC stand-alone system electrical and thermal response, load-tracking characteristics, and system efficiency is observed and discussed in the proposed control and optimization strategy. Experimental results verify the correctness of the theoretical analysis and the effectiveness of the proposed optimization and control strategy.

Keywords: solid oxide fuel cell (SOFC), direct-current (DC) microgrid, control and optimization, high efficiency, thermal safety, transient response

1 INTRODUCTION

In recent years, global energy demand is steadily increasing (Zhang et al., 2020a; Deng et al., 2021), which has made a great impact on various fields (Xie et al., 2020). Thus, the electric power generation, transmission, and distribution around the globe will be subjected to pre-eminent concern due to several reasons, such as the limited fossil fuel resources, incremental electric power consumption, global climatic change, legislation for integrating renewable energy sources (RESs), and stochastic properties of the RESs and their associated challenges (Zhang Y et al., 2019; Zhang et al., 2020b; Xu et al., 2022). Compared to the alternating-current (AC) microgrids, the direct-current (DC) microgrids have attracted both researchers and industrialists since it has many advantages, including the role to effectively solve the effects of the distributed generation accessing network to improve the power quality and transactive energy, and the elimination of the frequency and phase control (Eid, 2014; Srinivasan and Kwasinski, 2020). Among several novel propositions, the solid

oxide fuel cell (SOFC) is one of the most effective and efficient RESs, which can generate electricity directly from the electrochemical reaction with the least spread of pollution, compared to the conventional energy production methods (Wang et al., 2022). The SOFCs are excellent regional power supply equipment applied to the power supply systems and will be widely used in the DC microgrids, due to the advantages of fuel adaptability, high efficiency, noiseless operation, and high reliability. The advantages of the SOFC generation in the microgrids can be summarized as follows:

- Improving system efficiency. Fuel utilization of the SOFC system is more than 60% and its energy storage systems (ESSs) sometimes reached 90% (Barnes, 2002). Its power generation efficiency can be 55%–65%, which is generally more efficient than the combustion engines (Singhal and Kendall, 2003). This is important in the DC microgrids needed for the combined heat and power systems.
- Cutting down expenditures. The wide varieties of fuel can be used in the SOFC system, and the DC microgrids based on the RES can be easily interfacing with the distribution generation without the interlinking AC/DC-and DC/AC-interfacing transformation stages, just by realizing the combined power supply through a variety of the new energy sources. Its implementation is simple and cost-effective. It shows a good application perspective in majority of the rural areas that do not have access to reliable electricity (Gandini and Almeida, 2017; Hirsch et al., 2018).
- Strengthening capability of the regulating peak. The SOFCs can serve as the independent power generation as well as the grid-connected system, and many SOFC power stations connected to the power grid can greatly enhance the peak regulation capability of the microgrids.
- Improving security. The worldwide power grids mostly use the high-voltage electricity for long distance transmission, which result in low reliability and serious losses caused by large area blackouts due to accidents. The SOFCs can go on the power supply for the basic load to support the power grid.

However, the researches on the SOFC are still in their beginning stage, with the lack of dynamic load following, the poor thermal response, and gas starvation being some of the drawbacks of SOFC systems that need to be addressed (Mumtaz et al., 2018; Pranita et al., 2022), which requires constant improvement and development to large-scale commercial applications, especially the fast load tracking on the premise of the high efficiency and operating safety under optimal working conditions should be discussed in the DC microgrids. For the SOFC-based DC microgrid development, safe operating, including thermal management and gas starvation, is an important control task, especially the high operating temperatures may lead to the thermal gradient and local hot spots (Zhang et al., 2010). Moreover, the optimal operation points and optimized power switching strategies are discussed to improve the system efficiency. In addition, the load transients often involve significant peaks in power relative to steady-state load (Li et al., 2015). Thus, load tracking is another important control task in the DC microgrids.

In recent years, diverse thermal safety has been focused and discussed to ensure the SOFC system operating in the proper temperature. When the SOFC system temperature has been recognized to have the significant impact on the cell lifespan and operating safety; in order to mitigate temperature excursion and ensure thermal safety, the excess air for the cooling SOFC system is controlled to maintain the SOFC system within a safe range (Sorrentino et al., 2008; Huo et al., 2010; Hajimolana et al., 2013), including a proportional–integral (PI) controller, a variable structure controller, or a neural network predictive controller, which are suitable for the thermal safety control. However, these control schemes had not considered the maximum electrical efficiency operations of the SOFC stand-alone system. Moreover, the gas starvation problem, considering the operating safety, is also discussed (Carré et al., 2015), by developing a feed-forward control for the SOFC system with anode-off gas recycle. The results show that the control scheme was sensitive to external disturbance and produced a steady-state error.

Many optimal control methods have been raised and developed, which can provide an important reference on the swift response and dynamic load variation issues of the SOFC-based DC microgrids. The load-tracking studies under both stand-alone and grid-connected conditions have been discussed, especially the load-tracking and small signal stability issues pertaining to grid-connection is investigated (Padulles et al., 2000; Zhu and Tomsovic, 2002; Li et al., 2005). An adaptive control paradigm is portrayed for the swift response of the SOFC in a grid-connected microgrid (Awais et al., 2021). The aforementioned control strategies are restricted to its constraints or unable to tackle the sudden and large load variations, and the cooperative thermo-electrical control strategy has been not considered, which further needs to be discussed and analyzed.

The main goal of this article was to design the optimization and control strategy to ensure fast load following by comprehensively considering the high efficiency, fuel exhaustion, and thermal safety in the SOFC-based hybrid DC microgrids. This article is organized as follows: **Section 2** describes the hybrid power system architecture, including its essential operational requirements. **Section 3** deals with the modeling and validation and optimization and control strategy of the SOFC-based hybrid DC microgrid. Finally, the article ends with a conclusion in **Section 4**.

2 SOLID OXIDE FUEL CELL HYBRID DIRECT-CURRENT MICROGRID SYSTEM LAYOUT

2.1 System Architecture

The overall block diagram of the SOFC-based hybrid DC microgrid system is shown in **Figure 1**. It includes the SOFC stand-alone system, commonly used DC/DC boost converters, whose controlled output voltage is greater than the input voltage,

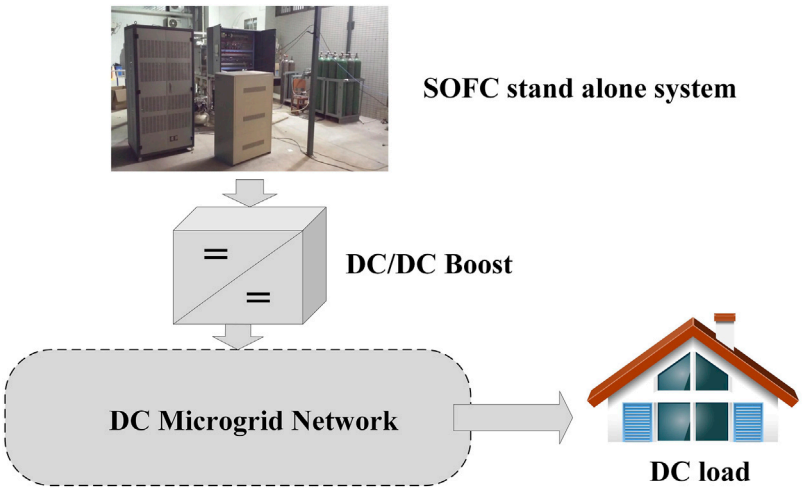


FIGURE 1 | Schematic diagram of the SOFC-based hybrid DC microgrid.

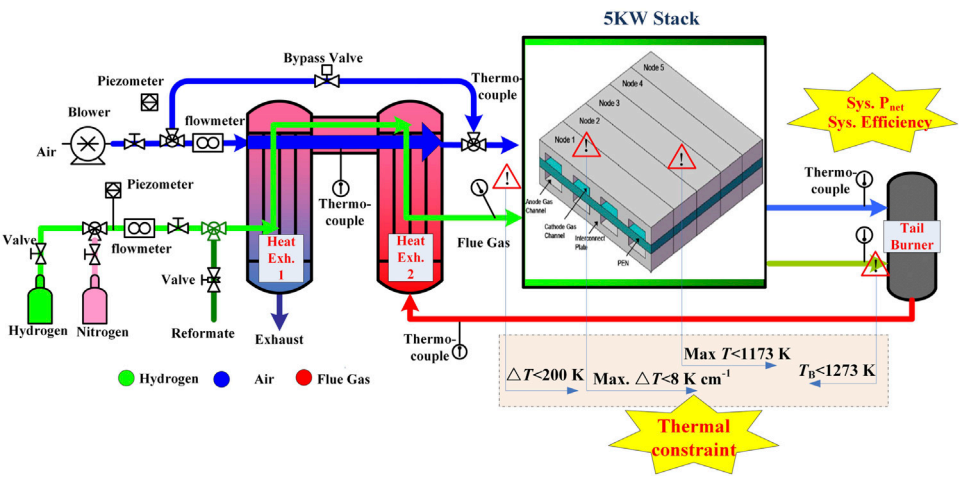


FIGURE 2 | Overall SOFC stand-alone system layout.

which can maintain the high reliability and load power supply (Ahmed and Blejis, 2013; Armghan et al., 2020; Liu et al., 2020). The investigated SOFC stand-alone system is shown in **Figure 2**. The system incorporates the stack and balance of plant (BOP), including the fuel feed pipes and valves, air feed pipes and valves, heat exchangers, SOFC stack, and tail-burner. The rated full power of the co-flow planar stack is 5 kW, which can provide the power to the external load through electrochemical reactions. Generally, a classic SOFC BOP includes the air/fuel feed pipes and valves, second air bypass manifold, fuel and air heat exchangers, and tail-burner. Then, the fuel and air heat exchangers are designed to minimize the inlet temperature difference of the SOFC stack. The tail-burner can promote the utility rate of the fuel by burning the tail gas. Meanwhile, its output hot gas is spilled into the heat exchanger as the heat source. Finally, the temperature of the SOFC system is convenient to control by adding the second air bypass manifold.

2.2 System Essential Operational Requirements

In addition to the fast power tracking, operating safety (including the thermal safety and avoiding fuel exhaustion) and high efficiency are the most important considerations for control and optimization. This article mainly discusses the thermal performance indices associated with thermal safety, fuel concentration indices associated with fuel exhaustion, and optimal operation points associated with high efficiency, which are depicted as follows:

TABLE 1 | OOPs of the SOFC stand-alone system.

SOFC net output power P_{net} (W)	Output current I_s (A)	Inlet air flow rate F_{air} (mol/s)	Inlet fuel flow rate F_{H2} (mol/s)	Bypass valve opening ratio *BP* (%)
1000	10	0.09920	0.00772	0.2
1500	14	0.13888	0.01135	0.2
2000	20	0.19841	0.01543	0.1
2500	26	0.27405	0.02006	0.05
3000	32	0.34390	0.02469	0
3500	38	0.37697	0.02932	0
4000	44	0.43649	0.03419	0
4500	50	0.49601	0.03997	0
5000	52	0.57538	0.04774	0
5500	58	0.63490	0.05428	0

2.2.1 Thermal Performance Indices

The high temperature or temperature gradient in the SOFC may cause material deformation or even damage, thus being the essential operational requirements. For example, as one of the most important components, the high temperature gradient in stack could result in large thermal stress and may cause stack deformation or even damage. In addition, stack temperature should be within the materials' bearing range. Under the aforementioned assumption, the temperature constraints in the SOFC can be shown as follows:

(1) Burner temperature $T_B \leq 1273K$.
(2) Maximum positive-electrode-electrolyte-negative (PEN) temperature $Max.T_{PEN} \in [873K, 1173K]$.
(3) Maximum PEN temperature gradient $Max.|\Delta T_{PEN}| \leq 8Kcm^{-1}$.
(4) Stack inlet temperature difference $\Delta T_{inlet} \leq 200K$.

2.2.2 Fuel Concentration

The external load changes, ensuring that adequate fuel supply is one of the most important prerequisites for the safe operating. The primary cause of fuel exhaustion is identified as fuel delay due to the slow dynamics in the fuel and air supply path. To observe the fuel exhaustion condition, the fuel concentration in the stack must be observed in real time. For avoiding the fuel exhaustion, the fuel concentration in the stack must meet the basic requirements $X_{H2} > 0$.

2.2.3 Optimal Operation Points

For the further analysis of SOFC system performance, considering the high efficiency of SOFC, the operating parameters, including the system inlet air and fuel flow rate and system current, are selected as the assemble-regulating variables in this article. The too low or high value of all these aforementioned operating parameters would lead to low efficiency or poor system performance, and they must be operated within their region under the different external load powers.

Based on the performance indices discussed previously, in addition to the thermal and electrical management and constraint enforcement, the optimal operating points (OOPs) are manipulated to achieve the optimal energy efficiency through the transverse optimization process, which have been conducted in our and other's previous works (Zhang et al., 2015a; Zhang et al., 2015b; Zhang et al., 2018; Zhang L et al., 2019). The OOPs can ensure thermal safety, and high efficiency in static can be obtained, which is shown in **Table 1**.

3 OPTIMIZATION AND CONTROL OF SOLID OXIDE FUEL CELL-BASED DIRECT-CURRENT MICROGRID

In the study of optimization and control of the hybrid SOFC system for the voltage stability, it is necessary to develop the full and quite accurate model of each subsystem component. This section first deals with the modeling of the system.

3.1 Mathematical Model

A lot of researches considering the SOFC stand-alone system model with the BOP were established for the simulation and optimization in our earlier works (Zhang et al., 2015a; Zhang et al., 2015b; Zhang et al., 2018; Zhang L et al., 2019), especially the exit temperatures, species molar fractions, and molar flow rates of each control volume, which can be divided into the gas phase and solid phase, as shown in **Figure 3**, have been obtained from the transient energy, species, and mass conservation equations, respectively. These conservation equations are of the same general form within various system components in **Table 2**. $\dot{Q}$ is the energy, T is the temperature, and V, ρ and C are the volume, density, and specific heat capacity of each control volume, respectively. W is the generating power, R_i is the reaction rate of the individual species i, U is the voltage, i is the current density, N is the molar flow rate, h is the gas enthalpy, S is the area, F is the gas flow rates or Faraday's constant, X is the molar fraction, R is the universal gas constant, C_v is the constant volume-specific heat capacity, C_p is the constant pressure-specific heat capacity, h_{gs} is the convection heat transfer coefficient, λ_g is the gas thermal conductivity, d is the radius, u is the velocity, μ is gas viscosity, and L is the distance between the control volume. The subscript S represents the solid control volume, and in and out represent the inlet and outlet, respectively. Cell denotes the fuel cell, and ref represents reference. PEN is the positive electrode-electrolyte-negative electrode, and cond and conv represent the heat conduction and thermal conversion, respectively.

3.1.1 Burner

Burner is the main component for the off-gas recovery and recycle. To reduce the computational burden, the SOFC burner is assumed to operate adiabatically, and the fuel is oxidized completely. It is assumed as the 0D lumped parameter model, and the heat transfer with the external

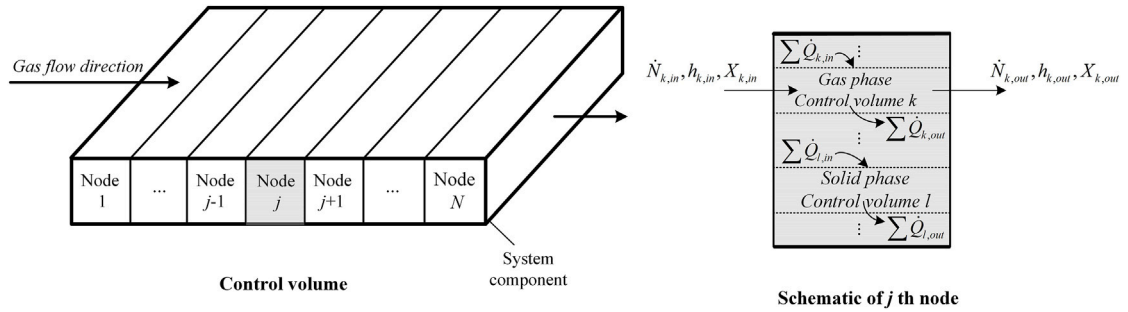


FIGURE 3 | Nodes and control volumes of the 1D model.

TABLE 2 | Universal SOFC system building equations.

Control volume	Variable	Equation
Solid phase	BOP temperature	$\sum \dot{Q}_{in} = \rho_s V_s C_s \frac{dT}{dt}$(1)
	PEN temperature	$\rho_{PEN} V_{PEN} C_{PEN} \frac{dT_{PEN}}{dt} = \sum \dot{Q}_{in,PEN} + \dot{Q}_{react} - \dot{W}_{out}$(2)
		$\dot{Q}_{react} = R_{H_2O} \cdot h_{H_2O}$(3)
		$\dot{W}_{out} = i \cdot S_{node} \cdot U_{cell}$(4)
Gas phase	Molar flow rate	$F_{out} = F_{in} + \sum R_i$(5)
		Stack: $R_{O_2} = \frac{1}{2} R_{H_2} = -\frac{1}{2} R_{H_2O} = -\frac{iS_{node}}{2F}$(6)
		Burner: $R_{O_2} = \frac{1}{2} R_{H_2} = -\frac{1}{2} R_{H_2O} = -F_{in} X_{H_2,in}$(7)
	Molar fraction	$N \frac{d(X_i)}{dt} = F_{in} X_{i,in} - F_{out} X_{i,out} + R_i$(8)
		$PV = NRT$(9)
	Temperature	$NC_v \frac{dT}{dt} = F_{in} h_{in} - F_{out} h_{out} + \sum \dot{Q}_{in}$(10)
		$C_v = \sum X_i C_{p,i}(T) - R,\ i \in \{H_2, O_2, H_2O, N_2\}$(11)
		$h = \sum X_i \left(\int_{T=Tref}^{T} C_{P,i}(T) dT + h_{ref,i} \right), i \in \{H_2, O_2, H_2O, N_2\}$(12)
Gas and solid adjacent components	Heat conduction	$\dot{Q}_{cond} = \frac{S_{area} \cdot k_{ss} \cdot (T_2 - T_1)}{L}$(13)
	Heat transfer	$\dot{Q}_{conv} = S_{area} \cdot h_{gs} \cdot (T_2 - T_1)$(14)
		$h_{gs} = 0.023 \frac{\lambda_g}{d} \left(\frac{du\rho}{\mu} \right)^{0.8} \left(\frac{C_p \mu}{\lambda_g} \right)^{0.4}$(15)

environment is ignored. Simultaneously, the catalytic combustion reaction time in the burner is in millisecond; thus, its reaction process can be neglected. The quasi-static model is selected to build the burner model. Its burning wall temperature can be determined according to **(1)**, and its outlet molar flow rate, molar fractions, and outlet temperature can be computed by **(5)** and **(7)** to **(12)**, respectively.

3.1.2 Fuel and Air Heat Exchangers

By introducing two heat exchangers in the SOFC system, the fuel and air can be preheated with the same hot stream from the burner at the same time. Thus, it can minimize the stack inlet temperature difference effectively. The structure of both the heat exchangers is assumed to be counter-current pipe heat exchangers. The integrated fuel and air heat exchangers are discretized into N nodes in the flow direction. As shown in **Figure 3**, each node includes the gas phase control volumes (air, fuel, and exhaust) and solid phase control volumes (the fuel tube, air tube, and exhaust tube). The temperature and species molar fractions of the gas phase control volumes are calculated by the conservation of the energy equation and the species conservation equation, respectively. The temperature of the solid phase control volumes is calculated from the dynamic solid-state energy conservation equation, as shown in **(1)**, whereas the temperature of the gas phase control volumes can be calculated from **(10)**.

3.1.3 Mass Flow Controllers

As shown in **Figure 2**, the air/fuel fed into the SOFC stack is manipulated by three pipes and valves, where the dynamics can be approximated by the first-order plus delay time, which is given by:

$$G(s) = \frac{1}{Ts + 1} e^{-T_d s}, \tag{16}$$

where T is the inertia time constant; and T_d is delay time or dead time, in which the delay time have a direct effect on the dynamic response, especially during the load tracing.

3.1.4 Solid Oxide Fuel Cell Stack

In this article, each single cell unit in the SOFC stack is assumed to operate identically; thus, a single cell unit can be taken as a

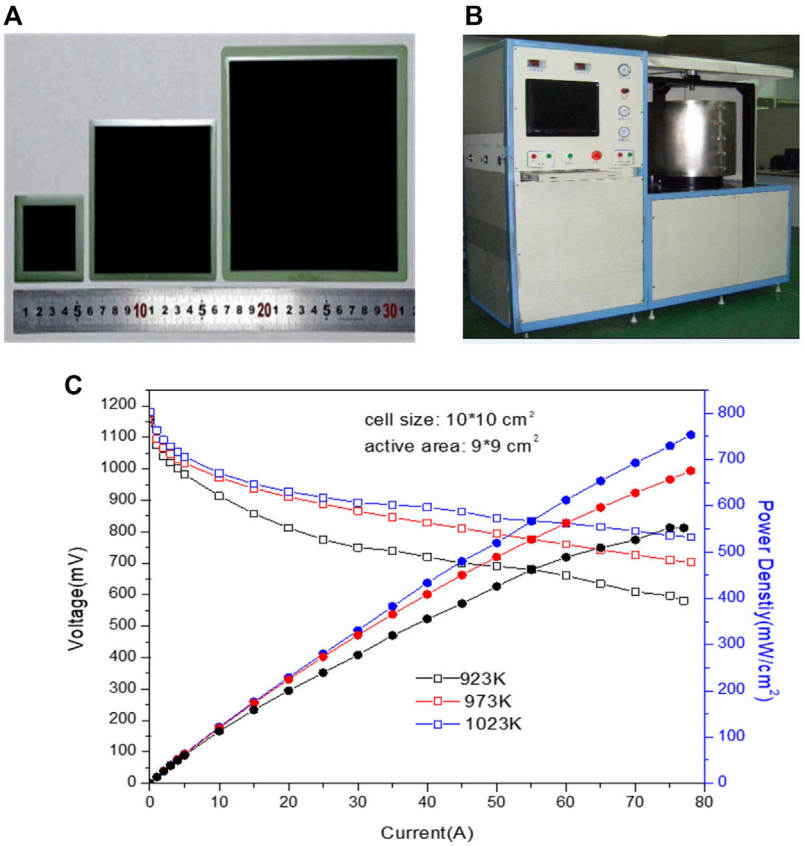


FIGURE 4 | Test of the fuel cells. **(A)** Fuel cell; **(B)** test furnace; and **(C)** electrical characteristics.

TABLE 3 | Parameter values of the equivalent resistance.

n	a_0	a_1	a_2	a_3	a_4
4	0.0028	0.2796	−0.9274	1.5921	−0.8775

representative of the entire stack performance. In addition, as shown in **Figure 3**, the fuel cell can be quasi-dimensionally discretized into five nodes in this article.

Four temperature layers (such as fuel channel, air channel, interconnector, and PEN) are considered for the fuel cell in the SOFC stack model. The temperatures in the PEN and interconnect plates are determined using the energy conservation equation, as shown in **(3)** to **(5)**. In addition, the temperatures and species mole fractions in the cathode and anode gas streams are calculated from the conservation of the energy equation and the species conservation equation, as shown in **(5)** to **(12)**, respectively. Conduction of heat transfer between the solid phase control volumes is calculated based on the Fourier's Law, as shown in **(13)**. Finally, the convection heat transfer between each gas and solid phase control volume is determined according to **(14)** and **(15)**.

As the polarization losses, ohmic loss and concentration loss are the function of the PEN temperature, gas pressure, and current density in the fuel cell. The fuel cell output voltage can be represented using a nonlinear algebraic equation, which is given by:

$$U_{cell} = f\left(i, p_{H_2}, p_{O_2}, p_{H_2O}, p_a, T_{PEN}\right), \tag{17}$$

where T_{PEN} represents the PEN temperature. For the accuracy of the model building, the electrical dynamic model of the fuel cell is obtained based on a lot of experimental data for the practical guidance. The equivalent resistance is shown using the partial derivative method, which is given by:

$$R_{tot} = \frac{\partial\left(U_{ohm} + U_{act} + U_{con}\right)}{\partial I}, \tag{18}$$

where I is the current; and U_{ohm}, U_{act}, and U_{con} represent the ohmic loss, activation loss, and concentration loss in the SOFC system, respectively, provided that:

$$R_{tot} = a_0 + a_1 x + \cdots + a_n x^n; \; x = 0.001 T_{PEN}/I_s, \tag{19}$$

where $a_0, \ldots\ldots a_{\mathrm{n}}$ represent the polynomial coefficient. Then, we can have:

$$U_{cell}=U_{OCV}-IR_{tot},$$
$$=U_{OCV}-I\left(a_0+a_1\cdot\left(\frac{0.001T_{PEN}}{I_s}\right)+\cdots+a_n\cdot\left(\frac{0.001T_{PEN}}{I_s}\right)^n\right). \tag{20}$$

The open circuit voltage can be adopted by the Nernst voltage, which is given by:

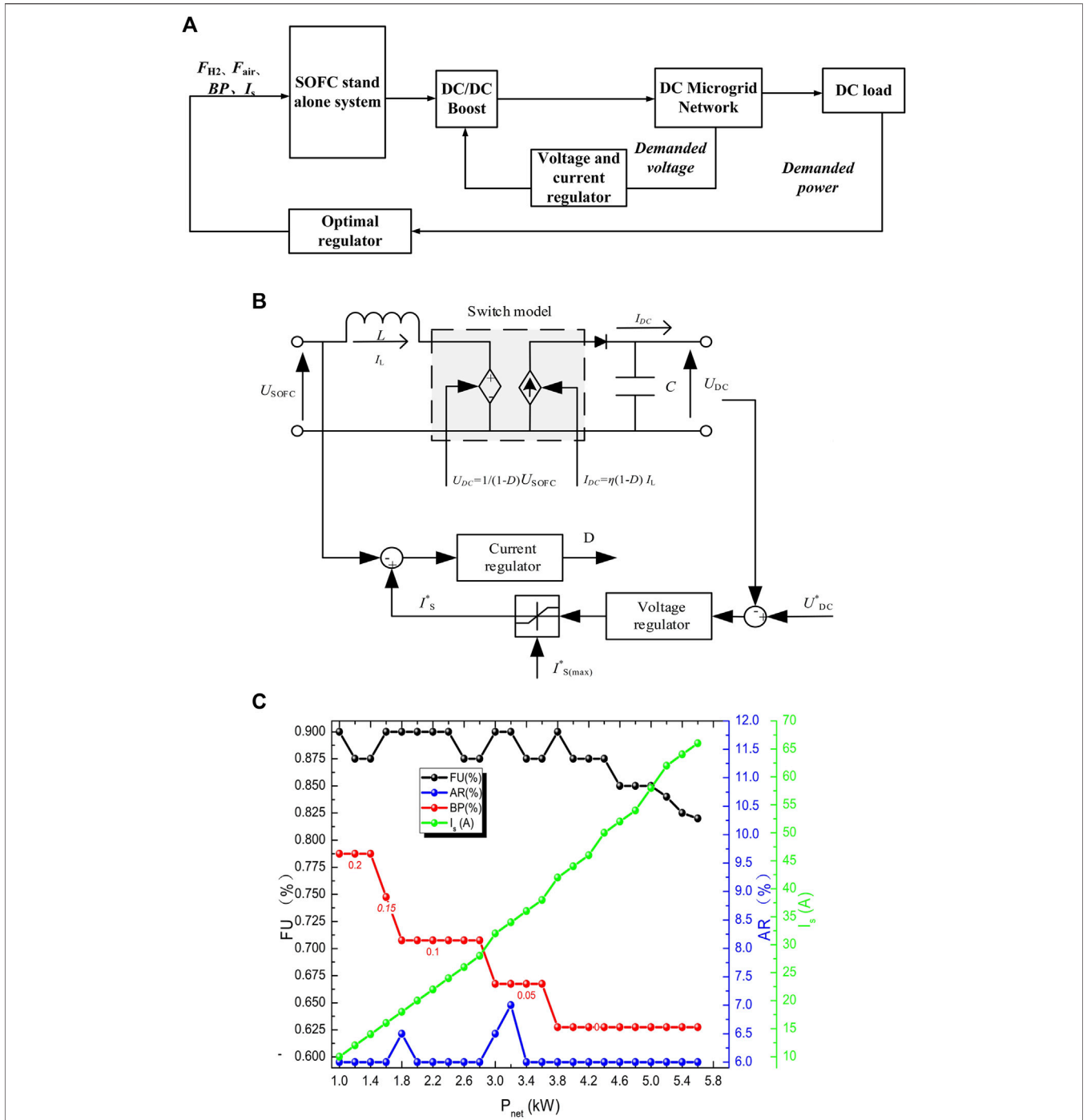


FIGURE 5 | Control and optimization of the SOFC DC microgrid. **(A)** Overall control schemes; **(B)** boost-type DC/DC converter model; and **(C)** steady-state optimal points for the optimal regulator.

$$U_{OCV} = E_0 + \frac{RT_{PEN}}{2F}\ln\left(\frac{P_{H_2}P_{O_2}^{0.5}}{P_{H_2O}}\right), \tag{21}$$

$$E_0 = -12.45058 + 26.3104\,(T_{PEN}/1000) - 12.4\,(T_{PEN}/1000)^2 - (2.7645e-4)/T_{PEN}. \tag{22}$$

To get the polynomial coefficient in **(19)**, the electrical characteristics of the single cell ($10^*10\ cm^2$) under the ideal conditions are conducted in an electrical furnace. As shown in **Figures 4A,B**, the voltage, current, and power density are investigated. The operating temperatures were set in the range from 923 to 973 K with the space of 50 K, especially the output power density and voltage are investigated under each temperature, and the open circuit voltage is calculated when the current is 0 A.

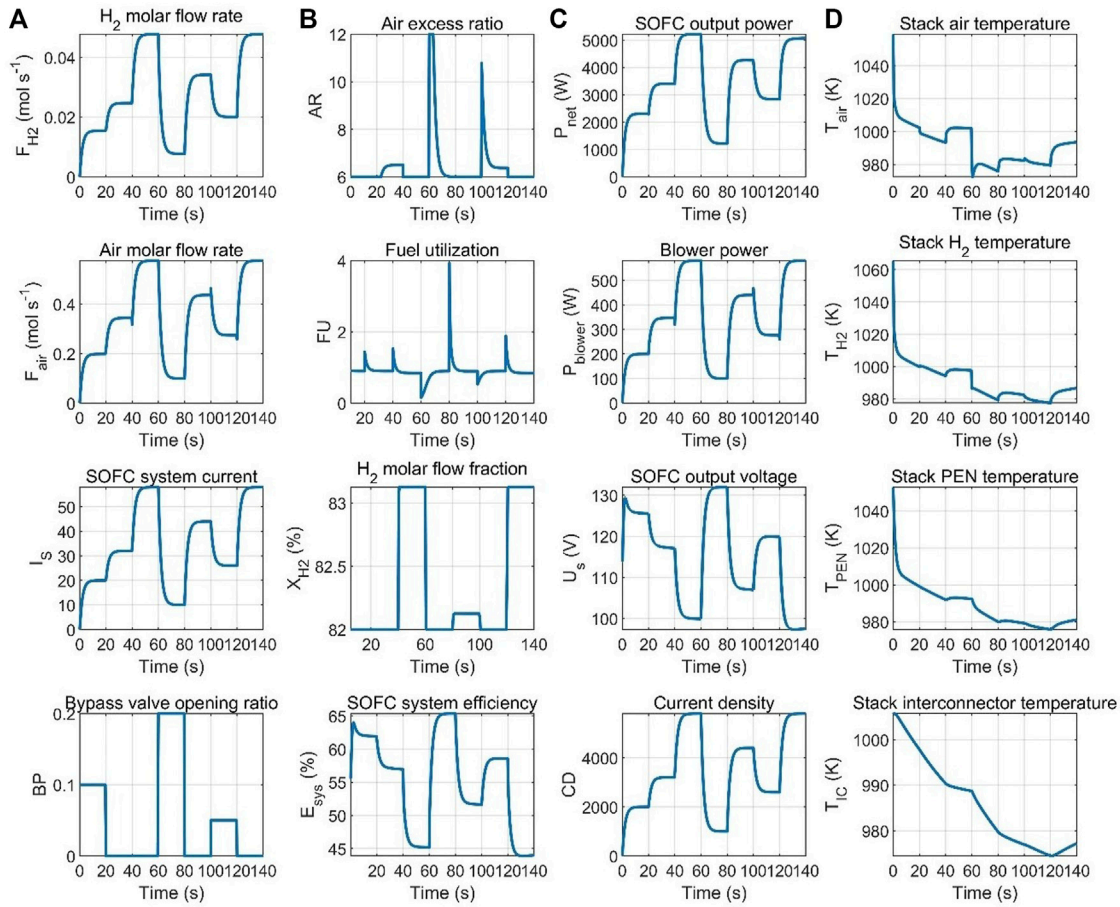


FIGURE 6 | Characteristics of the SOFC-based DC microgrid system. **(A)** Input characteristic and stack remain H_2 molar flow fraction; **(B)** related outputs considering efficiency and fuel exhaustion; **(C)** output electrical characteristic; and **(D)** Output thermal characteristic.

Then, the electrical characteristics curves of the voltage–current–power density are shown in **Figure 4C**; it showed that the open circuit voltage is about 1.12–1.18 V, and the power density is about 550–770 mW/cm^2. Based on the value obtained in our test, referring to **(20)**, the equivalent resistance can be calculated by the least square fitting; assuming the polynomial of order $n = 4$, the values of a_1-a_4 are shown in **Table 3**. Referring to our previous work, the Newton iteration algorithm is adopted to build the electrical characteristic model to ensure that the voltage of each fuel cell node is equal, which will not be addressed in this article. Moreover, the SOFC system structure parameters are confirmed by the physical facility in our group, and the physical parameters are determined by referring to the authoritative chemical handbook. More detail can be found in Zhang et al. (2015a) and Zhang et al. (2015b).

3.1.5 Blower

The blower supplies air to the SOFC system, which is the main parasitic losses in the SOFC stand-alone system; the output power of the different air flow can be expressed as:

$$P_{bl} = -\frac{1}{\tau_{bl}} * \frac{\gamma R T_{amb}}{\gamma - 1}\left[\left(\frac{p_{out}}{p_{amb}}\right)^{(\gamma-1)/\gamma} - 1\right] F_{air}, \tag{23}$$

where τ_{bl} is the effectiveness, γ is the specific heat ratio, and p is the pressure.

3.1.6 Solid Oxide Fuel Cell System Efficiency

The efficiency of the SOFC stand-alone system can be defined as:

$$\eta_{sys} = \frac{U_s \cdot I_s - P_{bl}}{F_{H2} \cdot LHV_{H2}} \times 100\%, \tag{24}$$

where U_s is the stack voltage and LHV_{H2} is the low heating value of H_2.

3.2 Optimization and Control

As shown in **Figure 5A**, the SOFC stand-alone system is connected to the DC/DC converter of the boost type. The design of the two control loops is made considering the voltage stability and SOFC dynamics in the DC microgrids. The voltage and current regulator are shown in **Figure 5B**, which can allow voltage conversion as well as the full control of the fuel cell current and DC bus voltage. The average value of the DC/DC converter models can be referred to Zakzouk et al. (2019) for this study. The optimal regulator mainly controls the load demanded power by referring to OOPs. **Figure 5C** shows the high fuel utilization (FU) and low air

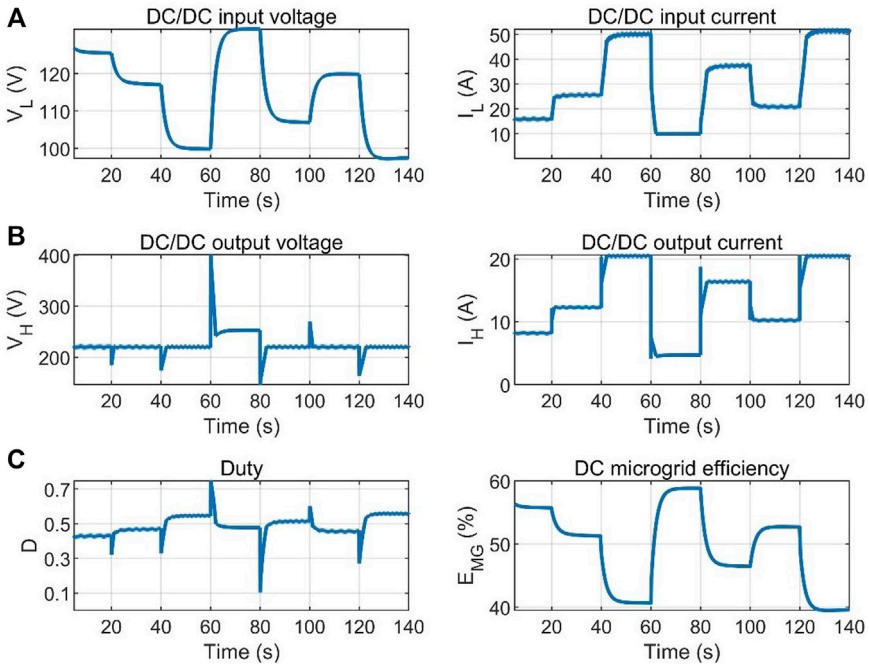


FIGURE 7 | Performances of the SOFC-based hybrid DC microgrid system. **(A)** DC/DC inputs; **(B)** DC/DC outputs; and **(C)** Duty and DC microgrid efficiency.

excess ratio (AR) under the OOPs in the steady state, which can ensure high efficiency. Moreover, the PI regulator uses the control method, and the demanded voltage of the DC microgrid is 220 V in this study. The demanded load power change is given as 2.5 kW→3.5 kW→5 kW→1.5 kW→4.5 kW→5 kW. The efficiency of the DC/DC boost converter in this study is assumed to be 90%.

The proposed control and optimization scheme for the SOFC-based DC microgrids is implemented in the Simulink/SimPower systems, and the characteristics of the SOFC-based DC microgrid system are shown in **Figure 6**. **Figure 6A** shows the system input characteristics and stack remain H_2 molar flow fraction in the proposed SOFC system. It is worth pointing that our previous study (Zhang L et al., 2019) has revealed that safety operating, whether ensuring the thermal safety or avoiding fuel exhaustion, should be in the cost of the fast load tracing, namely, the SOFC system current that should slowly change. The SOFC system current slope limitation is adopted to deal with the problem of fuel exhaustion and thermal safety. By observing the H_2 molar flow fraction in the SOFC stack, the large amount of hydrogen is shown in stack (≥82%). The output electrical characteristics of the proposed SOFC system are shown in **Figure 6C**, and the response time of the output electrical variables, including the SOFC output net power and voltage, is within tens of seconds. In addition, by the proposed optimal regulator with the OOPs, considering the parasitic loss of the blower power, the SOFC system efficiency shows a high value (45%–65%). Due to the current slow switching and BP regulation, the output thermal characteristics of the proposed SOFC system shown in **Figure 6D** are all within their temperature constraints.

The performances of the SOFC-based DC microgrid system is shown in **Figure 7**. The DC/DC inputs (**Figure 7A**) responses based on the SOFC system outputs and load changes, and the current is well limited to its maximum reference current (60 A). The DC/DC voltage outputs (**Figure 7B**) are well regulated at 220 V, and the power outputs can well meet the load requirements. **Figure 7C** shows the duty and efficiency to deal with voltage stability and the energy convert and transfer efficiency (40%–60%) in the proposed DC microgrid.

4 CONCLUSION

This article presented the optimization and control strategy for the SOFC-based hybrid DC microgrids from the perspective of high efficiency, thermal safety fuel exhaustion, and transient response. The structure of the SOFC-based DC microgrid is first introduced. Moreover, the hybrid system has been modeled, especially the SOFC stand-alone system core part stack is validated with the experiments for the more precise corresponding electrical characteristics in the DC microgrid. Considering the high efficiency, steady-state thermal safety, and load tracing, the optimization and control strategy based on the voltage and current regulator and OOPs for the SOFC-based DC microgrid is implemented. The performance shows that the proposed DC microgrid has the large amount of hydrogen in stack (≥82%), thermal safety, voltage stability (220 V), and high efficiency (40%–60%). An alternative is to design the optimization, control, and energy management strategy to optimize all the performance of the SOFC-based DC microgrid, such as the SOFC/battery/supercapacitor-based DC microgrid, which is the next topic for our further studies.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

LZ: experiments, research methods, data processing, writing—original draft, and funding acquisition. HX: writing—original draft, resources, supervision, project administration, and funding acquisition. WT: experiment guidance. FW: experiment guidance. CX: performed writing—review. WZ: performed writing—review. All authors have contributed to the manuscript and approved the submitted version.

FUNDING

This work was co-supported by the Guangdong Basic and Applied Basic Research Foundation (No. 2021A1515010768), by the Science and Technology Talents Foundation Project of Air Force Early Warning Academy (No. 2021KJY11), by the Fundamental Research Funds for the Central Universities, Sun Yat-sen University (No. 2022ZZ028), by the National Natural Science Foundation of China (No. 61801510), and by the National Science Foundation of Hubei Province (No. 2019CFB263).

ACKNOWLEDGMENTS

All authors would like to thank the handling editors and reviewers for their very competent comments and constructive suggestions to improve this article.

REFERENCES

Ahmed, O. A., and Bleijs, J. A. M. (2013). Power Flow Control Methods for an Ultracapacitor Bidirectional Converter in DC Microgrids-A Comparative Study. *Renew. Sustain. Energy Rev.* 26, 727–738. doi:10.1016/j.rser.2013.06.021

Armghan, H., Yang, M., Wang, M. Q., Ali, N., and Armghan, A. (2020). Nonlinear Integral Backstepping Based Control of A DC Microgrid with Renewable Generation and Energy Storage Systems. *Int. J. Electr. Power & Energy Syst.* 117, 105613. doi:10.1016/j.ijepes.2019.105613

Awais, M., Khan, L., Ahmad, S., and Jamil, M. (2021). Feedback-Linearization-Based Fuel-Cell Adaptive-Control Paradigm in a Microgrid Using a Wavelet-Entrenched NeuroFuzzy Framework. *Energies* 14 (71-17), 1850. doi:10.3390/en14071850

Barnes, M. (2002). *Fuel Cell System Explained*. 2nd ed. New York, NY, USA: John Wiley.

Carré, M., Brandenburger, R., Friede, W., Lapicque, F., Limbeck, U., and da Silva, P. (2015). Feed-Forward Control of A Solid Oxide Fuel Cell System with Anode Offgas Recycle. *J. Power Sources* 282 (15), 498–510. doi:10.1016/j.jpowsour.2015.02.053

Deng, Y., Zhang, Y., Luo, F., and Mu, Y. (2021). Operational Planning of Centralized Charging Stations Utilizing Second-Life Battery Energy Storage Systems. *IEEE Trans. Sustain. Energ.* 12 (1), 387–399. doi:10.1109/TSTE.2020.3001015

Eid, A. (2014). Utility Integration of PV-Wind-Fuel Cell Hybrid Distributed Generation Systems under Variable Load Demands. *Int. J. Electr. Power & Energy Syst.* 62 (11), 689–699. doi:10.1016/j.ijepes.2014.05.020

Gandini, D., and de Almeida, A. T. (2017). Direct Current Microgrids Based on Solar Power Systems and Storage Optimization, as A Tool for Cost-Effective Rural Electrification. *Renew. Energy* 111, 275–283. doi:10.1016/j.renene.2017.04.009

Hajimolana, S. A., Tonekabonimoghadam, S. M., Hussain, M. A., Chakrabarti, M. H., Jayakumar, N. S., and Hashim, M. A. (2013). Thermal Stress Management of A Solid Oxide Fuel Cell Using Neural Network Predictive Control. *Energy* 62 (1), 320–329. doi:10.1016/j.energy.2013.08.031

Hirsch, A., Parag, Y., and Guerrero, J. (2018). Microgrids: A Review of Technologies, Key Drivers, and Outstanding Issues. *Renew. Sustain. Energy Rev.* 90, 402–411. doi:10.1016/j.rser.2018.03.040

Huo, H.-B., Wu, Y.-X., Liu, Y.-Q., Gan, S.-H., and Kuang, X.-H. (2010). Control-oriented Nonlinear Modeling and Temperature Control for Solid Oxide Fuel Cell. *J. Fuel Cell Sci. Technol.* 7 (4), 0410051–0410059. doi:10.1115/1.3211101

Li, Y. H., Choi, S. S., and Rajakaruna, S. (2005). An Analysis of the Control and Operation of A Solid Oxide Fuel-Cell Power Plant in an Isolated System. *IEEE Trans. Energy Convers.* 20 (2), 381–387. doi:10.1109/TEC.2005.847998

Li, Y., Wu, Q., and Zhu, H. (2015). Hierarchical Load Tracking Control of a Grid-Connected Solid Oxide Fuel Cell for Maximum Electrical Efficiency Operation. *Energies* 8 (3), 1896–1916. doi:10.3390/en8031896

Lingchong, X., Ibragimova, A. V., and Shilova, K. I. (2022). Analysis of International Practice in the Use of Renewable Energy Sources (RES). *IOP Conf. Ser. Earth Environ. Sci.* 979, 012188. doi:10.1088/1755-1315/979/1/012188

Liu, Z., Zhao, J., and Zou, Z. (2020). Impedance Modeling, Dynamic Analysis and Damping Enhancement for DC Microgrid with Multiple Types of Loads. *Int. J. Electr. Power & Energy Syst.* 122 (1-12), 106183. doi:10.1016/j.ijepes.2020.106183

Mumtaz, S., Khan, L., Ahmed, S., and Badar, R. (2018). Correction: Indirect Adaptive Soft Computing Based Wavelet-Embedded Control Paradigms for WT/PV/SOFC in A Grid/Charging Station Connected Hybrid Power System. *PLoS ONE* 13 (4), e0195914–17. doi:10.1371/journal.pone.0195914

Padulles, J., Ault, G. W., and McDonald, J. R. (2000). An Integrated SOFC Plant Dynamic Model for Power Systems Simulation. *J. Power Sources* 86 (1), 495–500. doi:10.1016/S0378-7753(99)00430-9

Pranita, R., Sanjoy, K. M., and Sujit, K. B. (2022). Renewable Energy Generation System Connected to Micro Grid and Analysis of Energy Management: A Critical Review. *Int. J. Power Electron. Drive Syst.* 13 (1), 470–479. doi:10.11591/ijpeds.v13.i1

Singhal, S. C., and Kendall, K. (2003). *High Temperature Solid Oxide Fuel Cells: Fundamentals, Design, and Applications*. New York, NY, USA: Elsevier.

Sorrentino, M., Pianese, C., and Guezennec, Y. G. (2008). A Hierarchical Modeling Approach to the Simulation and Control of Planar Solid Oxide Fuel Cells. *J. Power Sources* 180 (1), 380–392. doi:10.1016/j.jpowsour.2008.02.021

Srinivasan, M., and Kwasinski, A. (2020). Control Analysis of Parallel DC-DC Converters in A DC Microgrid with Constant Power Loads. *Int. J. Electr. Power & Energy Syst.* 122 (4), 106207. (1-9). doi:10.1016/j.ijepes.2020.106207

Wang, Y., Li, S., Sun, H., Huang, C., and Youssefi, N. (2022). The Utilization of Adaptive African Vulture Optimizer for Optimal Parameter Identification of SOFC. *Energy Rep.* 8, 551–560. doi:10.1016/j.egyr.2021.11.257

Xie, H., Hu, J., Duan, K., and Wang, G. (2020). High-Efficiency and High-Precision Reconstruction Strategy for P-Band Ultra-Wideband Bistatic Synthetic Aperture Radar Raw Data Including Motion Errors. *IEEE Access* 8, 31143–31158. doi:10.1109/ACCESS.2020.2971660

Zakzouk, N. E., Khamis, A. K., Abdelsalam, A. K., and Williams, B. W. (2019). Continuous-Input Continuous-Output Current Buck-Boost DC/DC Converters for Renewable Energy Applications: Modelling and Performance Assessment. *Energies* 12 (11), 2208. doi:10.3390/en12112208

Zhang, L., Jiang, J., Cheng, H., Deng, Z., and Li, X. (2015a). Control Strategy for Power Management, Efficiency-Optimization and Operating-Safety of a 5-kW Solid Oxide Fuel Cell System. *Electrochimica Acta* 177, 237–249. doi:10.1016/j.electacta.2015.02.045

Zhang, L., Li, X., Jiang, J., Li, S., Yang, J., and Li, J. (2015b). Dynamic Modeling and Analysis of A 5-kW Solid Oxide Fuel Cell System from the Perspectives of Cooperative Control of Thermal Safety and High Efficiency. *Int. J. Hydrogen Energy* 40, 456–476. doi:10.1016/j.ijhydene.2014.10.149

Zhang, L., Shi, S., Jiang, J., and Li, X. (2019). Current-based MPC for Operating-Safety Analysis of a Reduced-Order Solid Oxide Fuel Cell System. *Ionics* 25 (4), 1759–1772. doi:10.1007/s11581-018-2654-8

Zhang, L., Shi, S., Jiang, J., Wang, F., Xie, H., Chen, H., et al. (2018). An Optimization and Fast Load-Oriented Control for Current-Based Solid Oxide Fuel Cell System. *J. Solid State Electrochem* 22 (18), 2863–2877. doi:10.1007/s10008-018-3996-x

Zhang, X., Chan, S. H., Li, G., Ho, H. K., Li, J., and Feng, Z. (2010). A Review of Integration Strategies for Solid Oxide Fuel Cells. *J. Power Sources* 195 (3), 685–702. doi:10.1016/j.jpowsour.2009.07.045

Zhang, Y., Meng, K., Luo, F., Yang, H., Zhu, J., and Dong, Z. Y. (2020a). Multi-agent-based Voltage Regulation Scheme for High Photovoltaic Penetrated Active Distribution Networks Using Battery Energy Storage Systems. *IEEE Access* 8, 7323–7333. doi:10.1109/ACCESS.2019.2962717

Zhang, Y., Xu, Y., Yang, H., and Dong, Z. Y. (2019). Voltage Regulation-Oriented Co-planning of Distributed Generation and Battery Storage in Active Distribution Networks. *Int. J. Electr. Power & Energy Syst.* 105, 79–88. doi:10.1016/j.ijepes.2018.07.036

Zhang, Y., Xu, Y., Yang, H., Dong, Z. Y., and Zhang, R. (2020b). Optimal Whole-Life-Cycle Planning of Battery Energy Storage for Multi-Functional Services in Power Systems. *IEEE Trans. Sustain. Energy* 11 (4), 2077–2086. doi:10.1109/TSTE.2019.2942066

Zhu, Y., and Tomsovic, K. (2002). Development of Models for Analyzing the Load-Following Performance of Microturbines and Fuel Cells. *Electr. Power Syst. Res.* 62 (1), 1–11. doi:10.1016/S0378-7796(02)00033-0

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Energy Research

METHODS
published: 17 August 2022
doi: 10.3389/fenrg.2022.952684

Multistage Bilevel Planning Model of Energy Storage System in Urban Power Grid Considering Network Reconfiguration

Zhongqi Cai[1], Kun Yang[1], Yong Chen[1], Ruixiong Yang[1], Yanxun Gu[1], Yu Zeng[2], Xi Zhang[2], Sashuang Sun[2], Sirong Pan[2], Youbo Liu[2] and Junyong Liu[2]*

[1]Zhuhai Power Supply Bureau of Guangdong Power Grid, Zhuhai, China, [2]College of Electrical Engineering, Sichuan University, Chengdu, China

The large-scale integration of renewable energy sources (RESs) and the rapid development of loads cause frequent transmission congestion in the urban power grid (UPG). Transmission system operators usually perform the high-voltage distribution network (HVDN) reconfiguration to mitigate the transmission congestion. However, as the loads and RESs change rapidly, the HVDN reconfiguration might be conducted frequently. This might cause severe security problem. An energy storage system (ESS) provides an effective way of alleviating the transmission congestion. If the ESS is installed and operated elaborately, the transmission congestion of UPG can be mitigated with a little HVDN reconfiguration. Hence, this study proposes a multistage bilevel planning model for the optimal allocation of ESS. The upper-level model aims at maximizing the annual comprehensive revenue of HVDN, and the lower-level model focuses on the minimization of the operational cost. Simulation results carried out on a real-world test system verify that the proposed method has the great potential of reducing the investment and operational cost while mitigating the transmission congestion.

Keywords: energy storage station, multistage planning, high-voltage distribution network, congestion management, network reconfiguration, load shedding

OPEN ACCESS

Edited by:
Yunfei Mu,
Tianjin University, China

Reviewed by:
Youwei Jia,
Southern University of Science and Technology, China
Wei Pei,
Institute of Electrical Engineering (CAS), China

****Correspondence:***
Yu Zeng
zy2416156976@163.com

Specialty section:
This article was submitted to Smart Grids, a section of the journal Frontiers in Energy Research

Received: *25 May 2022*
Accepted: *20 June 2022*
Published: *17 August 2022*

Citation:
Cai Z, Yang K, Chen Y, Yang R, Gu Y, Zeng Y, Zhang X, Sun S, Pan S, Liu Y and Liu J (2022) Multistage Bilevel Planning Model of Energy Storage System in Urban Power Grid Considering Network Reconfiguration. Front. Energy Res. 10:952684. doi: 10.3389/fenrg.2022.952684

1 INTRODUCTION

The acceleration of urbanization in many developing countries has caused the surge of the electricity load and renewable energy sources (RESs). Owing to the limited transmission capacity, the transmission congestion occurs more frequently than before. This limits the penetration of RESs (He et al., 2021). To guarantee the safe operation of the power system, the grid dispatchers have to curtail the generation of RESs and shed load as reported by Bird et al. (2016) and Goop et al. (2017).

Transmission switching is an efficient way to deal with the congestion management problems by elaborately selecting the transmission lines to be disconnected (Zhang et al., 2022). Extensive studies have been performed recently on this subject. In a study by Khanabadi et al. (2018), a decentralized transmission switching model is proposed to alleviate transmission congestion under credible contingencies. In a study by Salkuti (2018), a multiobjective-based congestion management methodology is established, considering the influences of network reconfiguration. The network topology is reconfigured to improve the hosting capacity of renewable generation and variable loads in a study by Haghighat and Zeng (2016). Many researchers have investigated the method that

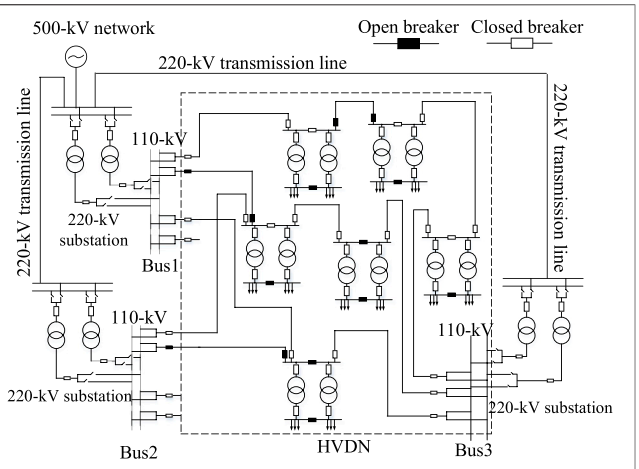


FIGURE 1 | Illustration of the urban power grid structure.

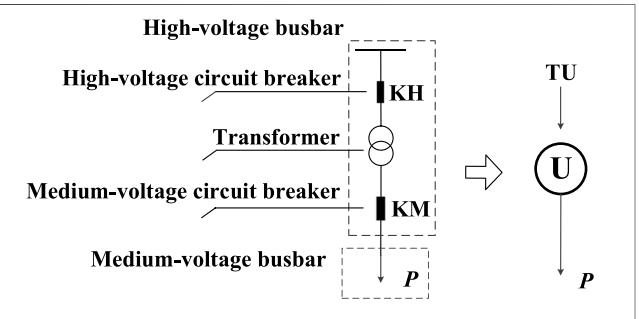


FIGURE 2 | Structure of transformer unit.

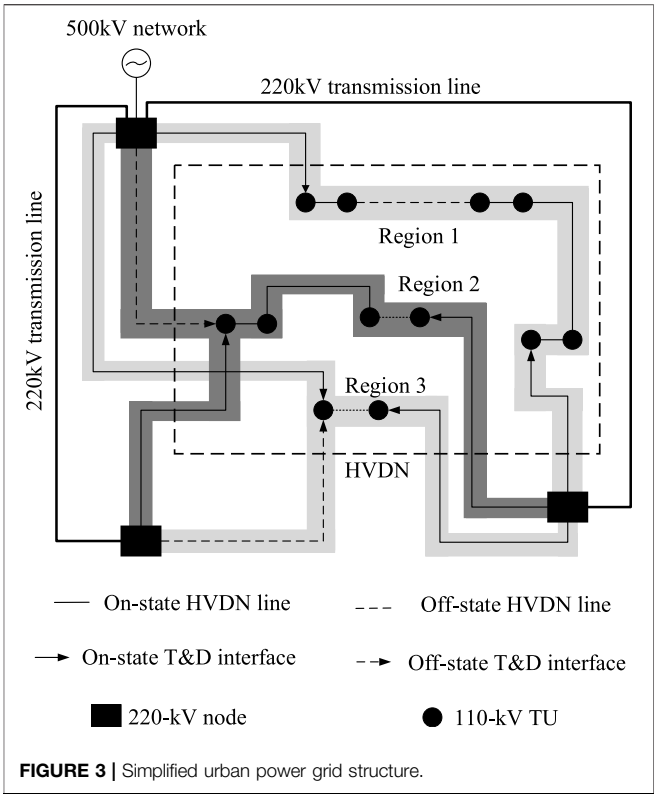


FIGURE 3 | Simplified urban power grid structure.

relieves the transmission congestion by performing high-voltage distribution network (HVDN) reconfiguration. The research by Zhang et al. (2020) and Hoffrichter et al. (2018) shows the great potential of implementing HVDN reconfiguration to mitigate the transmission congestion. However, implementing congestion management through frequent transmission switching could increase the risk of safe operation and reduce the power supply reliability.

The increasing penetration of RESs stimulates the installation of grid-side energy storage systems (ESSs), providing an effective solution to relieving the transmission congestion. On the grid side, the ESS plays the role of delaying the investment in transmission and distribution infrastructure (Hu et al., 2012; Macrae et al., 2014; Macrae et al., 2016), shaving peak load (Subramani et al., 2018), regulating the frequency (Dhundhara and Verma, 2018), and ensuring the safe operation (Nick et al., 2014). Several studies investigate the role of the ESS in increasing the transmission capacity of congested transmission networks (Del Rosso and Eckroad, 2014). In a study by Yan et al. (2020), a robust optimization model is designed to operate the ESS, considering uncertainties. The results show that the ESS could increase the system flexibility and mitigate the transmission congestion. In the research by Yang et al. (2021), a joint planning method of the ESS and transmission network is proposed to relieve the transmission congestion and reduce the curtailment of RES. Chen and Liu (2021) proposed a network reconfiguration integrated dynamic tariff–subsidy congestion management method, alleviating microgrid congestion caused by RESs and flexible demands through the ESS and network reconfiguration. In the research by Nick et al. (2017), the siting and sizing schemes of ESSs are optimized, considering the impact of network reconfiguration. In a study by Fiorini et al. (2017), the sizing and siting of large-scale ESSs are optimized in transmission grids to enhance the use of renewables. However, the investment cost of the ESS is relatively high owing to technical reasons and unreasonable planning methods. On the one hand, current ESS planning methods do not consider the development of the transmission network and the growth of the load, causing the excessive investment at the early stage (Cao et al., 2020). On the other hand, the planning model does not consider the role of HVDN reconfiguration in alleviating transmission congestion.

In order to fill up the gaps discussed above, this study proposes a multistage bilevel planning method for the ESS that considers the HVDN reconfiguration. The main contributions are shown as follows:

1) A collaborative scheduling strategy that co-optimizes the ESS operational strategy and the HVDN topological structure is developed, avoiding the frequent HVDN reconfiguration and load shedding while maximizing the operating benefits of the urban power grid (UPG).

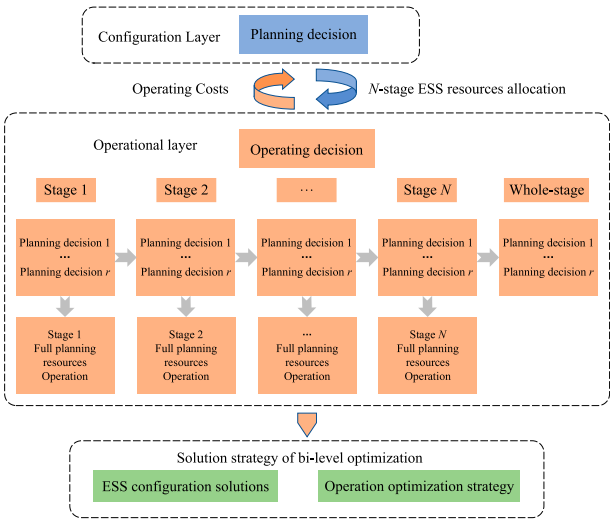


FIGURE 4 | Flow diagram of energy storage system's multistage planning.

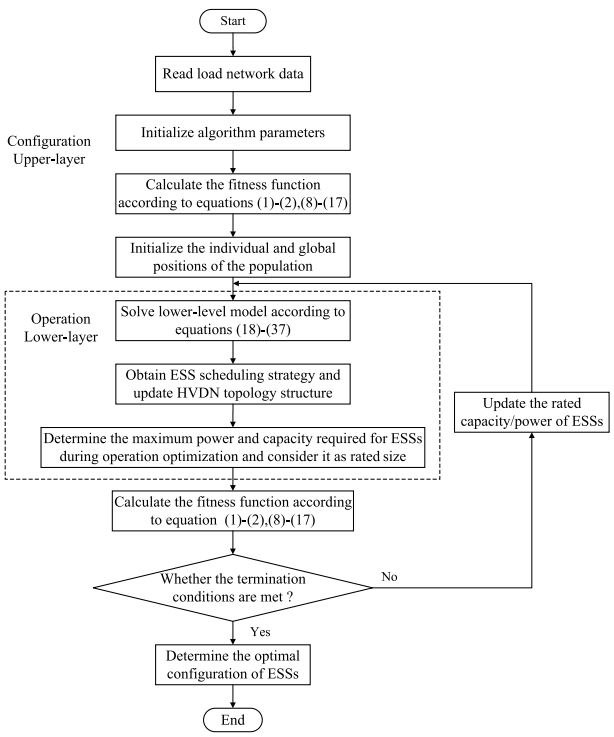


FIGURE 5 | Flowchart of bilevel optimization solution.

2) The proposed ESS planning procedure that consists of HVDN reconfiguration can effectively reduce the cost of ESS installation and improve the efficiency of the ESS.
3) The idea of multistage planning is incorporated into the optimal ESS allocation. On the one hand, it can maximize the utilization of the ESS. On the other hand, it can effectively avoid the waste of energy storage resources, reduce the investment and operation costs, and improve the revenue of the ESS.

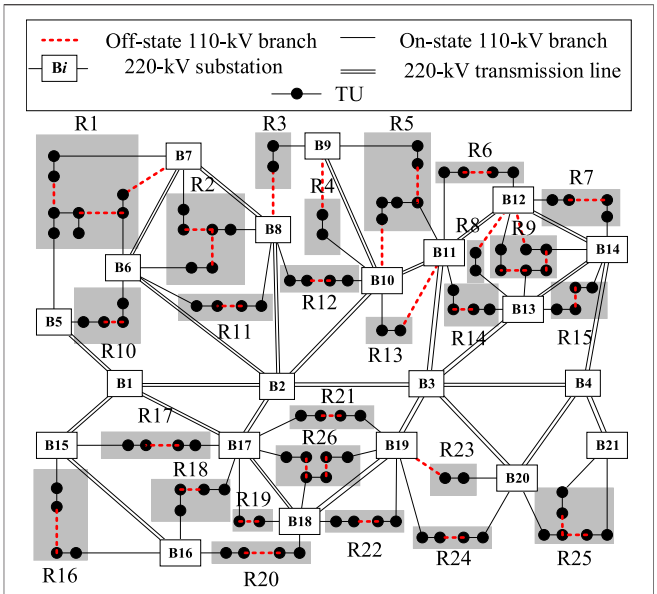


FIGURE 6 | Network configuration of the China 407-node urban power grid system.

2 TYPICAL STRUCTURE OF URBAN POWER GRID

In general, the UPG mainly consists of a 220-kV/500-kV transmission network and a 110-kV HVDN (Yuan and Hesamzadeh, 2017; Zhang et al., 2020) (as shown in **Figure 1**). Owing to the high density of urban loads and the limited transmission capacity, some transmission lines might be overload during system operations (Zhang et al., 2020).

To simplify the HVDN topology structure and improve the computation efficiency, the concept of a transformer unit (TU) is defined as shown in **Figure 2**, in which the letter *P* represents the active power transferred by TU. In order to simplify the problem, only the balance of the active load is considered.

Hence, the UPG structure in **Figure 1** can be simplified as shown in **Figure 3**.

3 MULTISTAGE PLANNING APPROACH

The flow chart of the multistage planning approach for the ESS in the UPG is shown in **Figure 4**.

Assume that the planning horizon of the ESS is *n* years. The planning stages are determined by the growing rate of the load. During the period of rapid load growth, the number of the stages should be more, and the length of the stages should be shorter. In this study, the planning horizon is divided into *N* stages as shown in **Eq. 1**:

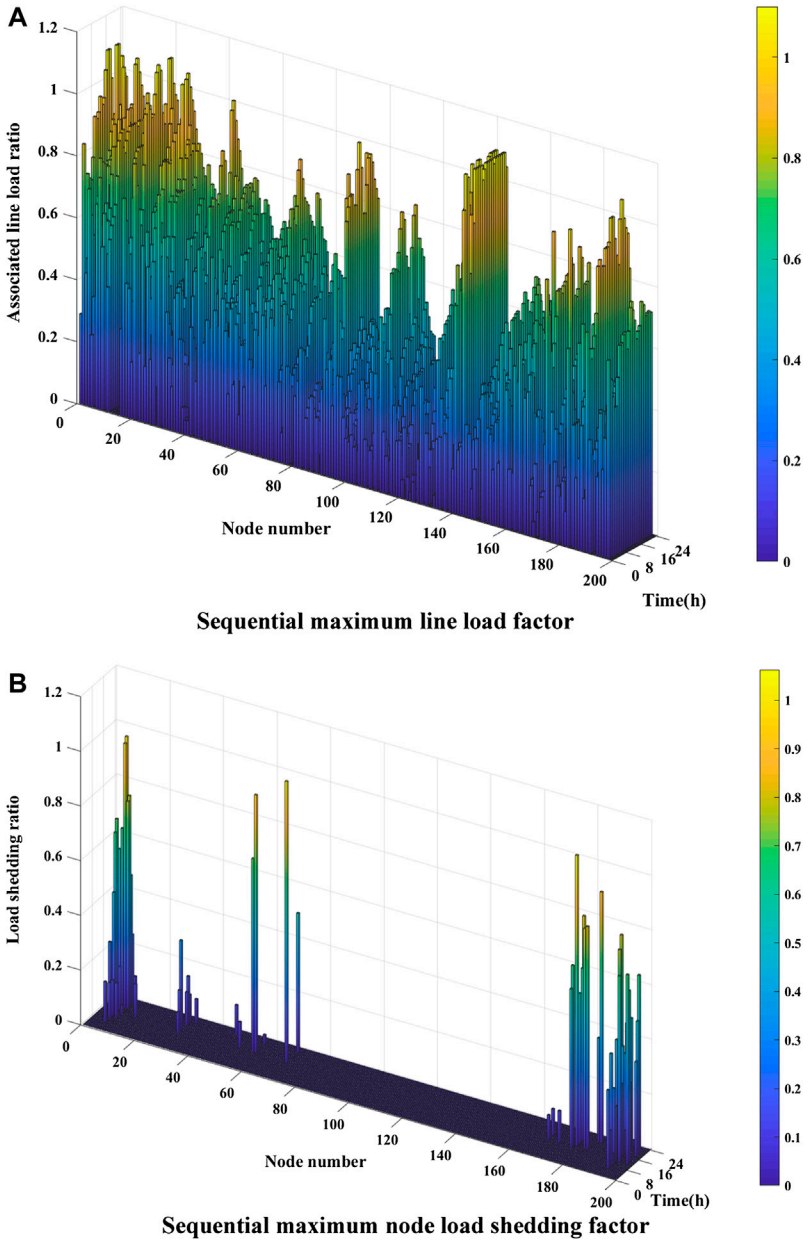


FIGURE 7 | Sequential technical indices for energy storage system siting. **(A)** sequential maximum line load factor. **(B)** sequential maximum node load shedding factor.

$$S = [S_1,\ S_2,\ \cdots,\ S_N] \tag{1}$$

where S_N represents the Nth planning stage.

The planning scheme of the ESS at different stages is denoted as

$$E_{\text{set}} = [E_{set1},\ E_{set2},\ \cdots,\ E_{setN}] \tag{2}$$

where $E_{\text{set}N}$ denotes the ESS planning scheme at the Nth stage.

Note that the planning scheme at the Nth stage is determined on the basis of the prior planning scheme $E_{\text{set},N\text{-}1}$.

4 PROBLEM FORMULATION

In order to reduce the difficulty of solving the ESS planning problem and improve the efficiency, the siting of the ESS is determined using a multiattribute comprehensive index evaluation model from the study by Guo et al., 2020, Song et al. (2019). The sizing problem is described using a bilevel mathematical model.

4.1 Siting Model

The siting of the ESS is determined using a multiattribute comprehensive index that comprises the line load rate and the

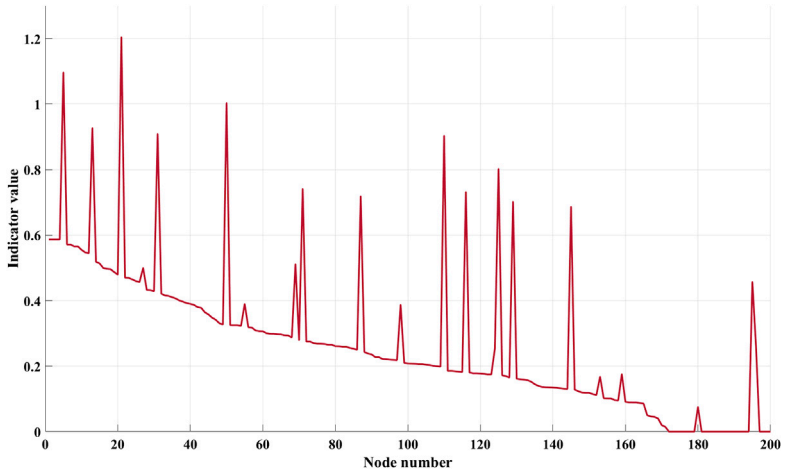


FIGURE 8 | Comprehensive indicators for energy storage system siting.

TABLE 1 | Energy storage system planning results of each stage of case I.

Num	Name	Stage 1		Stage 2		Stage 3		Total	
		P	E	P	E	P	E	P	E
1	ESS-5	14	16	14	14	14	19	41	49
2	ESS-14	5	21	17	15	16	12	38	49
3	ESS-24	9	19	19	15	19	20	45	55
4	ESS-38	14	19	18	14	16	20	48	52
5	ESS-74	15	18	19	17	17	12	51	46
6	ESS-114	16	22	14	13	13	14	44	50
7	ESS-125	13	16	18	20	16	17	47	53
8	ESS-173	10	20	14	17	19	13	43	50
	Total	96	151	133	125	130	127	357	404

Investment cost	Operation cost	Peak-shaving revenue	Lagged facility upgrades revenue	Net revenue
73,680	21,420	66,810	42,840	14,550

TABLE 2 | Energy storage system planning results of each stage of case II.

Num	Name	Stage 1		Stage 2		Stage 3		Total	
		P	E	P	E	P	E	P	E
1	ESS-5	46	64	—	—	—	—	46	64
2	ESS-14	43	51	—	—	—	—	43	51
3	ESS-24	39	56	—	—	—	—	39	56
4	ESS-38	50	68	—	—	—	—	50	68
5	ESS-74	51	57	—	—	—	—	51	57
6	ESS-114	35	58	—	—	—	—	35	58
7	ESS-125	60	49	—	—	—	—	60	49
8	ESS-173	31	47	—	—	—	—	31	47
	Total	355	449	0	0	0	0	355	449

Investment cost	Operation cost	Peak-shaving revenue	Lagged facility upgrades revenue	Net revenue
76,330	23,300	64,597.5	41,600	6567.5

TABLE 3 | Energy storage system planning results of each stage of case III.

Num	Name	Stage 1		Stage 2		Stage 3		Total	
		P	E	P	E	P	E	P	E
1	ESS-5	55	68	—	—	—	—	55	68
2	ESS-14	49	70	—	—	—	—	49	70
3	ESS-24	54	61	—	—	—	—	54	61
4	ESS-38	41	56	—	—	—	—	41	56
5	ESS-74	57	60	—	—	—	—	57	60
6	ESS-114	49	44	—	—	—	—	49	44
7	ESS-125	36	45	—	—	—	—	36	45
8	ESS-173	44	57	—	—	—	—	44	57
	Total	384	460	0	0	0	0	384	460

Investment cost	Operation cost	Peak-adjusted revenue	Lagged facility upgrades revenue	Net revenue
78,200	24,040	59,358	46,080	3,198

nodal load curtailment rate. The detailed descriptions of the index are shown as follows:

The line load rate is the ratio of the maximum line load rate to the line capacity. The line load factor matrix F_1 for each node is formed by taking the line load rate of each branch associated with the node as an indicator:

$$F_1 = \left[f_1(1), f_1(2), \cdots, f_1(m), \cdots, f_1(N)\right]^{\mathrm{T}} \tag{3}$$

where $f_1(k)$ is the line load rate of the branch associated with the node m. N is the number of the 110-kV nodes. $F_1(k)$ is modeled as follows:

$$f_1(k) = \max \frac{\max p_{k,mk,t}}{p_{mk}^{\max}}, t = 1, 2, \cdots, T \tag{4}$$

2) The nodal load curtailment rate is the ratio of the maximum load curtailment to the load at a node. This index reflects the overall reliability of the system and is used as an indicator to form the nodal load curtailment ratio matrix F_2:

$$F_2 = \left[f_2(1), f_2(2), \cdots, f_2(k), \cdots, f_2(N)\right]^{\mathrm{T}} \tag{5}$$

where $f_2(k)$ is the load curtailment rate of the node k. $f_2(k)$ is modeled as follows:

$$f_2(k) = \frac{\Delta p_{k,t}^{\mathrm{LCA}}}{p_{k,t}^{\mathrm{load}}}, t = 1, 2, \cdots, T \tag{6}$$

Thus, the comprehensive evaluation index of node k is formulated as follows:

$$r(k) = \alpha f_1(k) + \beta f_2(k) \tag{7}$$

where α and β are the weights for $f_1(k)$ and $f_2(k)$, respectively.

The index $r(k)$ for each node is ranked from the largest to the smallest, and the top S values are selected to install the ESS.

4.2 Sizing Model

4.2.1 Upper-Level Model

The upper-level (UL) problem takes the total net proceeds of the ESS within its life cycle as the objective. The detailed model is as follows:

$$\max \sum_{e\in E} \pi_e \sum_{t\in T} \left[C^{\mathrm{gain-peak}} + C^{\mathrm{lag-inv}}\right] - \sum_{t\in T} \left[C^{\mathrm{ess-inv}} + C^{\mathrm{ess-ope}}\right] \tag{8}$$

$$C^{\mathrm{gain-peak}} = \sum_{(i,k)\in C^{\mathrm{HVDN}}} C^{\mathrm{LCA}} \left(\Delta p_{e,ik,t}^{\mathrm{LCA,bef}} - \Delta p_{e,ik,t}^{\mathrm{LCA}}\right) \tag{9}$$

$$C^{\mathrm{lag-inv}} = \sum_{k\in \mathbf{B}_k^{\mathrm{E}}} C^r \eta \bar{p}_k \tag{10}$$

$$C^{\mathrm{ess-inv}} = \sum_{k\in \mathbf{B}_k^{\mathrm{E}}} \left(C^s \bar{s}_k + C^p \bar{p}_k\right) \tag{11}$$

$$C^{\mathrm{ess-ope}} = \sum_{k\in \mathbf{B}_k^{\mathrm{E}}} \left(\frac{1+i_r}{1+\varphi}\right)^{\tau} C^f \bar{p}_k \tag{12}$$

s.t.

$$0 \le \bar{s}_k \le s_{\max} \tag{13}$$

$$0 \le \bar{p}_k \le p_{\max} \tag{14}$$

$$C^{\mathrm{ess-inv}} \le C_{\max}^{\mathrm{ess-inv}} \tag{15}$$

Eq. 5 maximizes the total net proceeds of the ESS within its life cycle. It can be calculated using **Eqs. 9–12**. **Equation 9** calculates the peak-shaving revenue of the ESS. **Equation 10** calculates the revenue for delaying the investment on upgrading the grid infrastructure. **Equations 11** and **12** are the investment cost and the operational cost for the ESS, respectively. In **Eq. 11**, the parameters C^s and C^p can be calculated as follows (Pandi et al., 2015; Hassan and Dvorkin, 2018):

$$C^s = \hat{C}^s \frac{\varphi(1+\varphi)^{\omega}}{(1+\varphi)^{\omega} - 1} \cdot \frac{1}{N_D} \tag{16}$$

$$C^p = \hat{C}^p \frac{\varphi(1+\varphi)^{\omega}}{(1+\varphi)^{\omega} - 1} \cdot \frac{1}{N_D} \tag{17}$$

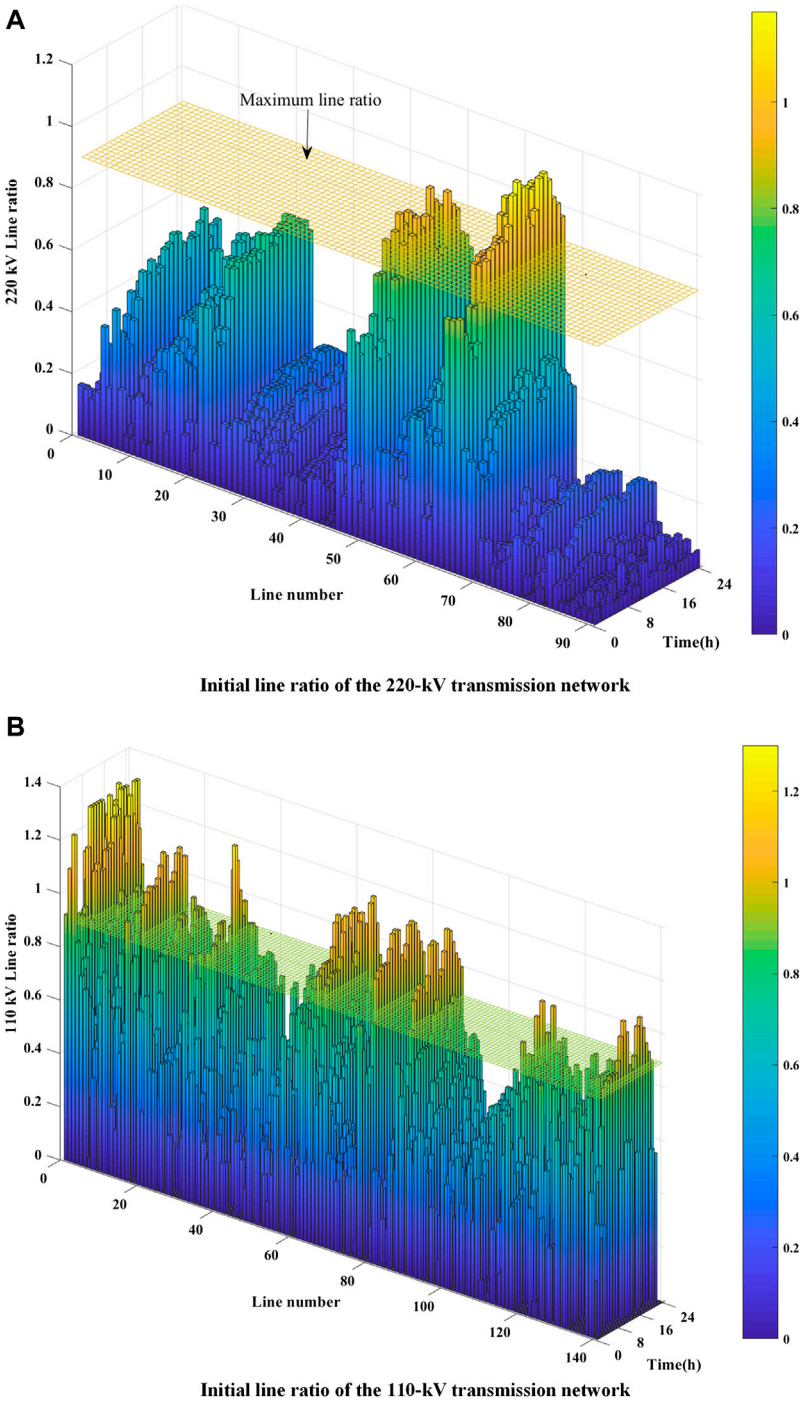


FIGURE 9 | System initial operation condition. **(A)** initial line ratio of the 220-kV transmission network. **(B)** initial line ratio of the 110-kV transmission network.

where $\hat{C}^s$ and $\hat{C}^p$ represent the energy- and power-related components of the ESS investment cost, respectively; φ represents the annual discount rate; ω represents the ESS lifetime; and N_D represents the number of days in the target year (Hassan and Dvorkin, 2018). **Equations 13** and **14** show the maximum capacity (power rating and energy) of the ESS that can be installed at each HVDN node. **Eq. 15** limits the total investment cost for the ESS.

4.2.2 Lower-Level Model

The lower-level model optimizes the topological structure of the HVDN, which aims at minimizing the curtailment of the load.

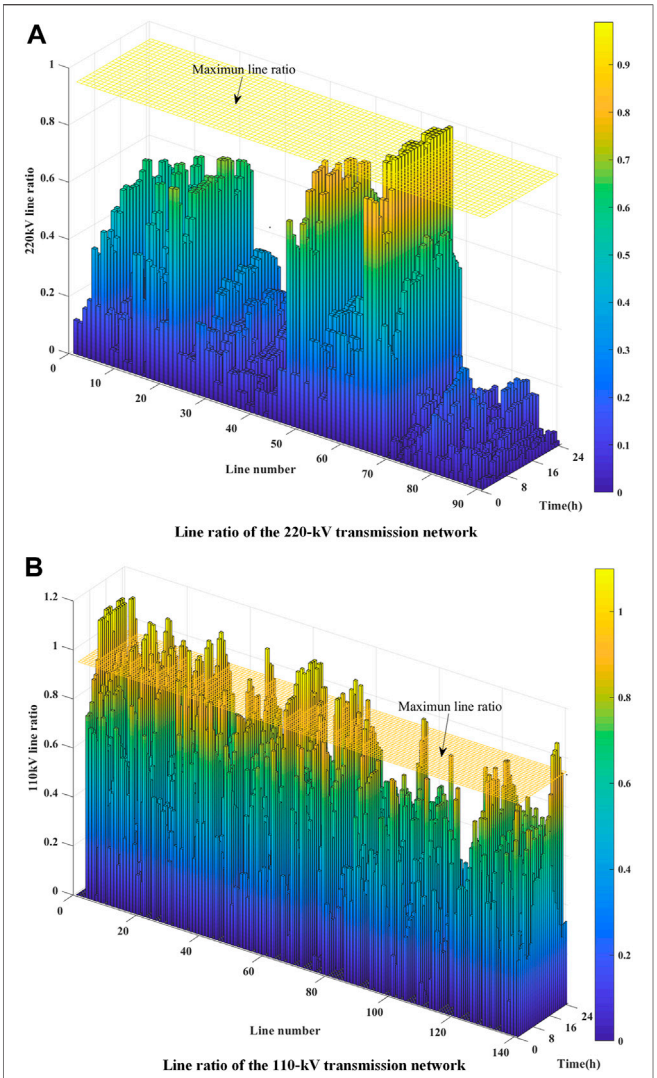


FIGURE 10 | System operation condition. **(A)** line ratio of the 220-kV transmission network. **(B)** line ratio of the 110-kV transmission network.

The detailed reconfiguration model of the HVDN in each typical day is represented as follows: $\forall e \in E,\ t \in T$:

$$\min f_n = \sum_{t\in T}\sum_{k\in \mathbf{B}^{\mathrm{H}}} C^{\mathrm{LCA}} \Delta p^{\mathrm{LCA}}_{e,k,t} \tag{18}$$

s.t.

$$p^{\mathrm{g}}_{e,i,t} - p^{\mathrm{load}}_{e,i,t} - \sum_{(i,j)\in \mathbf{L}^{\mathrm{T}}} p_{e,ij,t} = 0, \quad \forall i \in \mathbf{B}^{\mathrm{T}} \tag{19}$$

$$p^{\mathrm{load}}_{e,i,t} - \sum_{(i,k)\in \mathbf{L}^{\mathrm{T-H}}} p_{e,ik,t} = 0, \quad \forall i \in \mathbf{B}^{\mathrm{T}} \tag{20}$$

$$p^{\mathrm{load}}_{e,k,t} - \Delta p^{\mathrm{LCA}}_{e,k,t} = \sum_{(i,k)\in \mathbf{L}^{\mathrm{T-H}}} p_{e,ik,t} + \sum_{(m,k)\in \mathbf{L}^{\mathrm{H}}} p_{e,mk,t} + p^{\mathrm{ess}}_{e,k,t} - \sum_{(k,l)\in \mathbf{L}^{\mathrm{H}}} p_{e,kl,t}, \ \forall k \in \mathbf{B}^{\mathrm{H}} \tag{21}$$

$$0 \le z_{e,mk,t} + z_{e,km,t} \le 1, \quad \forall (m,k) \in \mathbf{L}^{\mathrm{H}} \tag{22}$$

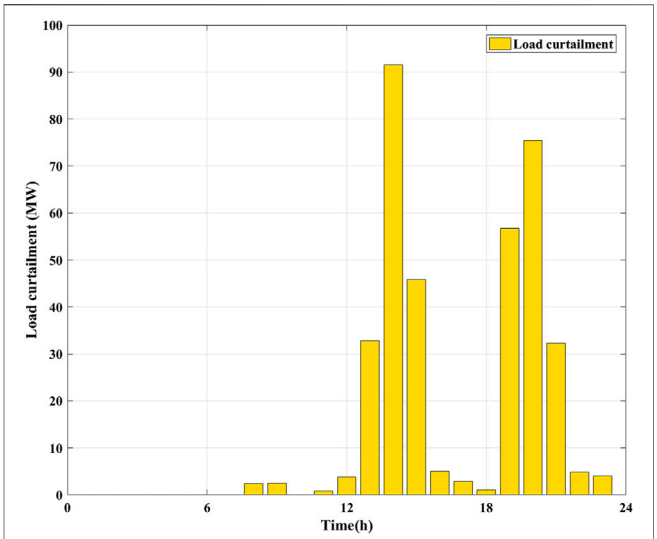


FIGURE 11 | Load curtailment.

$$\sum_{(i,k)\in \mathbf{L}^{\mathrm{T-H}}} z_{e,ik,t} + \sum_{(m,k)\in \mathbf{L}^{\mathrm{H}}} z_{e,mk,t} = 1 \tag{23}$$

$$z_{e,ki,t} \equiv 0, \quad k \in \mathbf{B}^{\mathrm{H}}, i \in \mathbf{B}^{\mathrm{U}}, \forall (i,k) \in \mathbf{L}^{\mathrm{T-H}} \tag{24}$$

$$0 \le \sum_{t\in T} \left(z_{e,km,t+1} - z_{e,km,t}\right) \le q, \ \forall (m,k) \in \mathbf{L}^{\mathrm{H}} \tag{25}$$

$$0 \le \sum_{t\in T} \left(z_{e,mk,t+1} - z_{e,mk,t}\right) \le q, \quad \forall (m,k) \in \mathbf{L}^{\mathrm{H}} \tag{26}$$

$$0 \le \sum_{(m,k)\in \mathbf{L}^{\mathrm{H}}} \left(z_{e,km,t+1} - z_{e,km,t}\right) \le \lambda \tag{27}$$

$$0 \le \sum_{(m,k)\in \mathbf{L}^{\mathrm{H}}} \left(z_{e,mk,t+1} - z_{e,mk,t}\right) \le \lambda \tag{28}$$

$$0 \le p_{e,mk,t} \le z_{e,mk,t} p^{\max}_{mk}, \quad \forall (m,k) \in \mathbf{L}^{\mathrm{H}} \tag{29}$$

$$0 \le p_{e,ik,t} \le z_{e,ik,t} p^{\max}_{ik}, \quad \forall (i,k) \in \mathbf{L}^{\mathrm{H}} \tag{30}$$

$$0 \le \Delta p^{\mathrm{LCA}}_{e,k,t} \le \Delta p^{\mathrm{LCA,\,max}}_{k}, \quad \forall k \in \mathbf{B}^{\mathrm{H}} \tag{31}$$

$$p^{\mathrm{ess}}_{e,k,t} = p^{\mathrm{dis}}_{e,k,t} - p^{\mathrm{ch}}_{e,k,t}, \quad \forall k \in \mathbf{B}^{\mathrm{E}}_{k} \tag{32}$$

$$S^{\mathrm{soc}}_{e,k,t} = S^{\mathrm{soc}}_{e,k,t-1} + \left(\alpha^{\mathrm{ch}}_{k} p^{\mathrm{ch}}_{e,k,t} - \alpha^{\mathrm{dis}}_{k} p^{\mathrm{dis}}_{e,k,t}\right)/\bar{s}_{k}, \quad \forall k \in \mathbf{B}^{\mathrm{E}}_{k} \tag{33}$$

$$0 \le p^{\mathrm{dis}}_{e,k,t} \le u^{\mathrm{dis}}_{e,k,t} \bar{p}_{k}, \quad \forall k \in \mathbf{B}^{\mathrm{E}}_{k} \tag{34}$$

$$0 \le p^{\mathrm{ch}}_{e,k,t} \le u^{\mathrm{ch}}_{e,k,t} \bar{p}_{k}, \quad \forall k \in \mathbf{B}^{\mathrm{E}}_{k} \tag{35}$$

$$S^{\mathrm{soc}}_{k,\min} \le S^{\mathrm{soc}}_{e,k,t} \le S^{\mathrm{soc}}_{k,\max}, \forall k \in \mathbf{B}^{\mathrm{E}}_{k} \tag{36}$$

$$u^{\mathrm{ch}}_{e,k,t} + u^{\mathrm{dis}}_{e,k,t} \le 1, \quad \forall k \in \mathbf{B}^{\mathrm{E}}_{k} \tag{37}$$

Equation 18 minimizes the curtailment of the load at HVDN node, **Eqs 19–21** are the power balance constraints, **Eqs 22–24** are radial constraints, **Eq. 25** and **Eq. 28** represent the limitation for the number of switching constraints, **Eq. 29** and **Eq. 30** represent the branch power constraints, **Eq. 31** is the load curtailment amount constraints at the HVDN node, and **Eqs 32–37** represent the ESS operation constraints.

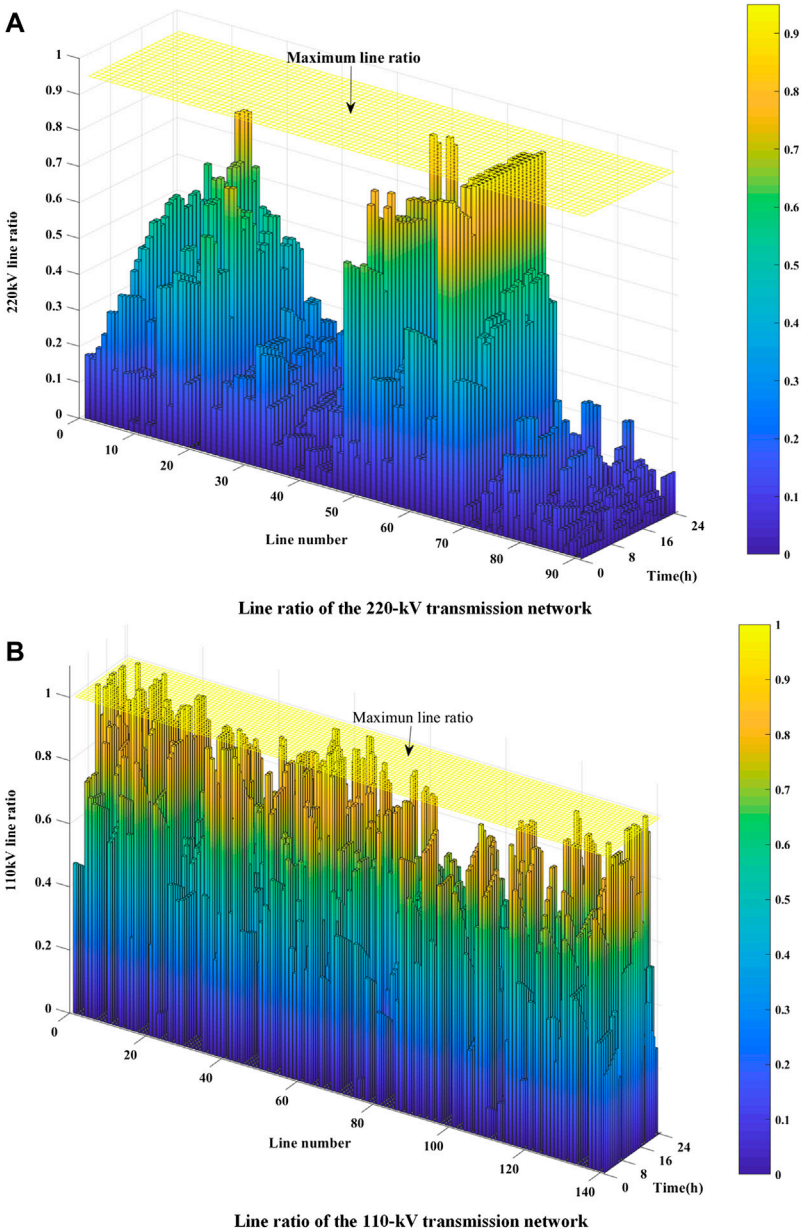


FIGURE 12 | System operation condition. **(A)** line ratio of the 220-kV transmission network. **(B)** line ratio of the 110-kV transmission network.

4.2.3 Solving Process

The UL problem features strong nonlinearity, and the LL problem is a mixed-integer linear programming problem. Therefore, a hybrid method is used in this study, including the CPLEX optimizer and the PSO (Song et al., 2019). The solving process is a procedure of alternating iterations between the UL model and the lower-level model through coupling variables. First, the UL model is solved to obtain the initial size for the ESS. Then, the lower-level model is solved to obtain the optimal collaborative operational strategy of the ESS and HVDN based on the ESS size yielded by the UL model. In the end, the updated size of each ESS is used to check whether the termination conditions are met. The flowchart of the process is shown in **Figure 5**.

5 CASE STUDY

5.1 Simulation Setup

The proposed method is tested on a real 407-node UPG system in China. The system has a peak load of 4,080 MW, and the transmission system comprises 50,220-kV substations and 91,220-kV lines. Each substation has two or three nodes that

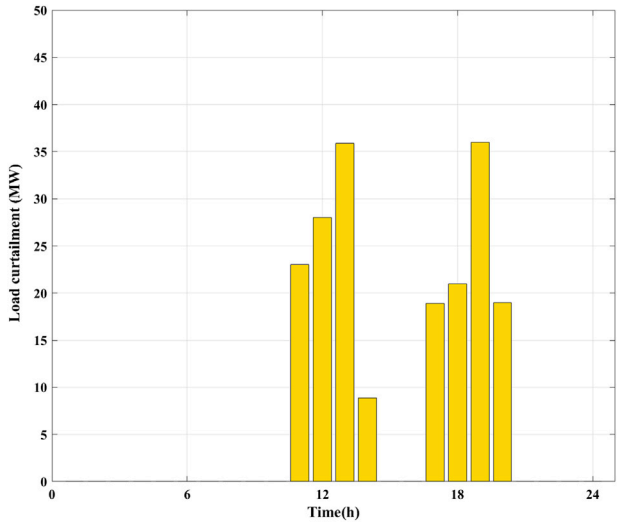


FIGURE 13 | Load curtailment. **(A)** energy storage system (ESS)-24 operational strategy. **(B)** ESS-125 operational strategy.

are connected with the HVDN. The HVDN comprises 68,110-kV substations and 138,110-kV lines. **Figure 6** shows part of the grid structure of the UPG. The ESSs are installed on the 110-kV DCU. The technical parameters of an ESS can be found in a study by Chen and Liu (2021). Other parameters are as follows: the construction life of the ESS is 2 a, its service life is 15 a, and its utilization days in a year are 280 days; the discount rate is taken as 8%; and the inflation rate is taken as 1.5%. Assuming that the ESS planning horizon n is 15 years and is divided into three stages, i.e., $N = 3$. The duration of each stage is 3, 5, and 7 years.

To show the advantages of the proposed method, three cases are set for comparison.

Case I: consider the multistage ESS planning scheme with the HVDN reconfiguration.
Case II: consider the single-stage ESS planning scheme with the HVDN reconfiguration.
Case III: consider the single-stage ESS planning scheme without the HVDN reconfiguration.

The numerical results are discussed as follows.

5.2 Simulation Results

The time-series data analysis results of the line load rate and the node load shedding rate are shown in **Figure 7**.

Figure 7A shows the maximum load factor of the lines that connect with the nodes during a day. It can be seen from **Figure 7A** that each node experiences a period of high load rate.

Figure 7B shows the node load shedding index during a day. It can be seen from **Figure 7B** that some nodes encounter the curtailment of the load during the peak time to meet the constraints.

According to the importance of the above indicators in the upper objective function, the weights of each indicator are set to form a comprehensive indicator curve as shown in **Figure 8**. As shown in **Figure 8**, the indicator value of eight nodes, i.e., nodes 5, 14, 24, 38, 74, 114, 125, and 173, surpass the threshold. Hence, we chose these nodes as the candidates for installing the ESS.

TABLE 4 | Comparison of results of different congestion management schemes.

Method	Optimization results	
Method I	Total number of switching	35 times
	Total load curtailment	361.89 MW
Method II	Total number of switching	19 times
	Total load curtailment	190.68 MW

The planning scenarios are solved separately for different cases, and the configuration of the ESS at each stage is shown in **Tables 1–3**.

1) As shown in **Table 1**, the capacity of the ESS keeps increasing at each planning stage to achieve the maximization of the objective of the UL model.

2) As shown in **Table 2**, regardless of the ESS construction sequence, the investment cost increases by 3.6% from RMB 736.8 million in case I to RMB 763.30 million in case II, and the operation cost increases by 8.8% from RMB 214.2 million in case I to RMB 233 million in case II. It is obvious that the proposed multistage planning method achieves a better operational economy. The waste of the resources caused by the overinvestment is avoided while the growing load demand is satisfied.

3) From **Tables 2** and **3**, it can be seen that the ESS planning without considering the HVDN reconfiguration not only requires a larger investment cost but also has a lower net benefit. The results demonstrate that the collaboration of the ESS operational strategy and HVDN reconfiguration can reduce the cost for ESS allocation.
4) As shown in **Tables 1–3**, the net benefits of the system are positive during the whole life cycle of the ESS. It shows that the investment of the ESS can effectively save the cost for constructing new transmission lines and substations.

The line ratio during a day is shown in **Figure 9**. To highlight the advantages of the proposed method, the method that only considers the HVDN reconfiguration (Haghighat and Zeng, 2016) is used to make a comparison with the proposed method. The numerical results are discussed as follows.

As shown in **Figure 10**, most of the line ratios are limited to the acceptable level after reconfiguring the HVDN topology. The load curtailment is shown in **Figure 11**.

In **Figure 11**, the load curtailment occurs at the 12th to 23rd time intervals and 361.89-MW load is curtailed in total. Although the transmission congestion is alleviated, the problems of frequent HVDN reconfiguration and the load shedding remain unsolved.

The optimized line ratio yielded by the proposed method is shown in Figure 12.

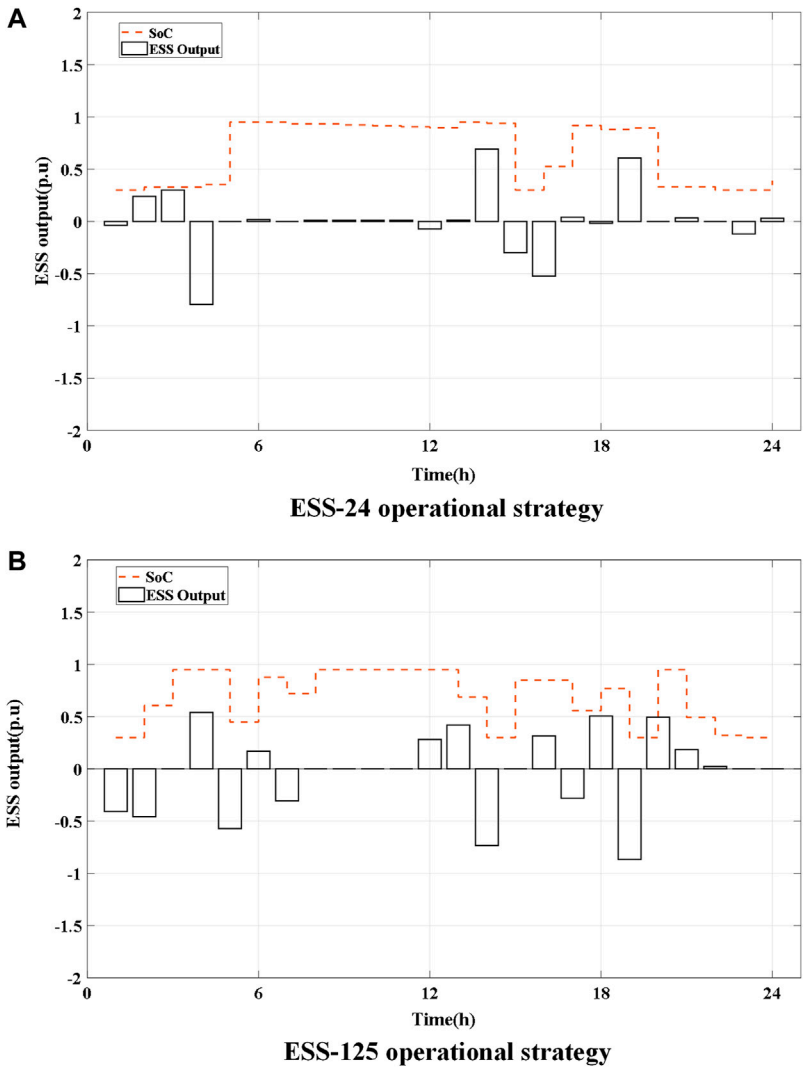


FIGURE 14 | Energy storage system (ESS) operational strategy. **(A)** ESS-24 operational strategy. **(B)** ESS-125 operational strategy.

As shown in **Figure 12**, all line ratios are limited to the acceptable levels after performing the optimal ESS charging/discharging strategy and HVDN reconfiguration. The operational results are shown in **Figures 13** and **14**.

As shown in **Figures 13** and **14**, the load shedding is reduced and the SoC of each ESS is within the specified range.

5.3 Results and Discussion

The optimization results of the two methods are shown in **Table 4**. The proposed method is represented as method II.

As shown in **Table 4**, the traditional method that only considers the HVDN reconfiguration requires 361.89-MW load curtailment, and the number of switch actions reaches 35 times. Based on the collaboration of the ESS operational strategy and the HVDN reconfiguration proposed in this study, the load shedding can be reduced to 190.68 MW. Besides, the number of switch actions reduces to 19 times compared with that in the conventional method.

6 CONCLUSION

To reduce the frequency of HVDN reconfiguration and mitigate the transmission congestion, this study proposes a multistage planning method for ESS allocation in the UPG, considering the influence of HVDN reconfiguration. From the numerical results, conclusions can be drawn, as follows:

1) Through co-optimizing the HVDN topology and ESS operational strategy, transmission congestion can be effectively mitigated while reducing the load curtailment and the frequency of the HVDN reconfiguration.
2) The proposed ESS planning procedure incorporates the HVDN reconfiguration, which can effectively reduce the cost of ESS installation.

3) The multistage planning method can improve the economic operation of the ESS compared with the single-stage planning method.

In future work, the effect of the 10-kV distribution system reconfiguration can be further incorporated in the ESS planning model. Besides, the uncertainties of the RES and electric vehicles can be also considered to improve the adaptiveness of the proposed method.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, and further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

ZC, KY, XZ, and YZ: conceptualization, methodology; YC, RY, YG, XZ, and YZ: writing original draft preparation; JL and YL: supervision; YZ, SS, and SP: writing reviewing and editing.

FUNDING

The authors declare that this study received funding from the guangdong Power Grid Corporation Electricity Planning Theme Research Project Grant(030400QQ00210003). The funder had the following involvement in the study: design, collection, analysis, interpretation of data, the writing of this article and the decision to submit it for publication.

REFERENCES

Bird, L., Lew, D., Milligan, M., Carlini, E. M., Estanqueiro, A., Flynn, D., et al. (2016). Wind and Solar Energy Curtailment: a Review of International Experience. *Renew. Sustain. Energy Rev.* 65, 577–586. doi:10.1016/j.rser.2016.06.082

Cao, Y., Mu, Y., Jia, H., Yu, X., Song, Y., Wu, K, et al. (2020). Multi-stage Planning of Park-Level Integrated Energy System Considering Construction Time Sequence[J]. *Proc. CSEE* 40 (21), 6815–6828. doi:10.13334/j.0258-8013.pcsee.200622

Chen, Y., and Liu, Y. (2021). Congestion Management of Microgrids with Renewable Energy Resources and Energy Storage Systems[J]. *Front. Energy Res.* 9, 708087. doi:10.3389/fenrg.2021.708087

Del Rosso, A. D., and Eckroad, S. W. (2014). Energy Storage for Relief of Transmission Congestion. *IEEE Trans. Smart Grid* 5 (2), 1138–1146. doi:10.1109/tsg.2013.2277411

Dhundhara, S., and Verma, Y. P. (2018). Capacitive Energy Storage with Optimized Controller for Frequency Regulation in Realistic Multisource Deregulated Power System. *Energy* 147, 1108–1128. doi:10.1016/j.energy.2018.01.076

Fiorini, L., Pagani, G. A., Pelacchi, P., Poli, D., and Aiello, M. (2017). Sizing and Siting of Large-Scale Batteries in Transmission Grids to Optimize the Use of Renewables[J]. *IEEE J. Emerg. Sel. Top. Circuits Syst.* 7 (2), 285–294. doi:10.1109/JETCAS.2017.2657795

Goop, J., Odenberger, M., and Johnsson, F. (2017). The Effect of High Levels of Solar Generation on Congestion in the European Electricity Transmission Grid. *Appl. Energy* 205, 1128–1140. doi:10.1016/j.apenergy.2017.08.143

Guo, W., Xiu, X., Li, W., and Li, J. (2020). Siting and Configuration Methods for Grid-Side Energy Storage System Considering Multi-Attribute Comprehensive Indices and Economy[J]. *Electr. Power Constr.* 41 (04), 53–62.

Haghighat, H., and Zeng, B. (2016). Distribution System Reconfiguration under Uncertain Load and Renewable Generation. *IEEE Trans. Power Syst.* 31 (4), 2666–2675. doi:10.1109/tpwrs.2015.2481508

Hassan, A., and Dvorkin, Y. (2018). Energy Storage Siting and Sizing in Coordinated Distribution and Transmission Systems[J]. *IEEE Trans. Sustain. Energy* 9, 1692–1701. doi:10.1109/TSTE.2018.2809580

He, H., Du, E., Zhang, N., Kang, C., and Wang, X. (2021). Enhancing the Power Grid Flexibility with Battery Energy Storage Transportation and Transmission Switching. *Appl. Energy* 290 (2), 116692. doi:10.1016/j.apenergy.2021.116692

Hoffrichter, A., Barrios, H., Massmann, J., Venkataramanachar, B., and Schnettler, A. (2018). Impact of Considering 110 kV Grid Structures on the Congestion Management in the German Transmission Grid. *Scigrid Int. Conf. Power Grid Model.* 977. doi:10.1088/1742-6596/977/1/012004

Hu, Z. C., Zhang, F., and Li, B. W. (2012). "Transmission Expansion Planning Considering the Deployment of Energy Storage Systems[C]," in *2012 IEEE Power and Energy Society General Meeting* (San Diego, CA: IEEE).

Khanabadi, M., Fu, Y., and Liu, C. (2018). Decentralized Transmission Line Switching for Congestion Management of Interconnected Power Systems. *IEEE Trans. Power Syst.* 33 (6), 5902–5912. doi:10.1109/tpwrs.2018.2838046

Macrae, C., Ozlen, M., and Ernst, A. (2014). "Transmission Expansion Planning Considering Energy Storage[C]," in *2014 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC)* (Ixtapa, Mexico: IEEE).

Macrae, C. A. G., Ernst, A. T., and Ozlen, M. (2016). A Benders Decomposition Approach to Transmission Expansion Planning Considering Energy Storage. *Energy* 112, 795–803. doi:10.1016/j.energy.2016.06.080

Nick, M., Cherkaoui, R., and Paolone, M. (2017). Optimal Planning of Distributed Energy Storage Systems in Active Distribution Networks Embedding Grid Reconfiguration[J]. *IEEE Trans. Power Syst.* 33 (2), 1577–1590. doi:10.1109/TPWRS.2017.2734942

Nick, M., Cherkaoui, R., and Paolone, M. (2014). Optimal Allocation of Dispersed Energy Storage Systems in Active Distribution Networks for Energy Balance and Grid Support. *IEEE Trans. Power Syst.* 29 (5), 2300–2310. doi:10.1109/tpwrs.2014.2302020

Pandi, H., Wang, Y., Qiu, T., Dvorkin, Y., and Kirschen, D. S. (2015). Near-optimal Method for Siting and Sizing of Distributed Storage in a Transmission Network. *IEEE Trans. Power Syst.* 30 (5), 2288–2300. doi:10.1109/TPWRS.2014.2364257

Salkuti, S. R. (2018). Congestion Management Using Optimal Transmission Switching. *IEEE Syst. J.* 12, 3555–3564. doi:10.1109/jsyst.2018.2808260

Song, T., Han, X., Jia, Y., Qin, W., Zhang, B., and Zhang, Q. (2019). "Multi-Stage Bi-level Planning of Energy Storage Considering Cycling Degradation," in *2019 IEEE Power & Energy Society General Meeting (PESGM)* (Atlanta, Georgia, United States: IEEE). doi:10.1109/pesgm40551.2019.8973797

Subramani, G., Ramachandaramurthy, V. K., and Vijyakumar, K. N. (2018). Optimal Sizing of Battery Energy Storage System (BESS) for Peak Shaving under Malaysian Electricity Tariff. *Adv. Sci. Lett.* 24 (3), 1861–1865. doi:10.1166/asl.2018.11177

Yan, X., Gu, C., Zhang, X., and Li, F. (2020). Robust Optimization-Based Energy Storage Operation for System Congestion Management. *IEEE Syst. J.* 14 (2), 2694–2702. doi:10.1109/jsyst.2019.2932897

Yang, X., Chai, G., Liu, X., Xu, M., and Guo, Q. (2021). Storage-Transmission Joint Planning Method to Deal with Insufficient Flexibility and Transmission Congestion. *Front. Energy Res.* 8, 612909. doi:10.3389/fenrg.2020.612909

Yuan, Z., and Hesamzadeh, M. R. (2017). Hierarchical Coordination of TSO-DSO Economic Dispatch Considering Large-Scale Integration of Distributed Energy Resources. *Appl. Energy* 195, 600–615. doi:10.1016/j.apenergy.2017.03.042

Zhang, X., Liu, Y., Gao, H., Wang, L., and Liu, J. (2020). A Bi-level Corrective Line Switching Model for Urban Power Grid Congestion Mitigation. *IEEE Trans. Power Syst.* 35 (4), 2959–2970. doi:10.1109/tpwrs.2019.2959586

Zhang, X., Liu, Y., Zeng, Y., Wu, G., and Liu, J. (2022). Prosumer-centric Energy Storage System and High Voltage Distribution Network Topology Co-optimisation for Urban Grid Congestion Management. *IET Smart Grid* 24 (7), 1–11. doi:10.1049/stg2.12061

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

NOMENCLATURE

Abbreviations

UPG Urban power grid

TS 220-kV transmission system (TS) in the UPG

HVDN 110-kV high-voltage distribution network

ESS Energy storage system

B. Indexes

i,j Node index in the TS

k,l,m Node index in the HVDN

b Node index in the HVDN

n Number of the regional HVDN

E Set of typical operational scenarios, indexed by e

T Set of time intervals, indexed by t

C. Sets

$\mathbf{B}^{\mathrm{T}}$, $\mathbf{L}^{\mathrm{T}}$ Node set and branch set in the TS

$\mathbf{B}^{\mathrm{H}}$, $\mathbf{L}^{\mathrm{H}}$ Node set and branch set in the HVDN

$\mathbf{L}^{\mathrm{T-H}}$ Branch set in the TS and HVDN interface

$\mathbf{B}_k^{\mathrm{E}}$ Set of nodes that connect with an ESS

$\mathbf{E}_{\mathrm{set}}$ Indicating the planning scheme of the ESS

S Set of the planning stages

D. Parameter

$C^s/C^p/C^f$ Unit prices for energy reservoir (C^s), power rating (C^f), and operational costs (C^f)

π_e Weight of a typical operational scenario indexed by e

n_E Number of the typical operational scenarios

$\alpha_k^{\mathbf{ch/dis}}$ Charging/discharging efficiency of an ESS

ω Energy storage lifetime

φ Annual discount rate

λ The maximum number of the operated HVDN switches during the time interval t

q The maximum number of times an HVDN switch can be operated in a day

i_r Inflation rate

N_D Number of days in a calendar year

C^{LCA} Cost for load curtailment

$\Delta p_k^{\mathrm{LCA,\,max}}$ Maximum load curtailment amount (LCA) at the bus k

$s_{\max}$ Maximum install energy of an ESS

$p_{\max}$ Maximum install power rating of an ESS

$S_{k,\min}^{\mathbf{soc}}$, $S_{k,\max}^{\mathbf{soc}}$ Minimum and maximum SoCs of the ESS at the bus k

$p_{ik}^{\max}$ Maximum active power through the branch (i,k)

$p_{mk}^{\max}$ Maximum active power through the branch (m,k)

$C_{\max}^{\mathbf{ess-inv}}$ Maximum ESS investment cost

E. Variables

$\Delta p_{e,k,t}^{\mathrm{LCA}}$ LCA at the node k during the time interval t

$p_{e,i,t}^{g}$ Active power injected into the node i during the time interval t

$p_{e,i,t}^{\mathbf{load}}$, $p_{e,k,t}^{\mathrm{load}}$ Active load demand at the node i and k during the time interval t

$p_{e,mk,t}$, $p_{e,kl,t}$ Active power through the branch (m,k) and (k,l) during the time interval t

$p_{e,ij,t}$, $p_{e,ik,t}$ Active power through the branch (i,j) and (i,k) during the time interval t

$p_{k,mk,t}$ Active power through the branch (m,k) to/from node k during the time interval t

$z_{e,ik,t}$, $z_{e,mk,t}$ Binary variable that indicates the power direction through the branch (i,k) and (m,k) during the time interval t

$p_{e,k,t}^{\mathbf{ch/dis}}$ Charging/discharging of storage at the bus k during the time interval t

$S_{e,k,t}^{\mathbf{soc}}$ SoC of ESS at the time interval t

$\bar{s}_k/\bar{p}_k$ Energy/power rating of the ESS at the bus k

$u_{e,k,t}^{\mathbf{ch/dis}}$ Binary variable indicating whether the ESS is installed at the bus k

frontiers | Frontiers in Energy Research

TYPE Original Research
PUBLISHED 26 August 2022
DOI 10.3389/fenrg.2022.959786

OPEN ACCESS

EDITED BY
Gaoqi Liang,
Nanyang Technological University, Singapore

REVIEWED BY
Peishuai Li,
Nanjing University of Science and Technology, China
Minghao Wang,
Hong Kong Polytechnic University, Hong Kong, SAR China
Youwei Jia,
Hong Kong Polytechnic University, Hong Kong, SAR China

*CORRESPONDENCE
Junhua Zhao,
zhaojunhua@cuhk.edu.cn

SPECIALTY SECTION
This article was submitted to Smart Grids, a section of the journal Frontiers in Energy Research

RECEIVED 02 June 2022
ACCEPTED 28 June 2022
PUBLISHED 26 August 2022

CITATION
Yang C, He B, Liao H, Ruan J and Zhao J (2022), Price-based low-carbon demand response considering the conduction of carbon emission costs in smart grids.
Front. Energy Res. 10:959786.
doi: 10.3389/fenrg.2022.959786

Price-based low-carbon demand response considering the conduction of carbon emission costs in smart grids

Chao Yang, Binghao He, Huanxin Liao, Jiaqi Ruan and Junhua Zhao *

School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, Shenzhen, China

The rapid development of the Industrial Internet-of-Things extends demand response (DR) research to the aspect of low-carbon emission in smart grids. This study proposed the concept of low-carbon DR (LCDR) in the electricity market as well as the price-based LCDR mechanism and its model. First, carbon cost conduction from the generation side to the demand side was analyzed, and then conduction function was quantifiably deduced. Second, the mechanism and model of price-based LCDR were proposed by considering three DR signals, namely, the electricity price, carbon price, and carbon emission intensity of the demand side, based on the traditional price-based DR (PBDR) mechanism. Third, the proposed LCDR mechanism was applied to the environmental–economic dispatch optimization problem. At last, case studies on the modified IEEE 39-bus system verified that the LCDR mechanism can reduce carbon emissions while maintaining the function of the traditional PBDR. Meanwhile, the applicability of LCDR was illustrated based on carbon emission sensitivity to LCDR model parameters. The proposed mechanism can guide participants in the electricity market in reducing electricity carbon emissions.

KEYWORDS

price-based low-carbon demand response, carbon emissions, carbon emission cost conduction, carbon emission intensity, environmental–economic dispatch

Introduction

The application of Industrial Internet-of-Things (IIoTs) (Sisinni et al., 2018) brings advanced measurement and communication technologies to the power grid, such as cyber networks (Sridhar et al., 2012) and smart meters (Depuru et al., 2011), making the power grid smarter than ever. The extensive use of smart equipment significantly enhances the observability and controllability of the demand side in the smart grid and provides strong hardware and data support for the development of demand response (DR) (Chen et al., 2020). DR refers to "changes in electric usage by customers from their normal consumption patterns in response to changes in the price of electricity over time, or

to incentive payments designed to induce lower electricity use at times of high wholesale electricity market prices or when system reliability is jeopardized" (Siano, 2014).

At present, carbon emission reduction has achieved a global consensus (United Nation, 2021). As consumption terminals, the power consumers in the demand side should undertake their carbon reduction responsibility. Based on DR, consumers can change their load in time and space with their power suppliers from high- to low-carbon emission generators, resulting in carbon emission reduction. In this context, DR can play a vital role in carbon neutrality based on the IIoTs.

By widely using smart meters, the energy consumption of the demand side can be effectively monitored. Using high-frequency meter data, more accurate characteristics of loads and carbon emissions can be obtained. Based on the advanced cyber networks, the consumers can receive more accurate and timely DR signals (Albadi and El-Saadany, 2007) from independent system operators or dispatching departments. These signals include electricity price, electricity carbon emission intensity (CEI), incentives, and control instructions. As a result, consumers can achieve carbon emission reduction (Stoll et al., 2014), which is beneficial for themselves and the power grid, by reducing their carbon consumption.

DR programs generally include the price-based demand response (PBDR) (Conejo et al., 2010) and the incentive-based DR (IBDR) (Lu and Hong, 2019). In PBDR programs, consumers respond to real-time prices based on the demand elasticity model (Kirschen et al., 2000) and modify their demands by shaving/shifting flexible load from peak times into valley times to flatten load curves. PBDR programs are usually applied to the day-ahead and real-time market clearing (Wu, 2013; (Li et al., 2022), economic dispatching (Dehnavi and Abdi, 2016), and power grid planning (Wang et al., 2020). A successful PBDR can help the power market set efficient electricity prices, improve economic efficiency, and reduce environmental costs and carbon emissions.

At present, carbon emission costs (Gillenwater and Breidenich, 2009) have been internalized into the electricity market. It is generally considered that carbon emission costs are directly included in the generation side (Newcomer et al., 2008). Through market clearing, carbon emission costs can be conducted from the generation side to demand side, accompanied by the increase in electricity prices (Wook et al. (2010); Panagiotis et al. (2021)). The energy consumption price that consumers pay becomes the integrated electricity–carbon price. Thus, the PBDR needs to take carbon emission costs into account; moreover, research on PBDR should be extended to the field of low-carbon emission.

Recent research on low-carbon–oriented PBDR mostly focuses on the impact of traditional PBDR programs on carbon emissions. He et al., (2020) proposed an environmental–economic dispatch (EED) model with a price-based integrated DR program considering the carbon trading scheme. In the study by Zeng et al. (2014), PBDR is used for planning distribution systems in a transition toward low-carbon sustainability. A new piecewise linear approximation method for calculating carbon emission factors was proposed by Fleschutz et al. (2021) to quantify the effect of PBDRs on operational carbon emissions in European countries. Song et al. (2014) and Dahl and Petersen (2016) developed new simulation models to investigate the joint influence of price and CO_2 signals in DR programs. However, existing research still have major problems. First, there is no clear analysis of the conduction process of carbon emission costs from the generation side to the demand side, making it difficult to quantify carbon emission costs for the demand side. Second, there is a lack of low-carbon–oriented DR mechanism to achieve better carbon emission reductions.

Therefore, this study proposed the concept of LCDR as well as the price-based LCDR mechanism considering the conduction of carbon emission costs in the smart grid. The main contributions of this study are as follows:

1) Quantified carbon emission cost conduction from the generation side to the demand side in the wholesale electricity market and deduced electricity carbon emission cost function and carbon emission cost conductivity (CECC) in the demand side.
2) Proposed the price-based LCDR mechanism and its model, which integrates three DR signals, namely, electricity price, carbon price, and CEI of the demand side.
3) Established an EED model considering the price-based LCDR model to verify the proposed mechanism and analyzed the applicability of the mechanism based on the sensitivity of carbon emission to the three main parameters of the LCDR model.

The remainder of the paper is organized as follows. Section 2 quantifies carbon emission cost conduction in the wholesale electricity market. Section 3 proposes the price-based LCDR mechanism and its corresponding model. Section 4 establishes an EED optimization model considering the LCDR mechanism. Section 5 presents case studies to show the function of carbon emission reduction in the LCDR mechanism and its application scenarios.

Quantified analysis of carbon emission cost conduction in the wholesale electricity market

In general, carbon emission cost is internalized into electricity cost and directly included in the generation side. It is directly passed to the demand side in the wholesale electricity market, resulting in the increase in electricity prices. According to the operation characteristic of the power grid and the market clearing method, the additional carbon emission cost per unit of

electricity consumption is different. That is, each node in the power grid has a different CECC.

Therefore, this study proposed a quantified analysis method for carbon emission cost conduction including three parts: 1) the carbon emission cost of electricity consumption, 2) the clearing mechanism of the electricity market, 3) the electricity CEI of the demand side.

Carbon emission cost of electricity consumption

Carbon price in the carbon trading market offers a generally accepted reference price for market participants, which can be used to calculate carbon emission cost. In general, carbon emission cost (Zhang et al., 2016) is the product of carbon price and carbon emissions:

$$C_t^c = \rho_t^c \cdot E_t \tag{1}$$

In the power grid, the carbon emission cost generated by power production and consumption can be accurately calculated using real-time carbon price, electricity CEI, and electricity power, benefiting from advanced smart meters:

$$C_t^c = \rho_t^c \cdot e_t \cdot P_t \tag{2}$$

where e_t and P_t are determined by the physical characteristic and the market clearing mechanism of the power grid.

Clearing mechanism of the electricity market

When demand-side loads are determined, the market clearing mechanism determines the operation state and output of the generator.

The security-constrained economic dispatch (Jabr et al., 2000) is a general spot market clearing model. The purpose of this model is to maximize the social welfare of power grids over a period of time. It needs to meet a series of physical operational constraints of the power grid. In general, the time period is divided into multiple time steps in the economic dispatch model. Then, the scheduling result of each time step is calculated using the optimal power flow (OPF) model (Dommel and Tinney, 1968). The OPF model aims to minimize the short-term generation costs, calculate the output of each generator, and determine the electricity price for consumers using the well-known locational marginal price method (Kirschen and Strbac, 2004). Using the obtained market clearing results, the electricity carbon emissions of generators and consumers can be calculated as follows:

$$E_{G,t} = e_G \cdot P_{G,t} \tag{3}$$

$$E_{D,t} = e_{D,t} \cdot P_{D,t} \tag{4}$$

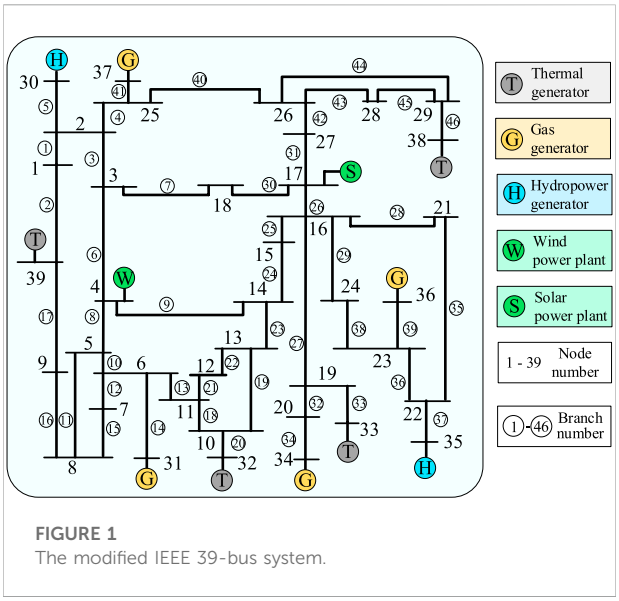


FIGURE 1
The modified IEEE 39-bus system.

where e_G is a performance parameter for generators, which is supplied by generator producers.

It can be seen that with a certain load and carbon price, the carbon emission cost of a consumer is only determined by $e_{D,t}$. Thus, $e_{D,t}$ can quantify carbon costs conducted from the generation side to the demand side. Therefore, the solution of $e_{D,t}$ is the key problem. The solution contains two parts, namely, the operation condition problem and the market clearing mechanism of the power grid. Based on the market clearing results, the carbon emission flow (CEF) method (Zhou et al., 2015) is introduced to calculate $e_{d,t}$ at time t in this study, which is presented in detail in the following section.

Electricity carbon emission intensity of the demand side

The introduced CEF model can track the footprint of carbon emissions accompanying power flow in the power grid. Then, the consumers' real-time electricity CEI can be calculated, and from which generator carbon emissions are coming at a certain time can be determined. According to Zhou et al. (2015), the electricity CEI of nodes can be calculated as follows:

$$\boldsymbol{e}_{D,t} = \left(\boldsymbol{P}_{D,t} - \boldsymbol{P}_{B,t}^{'T}\right)^{-1} \cdot \boldsymbol{R}_{C,t} \tag{5}$$

$$\boldsymbol{R}_{C,t} = \boldsymbol{P}_{G,t} \cdot \boldsymbol{e}_G \tag{6}$$

where $\boldsymbol{e}_{D,t}$ and $\boldsymbol{R}_{C,t}$ are *N-dimensional* column vectors, $\boldsymbol{P}_{D,t}$ and $\boldsymbol{P}'_{B,t}$ are $N \times N$-level matrices, $\boldsymbol{P}_{G,t}$ is an $N \times G$-level matrix, and $\boldsymbol{e}_{D,t}$ is a G-dimensional column vector.

We further derived the carbon emission cost conduction function for electricity consumptions and the equivalent

conducted carbon price for demand-side consumers according to Eqs. 2 and 5:

$$C^{c}_{D,t} = \rho^{c}_{t} \cdot e_{D,t} \cdot P_{D,t} \tag{7}$$

$$\rho^{c}_{D,t} = \frac{C^{c}_{D,t}}{P_{D,t}} = e_{D,t}\rho^{c}_{t} \tag{8}$$

Based on this, the concept of CECC was proposed in this study. The nodal CECC of the ith node is expressed as follows:

$$\begin{cases} \gamma^{c}_{D,i,t} = \dfrac{e_{D,i,t}}{e_{ACEI,t}} \\ e_{ACEI,t} = \dfrac{\sum_{g=1}^{G} P_{G,g,t} e_{G,g}}{\sum_{i=1}^{N} P_{D,i,t}} \end{cases} \tag{9}$$

$\gamma^{c}_{D,i,t}$ accurately reflects the difference in CECC for consumers caused by the operation characteristics of the power grid in time and space.

Mechanism of the price-based low-carbon demand response

Price-based demand response model

Based on the economic theory, the elasticity coefficient is generally used to describe the change rate between consumers' demand and price (Conejo et al., 2010) in PBDR programs, which is expressed as follows:

$$\varepsilon_{t,\tau} = \frac{\rho_{\tau}}{P_{D,t}} \frac{dP_{D,t}}{d\rho_{\tau}} \tag{10}$$

When t equals τ, $\varepsilon_{t,t}$ is defined as a negative self-elasticity. On the contrary, $\varepsilon_{t,\tau}$ is defined as a positive cross-elasticity. By introducing the multiperiod linear elastic load model, the response load amount is expressed as follows:

$$\Delta P_{D,t} = \sum_{\tau=1}^{T} \varepsilon_{t,\tau} P^{0}_{D,\tau} \left(\frac{\rho_{\tau}}{\rho^{0}_{\tau}} - 1 \right) \tag{11}$$

Price-based low-carbon demand response model

This study proposed a price-based LCDR model considering two factors that impact the demand of consumers, such as the integrated electricity–carbon price and the nodal CEI of the demand side.

First, consumers make demand changes in response to the integrated price that includes electricity and carbon costs conducted from the generation side. According to Eq. 11, the first part of the DR amount on the ith node is as follows:

$$\Delta P'_{D,i,t} = \sum_{\tau=1}^{T} \varepsilon_{i,t,\tau} P^{0}_{D,i,\tau} \left(\frac{\rho^{e}_{i,\tau}}{\rho^{e0}_{i,\tau}} - 1 \right) \tag{12}$$

Then, consumers continue to carry out the low-carbon response to their nodal CEIs. By translating the impact of CEIs into carbon costs, the second part of the DR amount is expressed as follows:

$$\Delta P''_{D,i,t} = \sum_{\tau=1}^{T} \varepsilon_{i,t,\tau} P^{0}_{D,i,\tau} \left(\frac{e_{D,i,\tau}}{e^{0}_{D,i,\tau}} + \frac{\rho^{c}_{\tau}}{\rho^{c0}_{\tau}} - 2 \right) \Bigg] \tag{13}$$

The detailed derivation of Eq. 13 is presented in the Appendix. Based on Eqs. 12 and 13, the total DR amount of the LCDR is derived as follows:

$$\begin{aligned} \Delta P_{D,i,t} &= \Delta P'_{D,i,t} + \Delta P''_{D,i,t} \\ &= \sum_{\tau=1}^{T} \varepsilon_{i,t,\tau} P^{0}_{D,i,\tau} \left[\left(\frac{\rho^{e}_{i,\tau}}{\rho^{e0}_{i,\tau}} - 1 \right) + \left(\frac{e_{D,i,\tau}}{e^{0}_{D,i,\tau}} + \frac{\rho^{c}_{\tau}}{\rho^{c0}_{\tau}} - 2 \right) \right] \end{aligned} \tag{14}$$

Furthermore, when using the single-period elastic load model (Zeng et al., 2014) for simplicity, the relevant parts of $\varepsilon_{i,t,\tau}$ in Eq. 14 can be omitted. At last, the designed LCDR model is expressed as follows:

$$\Delta P_{D,i,t} \approx \varepsilon_{i,t,t} P^{0}_{D,i,t} \left[\left(\frac{\rho^{e}_{i,t}}{\rho^{e0}_{i,t}} - 1 \right) + \left(\frac{e_{D,i,t}}{e^{0}_{D,i,t}} + \frac{\rho^{c}_{t}}{\rho^{c0}_{t}} - 2 \right) \right] \tag{15}$$

Environmental–economic dispatch optimization model considering the price-based low-carbon demand response

The EED model determines the operation state of power grids and the three DR signals. Thus, the proposed price-based LCDR is studied based on the EED model.

Objective function

The EED optimization model aims to minimize short-term generation costs within 1 day in this study. Generation costs are the sum of fuel and carbon emission costs, which is expressed as follows

$$\text{Min}: F = \sum_{t=1}^{T} \sum_{g=1}^{G} k \left[C^{e}_{G}\left(P_{G,g,t} \right) + C^{c}_{G}\left(P_{G,g,t} \right) \right] \tag{16}$$

where the fuel and carbon emission cost functions of thermal and gas generators are calculated as follows:

$$C^{e}_{G,g,t} = a_g P^{2}_{G,g,t} + b_g P_{G,g,t} + c_g \tag{17}$$

TABLE 1 Parameters of generators and power plants in the modified IEEE 39-bus system.

Location	Type	Rated active power (MW)	Ramping rate (MW/min)	CEI (tCO_2/MW)	Generation cost coefficients		
					a_g	b_g	c_g
4	Wind	--	--	0.0	0.00	0	0.0
17	Solar	--	--	0.0	0.00	0	0.0
30	Hydro	1,040.0	5.20	0.0	0.00	0	0.0
31	Gas	646.0	3.23	0.4	0.02	20	2.0
32	Thermal	725.0	3.63	1.0	0.01	5	0.2
33	Thermal	652.0	3.26	1.0	0.01	5	0.2
34	Gas	508.0	2.54	0.4	0.02	20	2.0
35	Hydro	687.0	3.44	0.0	0.00	0	2.0
36	Gas	580.0	2.90	0.4	0.04	20	2.0
37	Gas	564.0	2.82	0.4	0.04	20	2.0
38	Thermal	865.0	4.33	1.0	0.01	5	0.2
39	Thermal	1,100.0	5.50	1.0	0.01	5	0.2

$$C^{c}_{G,g,t} = \rho^{c}_{t} E_{G,g,t} = \rho^{c}_{t} e_{G,g,t} P_{G,g,t} \tag{18}$$

The short-term generation costs of renewable power plants, such as wind and solar powers, are assumed to be zero. The uncertainty and randomness of their outputs are also ignored in this study. Moreover, DR costs are not taken into account in the objective function. It supposes that there are no free carbon emission quotas for participants, such as electricity suppliers and consumers, in the electricity market in this study. That is, all electricity carbon emission costs must be paid based on the carbon price.

Constraints

The proposed EED model should meet the following equality and inequality constraints.

The power grid–related constraints are expressed as follows:

$$\sum_{g=1}^{G} P_{G,g,t} - \sum_{i=1}^{N} P_{D,i,t} = 0 \tag{19}$$

$$\begin{cases} P_{L,l,t} = (\theta_a - \theta_b)/x_l \\ \boldsymbol{\theta} = \boldsymbol{B}^{-1}\boldsymbol{P}_{Inj} \end{cases} \tag{20}$$

$$P^{min}_{L,l} \le P_{L,l,t} \le P^{max}_{L,l} \tag{21}$$

$$P^{min}_{G,g} \le P_{G,g,t} \le P^{max}_{G,g} \tag{22}$$

$$\begin{cases} P_{G,g,t} - P_{G,g,t-1} \le R^{up}_{g} & \text{if } P_{G,g,t} \ge P_{G,g,t-1} \\ P_{G,g,t-1} - P_{G,g,t} \le R^{down}_{g} & \text{if } P_{G,g,t-1} \ge P_{G,g,t} \end{cases} \tag{23}$$

$$0 \le P_{W,g,t} \le P^{pre}_{W,g,t} \tag{24}$$

$$0 \le P_{S,g,t} \le P^{pre}_{S,g,t} \tag{25}$$

The price-based LCDR-related constraints are as follows: Eqs. 15.

$$P_{D,i,t} = P^{0}_{D,i,t} + \Delta P_{D,i,t} \tag{26}$$

$$-D^{max}_{i} \le \Delta P_{D,i,t} \le D^{max}_{i} \tag{27}$$

$$\sum_{t=1}^{T} P^{0}_{D,i,t} = \sum_{t=1}^{T} P_{D,i,t} \tag{28}$$

Constraints (19–25) impose the power grid–related constraints. Equation 19 ensures the power balance, and Eq. 20 calculates the direct current power flow in branch l, where $\boldsymbol{\theta}$ and $\boldsymbol{P}_{Inj}$ are N-dimensional column vectors, and $\boldsymbol{B}$ is an $N \times N$-level matrix. Constraint (21) denotes the capacity limits of the power flow in branches, constraints (22) and (23) denote up and down output limits and generation ramping limits for generators, and constraints (24) and (25) are output limits for wind and solar power plants, respectively.

Equation 15 and constraints (26–28) explain the LCDR-related constraints. Equation 15 calculates the response load $\Delta P_{D,i,t}$ of consumers, which is mainly influenced by three DR signals. $\Delta P_{D,i,t}$ is the controllable variable of the demand side. Constraint (26) indicates the total load of consumers after the LCDR, constraint (27) sets limits on the amount of load shifting, and constraint (28) considers the total demand energy that remains unchanged within 1 day.

In addition, $\rho^{e0}_{i,t}$ is calculated according to the method proposed by Jin et al. (2013) in this study. The hours within a day are divided into day and night. $\rho^{e0}_{i,t}$ is expressed as the demand-weighted average price within all the same type hours of consumers on the ith node before implementing the LCDR:

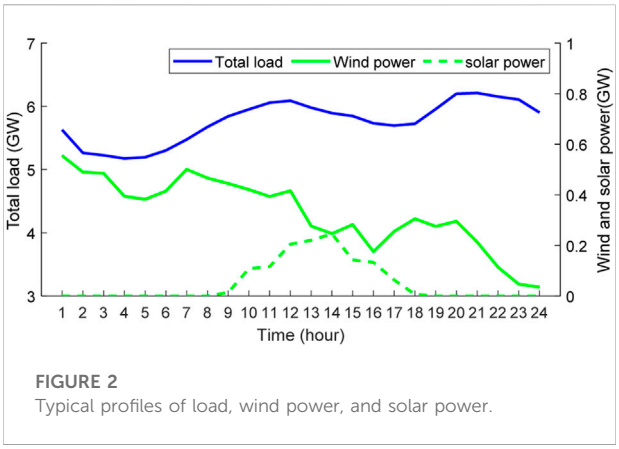


FIGURE 2
Typical profiles of load, wind power, and solar power.

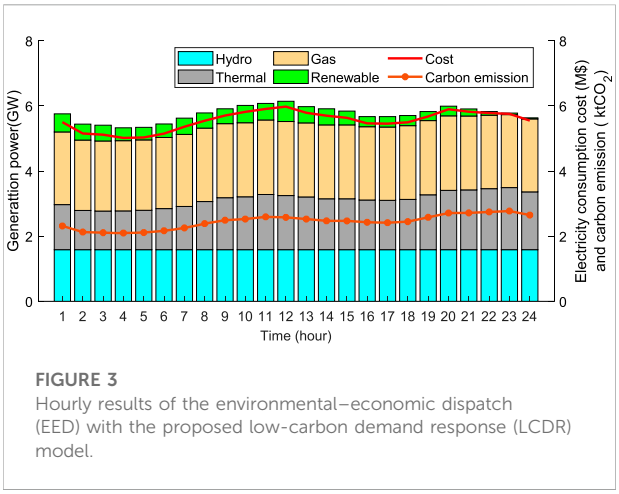


FIGURE 3
Hourly results of the environmental–economic dispatch (EED) with the proposed low-carbon demand response (LCDR) model.

$$\rho_{i,t}^{e0} = \frac{\sum_{t\in T^h} P_{D,i,t}^{0} \rho_{i,t}^{e,Pre}}{\sum_{t\in T^h} P_{D,i,t}^{0}} \tag{29}$$

In parallel, $e_{D,i,t}^{0}$ is also expressed as the demand-weighted CEI without implementing the LCDR:

$$e_{D,i,t}^{0} = \frac{\sum_{t\in T^h} P_{D,i,t}^{0} e_{D,i,t}^{pre}}{\sum_{t\in T^h} P_{D,i,t}^{0}} \tag{30}$$

Furthermore, ρ_t^{c0} is predicted based on day-ahead carbon prices. For simplicity, it is assumed as the closing price of the previous-day carbon market in this study.

Case study

The proposed EED model was simulated on a modified IEEE 39-bus system. Three cases were used to verify the superiority of the proposed price-based LCDR mechanism in the carbon emission reduction:

Case 1: EED model without DR as the benchmark.

Case 2: EED model with PBDR based on the integrated electricity–carbon price.

Case 3: EED model with the proposed price-based LCDR based on the integrated electricity–carbon price and nodal CEI.

The proposed EED optimization model can be established as an MINLP problem and solved using a modified MATLAB-MOST solver (Zimmerman et al., 2011).

Settings of modified IEEE 39-bus system

The modified IEEE 39-bus system is shown in Figure 1, which includes four thermal generators, four gas generators, two hydropower generators, one wind power plant, and one solar power plant.

Table 1 shows the detailed parameters of these generators and power plants, such as the type, rated active power, ramping rate, generation CEI, and short-term generation cost coefficients (U.S. Energy Information Administration, 2020). The wind and solar plants are placed on nodes 4 and 17, respectively, and their predicted output profiles are taken from Lee et al. (2021), as shown in Figure 2.

The hourly load data within a day were selected from an open-source synthetic representation of the Electric Reliability Council of Texas power grid (Wu et al., 2021). It was assumed that the load profiles are of the same typical waveform for each node in this study. The load amount of each node was set to 0.9 times the original data in the IEEE 39-bus benchmark system to better show the effect of the proposed LCDR mechanism. The total demand-side energy remained at 138.31 GW h throughout the studied day.

The values of self-elasticity coefficients of the demand side at different hours were set to −0.2 (Jin et al., 2013). The maximum DR amount of consumers participating in DR programs was both 20%, that is, the value of D_i^{max} was 0.2 p.u. The carbon price was assumed to be a fixed value of 50 \$/tCO_2 within a day. The bidding factor k was set to 1.5. Moreover, the power loss over branches was not considered in this study.

Results and analysis

The hourly results of the EED with the proposed LCDR model are illustrated in Figure 3, which include the hourly generation power composition, electricity consumption costs, and carbon emissions of the power grid. The results shows that terminal and gas generators output more power in the time periods of larger load, and there were higher average costs and ACEIs for the power grid.

TABLE 2 Comparison of electricity consumption costs and carbon emissions of the demand side for cases 1, 2, and 3.

Cases	Electricity consumption costs			Carbon emissions		
	Total (M$)	Unit (price) ($/MW·h)	Reduction (%)	Total (tCO_2)	Unit (ACEI) (tCO_2/MW)	Reduction (%)
1	13.398	96.871	0	59,045	0.4269	0
2	13.378	96.727	0.15	58,947	0.4262	0.16
3	13.327	96.358	0.53	58,669	0.4242	0.63

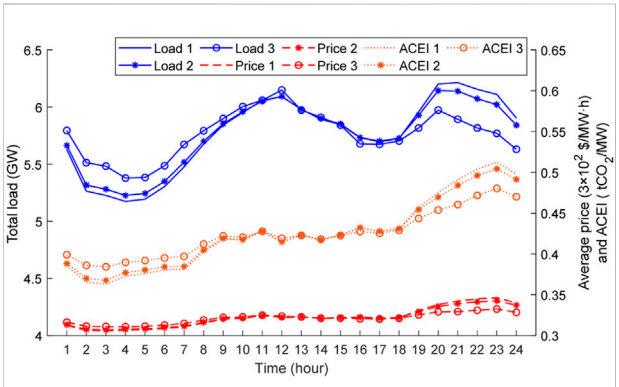


FIGURE 4
Hourly total load, integrated price, and average carbon emission intensity (ACEI) curves for cases 1, 2, and 3

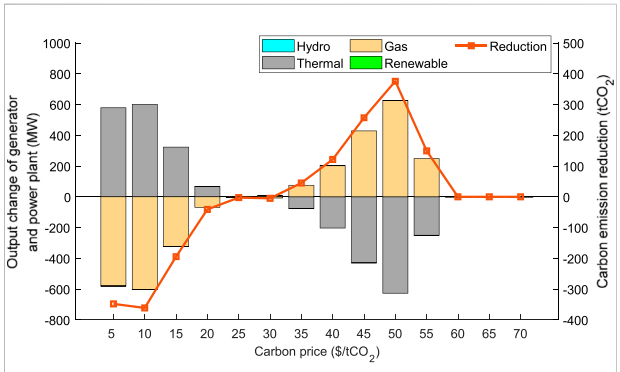


FIGURE 6
Carbon emission reductions and output changes of generators and power plants with different carbon prices.

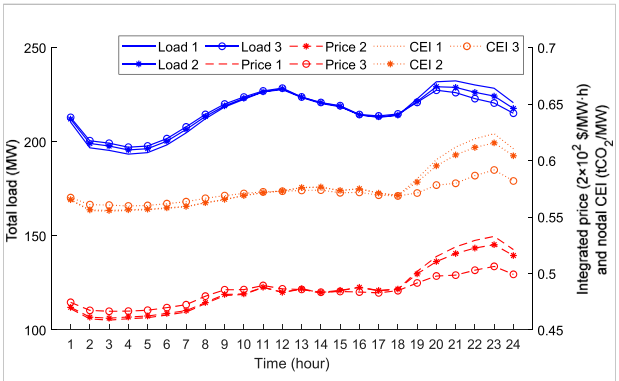


FIGURE 5
Hourly load, integrated price, and nodal carbon emission intensity (CEI) curves of node 7 for cases 1, 2, and 3

The results of these three cases are shown in Table 1. Both the proposed LCDR and the PBDR reduced the electricity consumption costs of the demand side and the carbon emissions of the power grid. The reductions caused by the LCDR were more significant than those caused by the PBDR. In particular, carbon emissions based on the LCDR were reduced by 0.63% (376 tCO2) in a day, which can accumulate into a large reduction amount in a long term. In addition, the electricity consumption costs were reduced by 0.53% with the LCDR, which was also better than the PBDR.

Figure 4 presents the hourly total load, hourly average integrated price of the demand side, and hourly ACEI curves of the power grid for cases 1, 2, and 3. The curve trends of the load, price, and ACEI are basically the same. The price and the ACEI both increase with the increase in load. However, they decrease when the wind and solar powers are high.

It can be observed from the three load curves that consumers mainly respond in the valley periods between 1:00 and 7:00 and the peak periods between 19:00 and 24:00 in case 1. Compared with the PBDR, the LCDR can cause a large range and quantity of load shifting by responding to the integrated price and nodal CEI. The reference values of price and CEI are low from 1:00 to 7:00 and high from 19:00 to 24:00. These lead to large demand changes as consumers pursue lower carbon emissions and electricity consumption costs based on Eq. 15 when implementing the LCDR. For example, the peak time of the load without DR occurs at 21:00, whereas it shifts to 12:00 with LCDR in case 3. Similar LCDR characteristics and demand changes can also be illustrated by the load, integrated price, and nodal CEI curves of node 7, as shown in Figure 5.

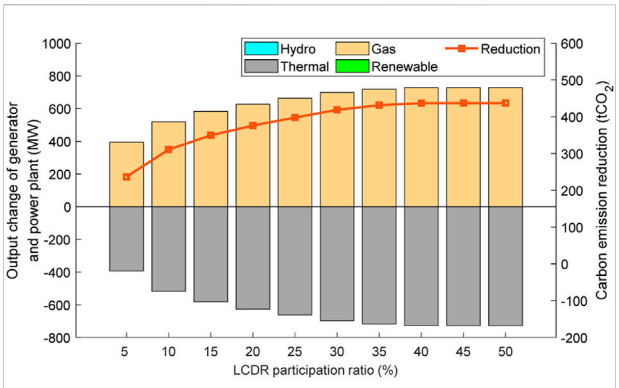


FIGURE 7
Carbon emission reductions and output changes of generators and power plants with different LCDR participation ratios.

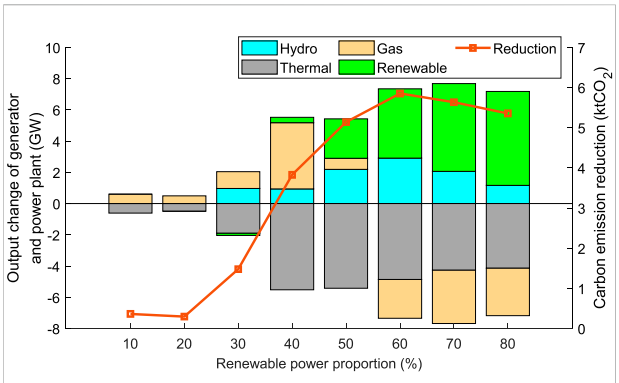


FIGURE 8
Carbon emission reductions and output changes of generators and power plants with different renewable power proportions.

Carbon emission sensitivity analysis

This study also analyzed the applicability of the proposed LCDR mechanism. Three variables that mainly affect the carbon emission reduction ability of the LCDR model were selected for carbon emission sensitivity analysis. The variables included carbon price, DR participation ratio of consumers, and the proportion of renewable power to total power. The carbon price was from \$5 to \$70 at intervals of \$5, the DR participation ratio was from 5 to 50% at intervals of 5%, and the proportion of renewable power was from 10 to 80% at intervals of 10%, respectively.

Carbon emission reductions with the change in the aforementioned three variables of case 3 compared with case 1 are shown in Figure 6, Figure 7, and Figure 8, respectively. Figure 6 demonstrates that carbon emission reduction initially increases with the increase in carbon price, then begins to decrease when carbon price exceeds \$50, and finally becomes 0. It is evident that only appropriate carbon prices can incent consumers to reduce carbon emissions positively when implementing LCDR programs. At low carbon prices, consumers do not pay enough attention to nodal CEI. Load shifting causes power to transfer from low- to high-CEI generators, leading to an increase in carbon emissions. At high carbon prices, carbon costs become the main part of electricity consumption costs, and low-CEI generators have priority to output more power. The power transfer caused by the LCDR is mainly between high-CEI thermal generators; thus, the emission reduction is very negligible.

The relationship between carbon emission reduction and LCDR participation ratio is shown in Figure 7. With the increase in the participation ratio, carbon reduction gradually increases, but the marginal reduction decreases and finally becomes 0. The maximum reduction is 436.83 tCO_2 when the ratio reaches 45%. It shows that only appropriate LCDR participation ratios can effectively reduce carbon emissions. Extensive flexible load participating in the LCDR induces limited reductions. This is because when the reference price and CEI remain unchanged, the maximum DR abilities of consumers are fixed. With the increase in LCDR participation ratios, the abilities of consumers are gradually released until their maximum values are reached and they stop growing.

We added several renewable power plants to analyze the sensitivity of carbon emissions to renewable power proportions. There are three more wind power plants located at nodes 1, 12, and 23 and two more solar power plants located at nodes 8 and 26. The different renewable power proportions were obtained by changing the outputs of all wind and solar plants. Figure 8 demonstrates that carbon emission reductions initially increase and then decrease with the increase in power proportions. This is because when proportions increase at an early period, the power transferred from high- to low-CEI generators or power plants continues to grow, resulting in increased reductions. Then, when proportions become very large, the outputs of renewable plants become very high. On the contrary, the outputs of carbon emission–produced generators continue to decrease, resulting in decreased carbon reductions.

Conclusion

Benefiting from the development of IIoTs in the smart grid, this study proposed a price-based LCDR mechanism based on the integrated electricity–carbon price and CEI signals. Then, the LCDR mechanism was verified using the EED optimization model. Four main conclusions were drawn based on the case study simulation: 1) Compared with traditional PBDRs, appropriate LCDR programs can incent the demand side to positively reduce electricity carbon emissions. 2) The carbon price needs to be suitable for the current stage of the low-carbon

transformation of power grids. Just with appropriate carbon prices can consumers effectively reduce carbon emissions by DR. 3) As the LCDR participation ratio increases, the marginal benefit for carbon emission reduction decreases. 4) With the increase in the renewable power proportion, carbon emission reductions caused by the LCDR initially increase and then decrease.

The proposed quantification method for carbon emission cost conduction in the power grid can be extended to IBDR and other fields in the electricity market. This study ignored the impact of consumers' utility function and also did not consider the DR costs, uncertainty of renewable power output, and change in daily carbon price. Therefore, further research on the LCDR considering the aforementioned issues needs to be conducted.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material; further inquiries can be directed to the corresponding author.

Author contributions

CY conceived the overall structure and framework of the article, proposed the research methods, and wrote the manuscript. BH wrote the code for experiments. HL and JR helped write the manuscript with constructive discussions. JZ provided guidance, supervision, and funding support for the manuscript.

Funding

This work is partially supported by the National Natural Science Foundation of China under Grant 72171206, the Guangdong Regional Joint Fund-Youth Fund under Grant 2021A1515110084, and the Shenzhen Institute of Artificial Intelligence and Robotics for Society.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Albadi, M. H., and El-Saadany, E. F. (2007). *Demand response in electricity markets: An overview*. Tampa, FL, USA, Jul: IEEE Power Engineering Society General Meeting, 1–5.

Chen, Y., Feng, X., Li, Z., Xu, Y., and Miragha, A. (2020). Multi-stage coordinated operation of a multi-energy microgrid with residential demand response under diverse uncertainties. *Energy Convers. Econ.* 1 (1), 20–33. doi:10.1049/enc2.12002

Conejo, A. J., Morales, J. M., and Baringo, L. (2010). Real-time demand response model. *IEEE Trans. Smart Grid* 1 (3), 236–242. doi:10.1109/tsg.2010.2078843

Dahl, K. M., and Petersen, S. (2016). Demand response potential of model predictive control of space heating based on price and carbon dioxide intensity signals. *Energy Build.* 125, 196–204. doi:10.1016/j.enbuild.2016.04.053

Dehnavi, E., and Abdi, H. (2016). Optimal pricing in time of use demand response by integrating with dynamic economic dispatch problem. *Energy* 109, 1086–1094, Aug. doi:10.1016/j.energy.2016.05.024

Depuru, S. S. S. R., Wang, L., Devabhaktuni, V., and Gudi, N. (2011). *Smart meters for power grid — challenges, issues, advantages and status," 2011*. Phoenix, AZ, USA: IEEE/PES Power Systems Conference and Exposition, 1–7.

Dommel, H. W., and Tinney, W. F. (1968). Optimal power flow solutions. *IEEE Trans. Power Apparatus Syst.* (10), 1866–1876. doi:10.1109/tpas.1968.292150

Fleschutz, M., Bohlayer, M., Braun, M., Henze, G., and Murphy, M. D. (2021). The effect of price-based demand response on carbon emissions in European electricity markets: the importance of adequate carbon prices. *Appl. Energy* 295, 117040. doi:10.1016/j.apenergy.2021.117040

Gillenwater, M., and Breidenich, C. (2009). Internalizing carbon costs in electricity markets: using certificates in a load-based emissions trading scheme. *Energy Policy* 37 (1), 290–299. doi:10.1016/j.enpol.2008.08.023

He, L., Lu, Z., Geng, L., Zhang, J., Li, X., Guo, X., et al. (2020). Environmental economic dispatch of integrated regional energy system considering integrated demand response. *Int. J. Electr. Power & Energy Syst.* 116, 105525. doi:10.1016/j.ijepes.2019.105525

Jabr, R. A., Coonick, A. H., and Cory, B. J. (2000). A homogeneous linear programming algorithm for the security constrained economic dispatch problem. *IEEE Trans. Power Syst.* 15 (3), 930–936. doi:10.1109/59.871715

Jin, S., Botterud, A., and Ryan, S. M. (2013). Impact of demand response on thermal generation investment with high wind penetration. *IEEE Trans. Smart Grid* 4 (4), 2374–2383. doi:10.1109/tsg.2013.2278882

Kirschen, D. S., Strbac, G., Cumperayot, P., and de Paiva Mendes, D. (2000). Factoring the elasticity of demand in electricity prices. *IEEE Trans. Power Syst.* 15 (2), 612–617. doi:10.1109/59.867149

Kirschen, D., and Strbac, G. (2004). *Fundamentals of power system economics*. Wiley.

Lee, K., Geng, X., Sivaranjani, S., Xia, B., Ming, H., Shakkottai, S., et al. (2021). Targeted demand response for mitigating price volatility and enhancing grid reliability in synthetic texas electricity markets. *IScience* 25 (2), 103723. doi:10.1016/j.isci.2021.103723

Li, J., Xu, D., Wang, J., Zhou, B., Wang, M., and Zhu, L. (2022). *P2P multi-grade energy trading for heterogeneous distributed energy resources and flexible demand*. IEEE Transactions on Smart Grid. (Early Access). [Online]. Available at: https://ieeexplore.ieee.org/abstract/document/9792462.

Lu, R., and Hong, S. H. (2019). Incentive-based demand response for smart grid with reinforcement learning and deep neural network. *Appl. energy* 236, 937–949. doi:10.1016/j.apenergy.2018.12.061

Newcomer, A., Blumsack, S. A., Apt, J., Lave, L. B., and Morgan, M. G. (2008). Short run effects of a price on carbon dioxide emissions from U.S. electric generators. *Environ. Sci. Technol.* 42 (9), 3139–3144. doi:10.1021/es071749d

Panagiotis, A., Biskas, P., and Liberopoulos, G. (2021). Evaluating the cost of emissions in a pool-based electricity market. *Appl. Energy* 298, 117253. doi:10.1016/j.apenergy.2021.117253

Siano, P. (2014). Demand response and smart grids—a survey. *Renew. Sustain. energy Rev.* 30, 461–478. doi:10.1016/j.rser.2013.10.022

Sisinni, E., Saifullah, A., Han, S., Jennehag, U., and Gidlund, M. (2018). Industrial internet of things: Challenges, opportunities, and directions. *IEEE Trans. Ind. Inf.* 14 (11), 4724–4734, Nov. doi:10.1109/tii.2018.2852491

Song, M., Alvehag, K., Widén, J., and Parisio, A. (2014). Estimating the impacts of demand response by simulating household behaviours under price and CO_2 signals. *Electr. power Syst. Res.* 111, 103–114. doi:10.1016/j.epsr.2014.02.016

Sridhar, S., Hahn, A., and Govindarasu, M. (2012). Cyber–physical system security for the electric power grid. *Proc. IEEE* 100 (1), 210–224. doi:10.1109/jproc.2011.2165269

Stoll, P., Brandt, N., and Nordström, L. (2014). Including dynamic CO_2 intensity with demand response. *Energy Policy* 65, 490–500. doi:10.1016/j.enpol.2013.10.044

United Nation (2021). *Report of the conference of the parties on its twenty- sixth session, held in glasgow from 31 october to 13 november 2021. Part one: Proceedings," UNFCCC. Conference of the parties*. Glasgow, UK: Session and meeting reports.COP 26, Nov

U.S. Energy Information Administration (2020). *Capital cost and performance characteristic estimates for utility scale electric power generating technologies*. Washington, DC: U.S. Department of Energy. [Online]. Available at: https://www.eia.gov/analysis/studies/powerplants/capitalcost/pdf/capital_cost_AEO2020.pdf.

Wang, Y., Qiu, J., Tao, Y., and Zhao, J. (2020). Carbon-oriented operational planning in coupled electricity and emission trading markets. *IEEE Trans. Power Syst.* 35 (4), 3145–3157, Jul. doi:10.1109/tpwrs.2020.2966663

Wook, K., Chattopadhyay, D., and Park, J. (2010). Impact of carbon cost on wholesale electricity price: a note on price pass-through issues. *Energy* 35 (8), 3441–3448, Aug. doi:10.1016/j.energy.2010.04.037

Wu, D., Zheng, X., Xu, Y., Olsen, D., Xia, B., Singh, C., et al. (2021). An open-source extendable model and corrective measure assessment of the 2021 texas power outage. *Adv. Appl. Energy* 4, 100056. doi:10.1016/j.adapen.2021.100056

Wu, L. (2013). Impact of price-based demand response on market clearing and locational marginal prices. *IET Gener. Transm. Distrib.* 7 (10), 1087–1095. doi:10.1049/iet-gtd.2012.0504

Zeng, B., Zhang, J., Yang, X., Wang, J., Dong, J., Zhang, Y., et al. (2014). Integrated planning for transition to low-carbon distribution system with renewable energy generation and demand response. *IEEE Trans. Power Syst.* 29 (3), 1153–1165. doi:10.1109/tpwrs.2013.2291553

Zhang, N., Hu, Z., Dai, D., Dang, S., Yao, M., Zhou, Y., et al. (2016). Unit commitment model in smart grid environment considering carbon emissions trading. *IEEE Trans. Smart Grid* 7 (1), 420–427. doi:10.1109/tsg.2015.2401337

Zhou, T., Chen, Q., Wang, J., Sun, Y., Xia, Q., and Yan, H. (2015). Carbon emission flow from generation to demand: a network-based model. *IEEE Trans. Smart Grid* 6 (5), 2386–2394. doi:10.1109/tsg.2015.2388695

Zimmerman, R. D., Murillo-Sanchez, C. E., and Thomas, R. J. (2011). Matpower: steady-state operations, planning and analysis tools for power systems research and education. *IEEE Trans. Power Syst.* 26 (1), 12–19. doi:10.1109/tpwrs.2010.2051168

Appendix

According to the proposed LCDR mechanism, consumers carry out their low-carbon response based on the nodal CEI. This study quantified the impact of nodal CEI on the demand by translating it into carbon emission cost. Based on Eqs. 8 and 10, the second part of the DR is expressed as follows:

$$dP_{\mathrm{D},t} = \frac{\varepsilon_{t,\tau} P_{\mathrm{D},t}}{\rho_\tau} d\rho_\tau = \frac{\varepsilon_{t,\tau} P_{\mathrm{D},t}}{\rho^{\mathrm{c}}_{\mathrm{D},t}} d\rho^{\mathrm{c}}_{\mathrm{D},t} = \frac{\varepsilon_{t,\tau} P_{\mathrm{D},t}}{e_{\mathrm{D},t}\rho^{c}_{t}} d\left(e_{\mathrm{D},t}\rho^{c}_{t}\right) \tag{A1}$$

By using the multiperiod linear elastic load model, Appendix Eq. A1 is derived as follows:

$$\begin{aligned}
\Delta P''_{\mathrm{D},i,t} &= \sum_{\tau=1}^{T} \varepsilon_{i,t,\tau} \frac{P^0_{\mathrm{D},i,\tau}}{e^0_{\mathrm{D},i,\tau}\rho^{\mathrm{c}0}_{\tau}} \left(\rho^{\mathrm{c}0}_{\tau} de_{\mathrm{D},i,\tau} + e^0_{\mathrm{D},i,\tau} d\rho^{\mathrm{c}}_{\tau}\right) \\
&= \sum_{\tau=1}^{T} \varepsilon_{i,t,\tau} P^0_{\mathrm{D},i,\tau} \left(\frac{\rho^{\mathrm{c}0}_{\tau} de_{\mathrm{D},i,\tau}}{e^0_{\mathrm{D},i,\tau}\rho^{\mathrm{c}0}_{\tau}} + \frac{e^0_{\mathrm{D},i,\tau} d\rho^{\mathrm{c}}_{\tau}}{e^0_{\mathrm{D},i,\tau}\rho^{\mathrm{c}0}_{\tau}}\right) \\
&= \sum_{\tau=1}^{T} \varepsilon_{i,t,\tau} P^0_{\mathrm{D},i,\tau} \left(\frac{e_{\mathrm{D},i,\tau} - e^0_{\mathrm{D},i,\tau}}{e^0_{\mathrm{D},i,\tau}} + \frac{\rho^{\mathrm{c}}_{\tau} - \rho^{\mathrm{c}0}_{\tau}}{\rho^{\mathrm{c}0}_{\tau}}\right) \\
&= \sum_{\tau=1}^{T} \varepsilon_{i,t,\tau} P^0_{\mathrm{D},i,\tau} \left(\frac{e_{\mathrm{D},i,\tau}}{e^0_{\mathrm{D},i,\tau}} + \frac{\rho^{c}_{\tau}}{\rho^{\mathrm{c}0}_{\tau}} - 2\right)\Bigg]
\end{aligned} \tag{A2}$$

Furthermore, omitting the relevant parts of $\varepsilon_{\mathrm{i,t,\tau}}$, the secondary demand deviation responding to nodal CEI can be finally denoted as follows:

$$\Delta P''_{\mathrm{D},i,t} \approx \varepsilon_{i,t,t} P^0_{\mathrm{D},i,t} \left(\frac{e_{\mathrm{D},i,t}}{e^0_{\mathrm{D},i,t}} + \frac{\rho^{c}_{t}}{\rho^{\mathrm{c}0}_{t}} - 2\right) \tag{A3}$$

Nomenclature

Indices and sets

a*, *b Index of two end nodes on branches

g*, *G Index and total number of generators and power plants

***h*,** T^h Index and set of diurnal hour types

i*, *N Index and total number of power grid nodes

k Bidding factor

l*, *L Index and total number of branches

t*, *τ*, *T Index and total number of time periods

Parameters

a_g , b_g, c_g First-, second-, and third-order cost coefficients of generator g

$D_i^{\max}$ Bound of the DR amount for node i

$e_{\mathbf{G}}$, $e_{\mathrm{G},g}$ Carbon emission intensity (CEI) for generators and generator g

$P_{\mathbf{G},g}^{\min}$, $P_{\mathrm{G},g}^{\max}$ Upper and lower output limits for generator g

$P_{\mathbf{L},l}^{\min}$, $P_{\mathrm{L},l}^{\max}$ Upper and lower capacity bounds for branch l

$R_g^{\mathbf{up}}$, R_g^{down} Ramp-up and ramp-down rate limits for generator g

Variables

$C_t^{\mathbf{c}}$ Carbon emission cost at time t

$C_{\mathbf{G},g,t}^{\mathbf{e}}$, $C_{\mathrm{G},g,t}^{\mathrm{c}}$ Fuel and carbon emission costs for generator g at time t

$C_{\mathbf{D},t}^{\mathbf{c}}$ Carbon emission cost for consumers at time t

e_t , $e_{\mathrm{ACEI},t}$ CEI and average CEI (ACEI) of the power grid at time t

$e_{\mathbf{D},t}$, $e_{\mathrm{D},i,t}$ Demand-side CEI for nodes and node i at time t

$e_{\mathbf{D},i,t}^0$, $e_{\mathrm{D},i,\tau}^0$ Reference CEI for node i at time t and τ

$e_{\mathbf{D},i,t}^{\mathbf{pre}}$ Predicted CEI for node i at time t based on EED without DR

E_t , $E_{\mathrm{G},t}$, $E_{\mathrm{D},t}$ Carbon emission, carbon emission for generators, and consumers at time t

P_t Power or load amount at time t

$P_{\mathbf{D},t}$, $P_{\mathrm{D},i,t}$ Demand-side load of nodes and node i at time t

$P_{\mathbf{D},i,t}^0$ Reference load of node i at time t

$P_{\mathbf{G},t}$, $P_{\mathrm{G},g,t}$ Power output of generators and generator g at time t

$P_{\mathbf{L},l,t}$ Power flow in branch l at time t

$P_{\mathbf{W},g,t}$, $P_{\mathrm{S},g,t}$ Output of wind and solar power plant g at time t

$P_{\mathbf{W},g,t}^{\mathbf{pre}}$, $P_{\mathrm{S},g,t}^{\mathrm{pre}}$ Predicted output of wind and solar power plant g at time t

$\Delta P_{\mathbf{D},i,t}$, $\Delta P'_{\mathrm{D},i,t}$, $\Delta P''_{\mathrm{D},i,t}$ Total and first and second parts of the DR amount of node i at time t

x_l Reactance of branch l

$\gamma_{\mathbf{D},i,t}^{\mathbf{c}}$ Carbon price conductivity for node i at time t

$\varepsilon_{t,t}$, $\varepsilon_{t,\tau}$ Self-elasticity and cross-elasticity coefficient at related time t and τ

$\varepsilon_{i,t,\tau}$, $\varepsilon_{i,t,t}$ Self-elasticity and cross-elasticity coefficient for node i at time t and τ

θ_a , θ_b Voltage-phase angle of nodes a and b

ρ_t , ρ_τ Price at time t and τ

$\rho_{\mathbf{D},t}^{\mathbf{c}}$ Conducted carbon price for consumers at time t

$\rho_t^{\mathbf{c}}$, ρ_τ^c Carbon price at time t and τ

$\rho_t^{\mathbf{c0}}$, ρ_τ^{c0} Reference carbon price at time t and τ

$\rho_{i,\tau}^{\mathbf{e}}$, $\rho_{i,t}^{\mathrm{e}}$ Integrated electricity–carbon price of node i at time t and τ

$\rho_{i,t}^{\mathbf{e0}}$, $\rho_{i,\tau}^{\mathrm{e0}}$ Reference integrated price of node i at time t and τ

$\rho_{i,t}^{\mathbf{e,Pre}}$ Predicted integrated price of node i at time t based on EED without DR

Matrices

$\mathbf{B}$ Node admittance matrix for power grids

$\mathbf{e}_{\mathbf{D},t}$ Vector for the CEI of consumers at time t

$\mathbf{e}_{\mathbf{G},t}$ Vector for the CEI of generators at time t

$\mathbf{P'}_{B,t}$ Matrix for branch power outflow distribution at time t

$\mathbf{R}_{\mathbf{C},t}$ Vector for the ejected CEF rate of generators at time t

$\mathbf{P}_{\mathbf{D},t}$ Matrix for the power demand of consumers at time t

$\mathbf{P}_{\mathbf{G},t}$ Matrix for the power output of generators at time t

$\mathbf{P}_{Inj}$ Vector for power injection

$\boldsymbol{\theta}$ Vector for nodal voltage phase

frontiers | Frontiers in Energy Research

TYPE Original Research
PUBLISHED 15 September 2022
DOI 10.3389/fenrg.2022.959333

OPEN ACCESS

EDITED BY
Alberto Benato,
University of Padua, Italy

REVIEWED BY
Jianzhou Wang,
Dongbei University of Finance and Economics, China
Jieming Ma,
Xi'an Jiaotong-Liverpool University, China

*CORRESPONDENCE
Gaoqi Liang,
lianggaoqi@cuhk.edu.cn
Junhua Zhao,
zhaojunhua@cuhk.edu.cn

SPECIALTY SECTION
This article was submitted to Smart Grids,
a section of the journal
Frontiers in Energy Research

RECEIVED 01 June 2022
ACCEPTED 16 August 2022
PUBLISHED 15 September 2022

CITATION
Liu G, Zhang S, Zhao H, Liu J, Liang G, Zhao J and Sun G (2022), Super-resolution perception for wind power forecasting by enhancing historical data. *Front. Energy Res.* 10:959333.
doi: 10.3389/fenrg.2022.959333

Super-resolution perception for wind power forecasting by enhancing historical data

Guolong Liu[1,2,3], Shuwen Zhang[1], Huan Zhao[4], Jinjie Liu[1,4], Gaoqi Liang[1,2]*, Junhua Zhao[1,2]* and Guangzhong Sun[3]

[1]School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, Guangdong, China, [2]Shenzhen Institute of Artificial Intelligence and Robotics for Society, Shenzhen, Guangdong, China, [3]University of Science and Technology of China, Hefei, Anhui, China, [4]Shenzhen Research Institute of Big Data, Shenzhen, China

As an important part of renewable energy, wind power is crucial to the realization of carbon neutrality. It is worth studying on how to accurately predict the wind output so that it can be integrated into the power grid as much as possible to enhance its utilization rate. In this article, a data enhancement method and a framework are proposed to assist wind power forecasting. The proposed method uses the super-resolution perception technology to first detect the completeness and correctness of historical meteorological and wind power data collected by industrial devices. Then, the detected errors are corrected and the missing data are recovered to make the data complete. The frequency of the data is then increased using the proposed method so that the data become complete high-frequency data. Based on the enhanced complete high-frequency data with more detailed characteristics, more accurate forecasts of wind power can be achieved, thereby improving the utilization rate of wind power. Experiments based on public datasets are used to demonstrate the effectiveness of the proposed method and framework. With the proposed method and framework, higher frequency data with more detailed information can be achieved, thereby providing support for accurate wind power prediction that was not possible before.

KEYWORDS

super-resolution perception, SRPWPN, deep learning, short-term wind power forecasting, artificial intelligence

1 Introduction

The emission of greenhouse gases (GHG) leads to the continuous increase in global temperature, and the resulting climate change with frequent extreme weather is attracting attention worldwide (Liu et al., 2022). To cope with global climate change, many countries have set targets to limit carbon dioxide emissions; for example, the European Union, China, and the United States announced plans to achieve carbon neutrality before 2050, 2060, and 2050, respectively. To reduce carbon dioxide emissions, it is necessary to reduce the use of fossil fuels. As an alternative to fossil energy, renewable energy will account for a

substantial increase in the proportion of electricity in the future. Wind power is a vital source of renewable energy and will be vigorously developed. According to the GWEC 2021 of the Global Wind Energy Council, the new global installed capacity in 2020 is 93 GW, with a 53% year-on-year increase rate, and the global installed wind power capacity exceeded 742 GW by 2020 (GWEC, 2021).

Due to the randomness of wind speed, the generation of wind power is intermittent, and its large-scale integration into power grids will bring great challenges to the security and stability of power grids (Wang et al., 2021). To deal with the instability of power generation, a certain proportion of reserve services needs to be configured, and most of the units providing these services are thermal power units (Welikala et al., 2017). Most of the time, these units operate under inefficient conditions, which is not only a huge waste of investment but also brings additional impacts such as more carbon dioxide emissions to the environment. Therefore, how to use historical data to accurately predict wind power forecasting is a very valuable task. Short-term wind power forecasting is to predict the short-term power generation of wind farms, thereby reducing the uncertainty of wind power generation (Khazaei et al., 2022). Therefore, the accurate prediction of short-term wind power is of great significance for increasing the proportion of wind power integration into power grids, ensuring the safety and stability of power grids, and mitigating climate change.

However, accurate short-term wind power forecasting is a difficult problem. Many meteorological factors, such as wind speed, wind direction, and temperature, may affect the generation of wind power. Therefore, the generation of wind power often presents nonlinear uncertainty. There are many time series and classic machine learning methods used for short-term wind power forecasting, including the autoregressive moving average (ARMA) (Gomes and Castro, 2012), autoregressive integrated moving average (ARIMA) (Azimi et al., 2016), seasonal autoregressive integrated moving average (SARIMA) (Liu et al., 2021), and the generalized autoregressive conditional heterosked (GARCH) model (Chen et al., 2018). With the application of machine learning in various fields, some classic machine learning algorithms including support vector regression (SVR) (Ranganayaki and Deepa, 2017), classification and regression tree (CART) (González et al., 2015), and Gaussian process regression (GPR) (Fang and Chiang, 2016) are also used for short-term wind power forecasting. Due to the limitation of their learning ability, these methods cannot meet the requirements of achieving high-frequency accurate wind power forecasting. With their strong nonlinear learning ability, deep neural networks have achieved remarkable results in image processing, speech recognition, natural language processing, *etc*. There are also some studies on short-term wind power forecasting, such as deep belief network (DBN) (Wang et al., 2018), recurrent neural network (RNN) (Kumar et al., 2021), long short-term memory (LSTM) (Shahid et al., 2020), convolutional neural network (CNN) (Yu et al., 2020), semi-supervised generative adversarial network (SSGAN) (Zhou et al., 2021), and spatiotemporal attention networks (SAN) (Fu et al., 2019). Among them, the SSGAN has better generalization ability than other deep neural networks, and it is a semi-supervised learning method, which requires less data and is more practical.

Existing studies are based on data collected at the inherent sampling frequency of terminal devices such as Internet-of-things (IoT) devices and do not consider the utilization of higher-frequency complete data with more detailed information. Higher-frequency complete data can not only provide more detailed information for accurate wind power forecasting but also help achieve a shorter forecast period, thereby facilitating wind power to be efficiently integrated into power grids. Collecting high-frequency data is a challenging task with several problems. First, collecting high-frequency data requires the installation of high-frequency meters to replace the installed low-frequency meters, which will be an additional investment and a resource waste. Second, the transmission of high-frequency data requires a large amount of bandwidth, which requires upgrading the existing communication network to improve the data transmission capacity. Third, even if high-frequency data are collected and transmitted to where it is needed (such as a data center), massive storage space is required to store the data. Therefore, it is a more practical solution to recover high-frequency data from existing low-frequency data. There are some studies on improving the frequency of data, such as linear interpolation, binary interpolation, ARIMA, and backpropagation-artificial neural network (BP-ANN), but they have problems such as large errors, poor quality, or low computational efficiency (Liu et al., 2020; Liang et al., 2020). SRP is a technology that uses advanced artificial intelligence technologies to recover low-frequency incomplete data to obtain high-frequency complete data; its effectiveness has been verified in many areas. For example, in the study by Liang et al. (2020), super-resolution perception net for state estimation (SRPNSE) is proposed for improving data completeness in smart grid state estimation. A super-resolution perception convolutional neural network (SRPCNN) is proposed in the study by Liu et al. (2020) to generate high-frequency load data from low-frequency data collected by smart meters. A monthly-super-resolution perception convolutional neural network (M-SRPCNN) is proposed in the study by de-Paz-Centeno et al. (2021) to up-sample monthly energy consumption measured at hourly resolution. Compared with other data quality improvement methods, the SRP method has the advantages of higher efficiency, better quality, and richer information. Therefore, the application of SRP technologies to enhance wind power forecasting is a very valuable research topic.

Against this background, super-resolution perception wind power net (SRPWPN) is proposed based on SRP in this article to enhance the completeness and frequency of data used in wind power forecasting. The proposed method combines the attention mechanism and residual network, which can provide accurate data support for wind power forecasting. The main contributions of this article are described as follows:

1) This article is the first to formalize the SRP problem for short-term wind power forecasting and propose a corresponding framework to solve the proposed problem.
2) This article is the first to propose the SRPWPN to improve the quality of historical data to obtain complete high-frequency data, thereby providing more detailed information for more accurate and higher-frequency wind power forecasting. The SRPWPN provides more detailed information and higher quality data for the current short-term wind power forecasting task.
3) Based on public datasets, the effectiveness of the proposed method and framework is verified using experiments, demonstrating their data enhancement capability for short-term wind power forecasting. In addition, experiments also demonstrate that the proposed method can provide complete data with higher frequency, more detailed information, and higher quality.

The rest of this article is organized as follows: Section 2 formulates the problems of SRP for short-term wind power forecasting and short-term wind power forecasting. The methodology is introduced in Section 3. Experiments and numerical simulation results are reported and discussed in Section 4. Finally, Section 5 concludes this article.

2 Problem formulation

2.1 Problem formulation of SRP for short-term wind power forecasting

The purpose of SRP for wind power forecasting is to recover high-frequency complete data from the incomplete low-frequency data collected by terminal devices, thereby supporting more accurate wind power forecasting. The historical incomplete low-frequency data ILF includes features such as wind power, wind direction, wind speed, temperature, pressure, and density, which are expressed as follows:

$$ILF = \{ilf_0^0, ilf_0^1, \ldots, ilf_0^n, ilf_1^0, ilf_1^1, \ldots, ilf_1^n, \ldots, ilf_t^n\}, \quad (1)$$

where t represents the time index and n represents the number of features. Compared with ILF, complete high-frequency data CHF are more densely indexed in temporal dimension. The relationship between complete high-frequency data CHF and incomplete low-frequency data ILF is expressed as follows:

$$ILF = \downarrow_\alpha CHF + e, \quad (2)$$

where $\downarrow_\alpha$ represents the degradation function, α is the down-sampling factor, and e represents noise caused by sampling devices. The goal of SRP is to find a function $f(\cdot)$ such that its output $\widehat{CHF}$ is as close as possible to the complete high-frequency data CHF, which can be expressed as follows:

$$\widehat{CHF} = f(ILF) = \uparrow_\beta CLF \quad (3)$$

,

where $\uparrow_\beta$ is the SRP function which is implemented by a deep neural network and β is the SRP factor. For example, given incomplete low-frequency data ILF with a sampling interval of 15 min, when the SRP factor β is three, SRP can obtain complete high-frequency data $\widehat{CHF}$ with a sampling interval of 5 min.

2.2 Problem formulation of short-term wind power forecasting

Short-term wind power forecasting is based on historical data to predict wind power generation for some time in the future, where the historical data X with t time steps and n features can be expressed as follows:

$$X = \{x_0^0, x_0^1, \ldots, x_0^n, x_1^0, x_1^1, \ldots, x_1^n, \ldots, x_t^n\}. \quad (4)$$

The wind power forecasting task for k time steps in the future at time t can be expressed as follows:

$$\hat{Y} = g(X|\theta) + \varepsilon \quad (5)$$

,

where $\hat{Y}$ is the predicted wind power generation with k time steps in the future at time t, $g(\cdot|\theta)$ represents the relationship function described by the model parameter set θ that uses historical data to predict future wind power generation, and ε is the forecasting error. Therefore, the goal of model $g(\cdot|\theta)$ is to make prediction result $\hat{Y}$ as close as possible to actual data Y. Actual data Y are expressed as follows:

$$Y = \{x_{t+1}^p, x_{t+2}^p, \ldots, x_{t+k}^p\}. \quad (6)$$

3 Methodology

3.1 The framework of short-term wind power forecasting with SRP

The goal of SRP is to enhance the frequency and quality of historical data to achieve more accurate short-term wind power forecasting, so a framework for short-term wind power forecasting based on SRP is proposed to achieve the above goal. The framework is shown in Figure 1. First, the historical data containing six features such as wind speed and wind power are used as the input of the data preprocessing part as incomplete low-frequency data. Those missing values, duplicate values,

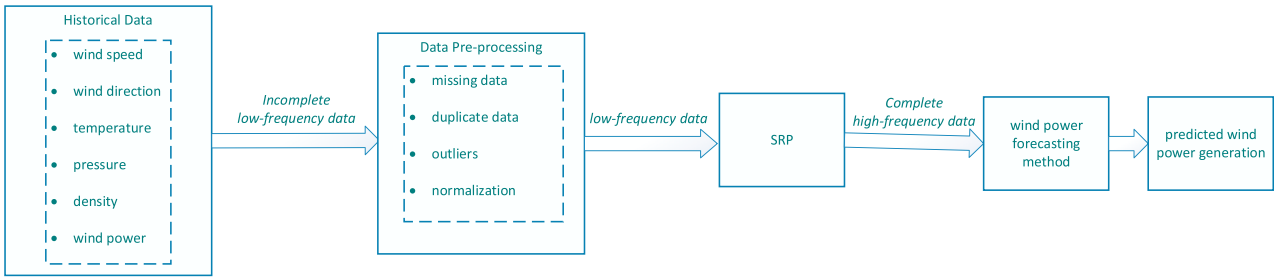


FIGURE 1
Framework of short-term wind power forecasting with SRP.

outliers, *etc*., are all processed, and the data are also normalized to obtain low-frequency data. Then the low-frequency data are used as the input of the SRP part, and the complete high-frequency data are obtained by enhancing the historical data through the SRP part. Then based on the complete high-frequency data obtained by SRP, the wind power forecasting method can predict the target wind power generation.

3.2 Super-resolution perception wind power net

SRPWPN is proposed to enhance the historical data, and its structure is shown in Figure 2. Data $CHF\downarrow\uparrow$ represent the data obtained by data CHF first through bicubic down-sampling and then bicubic up-sampling. $ILF\uparrow$ represents the data up-sampled by the bicubic function. Incomplete low-frequency data ILF are first extracted by three two-dimensional (2D) convolutional layers to achieve feature extraction and then used as the input for the next 16 super-resolution perception blocks (SRPB). In SRPB, the data are used as the input of the 2D convolutional layer, and then the corresponding output is normalized, and then the above process is repeated. F represents the identity mapping, which is added to the previous calculation results, and the rectified linear unit (ReLU) function is used for activation (Agarap, 2018).

CHF, $CHF\downarrow\uparrow$, and $ILF\uparrow$ are mapped to V, K, and Q by three 2D convolutional layers, respectively. To calculate the similarity between Q and K, Q and K are first sliced into patches denoted as q_i and k_j, then these patches are normalized by Eqs 7, 8, respectively. The similarity between q_i and k_j is obtained by Eq. 9 to form the correlation $Similarity_{ij}$. For the hard attention part, element b_i of hard attention map B is calculated by Eq. 10, and then used together with V as the input of the hard attention operation. In the hard attention operation, the index selection operation is shown in Eq. 11, where d_i is the element of D and v_i is the element of V. After the hard attention operation, D and output A of SRPB are concatenated as the input E of soft attention operation. In the soft attention part, element c_i of the hard attention map C is calculated by Eq. 13, and then also used as the input of the soft attention operation. Then, the convolution operation is performed on E to get the result E_{cov}. The results E_{cov} and C are element-wise multiplicated directly through Eq. 13, where $\odot$ represents the element-wise multiplication. Finally, EC and A are added together to get the final output $\widehat{CHF}$.

$$q_i^{norm} = \frac{q_i}{\|q_i\|} \tag{7}$$

$$k_j^{norm} = \frac{k_j}{\|k_j\|} \tag{8}$$

$$Similarity_{ij} = \left\langle \frac{q_i}{\|q_i\|}, \frac{k_j}{\|k_j\|} \right\rangle \tag{9}$$

$$b_i = \underset{j}{\operatorname{argmax}}\, Similarity_{ij} \tag{10}$$

$$d_i = v_{b_i} \tag{11}$$

$$c_i = \max_j Similarity_{ij} \tag{12}$$

$$EC = E_{cov} \odot C \tag{13}$$

4 Experiments

The experiments are divided into two parts: one part is to use SRPWPN to enhance historical data, and the other part is to perform short-term wind power forecasting on the recovered data and original data. For the first part, the following three experiments were performed:

1) The sampling interval of incomplete low-frequency data is 15 min, and the SRP factor β is three.
2) The sampling interval of incomplete low-frequency data is 10 min, and the SRP factor β is two.
3) The sampling interval of incomplete low-frequency data is 1 h, and the SRP factor β is six.

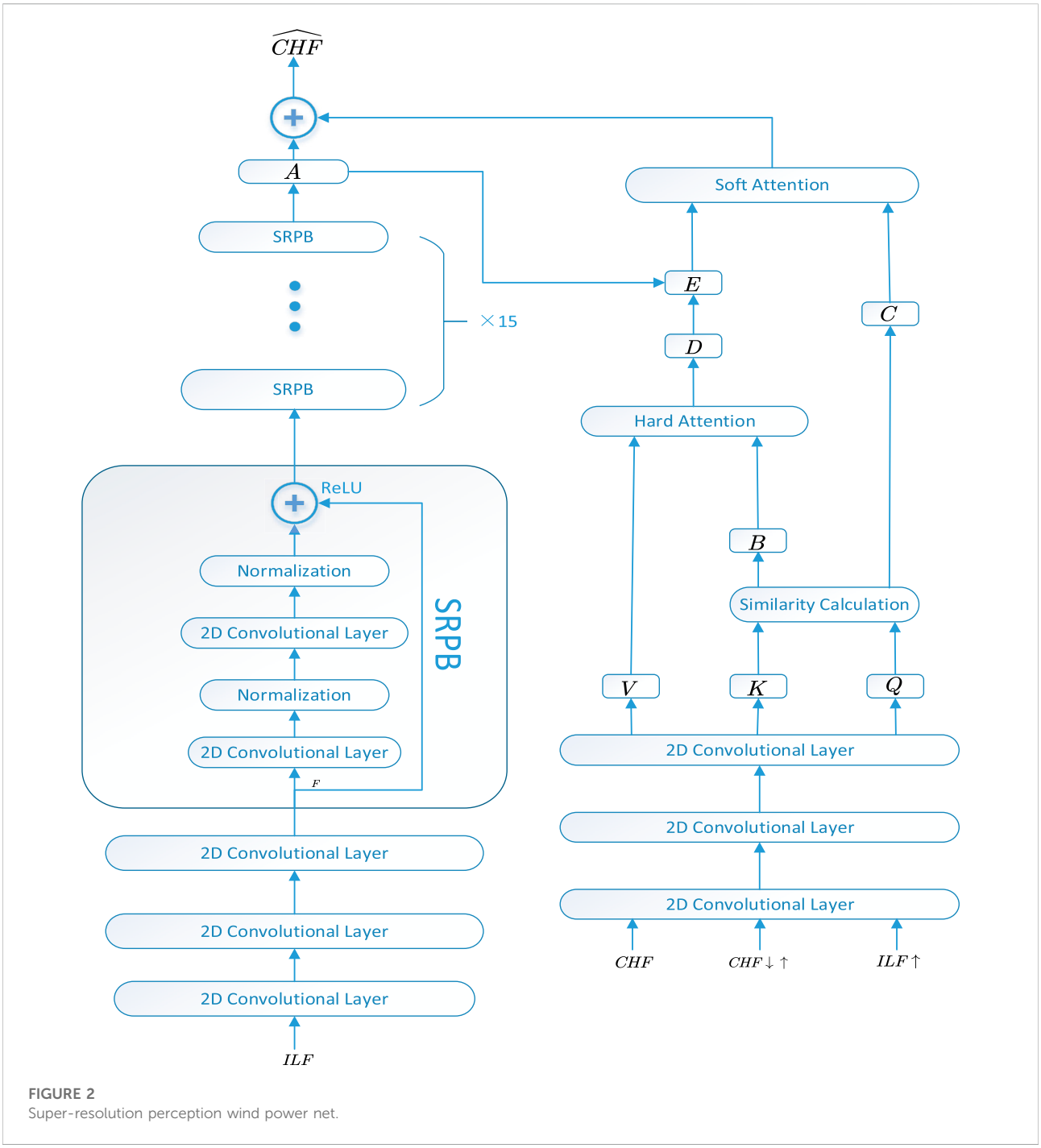


FIGURE 2
Super-resolution perception wind power net.

Experiments Eqs 1, 2 are conducted on the dataset National Renewable Energy Laboratory (NREL), and experiment (3) is conducted on the dataset The La Haute Borne (TLHB). For the second part, three short-term wind power forecasting methods, CNN, LSTM, and SSGAN, perform short-term wind power forecasting on the complete high-frequency data recovered by SRP and the original high-frequency data, respectively.

4.1 Dataset description

Two datasets were used in the experiments. The first dataset is from the National Renewable Energy Laboratory (NREL). There are six wind farms used for the experiments in NREL, their site IDs are 123229, 123815, 123978, 124043, 124044, and 124045, respectively. NREL contains meteorological information such as wind direction, wind speed, temperature, pressure, and density, as

TABLE 1 Information of six wind farms.

Index	Site ID	Latitude	Longitude	Capacity	Capacity factor
1	123229	48.716766	−101.827454	16	0.437
2	123815	48.870552	−101.73111	16	0.426
3	123978	48.895412	−101.98913	16	0.435
4	124043	48.91811	−101.90655	16	0.388
5	124044	48.91947	−101.87832	16	0.398
6	124045	48.92083	−101.85009	16	0.403

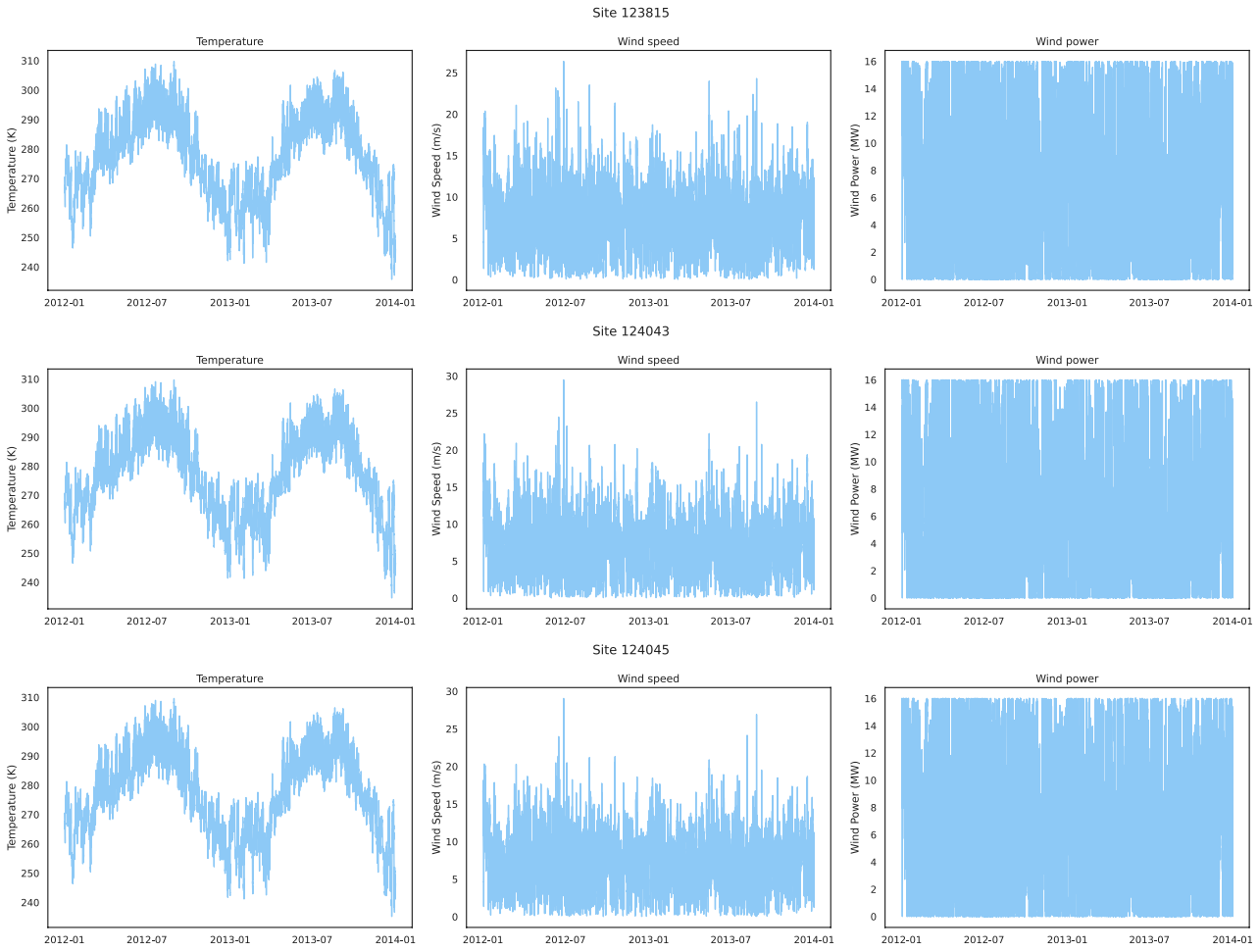


FIGURE 3
Graphs of historical data for three different wind farms in NREL.

well as wind power for each wind farm. The time range of the data is from 2012-01-01 00:00:00 to 2013-12-31 23:55:00, where the time interval between two consecutive points is 5 min. The abstract information of the used six wind farms is shown in Table 1, where capacity factor represents the average power output divided by the wind turbine's maximum power capability. The second dataset is the La Haute Borne (TLHB) wind farm, which is located in the Grand Est of northeastern France. There are four wind turbines in TLHB, their site IDs are R80711, R80721, R80736, and R80790, respectively. Other variables, such as wind speed, wind direction, and temperature, are also included in this dataset. The time range of the data is from

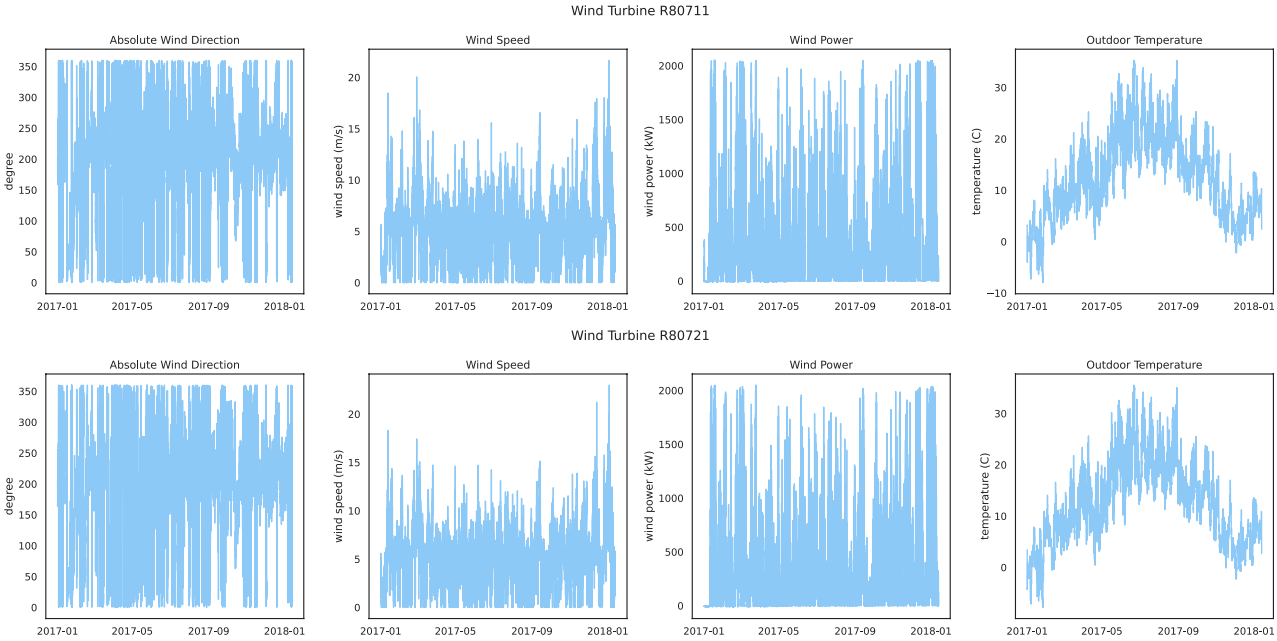


FIGURE 4
Graphs of historical data for two different wind farms in TLHB.

TABLE 2 Experimental results of different methods on NREL (MAPE on 15 min/10 min).

Site ID	LI	BI	ARIMA	BP-ANN	SRPCNN	SRPWPN
123229	27.92%/26.35%	31.83%/29.13%	26.63%/24.41%	22.72%/21.95%	9.88%/6.37%	**3.26%/2.04%**
123815	26.15%/25.03%	30.95%/27.84%	22.95%/20.19%	21.03%/19.94%	10.27%/8.52%	**3.34%/2.10%**
123978	25.98%/24.59%	28.72%/24.72%	24.86%/21.74%	22.65%/21.59%	9.15%/6.76%	**3.23%/1.93%**
124043	28.13%/26.87%	29.70%/25.28%	23.43%/21.26%	21.46%/20.73%	10.53%/7.52%	**3.29%/2.08%**
124044	26.34%/23.05%	30.18%/26.31%	25.48%/22.83%	20.87%/18.65%	8.84%/7.91%	**3.35%/1.97%**
124045	24.82%/21.86%	31.99%/28.63%	22.19%/20.04%	21.32%/19.48%	9.77%/7.15%	**3.20%/1.99%**

The bold characters is to emphasis the results which indicates the results are better than other methods.

2017-01-01 00:00:00 to 2018-01-13 00:00:00, where the time interval between two consecutive points is 10 min.

4.2 Exploratory data analysis

For NREL, three wind farms with site IDs of 123815, 124043, and 124045 are selected for exploratory data analysis, where temperature, wind speed, and wind power are selected as represented features for visualization. Figure 3 shows the historical data graphs of the represented features for the three wind farms. It can be seen that the temperature has obvious periodicity due to seasonal changes, while the wind speed does not have a similar periodicity as the temperature. The wind power related to the strong wind speed has almost the same pattern in every season. In addition, the maximum value of wind power is 16 MW, even if the wind speed does not reach the maximum value at the corresponding time. The reason is that when the wind speed exceeds the rated wind speed of the wind turbine and is less than the cut-out wind speed, and the wind turbine will generate constant power at the rated power. The wind power is zero at certain time points because the wind speed at those time points is lower than the cut-in wind speed of the wind turbine. For TLHB, two wind farms with site IDs R80711 and R80721 are selected, where absolute wind direction, wind speed, wind power, and outdoor temperature are selected represented features. The historical data of the represented features are shown in Figure 4. The wind power of the two wind farms is between 0 and 2,000 kW, and is not significantly affected by the season. Therefore, periodicity is not considered in the experiments, but these features are used directly.

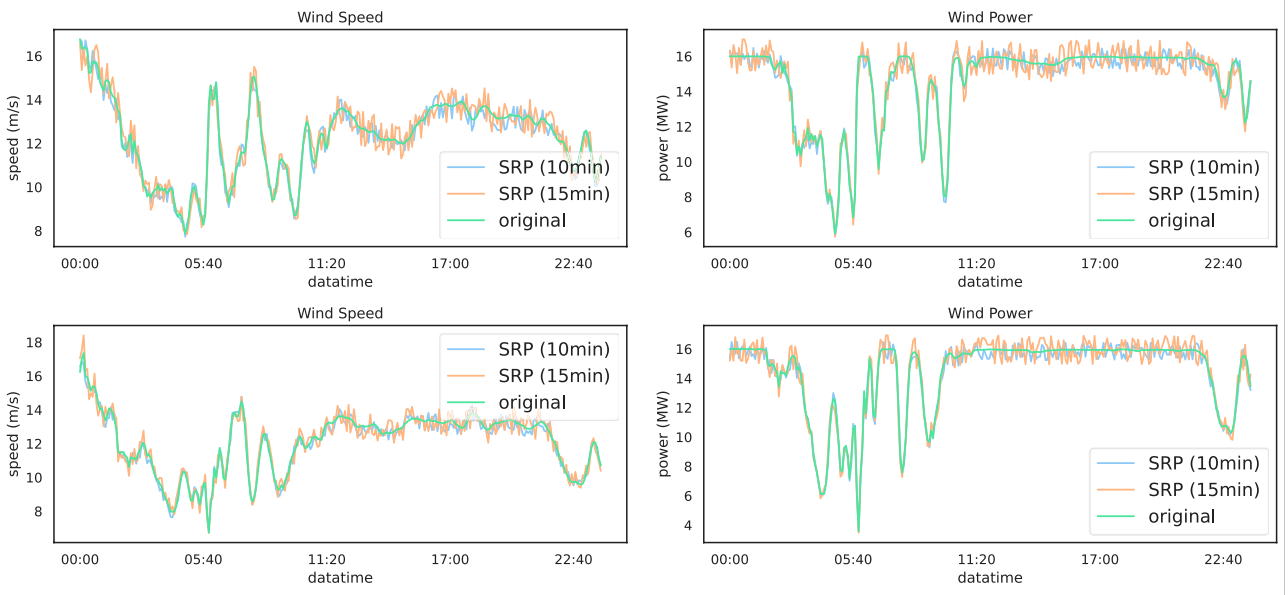


FIGURE 5
SRPWPN on historical data with site IDs 123229 and 123815 in NREL.

TABLE 3 Experimental results of different methods on TLHB.

Site ID	LI (%)	BI (%)	ARIMA (%)	BP-ANN (%)	SRPCNN (%)	SRPWPN (%)
R80711	35.58	42.47	29.05	25.82	15.49	**8.08**
R80721	36.04	41.53	30.82	26.93	16.85	**7.75**
R80736	34.65	39.49	31.37	24.18	14.64	**8.49**
R80790	33.60	38.35	28.64	25.04	15.06	**7.37**

The bold characters is to emphasis the results which indicates the results are better than other methods.

FIGURE 6
SRPWPN on historical data with site IDs R80711 and R80721 in TLHB.

TABLE 4 MAPE results of short-term wind power forecasting on TLHB.

Index	Site ID	CNN (%)	CNN_SRP (%)	LSTM (%)	LSTM_SRP (%)	SSGAN (%)	SSGAN_SRP (%)
1	123229	6.59	7.52	4.18	5.27	3.28	4.67
2	123815	6.61	7.75	4.09	5.09	3.24	4.93
3	123978	6.45	7.50	4.18	5.18	3.31	4.74
4	124043	6.35	7.62	4.37	5.06	3.15	4.73
5	124044	6.63	7.47	4.29	5.16	3.25	4.98
6	124045	6.47	7.45	4.39	4.61	3.59	4.98

TABLE 5 MAPE results of short-term wind power forecasting on NREL.

Index	Site ID	CNN (%)	CNN_SRP (%)	LSTM (%)	LSTM_SRP (%)	SSGAN (%)	SSGAN_SRP (%)
1	R80711	8.54	8.93	6.49	6.96	5.47	6.14
2	R80721	9.73	10.92	7.82	9.24	5.03	6.32
3	R80736	8.39	9.85	7.93	8.59	6.71	5.85
4	R80790	9.04	10.47	6.51	7.88	5.23	5.96

4.3 Data preprocessing

The historical data in NERL are down-sampled by Eq. 14 to obtain data with sampling intervals of 15 and 10 min, respectively. x_{hf} is the original high-frequency data, and the down-sampling factors α are three and two, respectively. Similarly, for TLHB, the sampling interval of down-sampled data is 1 h and the down-sampling factor α is six. Then the data are checked for missing data, and if there are missing data, then they are filled with the mean of the corresponding feature. Next, duplicate data are checked according to timestamps, and deleted if they exist. Then, each feature is detected and processed for outliers according to the method in reference (Liu et al., 2022). After that, each feature in the historical data is normalized by Eq. 15, where x^{norm} denotes the normalized result, x include the original data, x_{min} denotes the minimum value of this feature, and x_{max} denotes the maximum value of this feature.

$$x = \downarrow_{\alpha} x_{hf} \tag{14}$$

$$x^{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{15}$$

4.4 Experimental results of SRP for enhancing historical data

SRPWPN performs SRP on historical data from NREL with sampling intervals of 15 and 10 min, resulting in complete high-frequency data with a sampling interval of 5 min after enhancement. Similarly, SRP recovers data with a sampling frequency of 1 hour in TLHB to obtain data with a frequency of 10 min; 75% of the historical data is used for model training, 5% is used for model validation, and 20% is used to evaluate model performance. The loss function used by SRPWPN is defined as follows:

$$loss = \left\| y - \hat{y} \right\|_2^2, \tag{16}$$

where y represents the real data and $\hat{y}$ represents the recovered data by SRPWPN. For model training of SRPWPN, Adam was chosen as the optimizing algorithm (Kingma and Ba, 2014). The mean absolute percentage error (MAPE) is used as the metric for evaluating the performance of SRPWPN, which is shown as follows:

$$MAPE = \frac{1}{ND} \sum_{nd=1}^{ND} \left| \frac{y_{nd} - \hat{y}_{nd}}{y_{nd}} \right| \times 100\%, \tag{16a}$$

where ND denotes the number of data points, y_{nd} is the real value of nd-th data point, and $\hat{y}_{nd}$ is the recovered value of nd-th data point. To verify the experimental effect of SRPWPN, linear interpolation (LI), binary interpolation (BI), ARIMA, BP-ANN, and SRPCNN are added to the experiment as comparative methods.

The experimental results of SRPWPN on NREL are shown in Table 2. It can be seen that SRPWPN has a good performance on the historical data of the six wind farms in NREL, and the minimum MAPE is 1.93%, which means that the minimum error does not exceed 2%. In addition, the maximum MAPE is only 3.35%, which means that SRPWPN has very stable and excellent performance in enhancing historical data. The performance of SRPWPN on data with a sampling interval of 10 min is better than that with

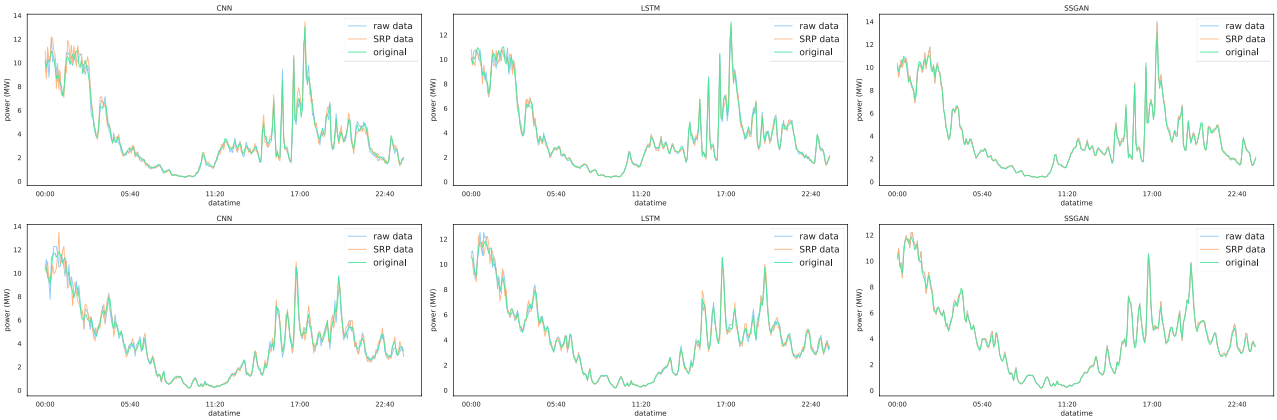


FIGURE 7
SSGAN on site IDs 123229 and 123815 in NREL.

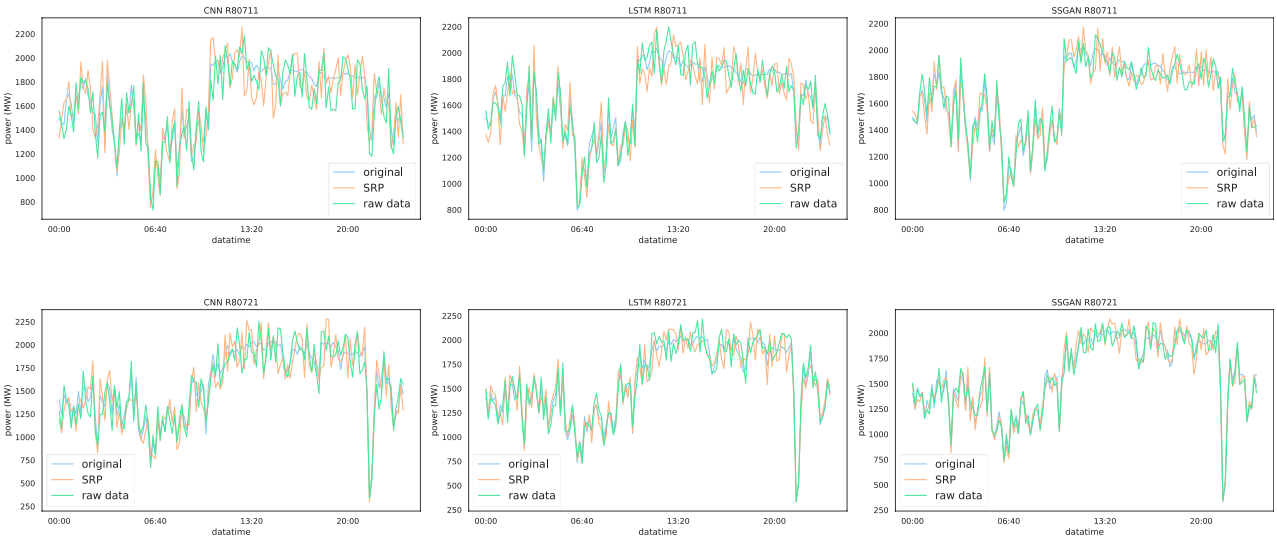


FIGURE 8
SSGAN on site IDs R80711 and R80721 in TLHB.

a sampling interval of 15 min, which means that a larger SRP factor needs to recover more detailed information and is more challenging. Figure 5 shows the actual effect of SRPWPN on the historical data of NREL with site IDs 123229 and 123815. The subfigures in the first row show the SRP effect of wind speed and wind power with site ID 123229, and the second row shows the SRP effect of wind speed and wind power with site ID 123815. It can be seen that most of the detailed information is recovered in SRPWPN, and the experiments with a smaller SRP factor have better results. Table 3 shows the experimental results of SRPWPN on TLHB. The minimum MAPE is 7.37%, which means that the minimum error does not exceed 7.5% and is slightly worse than the result on NREL. Since the maximum MAPE does not exceed 8.5%, this proves that SRPWPN has a similarly excellent performance. The actual effect of SRPWPN on TLHB with site IDs R80711 and R80721 is shown in Figure 6. The subfigures in the first row show the SRP effect of absolute wind direction, wind speed, wind power, and outdoor temperature with site ID R80711, and the second row shows the SRP effect of wind speed and wind power with site ID R80721. When the actual value fluctuates greatly, SRPWPN can also learn its internal

relationship well and reconstruct the corresponding information accurately. Taking the wind power of R80711 as an example, although its value decreased from near 1,900 KW to around 300 KW at around 10 pm, SRPWPN still recovered it accurately. It can be seen that most of the detailed information is recovered in SRPWPN.

4.5 Experimental results of short-term wind power forecasting with super-resolution perception wind power net

The experimental results of different short-term wind power forecasting methods on real high-frequency data and high-frequency data recovered by SRPWPN are compared to demonstrate that SRPWPN can provide almost the same information as high-frequency data. To verify the effectiveness of the complete high-frequency data recovered by SRPWPN, three short-term wind power forecasting methods including CNN, LSTM, and SSGAN are used in the experiments. For NREL, the three methods perform short-term wind forecasts on complete high-frequency data with a real sampling interval of 5 minutes and data with a sampling interval of 5 minutes recovered by using SRPWPN from incomplete low-frequency data with a sampling interval of 10 min. For TLHB, the three methods perform short-term wind forecasts on complete high-frequency data with a real sampling interval of 10 min and on data with a sampling interval of 10 min recovered by using SRPWPN from incomplete low-frequency data with a sampling interval of 1 h. In the experiments, the historical data of the past 7 days are used to predict the wind power generation of the next day. MAPE is used as the evaluation metric, and the experimental results are shown in Tables 4, 5. The column with the suffix SRP in the table represents the error of the short-term wind power forecasting results of the prediction method on the data recovered from the SRPWPN. Although the results of the three methods on the data recovered by SRPWPN are slightly inferior to the actual data, the biggest difference is not more than 2%. The worst of the three methods on the data recovered by SRPWPN is CNN, whose MAPE does not exceed 11% in TLHB, and the best is SSGAN, whose MAPE does not exceed 7% in the two datasets. It can be considered that the data recovered by SRPWPN are very close to the effect of real data in practical applications. As the best short-term wind power forecasting method, SSGAN is selected as the visualization method. CNN and LSTM are chosen as the comparative methods. Figure 7 shows SSGAN on the raw historical data and data recovered by SRPWPN with site IDs 123229 and 123815 in NREL. The subfigures in the first row show the performance of SSGAN on site ID 123229, and the second row shows the performance of SSGAN on site ID 123815. Original represents the real wind power of the two sites, SRP data represent the predicted results of SSGAN on complete high-frequency data recovered by SRPWPN, and raw data represent the results of SSGAN on the real data with a sampling interval of 5 minutes. Figure 8 shows SSGAN on the raw historical data and data recovered by SRPWPN with site IDs R80711 and R80721 in TLHB. The subfigures in the first row show the performance of SSGAN on site ID R80711, and the second row shows the performance of SSGAN on site ID R80721. Original represents the real wind power of the two sites, SRP data represent the predicted results of SSGAN on complete high-frequency data recovered by SRPWPN and raw data represent the results of SSGAN on the real data with a sampling interval of 10 min. The results predicted by SSGAN are almost the same as the actual results, which proves that the data recovered by SRPWPN can be well utilized by short-term wind power forecasting methods, thereby achieving higher frequency accurate wind power forecasting, and the recovered information is sufficient for short-term wind power forecasting to use.

5 Conclusion

A deep SRP network named SRPWPN is proposed for short-term wind power forecasting, and an SRP-based short-term wind power forecasting framework is proposed to achieve accurate short-term wind power forecasting at higher frequencies. In the proposed framework, SRPWPN is used to recover complete high-frequency data from incomplete low-frequency data, thereby enhancing historical data and then using short-term wind power forecasting methods on the enhanced data to achieve higher-frequency accurate short-term wind power forecasting. In the SRPWPN, incomplete low-frequency historical data are used as the input, and then detailed information is reconstructed through structures such as attention mechanism and SRPB, thereby obtaining complete high-frequency data for higher-frequency short-term wind power forecasting. The experimental results show that the SRPWPN can recover the most lost detailed information, and its maximum MAPE does not exceed 8.5%. The maximum MAPE of short-term wind power forecasting experiments on the complete high-frequency data recovered based on SRPWPN is less than 11%. The MAPE of the best short-term wind power forecasting method SSGAN is less than 6.5%, which proves the effectiveness of the proposed SRP-based wind power forecasting framework. Regardless of whether the sampling frequency is 10 min, 15 min, or 1 h of historical data, the proposed method enhances the data well. With the proposed data enhancement method, SRPWPN and SRP-based short-term wind power forecasting framework, accurate short-term wind power forecasting at higher frequencies can be achieved.

Data availability statement

Publicly available datasets were analyzed in this study. These data can be found here: https://www.nrel.gov/grid/eastern-western-wind-data.html.

Author contributions

GL: experiments, research methods, and writing the original draft; SZ: data processing, experiments, and graphic drawing; HZ: research methods and literature review. JL: resources and guide experiments. GL: supervision and project administration. JZ: project administration and fund acquisition. GS: project administration.

Funding

This work was supported in part by the National Natural Science Foundation of China (Key Program 71931003, 72061147004, 72171206, 72192805, and 42105145) and in part by the Shenzhen Institute of Artificial Intelligence and Robotics for Society.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Agarap, A. F. (2018). Deep learning using rectified linear units (relu). *arXiv preprint arXiv:1803.08375*. Available at:

Azimi, R., Ghofrani, M., and Ghayekhloo, M. (2016). A hybrid wind power forecasting model based on data mining and wavelets analysis. *Energy Convers. Manag.* 127, 208–225. doi:10.1016/j.enconman.2016.09.002

Chen, H., Li, F., and Wang, Y. (2018). Wind power forecasting based on outlier smooth transition autoregressive GARCH model. *J. Mod. Power Syst. Clean. Energy* 6, 532–539. doi:10.1007/s40565-016-0226-3

de-Paz-Centeno, I., Garcia-Ordás, M. T., Garcia-Olalla, O., Arenas, J., and Alaiz-Moretón, H. (2021). M-SRPCNN: A fully convolutional neural network approach for handling super resolution reconstruction on monthly energy consumption environments. *Energies* 14, 4765. doi:10.3390/en14164765

Fang, S., and Chiang, H.-D. (2016). A high-accuracy wind power forecasting model. *IEEE Trans. Power Syst.* 32, 1–1590. doi:10.1109/tpwrs.2016.2574700

Fu, X., Gao, F., Wu, J., Wei, X., and Duan, F. (2019). "Spatiotemporal attention networks for wind power forecasting," in 2019 International Conference on Data Mining Workshops (ICDMW), Beijing, China, November 8–11, 2019, 149–154.

Gomes, P., and Castro, R. (2012). Wind speed and wind power forecasting using statistical models: Autoregressive moving average (ARMA) and artificial neural networks (ANN). *Int. J. Sustain. Energy Dev.* 1, 41–50. doi:10.20533/ijsed.2046.3707.2012.0007

González, C., Mira-McWilliams, J., and Juárez, I. (2015). Important variable assessment and electricity price forecasting based on regression tree models: Classification and regression trees, Bagging and Random Forests. *IET Gener. Transm. &amp. Distrib.* 9, 1120–1128. doi:10.1049/iet-gtd.2014.0655

GWEC (2021). *Global wind report*. Brussels, Belgium: GWEC.

Khazaei, S., Ehsan, M., Soleymani, S., and Mohammadnezhad-Shourkaei, H. (2022). A high-accuracy hybrid method for short-term wind power forecasting. *Energy* 238, 122020. doi:10.1016/j.energy.2021.122020

Kingma, D. P., and Ba, J. (2014). Adam: A method for stochastic optimization. Available at: http://doi.acm.org.ezproxy.lib.ucf.edu/10.1145/1830483.1830503. 1–15. doi:10.1145/1830483.1830503

Kumar, D., Mathur, H. D., Bhanot, S., and Bansal, R. C. (2021). Forecasting of solar and wind power using LSTM RNN for load frequency control in isolated microgrid. *Int. J. Model. Simul.* 41, 311–323. doi:10.1080/02286203.2020.1767840

Liang, G., Liu, G., Zhao, J., Liu, Y., Gu, J., Sun, G., et al. (2020). Super resolution perception for improving data completeness in smart grid state estimation. *Engineering* 6, 789–800. doi:10.1016/j.eng.2020.06.006

Liu, G., Gu, J., Zhao, J., Wen, F., and Liang, G. (2020). Super resolution perception for smart meter data. *Inf. Sci.* 526, 263–273. doi:10.1016/j.ins.2020.03.088

Liu, G., Liu, J., Zhao, J., Qiu, J., Wu, Z., Mao, Y., et al. (2022). Real-time corporate carbon footprint estimation methodology based on appliance identification. *IEEE Trans. Ind. Inf.*, 1. doi:10.1109/tii.2022.3154467

Liu, X., Lin, Z., and Feng, Z. (2021). Short-term offshore wind speed forecast by seasonal ARIMA-A comparison against GRU and LSTM. *Energy* 227, 120492. doi:10.1016/j.energy.2021.120492

Ranganayaki, V., and Deepa, S. N. (2017). Svm based neuro fuzzy model for short term wind power forecasting. *Natl. Acad. Sci. Lett.* 40, 131–134. doi:10.1007/s40009-016-0521-6

Shahid, F., Zameer, A., Mehmood, A., and Raja, M. A. Z. (2020). A novel wavenets long short term memory paradigm for wind power prediction. *Appl. Energy* 269, 115098. doi:10.1016/j.apenergy.2020.115098

Wang, K., Qi, X., Liu, H., and Song, J. (2018). Deep belief network based k-means cluster approach for short-term wind power forecasting. *Energy* 165, 840–852. doi:10.1016/j.energy.2018.09.118

Wang, Y., Zou, R., Liu, F., Zhang, L., and Liu, Q. (2021). A review of wind speed and wind power forecasting with deep neural networks. *Appl. Energy* 304, 117766. doi:10.1016/j.apenergy.2021.117766

Welikala, S., Dinesh, C., Ekanayake, M. P. B., Godaliyadda, R. I., and Ekanayake, J. (2017). Incorporating appliance usage patterns for non-intrusive load monitoring and load forecasting. *IEEE Trans. Smart Grid* 10, 448–461. doi:10.1109/tsg.2017.2743760

Yu, Y., Han, X., Yang, M., and Yang, J. (2020). Probabilistic prediction of regional wind power based on spatiotemporal quantile regression. *IEEE Trans. Ind. Appl.* 56, 6117–6127. doi:10.1109/tia.2020.2992945

Zhou, B., Duan, H., Wu, Q., Wang, H., Or, S. W., Chan, K. W., et al. (2021). Short-term prediction of wind power and its ramp events based on semi-supervised generative adversarial network. *Int. J. Electr. Power &amp. Energy Syst.* 125, 106411. doi:10.1016/j.ijepes.2020.106411

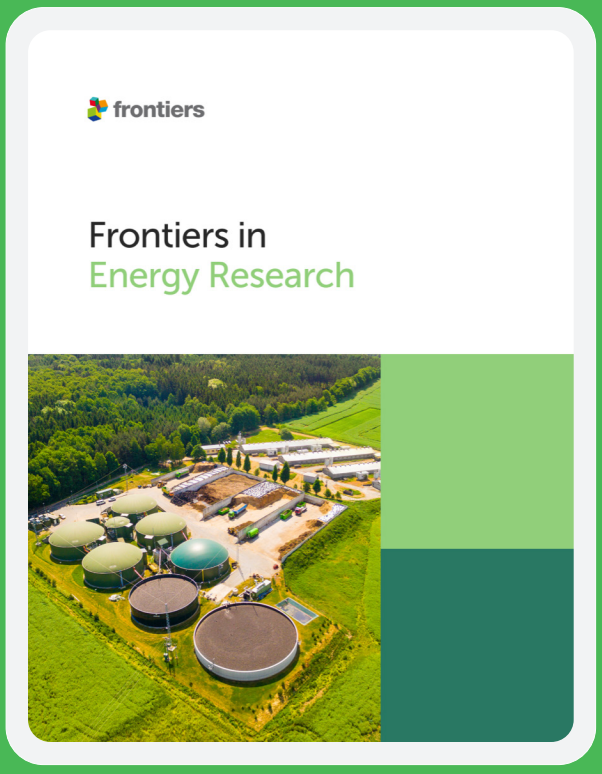
Frontiers in Energy Research

Advances and innovation in sustainable, reliable and affordable energy

Explores sustainable and environmental developments in energy. It focuses on technological advances supporting Sustainable Development Goal 7: access to affordable, reliable, sustainable and modern energy for all.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact