



OPEN ACCESS

EDITED BY
Nantian Huang,
Northeast Electric Power University,
China

REVIEWED BY
Leijiao Ge,
Tianjin University, China
Tao Huang,
Politecnico di Torino, Italy

*CORRESPONDENCE
Chengwei Peng,
1221056301@njupt.edu.cn

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

RECEIVED 26 October 2022
ACCEPTED 08 November 2022
PUBLISHED 17 January 2023

CITATION
Liu K, Zeng Y, Xu J, Jiang H, Huang Y and
Peng C (2023), Multi-source electricity
information fusion methods: A survey.
Front. Energy Res. 10:1080882.
doi: 10.3389/fenrg.2022.1080882

COPYRIGHT
© 2023 Liu, Zeng, Xu, Jiang, Huang and
Peng. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Multi-source electricity information fusion methods: A survey

Kunling Liu^{1,2}, Yu Zeng^{1,2}, Jia Xu^{1,2}, He Jiang^{1,2}, Yan Huang³ and Chengwei Peng^{4*}

¹Sichuan New Electric Power System Research Institute, Chengdu, China, ²State Grid Sichuan Information and Communication Company, Chengdu, China, ³NARI-TECH Nanjing Control Systems Co., Ltd., Nanjing, China, ⁴College of Automation & College of Artificial Intelligence, Nanjing University of Posts and Telecommunications, Nanjing, China

With the vigorous development of the global economy, the demand for electricity quality from all walks of life is also increasing, so it is essential to ensure the electric power grid's safe, stable, and efficient operation. Multi-source electric power information fusion, as the core technology of electric power grid data processing, has become the foundation to promote the intelligent and automatic development of the electric power grid. This paper presents the first work on the survey of the methods of electricity information fusion. It first gives an overview of the process of electricity information fusion and shows the types of electricity data. Then, we provide different classifications of existing methods in view of communication annotation and electric power data, and conduct a thorough comparison and analysis of them. Moreover, we introduce the relevant data sets and evaluation criteria of electric power information and summarize the corresponding evaluation scenarios. Finally, we conclude the maturity of existing works and provide an outlook on future multi-source electric power information fusion methods.

KEYWORDS

smart grid, electricity information, communication standard fusion, electricity data fusion, representation learning

1 Introduction

Electricity is an essential foundation for social development. As the global energy situation is becoming increasingly severe, every country has devoted itself to developing and practicing smart grids (Yao and Lai, 2010; Zhang et al., 2013). With the rapid development of the global economy, the demand for quality electricity in all walks of life is also increasing. Therefore, it is essential to ensure the electric power grid's safe, stable, and efficient operation.

There are many kinds of electric power data, including structural attributes (e.g., signal indicators, charging equipment, and external environment), unstructured text descriptions (e.g., operating instructions, operating mechanisms, and principles), and various topological graphs (e.g., electric power station topological information

and internal lines of equipment). Therefore, the challenges faced by data processing technology in data storage, processing, and display of smart grids have become the constraints of the intellectual development of smart grid (Xue and Lai, 2016), in which the method of multi-source electric power information fusion has become the core research topic in this field.

Several fusion methods for multi-source electric power information have been proposed and achieved some encouraging results. For example, Han et al. (2019) formulated a series of standards for substation communication networks and systems according to the International Electrotechnical Commission (IEC), called the IEC 61850 communication protocol, by which the standardization of electric power data in related systems was established. Furthermore, researchers explored the fusion methods in the view of the physical topology node (Han et al., 2019; Wang et al., 2020) and external protocol (He et al., 2019; Kong et al., 2021) based on IEC 61850. Benefiting from these fusion methods, fault detection, and information interconnection requirements could be solved to some extent. On the other hand, several intelligent algorithms have been employed to integrate and optimize information from the electric power data level to meet business needs for electricity, such as badness data identification (Pan et al., 2022), fusion efficiency improvement (Xia et al., 2022), and electric power consumption prediction (Shao et al., 2020).

Although current works of multi-source electric power information fusion have obtained significant progress, the complexity of electric power system not only lies in its massive data rules and attributes, the non-linearity of the topological electric power grid structure but also depends on temperature, humidity, time, space and many physical quantities collected in the environment (Xue and Lai, 2016). Therefore, no single fusion method can flexibly satisfy all the practical needs and overcome the challenges of smart grid data processing.

In this paper, we comprehensively survey the research works of electric power information fusion methods, especially the research progress of machine learning and deep learning for multi-source electric power information fusion in recent years. Precisely, we first present the overview of the fusion process for electric power information and list the related data types. Then, we provide two classifications of existing methods in view of communication annotation and electric power data, further introduce their critical ideas, and thoroughly compare them. Moreover, we introduce the relevant data sets and evaluation metrics of electric power information and summarize the corresponding evaluation scenarios. Finally, we conclude the maturity of existing works and give an outlook to future methods for multi-source electric power information fusion.

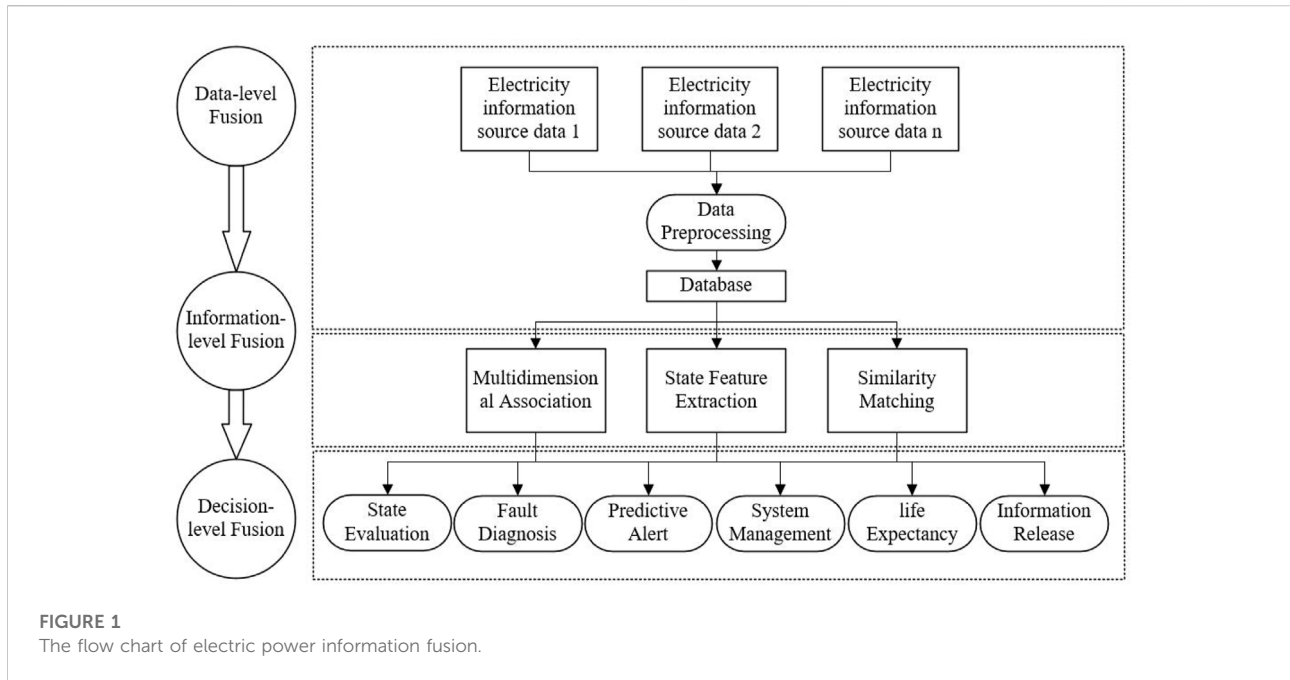
2 Background knowledge of electric power information fusion

2.1 Electric power information fusion process

Currently, electric power grid enterprises worldwide are starting to build integrated data platforms, and digital electric power grids as lots of measurement data can be immediately acquired and quickly shared, which can provide multi-source and heterogeneous information sources for fault diagnosis and other applications. In electric power systems, tedious data and a large amount of information are inconvenient for dispatchers to analyze and operate, so it is important to employ data fusion method to solve these issues. When some failures happen in the electric power grid, its data information reflects the abnormal changes in electrical quantity, protection, and circuit breaker, which can provide valid data information for electric power grid fault diagnosis. The reason is that the methods based on multi-source information fusion can make a comprehensive diagnosis according to the switching information and electrical information provided by different data sources, which can overcome the problem of fault diagnosis error caused by the uncertainty of fault information compared to employing a single data source. In addition, effective integration of multi-source can not only realize the standardization and unification of interfaces and real-time data-sharing requirements under the interoperability of different devices and systems but also aggregate distributed energy storage with similar controllable potential and functional space. Benefiting from the fusion methods, the demand response capability of distributed energy storage can be fully dispatched in the scheduling process. Although the fault diagnosis algorithms of electric power grid based on multi-source information fusion are still in the stage of rapid development, they still have several limitations in practical application scenarios.

Shui et al. (2013) applied the information fusion method to the intelligent warning system of electric power, in which three data fusion architectures are proposed. According to the information sharing, interactivity, and high efficiency of the electric power system, the general integration framework is constructed, by which the electric power system can map data three-tier structure to one of the electric power systems. The data layer corresponds to the sensing measurement layer, the characteristic layer corresponds to the electric power data management layer, and the decision layer corresponds to the electric power system application layer. Moreover, Li et al. (2016) employed the three-layer structure for the massive monitoring data of the energy Internet. They further proposed a data fusion schema based on a multi-layer mode.

As shown in Figure 1, we give a flow chart of electric power information fusion. We combine the electric power system to map the unique three-layer data fusion structure to the electric power



system’s three-layer structure. The fusion process has three levels, including data level, information level, and decision level. The data-level fusion is employed to the sensor measurement layer, and the electric power data measured by the sensor is transmitted to the data fusion center through the network line to complete the analysis, processing, and storage. The information-level fusion is utilized to the electric power data management layer, which extracts the state feature phasor from the original data source and then performs correlation analysis and similarity matching with the primary fusion feature extracted by the previous layer. The decision-level fusion is used to the electric power system application layer. The obtained decision vector can be combined with related algorithms to make classification, reasoning, identification, judgment, and other applications.

2.2 The flow chart of electric power information fusion the data type of electric power fusion

The data collected by the system has the characteristics of a diversified structure due to the distribution network’s different types of power equipment. It is the basis of data fusion to understand the distribution of heterogeneous data in the electric power system. According to the internal structure type, the types of electric power data can be divided into structured, unstructured, and topological data. More precisely, the structured data includes photovoltaic, energy storage, voltage, electric power, electricity, illuminance, temperature, and other structural attributes of charging piles, as well as monitoring data of action signal switches. Unstructured data

mainly contains unstructured text descriptions such as user manuals, operation introductions, maintenance records, operation mechanisms, and principles. Topological data mainly includes topological information on plants and stations and wiring diagrams of internal equipment, as shown in Table 1.

3 The methods of multi-source electric power information fusion

According to the characteristics and applications of existing methods, we mainly divide them into communication standard-based fusion methods and electric power data-based fusion methods, as shown in Figure 2.

In the upper classification of Figure 2, electric power information fusion methods are divided based on communication standards, which can be divided into the ones based on physical topology nodes, fusion the ones based on external protocol standards, and the ones based on model feature extraction as follow.

Fusion methods based on physical topology nodes: It models the topological structure of physical nodes with IEC 61850–90-6 protocol. These logical nodes are employed for fault location.

Fusion methods based on external protocol standards: It establishes a unified information model from the standard model level and the data type level.

Fusion methods based on model feature extraction: It integrates the related data from the electric power platform to extract the critical features of electrical equipment.

From another point of view in Figure 2, existing fusion methods can also be classified based on electric power data,

TABLE 1 Data type of electric power fusion.

Types	The list of electric power data type
Structural data	Voltage, electric power, electric quantity, irradiance, temperature, and other data of photovoltaic, energy storage, charging pile, etc., electric quantity information, operation state monitoring data such as protection action signal and switch tripping information, control data such as control instruction etc.
Unstructured data	Various text data include report data, video/audio data, image data, sensor flow data, instructions for use, technical documents, operation and maintenance records, operation mechanisms and principles etc.
Topological data	Plant topology information diagram, equipment internal wiring diagram etc.

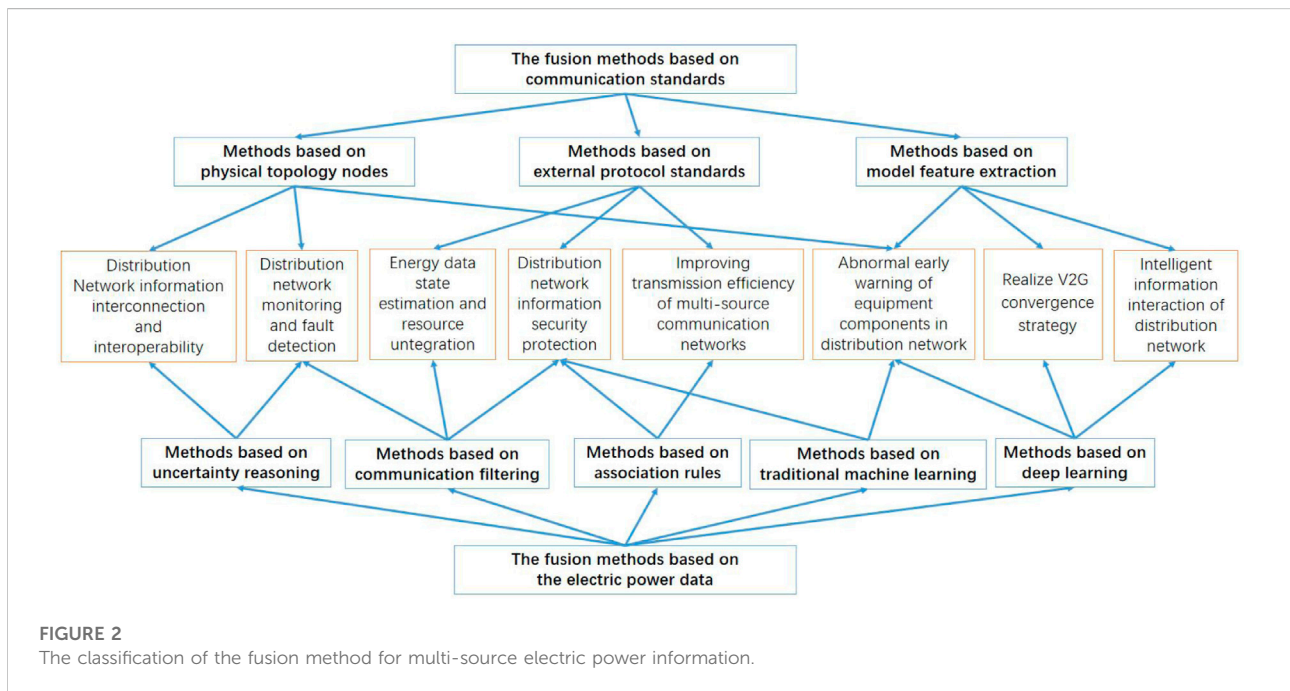


FIGURE 2 The classification of the fusion method for multi-source electric power information.

including the ones based on association rules, the ones based on communication filtering, the ones based on uncertainty reasoning, the ones based on traditional machine learning and the ones based on deep learning.

Fusion method based on association rules: It employs the clustering method to convert data into the one that is suitable for association rules, and the periodic association among converted data will be obtained, and then the data will be fused through ensemble methods (e.g., random forest).

Fusion methods based on communication filtering: It compresses multi-global information so that the error of heterogeneous data fusion and delay time can be reduced.

Fusion methods based on uncertainty reasoning: It employs Dempster-Shafer to converge the components' physical model and fault characteristics so that the influence of information quality changes can be overcome.

Fusion methods based on traditional machine learning: It uses machine learning methods (e.g., Bayesian-based method) to learn the feature from the limited annotation, which can ensure the reliability and fusion efficiency of the results.

Fusion methods based on deep learning: It employs deep learning to integrate and learn the features from the platform data, by which the learning features are more robust for fusion.

Next, we will introduce these methods because of communication annotation and electric power data.

3.1 The fusion methods based on a communication standard

Data mapping is the foundation of communication mapping. Information fusion needs to ensure the safe transmission of information under specific criteria. To promote the

standardization of intelligent substations, IEC TC57 has formulated standards for substation communication networks and systems-IEC 61850. The release of the IEC 61850 standard solves the problem that the traditional telecontrol communication protocol is complicated to realize interconnection and interoperability, which needs the heavy workload of installation and debugging between the distribution terminals of different manufacturers. It realizes the standardization of the distribution automation information model (Han et al., 2019). IEC 61850 series standards play an essential role in substation automation informatization, whose goal is realizing unified modeling and seamless communication of electric power systems. With the migration of public utilities to substations and other network solutions, IEC61850 has become the preferred protocol which is the first standardization work to solve the communication problems of intelligent electronic devices (IEDs) (Xyngi and Popov., 2010).

At present, there are lots of studies related to the application of the IEC 61850 standard. Liu D et al. (2020) studied the source-side maintenance technology of substation autonomous systems according to the IEC 61850 standard to improve the real-time information exchange between the electric power grid. Shantanu et al. (2021) evaluated the feasibility of applying unconventional high-voltage transformers in future digital substations under the IEC 61850 standard. The simulation results showed that unconventional high-voltage transformers were better than conventional transformers in key performance indicators such as ETE, time delay, DC offset, and frequency response. With the introduction of the substation automation system and advanced network and communication technology, the complexity of the electric power system increases dramatically, which may make the whole electric power grid vulnerable to hackers. To solve this problem, Suleman et al. (2021) proposed a network model developed in OPNET, demonstrating the results of various denial of service (DoS) attacks on digital substations based on IEC 61850. It was of great significance to understand the influence of these factors on the performance of digital substations.

Zhu et al. (2017) aimed at the configuration problem of a distributed intelligent application using the IEC 6185 standard in a distribution automation system. The authors presented a configuration solution from two aspects: the semantic model and the processing method. Taha and Suhail. (2020) proposed a communication technology based on IEC 61850 and XMPP. They developed the IEC 61850 information model of the UPFC controller to coordinate the stable operation of UPFC and DERs in the microgrid. With the international standard IEC 61850 and IEEE 2030 reference models, Leitea et al. (2016) proposed a voltage regulation optimization method based on the communication architecture model that coordinated the interaction between DGPV units to meet the connectivity and interoperability requirements.

Therefore, studying electric power information fusion technology based on communication standards is imperative. We divide the electric power information fusion technology based on communication standards into three categories: fusion based on physical topology nodes, fusion based on external protocol standards, and fusion based on model feature extraction.

Fusion methods based on physical topology nodes: Han et al. (2019) employed the IEC 61850–90-6 standard to divide the fault indication into two parts according to the mapping of logical nodes: fault detection and fault indication. The release of the IEC 61850 standard solved the problem that the traditional telecontrol communication protocol is hard to realize the interconnection, interoperability, and heavy workload of installation and debugging between distribution terminals of different manufacturers and the central station and distribution terminals so that the standardization of distribution automation information model is realized. Taking fault indication, fault location isolation, and electric power supply restoration as examples, the author analyzed the establishment process of fault detection and protection information model of the distribution network. Fault indication information was needed to convert the transient impulse control signal output by fault detection information into a continuous position indication signal. New logic nodes needed to be added to a unique distribution automation function. The new logical nodes added in IEC 61850–90-6 are shown in Table 2.

Wang et al. (2020), aimed at the problem that there was no physical link model in the IEC 61850 protocol, modeled the corresponding physical information nodes in the secondary communication of the process layer and constructed the physical topology of the secondary device communication. The authors put forward an intelligent warning and fault diagnosis schema for the secondary circuit of the smart substation and established a fault diagnosis method that combined a virtual circuit with an actual physical link according to the information flow characteristics in the smart substation. Through comprehensive analysis of configuration file information in the substation, link alarm information of protection measurement, and status information, of the control device, the probability of all the possible fault points would be calculated when the communication abnormality occurs is given. According to the failure point probability given by the system, the operation and maintenance engineer can find out the failure point by troubleshooting the corresponding equipment in turn, which could improve the operation and maintenance efficiency. This method makes full use of the state information given by each piece of equipment in the station and various data flowing on the network and automatically and intelligently analyzes and judges the fault points in the secondary circuit by relying on the correlation and coupling characteristics in each virtual circuit. It satisfied the

TABLE 2 New logical nodes for fault indication in IEC 61850–90-6.

Functional requirements	Logical node
Error indication	SFPI
Voltage indication generates a persistent “presence” signal and a “non-presence” signal according to the input from PTOV and PTUV	SVPI
Current indication generates a persistent “presence” signal and a “non-presence” signal according to the input from PTOC and PTUC	heSCPI
Calculation and statistics of faults in a given period	SFST

IEC, 61850–90-6 mainly solves the problems of fault location and isolation of distribution lines (feeders) and the establishment of electric power supply recovery logic nodes in non-fault sections in distribution automation. The logic nodes added for electric power supply restoration in IEC, 61850–90-6 are shown in Table 3.

TABLE 3 New logic nodes for electric power supply restoration in IEC 61850–90-6.

Functional requirements	Logical node
Detect the faults on the feeder based on overcurrent or Undervoltage	ASWI
Detects that the voltage on one side disappears and closes	AATS
Error indication	SFPI
Voltage presence indication	SVPI
When no voltage is detected on both sides of the segmentary, turn on the segmentary. After a voltage is detected on one side, the segmentary is closed after a specified delay	
If the voltage is not detected within the predetermined blocking time, it will block the segmentary and not close it again	RRFV
Current presence indication	SCPI
Fault status	AFSL
Fault isolation after indication completed	AFSI
Control of the electric power during the recovery process	ASRC

requirement of standardized information transmission and sharing in intelligent substations, and the developed system could achieve good results in real applications.

Fusion methods based on external protocol standards: He et al. (2019) analyzed the hybrid measurement architecture of an intelligent distribution network and compared the differences between D-PMU data, AMI data, and SCADA data in data composition, data accuracy, and time scale information. Based on the IEC 61850 standard and IEC 61968–301 (CIM) static mapping and dynamic mapping of SCADA interaction with new systems, the authors put forward a unified information model, which unified the equipment descriptions of IEC 61850 standard and IEC 61968–301 (CIM) standard, made up for the defect that the modeling standards followed by the wide-area measurement and control systems of smart distribution networks were not uniform in construction, and realized the data exchange among all systems. Kong et al. (2021) proposed an external protocol standard of IEC 61850 communication protocol called MQTT, which could be adopted in the cloud edge communication of distribution Internet of Things. The MQTT protocol is regarded as the application layer communication protocol for the information interaction between the cloud master station and

edge devices, which could enhance interoperability and solve the problems of standardization of data transmission and data description model between the cloud master station and the edge devices of distribution Internet of Things. The communication mapping methods from the communication service subset to the MQTT protocol are summarized, including the direct and indirect mapping methods. The authors concluded that the direct mapping method was more applicable and economical than complex ones on cloud-side communication of distribution Internet of Things.

Fusion methods based on model feature extraction: Zhao et al. (2021) proposed a fault diagnosis method based on the confidence fusion of the Dempster-Shafer theory, which could effectively and comprehensively utilize the redundant features of multi-source data information and thoroughly mined the fault features of related switching information and electrical information. The authors implemented the effective integration of multi-source alarm information, which was of great help in improving the accuracy of fault diagnosis and quickly identifying fault components. Lu et al. (2020) constructed the fault optimization detection model of substation equipment, combined with the collaborative fault

method, to improve the fault optimization diagnosis ability of substation equipment. They put forward the collaborative fault diagnosis method of substation equipment based on the information fusion technology of the sampling and feature extraction model. The experimental analysis showed that this method could obtain good synergies in fault diagnosis of substation equipment operation and own a high precision of detecting fault features. GOOSE message is a vital part of the IEC 61850 protocol, which embeds selection logic and simulation data. Li J et al. (2021) studied the measurement method based on the specification part of the IEC 61850 protocol. They analyzed the behavior characteristics of CPS from many GOOSE and manufacturing message specifications based on a digital substation network and system management scheme. To solve the problem of low accuracy of existing detection methods, the authors proposed an anomaly detection method based on difference sequence variance combined with the message characteristics of a digital substation, including the determination of membership function of traffic anomaly and CPS parameters of fusion.

3.2 The fusion methods based on electric power data

With the development and maturity of global electric power big data multi-source information fusion technology, fusion algorithms are gradually diversified. The multi-modal data fusion is an effective way to implement the collaborative analysis of multiple heterogeneous networks that pushes forward observable and controllable power grids (Wang B et al., 2022). We divide the electric power information fusion methods based on electric power data into five categories: the ones based on association rules, the ones based on communication filtering, the ones based on uncertainty reasoning, the ones based on traditional machine learning, and the ones based on deep learning.

Fusion methods based on association rules: Pan et al. (2022) applied the data mining method of association rules to the energy system and employed random forest to establish the training network of big energy data for data fusion. Inspired by incremental learning and offline learning, the authors proposed the MCS-RF framework of energy big data. By converting discrete data into data suitable for association rules, the accuracy of wrong data identification and energy data state estimation was improved. Compared with the traditional algorithms based on residual error, the proposed method can save time in identifying lousy energy data and obtain higher accuracy. Similarly, Liu B et al. (2021) adopted the extensive data analysis method based on random forest. The authors proposed a regional priority business screening model for the multi-station fusion project that could satisfy the concept of sustainable development and reliability.

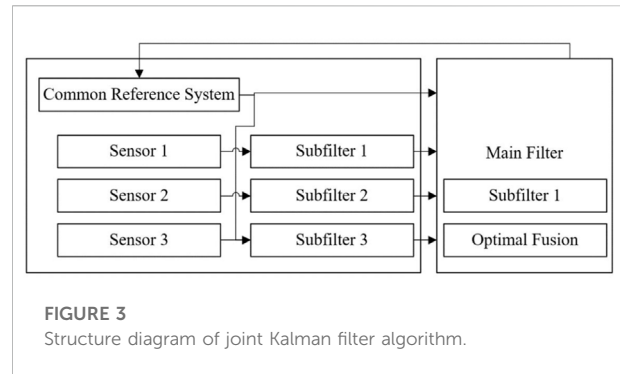


FIGURE 3
Structure diagram of joint Kalman filter algorithm.

Fusion methods based on communication filtering: Wen and Li (2018) proposed an information fusion method combining compressed sensing with global forwarding data. According to the communication range of network nodes, the clusters are divided into multiple areas. Each node selects the communication mode to transmit global information to the communication cluster head according to its location to reduce the global information fusion delay. In the process of global information collection in the multi-source communication network, the model employed the compressed sensing model to compress the global information and forward the global information according to the number of sub-nodes of network nodes. The fusion of global information in a multi-source communication network effectively improves its transmission efficiency.

As the traditional methods do not register the time during the collection of multi-source heterogeneous data in the distribution network, it leads to significant errors and low efficiency for the fusion of multi-source heterogeneous data. Xia et al. (2022) proposed a multi-source heterogeneous data fusion method in a distribution network based on a joint Kalman filter to solve this issue. The authors employed the joint Kalman filtering algorithm to implement heterogeneous data fusion. The structure diagram of the joint Kalman filtering algorithm is shown in Figure 3. We can observe that the data process of joint Kalman filtering algorithm is mainly divided into two ones: the primary filter and the local sub-filter. In the calculation, each sub-filter works independently. That is, the time update and the measured value update run separately. This method can improve data fusion efficiency and reduce data fusion error compared with traditional methods.

3.3 Structure diagram of joint kalman filter algorithm

Fusion methods based on uncertainty reasoning: Zhao et al. (2021) proposed a multi-layer fault diagnosis model of the electric power grid and used the Dempster-Shafer theory to

analyze the data information in the multi-source information fusion diagnosis layer. The authors considered various information, such as switching value and electrical quantity, which provided the basis for fault diagnosis to obtain the probability value of the possible fault components for each fault component. In addition, they further analyzed the related protection and the action of circuit breakers. Traditional fusion methods for fault diagnosis may reduce the confidence of fault components and affect the fault diagnosis results. Therefore, the author improved the multi-source information fusion diagnosis method based on the Dempster-Shafer theory. It could solve the problem that the uncertain factors (e.g., misoperation, refusal of protection, circuit breaker, transmission error of alarm information) could affect the accuracy of fault diagnosis results.

Fusion methods based on traditional machine learning: Electricity data fusion methods have broad applications in electric vehicles and the grid. Rik and Willett (2008) put forward a model of electric vehicles with three control types. The authors added electric and V2G vehicles into the energy system to integrate higher-level wind electric power without generating excess electric power and significantly reduce carbon dioxide emissions. Wang Y et al. (2022) employed the entity alignment of the Bayesian model to implement the fusion method for related attribute mapping. It was a highly reliable and low-complexity knowledge fusion method that combined a concept drift detection algorithm with an unsupervised reverse verification algorithm. The experiments showed that the proposed method was superior to the conventional machine learning algorithm regarding knowledge fusion efficiency and algorithm complexity.

Fusion methods based on deep learning: The forecast of electric power consumption is an essential task of smart grid construction. Related works pay attention to weather, holidays, and temperature for electric power forecasts. It is necessary to use lots of sensors to collect these data, which increases the cost of time and resources. Darudi et al. (2015) proposed a new data fusion algorithm based on an artificial neural network and adaptive neuro-fuzzy inference system, modified ordered weighted average (OWA). Shaxiaorui et al. (2020) proposed a hybrid depth prediction model based on CNN and LSTM, which could learn fusion features in parallel. As corresponding statistics were considered, the method could obtain more robust features even if some original information was lost. To predict electric power consumption, the authors incorporated the advantages of each model. Similarly, Liu and Meng (2020) adopted the depth-limited Boltzmann machine to encode all data into the same vector space. They applied the time series method to implement the effective fusion of power network communication service data and improve the accuracy of fusion results. Wang et al. (2021) proposed a universal fusion framework suitable for structured multiple time series and unstructured images, which could achieve the deep fusion for heterogeneous multi-

parameter under the power Internet of Things. Li G et al. (2021) proposed a multi-source log comprehensive feature extraction method based on Restricted Boltzmann Machine (RBM) to excavate security threats in the electric power grid by entirely using heterogeneous data sources in the electric power information system.

3.4 Method review

This section summarizes the fusion methods of multi-source electric power information listed in Table 4, including the core models, characteristics, and limitations.

For the fusion methods of electric power information based on a communication standard, the methods based on physical topology mainly depend on IEC 61850. These methods establish new logical and line nodes so that loops can be detected for the equipment in the station (Han et al., 2019; Wang et al., 2020). Although these methods have strong engineering, the hardware and maintenance costs for establishing physical topology nodes are high in the early stage. The fusion methods based on external protocol standards mainly construct the unified information model from the levels of a standard model and the data type. For these methods, the communication protocol standards are layered, and different mapping strategies are formulated according to the characteristics of protocol files (He et al., 2019; Kong et al., 2021). However, these methods lack scalability and depend on manual construction and maintenance. For the methods based on model feature extraction, the associated data in the data platform are integrated to extract the critical feature quantities of electrical equipment. The statistical probability is automatically calculated for the fault diagnosis or abnormal electrical equipment data (Lu et al., 2020; Li Y et al., 2021; Zhao et al., 2021). Although the methods are automatic and real-time, they are only suitable for single-task scenarios.

For the fusion methods of electric power information based on electric power data, the methods based on association rules are inspired by unsupervised learning, in which the clustering method is employed to transform the data to mine the association rules. Then, the periodic correlations among data are obtained. This way, random forest can further fuse the transformed data (Pan et al., 2022). Although the accuracy of related tasks is high, these methods can only deal with discrete data, and the subjective factors might influence the clustering results. The fusion methods based on communication filtering employ collaborative Kalman filtering (Wen and Li, 2018) and compressed sensing (Xia et al., 2022). These methods can effectively compress global information and reduce errors and delays during the process of heterogeneous data fusion. Therefore, they can improve the efficiency of signal fusion and forwarding. However, high requirements exist for the hardware storage space of network acquisition nodes, which sacrifice the fusion accuracy. The methods based on uncertain reasoning employed the Dempster-Shafer theory to model the

TABLE 4 Comparison of multi-source electric power information fusion methods.

Method name	Core ideas/models	Method characteristics	Limitations	Representative literature
Fusion methods based on physical topology nodes	Logical nodes in fault location, isolation, and non-fault areas are established based on IEC 61850–90-6 protocol, and the topological structure of physical nodes is modeled	Detecting the equipment in the station from the line nodes and loops has strong engineering practicability	The hardware and maintenance costs for establishing physical topology nodes are high in the early stages	Han et al., (2019); Wang et al., (2020)
Fusion methods based on external protocol standards	Establish a unified information model from the levels of the standard model and the data type	The communication protocol standards are layered, and different mapping strategies are drawn up according to the characteristics of protocol files	The scalability of the method is weak, and they depend on manual construction and maintenance	He et al., (2019); Kong et al., (2021)
Fusion methods based on model feature extraction	Integrate the related data of the data platform to extract the critical characteristic quantities of electrical equipment	The statistical probability can be automatically calculated for fault diagnosis or abnormal electrical equipment data	They are only suitable for the of single-scenarios	Zhao et al., (2021); Lu et al., (2020); Li Y et al., (2021)
Fusion methods based on association rules	K-Means Periodic Association Rule Random Forest Model	The clustering method transforms the data for mining the association rules. Data fusion is carried out through the periodic association rule and the random forest	Only discrete data can be processed, and subjective factors might influence clustering results	Pan et al., (2022)
Fusion methods based on communication filtering	Joint Kalman Filtering Compressed Sensing	The global information can be effectively compressed, and the efficiency and real-time performance of signal fusion and forwarding can be improved	High requirements for the hardware storage space of network acquisition nodes sacrifice the fusion accuracy	Wen and Li, (2018); Xia et al., (2022)
Fusion methods based on uncertainty reasoning	Dempster-Shafer theory	Modeling the physical model and fault characteristics of components based on Dempster-Shafer. It can overcome the influence of information quality change and realize uncertain reasoning	Limited by how objects are fused pairwise, their time complexity is extensive	Zhao et al., (2021)
Fusion methods based on traditional machine learning	Bayesian Estimation	It can ensure results' reliability and fusion efficiency with limited labeled data	They cannot realize the effective combination of features and further mine the hidden feature association between data	Wang B et al., (2022)
Fusion methods based on deep learning	Recurrent Neural Network Convolutional Neural Networks Boltzmann Machine	They can make the fusion features obtained by representation learning more robust and perform better	They heavily rely on lots of labeled data and the actual hardware	Jiang et al., (2019); Lin et al., (2020)

components' physical model and fault characteristics, which can overcome information quality change and realize uncertain reasoning (Zhao et al., 2021). Nevertheless, these methods are limited by how objects are fused pairwise, so their time complexity is extensive. Traditional machine learning of fusion methods is represented by Bayesian estimation (Wang B et al., 2022). Although these methods can ensure the reliability and fusion efficiency of results with limited labeled data, they cannot realize the effective combination of features and further mine the hidden feature association between data. Therefore, there exists a bottleneck to improving the accuracy of specific tasks such as fault detection. The methods based on deep learning mainly utilize the recurrent neural network. This kind of method makes the fusion features obtained by representation learning more robust and perform better. Nevertheless, these methods heavily rely on lots of labeled data and the actual hardware.

After a comprehensive comparison of the above existing methods, we discover that the fusion method of electric power information based on a communication standard is a well-established methodology because the IEC 61850 standard has been published for nearly 20 years. The requirement of information interconnection for fault detection and electric power consumption prediction is still a hot research topic. This kind of research works already accounts for about 30% of the fusion works. One of the main reasons is the business's need for electricity. In recent years, more than half of fusion works have been devoted to exploring machine learning and deep learning to learn rich features and meet the requirements of smart grid. Nevertheless, there is room for making progress in optimizing existing solutions. We conclude the test datasets of electric power information fusion and their data processing methods listed in Table 5.

TABLE 5 Datasets and data processing methods.

Datasets	Dataset size	Data processing method
Electric power flow Dataset system (Pan et al., 2022)	1,440 samples	The clustering method is employed discretization
Phasor Measurement Unit (PMU) Dataset (Xia et al., 2022)	Thousands of meteorological data, equipment data, and photovoltaic data	The original data is stored in a temporary table, and a processing thread corresponding to the data type is established
American Electric Company (AEP) Dataset ¹ (Shao et al., 2020)	121,273 pieces of data from December 2004 to January 2018	Normalization processing
Commonwealth Edison (COMED) Dataset ² (Shao et al., 2020)	66,497 data from December 2011 to January 2018	Normalization processing
Dayton Electric power and Light Company (DAYTON) Dataset ³ (Shao et al., 2020)	121,275 pieces of data from December 2004 to January 2018	Normalization processing
Public Data Set of University of California Machine Learning and Intelligent Systems Center (Wang Y et al., 2022)	Environmental data of 12 areas of Beijing from 2014 to 2017 including air quality, temperature/dew point temperature, wind direction/speed, air pressure, rainfall etc.	Ontology association
A data set composed of defective texts of an electric power company in a China province (Jiang et al., 2019)	23,409 data from 2015 to 2017, all from the operation and maintenance records of AC 220 kV oil-immersed transformers	Convert original text data into continued vectors

Evaluation metrics of electric power information fusion method.

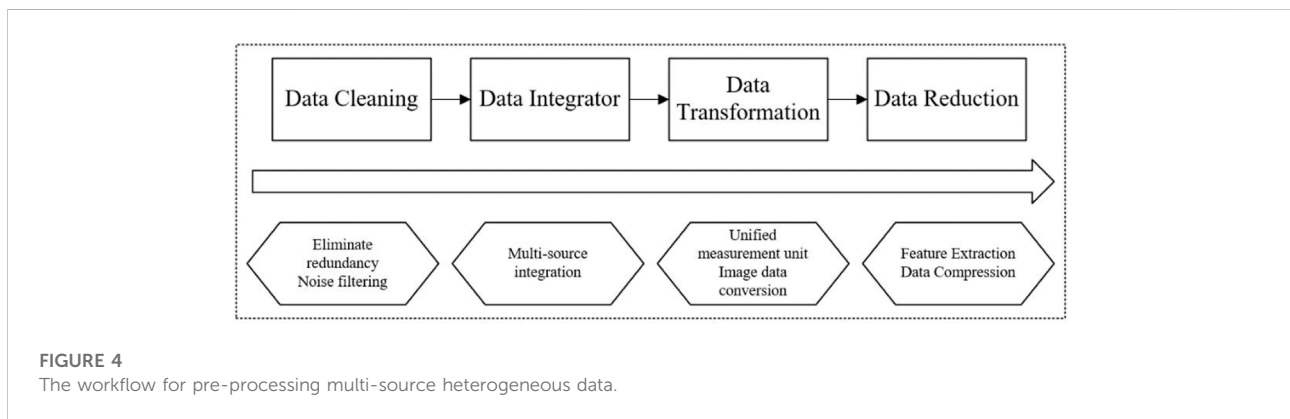


FIGURE 4 The workflow for pre-processing multi-source heterogeneous data.

4 Evaluation of electric power information fusion methods

4.1 Test dataset for electric power information fusion method

This section will introduce the test¹ datasets of electric power information fusion in detail. First, the standard pre-processing techniques for datasets are introduced in this paper². Then, we list the statistic of corresponding datasets for electric power information fusion.

There are several kinds and quantities of electric power data, but their interrelated relationships are vague as the initially collected data has high redundancy and noise content. Therefore, preliminary signal processing and data classification are essential. The function of data fusion is to ensure that the electric power information acquisition equipment can accurately provide feedback on the current electric power network data and store³ the historical data completely. In this way, the original data of heterogeneous electric power information systems can be processed by eliminating invalidity, filtering redundancy, and data interpolation. Afterward, the researchers could carry out the operations of feature extraction and feature matching operations among heterogeneous datasets, which facilitate the next big data fusion

1 https://www.kaggle.com/datasets/robikscube/hourly-energy-consumption?select=AEP_hourly.csv

2 https://www.kaggle.com/datasets/robikscube/hourly-energy-consumption?select=COMED_hourly.csv

3 https://www.kaggle.com/datasets/robikscube/hourly-energy-consumption?select=DAYTON_hourly.csv

and decision analysis. Figure 4 shows the concrete workflow for pre-processing multi-source heterogeneous data.

4.2 The workflow for pre-processing multi-source heterogeneous data

Several pre-processing methods exist for electric power data, such as the average, the least square methods, and so on.

Dai et al. (2017) proposed a non-parametric method (kernel density-mean). By comparing the average and the least square methods, the calculation results show that the kernel density-mean method could obtain better results in terms of the robustness of sampling frequency.

Pan et al. (2022) regarded the electric power flow data in the SCADA system as the samples, for which the standard deviation of the measured value was set to 0.02 and the standard deviation of the phase angle was set to 0.005. The authors simulated and analyzed the electric power flow data change from April to June. These data were discretized by the clustering method to obtain the discrete level of active electric power.

Xia et al. (2022) synchronized the dataset with Phasor Measurement Unit data, meteorological data, equipment data, and photovoltaic data. The pre-processing of data contains two steps. Firstly, extracted data from the data source and stored the original data in the corresponding temporary table. Therefore, it must establish a processing thread corresponding to the data type. Then, traverse the acquired data from the temporary table, and judge whether there is any abnormal situation in the data according to the rectification rules. If any abnormal situation exists, it needs to process the data and generate rectification records.

Zhao et al. (2021) employed electrical data and switching records as samples. Furthermore, the authors measured the confidence of samples after data and records were normalized.

Lin et al. (2020) analyzed that the same type of multi-source heterogeneous data fusion is meaningful for correlation analysis. Therefore, the authors concluded that Euclidean distance and Pearson correlation coefficient were the most common metrics methods for correlation analysis. The Euclidean distance of any two points x_i, x_j in m-dimensional space was defined in Eq. 1:

$$d(x_i, x_j) = \sqrt{\sum_{k=1}^m (x_{ik} - x_{jk})^2} \tag{1}$$

where x_{ik}, x_{jk} are the coordinates of two points in m-dimensional space.

Zhao et al. (2021) employed a local electric power system for simulation analysis. The confidence was calculated by synthesizing evidence based on the Dempster-Shafer theory. Then, the effectiveness of the proposed method was verified

by comparing the confidence levels as shown in Eq. 2, where p is the probability value of the i-th possible faulty element, and the fault confidence value of the possible faulty element p_i^* is calculated in Eq. 3. After then, the authors would set a confidence threshold and compare the predicted results with the real samples.

$$\sum_{Fault, end} = \{p_1, p_2, \dots, p_n\} \tag{2}$$

$$p_i^* = p_i / \max \{p_i | p_i \in \sum_{Fault, end}\} \tag{3}$$

To fairly evaluate the effectiveness of the proposed models based on deep learning, Shao et al. (2020) adopted Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) composed of multiple evaluation indicators for evaluation. The specific formulas are as follows Eqs 4-6:

$$RMSE = \sqrt{\frac{\sum_{n=1}^N (forecast_n - real_n)^2}{N}} \tag{4}$$

$$MAE = \frac{\sum_{n=1}^N |forecast_n - real_n|}{N} \tag{5}$$

$$MAPE = \frac{100\%}{n} \sum_{n=1}^N \left| \frac{forecast_n - real_n}{real_n} \right| \tag{6}$$

Where N is the number of test samples, f is the predicted value, and $real$ is the ground truth. RMSE evaluated the model by the standard deviation of the residuals between the ground truth and predicted values; MAE is the average vertical distance between the ground truth and predicted values, which is more robust to more significant errors than RMSE. RMSE and MAE increase significantly and rapidly when using extensive data for model training and evaluation. Hence, the authors employ MAPE to evaluate the fusion methods comprehensively.

Wang Y et al. (2022) used the calculation of Precision, Recall, and F1 to verify the feasibility of the proposed method for the evaluation of heterogeneous datasets. The calculation formula is shown in Eqs 7-9:

$$Precision = \frac{TP}{TP + FP} \tag{7}$$

$$Recall = \frac{TP}{FP + FN} \tag{8}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{9}$$

where TP (True Positive), FP (False Positive), and FN (False Negative) represent the entity alignment indicators of the fused ontology.

In addition, Jiang et al. (2019) found that there are three types of fault defect include general defects, severe defects, and critical defects, and they used macro-P macro precision (represented by $A_{macro-P}$), macro-R macro recall (represented by $A_{macro-R}$), macro-F1 macro comprehensive index (represented by

$A_{macro-F1}$) to evaluate. Eq. 10 lists their expression, where the precision rate p and recall rate R are the two-class evaluation indexes. Their definitions of p and R are shown in Eq. 11, where P_T represents the number of positive cases that are correctly judged as positive cases; N_F indicates the number of positive examples that are wrongly judged as counterexamples; P_F represents the number of negative examples that are wrongly judged as positive examples. For the three-classification problem, there exist three different positive examples and three values that correspond to p and R .

$$\left\{ \begin{array}{l} A_{macro-P} = \frac{1}{n} \sum_{i=1}^n P_i \\ A_{macro-R} = \frac{1}{n} \sum_{i=1}^n R_i \\ A_{macro-F1} = \frac{2A_{macro-R} \times A_{macro-P}}{A_{macro-R} + A_{macro-P}} \end{array} \right. \quad (10)$$

$$\left\{ \begin{array}{l} P = \frac{P_T}{P_T + P_F} \\ R = \frac{P_T}{P_T + N_F} \end{array} \right. \quad (11)$$

Wen and Li (2018) utilized the global information reconstruction error ε to evaluate the feasibility of the proposed methods. The defined formula is as follows:

$$\varepsilon = \|\bar{f} - f\|_2 / \|f\|_2 \quad (12)$$

f and \bar{f} represent the original and restored global information sequences, respectively.

4.3 Evaluation scenarios of electric power information fusion method

For the above test data sets and evaluation metrics, we summarize three types of test scenarios listed as follows:

Information interconnection of Electric power grid: It is suitable to evaluate the fusion methods based on communication standards. Through the mapping operation and standard specification between different protocols, the multi-type protocol signal transmission can be employed to verify the proposed model's compatibility and delay of signal transmission (Han et al., 2019; He et al., 2019; Kong et al., 2021).

Anomaly detection of substation data: It is also suitable for evaluating the fusion methods based on communication standards, which can verify the performances of the model in terms of accuracy, and missed detection rate by labeled signal datasets based on specific standard protocol (Wang et al., 2020; Li Y et al., 2021).

Fault detection of charging equipment: It is suitable for evaluating fusion methods based on electric power data. On the one hand, the effect of the models can be verified from the

waveform of the running equipment (Lu et al., 2020). On the other hand, this scenario can be modeled as a classification problem given machine learning. The methods can be verified by precision, recall, and F1 value and the corresponding variants of these metrics (Wang et al., 2022).

5 Conclusion and future work

Multi-source electric power information fusion, as the core technology of electric power grid data processing, has become the foundation to promote the intelligent and automatic development of the electric power grid. However, there does not exist survey of existing methods to discuss their limitation for this topic. This paper overviewed information fusion methods for multi-source electricity information. We thoroughly reviewed current methods regarding communication annotation and electric power data and further introduced their critical ideas. Then, we introduced the relevant datasets and evaluation criteria of electric power information and summarized the corresponding evaluation scenarios. We further evaluated the maturity of existing fusion methods and analyzed that the fusion method of electric power information based on a communication standard was a well-established methodology. Relatively, there is room for making progress on optimizing fusion methods of electric power information based on machine learning and deep learning, which could achieve the goal of information interconnection for satisfying the requirements in smart grid (e.g., fault detection, electric power consumption prediction). We believe that future research for multi-source electric power information fusion can be carried out from the following four aspects:

The fusion methods are based on multi-modal. Currently, the existing models mainly refine the features from the properties of electric systems or communication standards. However, these methods are not mature enough to contain text information and image data such as fault descriptions recorded by engineers and equipment pictures. The methods based on multi-modal are helpful to improve further the accuracy of electric power information fusion and other related tasks (Yang et al., 2021). On the one hand, the relation extraction models can extract more useful information from the textual descriptions, which can assist the electric power information fusion method; on the other hand, the image recognition technology can be used to improve the granularity for equipment fault type identification. Then, the performances of electric power information fusion can be further enhanced.

The fusion methods are based on federated learning. As the system data associated with charging piles and photovoltaic equipment has a certain degree of privacy, a set of electric power specification systems can be formed based on all levels through consultation among various electric

power companies and related platforms. Then, the high-quality data in electric power companies can be encrypted and shared using federal learning (Liu & Meng, 2020). Finally, a high-quality mapping relationship is constructed among electric power grid systems to satisfy more smart grid requirements.

The incremental dynamic information fusion strategy. As too many standards and attributes are increasing in the electric power grid system, it is hard for existing methods to immediately respond to the requirements for updating and iterating electric power systems because the cost of training and fusion of these models is high. In addition, it may not better satisfy the real-time tasks of electric power information. Therefore, incremental dynamic fusion technology (Yu et al., 2020) can be explored to reduce the cost of information fusion and ensure the real-time performance of the system.

The fusion methods are based on an interactive strategy. Considering application scenarios such as equipment fault monitoring requiring high precision, follow-up correction by domain experts is indispensable. Therefore, the interactive fusion strategy (Liu F et al., 2021c; Liu Q et al., 2021) could be adopted so that domain experts can screen the results of fusion methods in each iteration. This way, experts could mark the results that can be reused as training samples for feeding the models. It can not only ensure the reliability of the model results but also improve the fusion accuracy of the electric power information.

References

- Dai, J., Cao, J., Zhang, F., Liu, D., and Shen, X. (2017). Data pre-processing method and its evaluation strategy of SCADA data from wind farm. *Acta Energetica Solaris Sin.* 38 (9), 2597–2604. doi:10.19912/j.0254-0096.2017.09.038
- Darudi, A., Bashari, M., and Javidi, M. H. (2015). Electricity price forecasting using a new data fusion algorithm. *IET Gener. Transm. & Distrib.* 9 (12), 1382–1390. doi:10.1049/iet-gtd.2014.0653
- Han, G., Chen, Y., and Xu, B. (2019). Information model for distribution network fault detection and protection based on IEC 61850. *Distribution Util.* 36 (07), 8–12. doi:10.19421/j.cnki.1006-6357.2019.07.002
- He, X., Tu, C., and Li, P. (2019). Research on multi-source data fusion for smart distribution network. *South. Power Syst. Technol.* 13 (04), 42–47. doi:10.13648/j.cnki.issn1674-0629.2019.04.007
- Henrik, L., and Willett, K. (2008). Integration of renewable energy into the transport and electricity sectors through V2G. *Energy policy* 36 (9), 3578–3587. doi:10.1016/j.enpol.2008.06.007
- Jiang, Y., Peng, M., Ma, K., and Li, L. (2019). Evaluation method for power transformer conditions based on multi-source heterogeneous data fusion. *Guangdong Electr. Power* 32 (09), 137.
- Kong, C., Chen, Y., and Zhao, Q. (2021). Research on cloud-side communication mapping of the distribution internet of things based on MQTT protocol. *Power Syst. Prot. Control* 49 (08), 168–176. doi:10.19783/j.cnki.pspc.200775
- Leitea, L., Boaventuraa, W., Erricoa, L., Cardosoa, E., Dutra, R., and Lopes, B. (2016). Integrated voltage regulation in distribution grids with photovoltaic distribution generation assisted by telecommunication infrastructure. *Electr. Power Syst. Res.* 136, 110–124. doi:10.1016/j.epsr.2016.02.016
- Li, G., Yang, L., Liu, F., Yu, M., Song, Y., and Wen, F. (2016). A mutual information method for associated data fusion in energy internet. *Electr. Power Constr.* 37 (09), 22.
- Li, G., G., Zhang, X., Wu, T., and Zhang, F. (2021). Research on the maintenance technology of substation autonomous system source end based on IEC 61850. *Process Autom. Instrum.* 42 (07), 68–72. doi:10.16086/j.cnki.issn1000-0380.2020090028
- Li, J., Li, X., Gao, T., Zhang, J., and Zhang, B. (2021). Research and application of fault handling based on power grid multivariate information knowledge graph. *Electr. Power Inf. Commun. Technol.* 19 (11), 30–38. doi:10.16543/j.2095-641x.electric.power.ict.2021.11.005
- Li, Y., Y., Gao, B., Xu, L., Ding, J., Tang, H., and Shan, R. (2021). An anomaly detection method for digital substation abnormal data based on fusion of difference sequence variance and CPS. *Power Syst. Clean Energy* 37 (02), 30.
- Lin, Y., Chen, R., and Jin, T. (2020). Research on multi-source heterogeneous data fusion technology for complex information system. *China Meas. Test* 46 (07), 1.
- Liu, J., and Meng, X. (2020). Survey on privacy-preserving machine learning. *J. Comput. Res. Dev.* 57 (2), 346. doi:10.7544/issn1000-1239.2020.20190455
- Liu, B., Scells, H., Zuccon, G., Hua, W., and Zhao, G. (2021). *ActiveEA: Active learning for neural entity alignment*. arXiv preprint arXiv:2110.06474.
- Liu, D., D., Kong, D., Chang, Y., Ma, L., and Wang, R. (2020). Multi-source log comprehensive feature extraction based on restricted Boltzmann machine in power information system. *Comput. Syst. Appl.* 29 (11), 210–217. doi:10.15888/j.cnki.csa.007667
- Liu, F., F., Chen, M., Roth, D., and Collier, N. (2021). Visual pivoting for (unsupervised) entity alignment. *Proc. AAAI Conf. Artif. Intell.* 35 (5), 4257–4266. doi:10.1609/aaai.v35i5.16550
- Liu, Q., Q., Liu, P., Sun, Y., and Zhang, C. (2021). Screening mechanism for priority business of multi-station integration project based on random forest algorithm. *Smart Power* 49 (06), 32

Author contributions

KL and YZ contributed to the conception and design of the study. KL and JX wrote parts of the manuscript. HJ gave suggestions to technical support. YH and CP contributed to the improvement of the study design.

Conflict of interest

Authors KL, YZ, JX, and HJ were employed by State Grid Sichuan Information and Communication Company. Author YH was employed by NARI-TECH Nanjing Control Systems Co., Ltd.

The remaining author declares 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.

- Liu, Q. Q., Liu, X., Tang, W., Jin, H., and Yuan, H. (2020). Analysis of fusion technology of grid communication service data. *Inf. Technol.* 44 (03), 153–158. doi:10.13274/j.cnki.hdzj.2020.03.030
- Lu, Z., Yang, Z., Du, C., and He, G. (2020). Research cooperative diagnosis method of substation equipment operation fault based on information fusion technology. *Automation Instrum.* 10, 207–210. doi:10.14016/j.cnki.1001-9227.2020.10.207
- Pan, J., Zhang, F., Wang, L., Zhang, J., and Hao, B. (2022). Research on energy data processing technology based on multi-source heterogeneous. *Electron. Des. Eng.* 30 (16), 143–147. doi:10.14022/j.issn1674-6236.2022.16.031
- Shantanu, K., Ahmed, A., Narottam, D., and Syed, I. (2021). Toward a substation automation system based on IEC 61850. *Electronics* 10 (3), 310. doi:10.3390/electronics10030310
- Shao, X., Kim, C. S., and Sontakke, P. (2020). Accurate deep model for electricity consumption forecasting using multi-channel and multi-scale feature fusion CNN-LSTM. *Energies* 13 (8), 1881. doi:10.3390/en13081881
- Shen, J., Cao, R., Su, C., Cheng, C., Li, X., Wu, Y., et al. (2019). Big data platform architecture and key techniques of power generation scheduling for hydro-thermal-wind-solar hybrid system. *Proc. CSEE* 39 (1), 43–55. doi:10.13334/j.0258-8013.pcsee.181546
- Shui, Y., Lv, L., and You, B. (2013). Application of data fusion to power system intelligent early warning. *East China Electr. Power* 41 (03), 554.
- Suleman, A., Mohammad, H., Haris, M., and Muyeen, S. (2021). Denial-of-service attack on IEC 61850-based substation automation system: A crucial cyber threat towards smart substation pathways. *Sensors* 21 (19), 6415. doi:10.3390/s21196415
- Taha, S., and Suhail, S. M. H. (2020). IEC 61850 Modeling of UPFC and XMPP communication for power management in microgrids. *IEEE Access* 8, 141696–141704. doi:10.1109/access.2020.3013264
- Wang, H., Wang, B., Dong, X., Yao, L., Zhang, R., Ma, F., et al. (2021). Heterogeneous multi-parameter feature-level fusion for multi-source power sensing terminals: Fusion mode, fusion framework and application scenarios. *Nat. Cell. Biol.* 36 (07), 1314–1328. doi:10.1038/s41556-021-00796-6
- Wang, T., Liu, H., Shao, Q., and Yu, B. (2020). Research on intelligent early warning and fault diagnosis technology for the secondary loop of smart substation. *Electr. Meas. Instrum.* 57 (08), 59–63+98. doi:10.19753/j.issn1001-1390.2020.08.010
- Wang, B. B., Wang, H., Yao, L., Dong, X., Ma, H., and Ma, F. (2022). Multi-modal data fusion mode for power system and its key technical issues. *Automation Electr. Power Syst.* 46 (19), 188.
- Wang, Y. Y., Wang, X., Zhang, S., Zheng, G., Zhao, L., and Zheng, G. (2022). Research on efficient knowledge fusion method for heterogeneous big data environments. *Comput. Eng. Appl.* 58 (06), 142.
- Wen, H., and Li, Z. (2018). Simulation of global information fusion for multi-source communication networks. *Comput. Simul.* 35 (01), 188.
- Xia, W., Cai, W., Liu, Y., and Li, H. (2022). Multi-source heterogeneous data fusion of a distribution network based on a joint Kalman filter. *Power Syst. Prot. Control* 50 (10), 180–187. doi:10.19783/j.cnki.pspc.211485
- Xue, Y., and Lai, Y. (2016). Integration of big energy thinking and big data thinking (1) big data and electric power big data. *Automation Electr. Power Syst.* 40 (01), 1. doi:10.7500/AEPS20151208005
- Xyngi, I., and Popov, M. (2010). IEC61850 overview-where protection meets communication.” in 10th IET International Conference on Developments in Power System Protection (DPSP 2010). Managing the Change IET.
- Yang, Y., Zhan, D., Jiang, Y., and Xiong, H. (2021). Reliable multi-Modal learning: a survey. *J. Softw.* 32 (4), 1067–1081.
- Yao, J., and Lai, Y. (2010). The essential cause and technical requirements of the smart grid. *Automation Electr. Power Syst.* 34 (02), 1.
- Yu, H., He, D., Wang, G., Li, J., and Xie, Y. (2020). Data for intelligent decision making. *Acta Autom. Sin.* 46 (05), 878. doi:10.16383/j.aas.c180861
- Zhang, D., Yao, L., and Ma, W. (2013). Development strategies of smart grid in China and abroad. *Proc. CSEE* 33 (31), 1–15. doi:10.13334/j.0258-8013.pcsee.2013.31.001
- Zhao, W., Xiong, N., Ning, N., Li, Y., Gu, Z., and Li, H. (2021). Multi-layer intelligent fault diagnosis method of power grid based on multi-source information fusion. *South. Power Syst. Technol.* 15 (09), 9–15. doi:10.13648/j.cnki.issn1674-0629.2021.09.002
- Zhu, Z., Xu, B., Christoph, B., Tony, Y., and Yu, C. (2017). IEC 61850 configuration solution to distributed intelligence in distribution grid automation. *Energies* 10 (4), 528. doi:10.3390/en10040528