Intelligent Root Cause Analysis and Monitoring System for Smart Grid Full-Link Using Big Data Algorithms

Na Xiao[1*], De Meng[1], Yushan Zhao[1]

1 Platform Operation and Security Center, Information and Communication Branch of State Grid Jibei Electric Power Co., Ltd., Beijing 100053, China

Email Information:

Na Xiao: xnpaper2023@163.com

De Meng: mengde8801@163.com

Yushan Zhao: zzzyyy96@126.com

***Corresponding author, Email: xnpaper2023@163.com, xiao_na69@outlook.com**

Abstract

As the scale and complexity of the smart grid continue to expand, the issue of faults in power grid operations becomes increasingly prominent. This work aims to address the problem of fault localization in power grid operations and improve the efficiency of fault resolution in smart grid systems. By combining practical cases and collecting, processing, and analyzing data from various sources, a data model for the smart grid full-link is established. Multiple algorithms, including neural networks, genetic algorithms, wireless sensor network (WSN), and association rule mining, are comprehensively employed to construct an intelligent root cause analysis and monitoring system for the smart grid full-link. First, a data model for the smart grid full-link is established. Data from all aspects of the smart grid are unified and integrated and then preprocessed through data mining techniques. Then, a combination of neural network models and genetic algorithms is used for intelligent identification and classification of smart grid faults. Different monitoring strategies are implemented based on the types of faults. Furthermore, WSN and association rule mining algorithms are employed to deeply explore the causes of smart grid faults, identify potential root causes, and propose corresponding optimization solutions. The research results demonstrate that the intelligent root cause analysis and monitoring system for the smart grid full-link can effectively improve the accuracy and efficiency of fault diagnosis in the smart grid. The system achieves a localization accuracy of over 90% and reduces fault handling time by more than 50%. The proposed intelligent root cause analysis system for the smart grid full-link enables smart grid companies to quickly and accurately detect faults and conduct root cause analysis and provides strong support for the safe and stable operation of the smart grid.

Keywords: big data algorithm; grid scale; fault location issues; smart grid; wireless sensor network

1. Introduction

With the rapid development of the smart grid, the scale of the power grid is expanding the complexity is increasing, and the fault problem in power grid operation is becoming increasingly prominent. These faults pose a serious threat to the normal operation and reliability of the power supply system. Faults in power grid operations can cause power supply disruptions, revenue loss, equipment damage, safety hazards, grid instability, and cybersecurity threats. These threats can affect businesses, industries, and critical services such as healthcare and transportation. Proactive maintenance, grid modernization, cybersecurity measures, and investment in resilient infrastructure are necessary to address these threats.

Therefore, locating and solving power grid faults quickly and accurately has become one of the key problems to be solved urgently in the field of smart grids [1~3]. To improve the efficiency and accuracy of power grid fault handling, this work aims to carry out research based on big data algorithms and is committed to building a full-link intelligent root cause analysis and monitoring system. By collecting, processing, and analyzing all aspects of data, and using advanced algorithms and technologies, a comprehensive smart grid data model will be established to realize intelligent identification, classification, and root cause analysis of power grid faults [4~6]. Translating smart grid data into actionable recommendations involves analyzing data, identifying areas for improvement, developing recommendations, prioritizing them, planning implementation, monitoring and evaluating progress, and continuously refining strategies. This requires a collaborative approach and a commitment to data-driven decision-making, ultimately leading to optimized grid operations, enhanced customer experiences, and innovation in the smart grid industry. This work will provide strong support for smart grid companies, help them to detect faults, and take corresponding measures to ensure the safe and stable operation of smart grid quickly and accurately. The smart grid data model achieves scalability and adaptability through distributed computing, flexible data architectures, real-time processing, adaptive learning, interoperability standards, modular architecture, and collaborative ecosystems. By leveraging these techniques, the smart grid data model can evolve and scale to meet the demands of a rapidly changing energy landscape while delivering reliable, efficient, and sustainable grid services.

Smart grid is the development trend of modern power systems. It uses advanced scientific and technological means such as information technology, communication technology, and IoT technology to upgrade the traditional power grid into a network with the characteristics of self-repair, efficient operation, interactive service, and prevention of hacker attacks. An intelligent root cause analysis and monitoring system is proposed for the smart grid full-link, which combines various algorithms and achieves a localization accuracy of over 90%, reducing fault handling time by more than 50%. However, with the continuous expansion of the smart grid and the improvement of its complexity, the problem of power grid failure has gradually emerged, which not only affects the stable operation of the power grid but also has a certain impact on the social economy. Therefore, solving these problems and providing an accurate solution has become an important topic people are facing. Big data technology makes it possible to deal with this problem. Big data technology for power grid operations embraces scalability, distributed computing, cloud computing, containerization, real-time processing, AI integration, and open standards. By leveraging these capabilities, organizations can optimize resource utilization, adapt to changing demands, and improve grid reliability, resilience, and performance. It can collect the data of all links, including the operation status of power grid equipment, power load demand, user's electricity consumption behavior, etc., and conduct in-depth excavation and analysis to help people find the root cause of power grid failure. Meanwhile, the combination of big data technology and AI can accurately locate and predict power grid faults more effectively. Full-link intelligent analysis and monitoring system is to build a perfect fault diagnosis and early warning system based on big data and AI. The system can monitor the running state of the power grid in real-time, and once an abnormality occurs, it can quickly identify and locate the cause of the fault, thus making the best emergency treatment plan in time to ensure the stable operation of the power grid. In addition, full-link intelligent analysis can also deeply analyze the operation data of the power grid, find potential risks, and give early warning, thus effectively preventing the occurrence of power grid accidents.

This work adopts a comprehensive method to realize intelligent root cause analysis and monitoring of the whole link of the power grid [7~9]. Firstly, the data model of the whole link of the smart grid is established, the data from all aspects are integrated, and the data mining technology is used for preprocessing. Then, neural networks, genetic algorithms, wireless sensor networks (WSN), and association rule mining are comprehensively used to realize intelligent identification, classification, and root cause analysis of power grid faults. By combining these methods, this work aims to improve the accuracy of power grid fault location and implement corresponding monitoring strategies for different types of faults [10~13].

The research method in this work has important contributions and innovations in the field of smart grid fault analysis. The intelligent analysis and root cause location of power grid faults can be realized by using neural networks, genetic algorithms, WSN, and association rule mining. Finally, the research results show that the proposed intelligent root cause analysis and monitoring system of the smart grid can effectively improve the accuracy and efficiency of fault diagnosis, the positioning accuracy can reach over 90%, and the fault handling time can be reduced by over 50%. This work provides strong support for smart grid companies to detect faults and analyze root causes, which is of great significance to the safe and stable smart grid operation quickly and accurately. Overall, Section 2 provides a detailed literature review of related work. Section 3 presents the methodology used in this work, including the data model for the smart grid full-link, data mining techniques, and multiple algorithms employed in constructing the intelligent root cause analysis and monitoring system. Section 4 presents the research results, including the accuracy and efficiency of fault diagnosis and the reduction of fault handling time. Section 5 discusses the implications of the research and its potential future applications. Finally, Section 6 concludes the paper with a summary of the main contributions and future directions for research.

2. Literature Review

In the past few years, smart grid technology has made remarkable progress and has become an important means to achieve efficient use of energy and sustainable development. With the rise of big data technology, more and more studies focus on how to improve the efficiency and reliability of the power grid through intelligent and data-driven methods. Under this background, this work aims to explore the intelligent root cause analysis and monitoring system of power grid full link based on big data algorithm, to realize comprehensive monitoring of power grid operation state and accurate root cause analysis of problems. Abujubbeh et al. (2019) [13] summarized the application of software-defined WSNs in a smart grid. The main methods include combing and summarizing the relevant literature. The research results showed that the software-defined WSN had a wide application prospect in the smart grid. Dragičević et al. (2019) [14] comprehensively reviewed the future application of wireless networks in smart grids. Research methods include comprehensive analysis and induction of relevant literature. The research results showed that wireless networks were of great significance to the development of smart grid. Manoharan et al. (2020) [15] used the binary logistic regression method to monitor the smart grid through wireless sensors. The research results showed that the monitoring system based on a wireless sensor could effectively improve the performance and reliability of the smart grid. Bagdadee et al. (2020) [16] proposed a WSN based on the Internet of Things (IoT) for power quality control in the smart grid. The research results showed that the system could effectively monitor and control the power quality. Kamruzzaman et al. (2022) [17] proposed an energy-efficient sustainable wireless body area network design using network optimization, smart grid, and renewable energy system. The research results showed that the design could achieve efficient use of energy and sustainable development. Bolurian et al. (2023) [18] proposed a two-tier energy management model considering user behavior, which was applied to the WSN platform in the smart grid. The research results showed that the model could effectively improve the energy management efficiency of the smart grid.

Atitallah et al. (2020) [19] reviewed the methods and future directions of using deep learning and big data analysis of the IoT to support the development of smart cities. The research results showed that deep learning and big data analysis of the IoT had broad application potential in the field of smart cities. Kaffash et al. (2021) [20] analyzed the big data algorithm and its application in intelligent transportation systems (ITS) and made a bibliometrics analysis. The research results showed that big data algorithms had the potential to improve the efficiency and safety of transportation systems. Wang et al. (2022) [21] studied the application of big data analysis in intelligent manufacturing systems to improve production efficiency and quality. Big data optimization methods improve the efficiency of the smart grid data model through enhanced data processing speed, scalability and flexibility, advanced analytics capabilities, real-time monitoring, optimized resource allocation, and continuous improvement. By leveraging these methods, smart grid operators can optimize grid operations, unlock new insights, and improve overall performance and reliability. They comprehensively analyzed and summarized the relevant literature, which revealed the important role of big data analysis. Based on deep learning, Li et al. (2022) [22] conducted a big data analysis on the digital twins of the IoT in smart cities. Deep learning technology could realize efficient processing and intelligent decision-making of massive data in smart cities, and provide support for the development of smart cities. Smart grid operators can enhance cybersecurity by identifying potential risks, integrating security protocols, establishing access controls, deploying monitoring tools, developing incident response plans, providing cybersecurity training, and complying with regulatory requirements. These strategies can help safeguard critical infrastructure and services against cyber threats. Venkatachalam et al. (2022) [23] proposed a diabetes monitoring method based on a deep confidence neural network, combined with big data on the edge of IoT. This method used a deep learning algorithm to analyze and predict large-scale data and could realize accurate diabetes monitoring and management. Garrett et al. (2022) [24] studied the impact of climate change on pathogen outbreaks and used artificial intelligence (AI) to analyze big data to mitigate its impact. By analyzing big data, AI can help predict and control the outbreak of pathogens and provide effective coping strategies for disease prevention and control.

To sum up, the development of smart grid technology provides an important means to realize energy-efficient utilization and sustainable development, while the rise of big data algorithms provides new ideas and methods for smart grid monitoring and root cause analysis. Big data algorithms improve fault handling in power grids by enabling real-time monitoring, predictive maintenance, fault localization, optimization of response procedures, and continuous learning. These algorithms collect and analyze data from various sources, detect anomalies, pinpoint the exact location and nature of grid faults, predict potential equipment failures, and optimize response strategies. By leveraging advanced analytics, machine learning, and predictive modeling techniques, these algorithms empower grid operators to mitigate grid faults more effectively, ultimately improving grid reliability, resilience, and performance.

Therefore, the research on intelligent root cause analysis and monitoring systems of power grid full link based on big data algorithms has important practical significance and theoretical value. To understand significant discoveries, identify key ones, and analyze their practical significance. Consider their effects from multiple perspectives, and explore their practical applications while acknowledging limitations. Lastly, highlight future directions for further research and development.

3. Intelligent root cause analysis and monitoring system of power grid full link based on big data algorithm

3.1 Full link data model construction and big data optimization of smart grid

The research is devoted to building the data model of the whole link of the smart grid and applying the big data optimization method to improve its efficiency [25~28]. The establishment of this data model aims to integrate data from all aspects of the smart grid and preprocess it through data mining technology. Data mining technology is an essential tool in preprocessing integrated data within the smart grid data model. It contributes to data cleaning, integration, feature selection, anomaly detection, pattern recognition, predictive modeling, visualization, and interpretation. By leveraging data mining techniques, stakeholders can unlock the full potential of smart grid data, driving innovation, optimization, and decision-making in the modern energy landscape. The following is the specific construction process: Firstly, the data from various sources involved in the actual case are collected unified, and integrated. The system optimizes its algorithms and predictive models through real-time monitoring, data analytics, predictive modeling, adaptive learning, automated updates, collaboration, and continuous improvement. This allows it to adapt to changing grid conditions and operational requirements, enhancing grid reliability, resilience, and performance. Data from multiple data sources, including sensors, monitoring equipment, and measuring instruments, are integrated into a unified data platform [29, 30]. Next, the big data optimization method is applied to process the collected data. Big data optimization was critical in managing the massive volumes of data generated by smart grid sensors, meters, and devices. Techniques such as scalable and distributed data storage, real-time data processing, parallel processing, data compression, machine learning, resource optimization, and automated data quality assurance were utilized to derive actionable insights, optimize grid operations, and enhance overall performance and reliability in the smart grid ecosystem. This includes data cleaning, denoising, and outlier processing to ensure the quality and accuracy of data. Data cleaning involves identifying and correcting dataset errors, inconsistencies, and inaccuracies. Denoising is the process of removing noise and irrelevant information from a dataset. Outlier processing involves identifying, analyzing, and handling outliers or anomalies in a dataset that deviate significantly from the expected or normal behavior. These concepts are essential for ensuring data quality, accuracy, and reliability in various domains, including smart grid analytics [31, 32]. Through the above steps, the data model of the whole link of the smart grid is successfully built. Integrating and preprocessing data from the smart grid poses several challenges, including data heterogeneity, volume and velocity, quality and consistency, privacy and security, integration and interoperability, governance and ownership, latency and real-time processing, and governance and ethics. Addressing these challenges requires a holistic approach to data management, including robust infrastructure, data governance frameworks, data quality assurance processes, and stakeholder collaboration. Although, this data model not only integrates data from all aspects but also is optimized by big data, which makes it possible to analyze and deal with faults in smart grids more efficiently [33~34]. Big data algorithms are designed to efficiently handle vast volumes of diverse and complex data while extracting valuable insights. Some key competencies include scalability, parallel processing, distributed computing, complex data structures, machine learning and AI capabilities, real-time processing, robustness and fault tolerance, and interoperability and compatibility. These competencies enable organizations to derive valuable insights from large and complex datasets to drive informed decision-making, innovation, and competitive advantage. After constructing the data model of the smart grid with the complete link, a big data algorithm will be applied for deep mining and analysis. Big data algorithms are highly scalable, adaptable, and resilient. They automate data mining and analysis, support complex analytics, and enable real-time streaming data analysis. They can handle diverse data types, including structured, semi-structured, and unstructured. Big data algorithms are essential for extracting valuable insights from smart grid data and optimizing grid operations in the modern energy landscape. Deep learning has several strengths that make it a powerful tool for big data analytics. It can automatically learn hierarchical representations of data features, capture non-linear relationships, scale effectively, handle diverse and dynamic datasets, learn to represent data in high-dimensional feature spaces, support transfer learning, and enable end-to-end learning. These capabilities allow deep learning algorithms to uncover valuable insights and patterns that drive informed decision-making and innovation across various domains and industries. By using techniques such as machine learning, neural networks, and deep learning, these algorithms can extract key information from huge data and identify potential power grid fault modes and trends. The goal of identifying potential power grid fault modes and trends through big data algorithms is to enhance power grid operations' reliability, resilience, and efficiency while mitigating risks. Big data algorithms enable proactive identification of fault modes, detection of anomalies and abnormal patterns, prediction of fault occurrence, identification of trends and patterns, and optimization of maintenance and asset management. The objective is to empower operators with actionable insights, early warning capabilities, and decision support tools to anticipate, prevent, and mitigate the impact of faults on power grid operations. Furthermore, predictive analysis can be used to predict possible future faults and to take preventive measures in advance, which greatly improves the operation efficiency and safety of the power grid. Organizations can implement proactive measures based on predictive analysis insights to prevent future faults. These measures include developing preventive maintenance programs, optimizing operational parameters, and enhancing redundancy and resilience within the system. Personnel should receive training to enhance their competence in detecting and addressing potential faults. Real-time monitoring systems and alerting mechanisms should be deployed for timely intervention or corrective action. Collaboration with partners can foster innovation in predictive maintenance technologies and practices. The overall structure of the full link data model of a smart grid is analyzed, and the model structure diagram is shown in Figure 1:

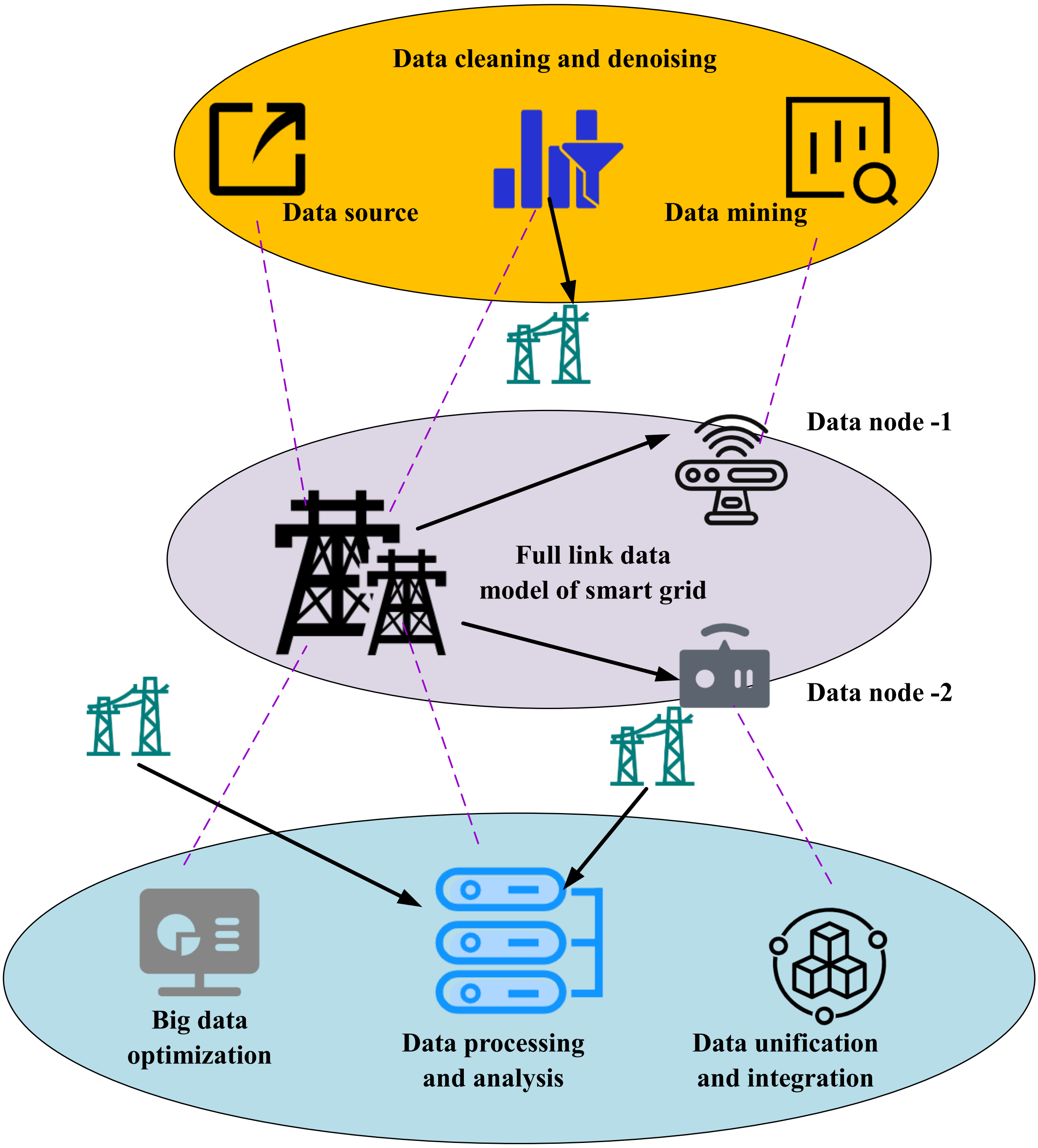

Figure 1 Structure diagram of full link data model of smart grid

3.2 Fault identification and intelligent root cause analysis of smart grid

By combining a neural network model and genetic algorithm, intelligent fault identification and classification of smart grid systems are realized, and root cause analysis is further carried out [35~39]. In the stage of fault identification, a model based on a neural network is constructed, and its powerful learning ability and feature extraction ability are used to identify different types of faults. To identify faults in a power grid, data is collected from various sources and pre-processed to ensure quality. Relevant features are extracted, and a machine-learning model is trained and evaluated using the data.

The model's performance is optimized and deployed in a real-time monitoring system for fault detection [40~42]. By training the model and using the marked data set for supervised learning, the model can accurately predict the fault types in the power grid [43,44]. In addition, in the intelligent root cause analysis and monitoring system of the power grid full link based on a big data algorithm, the decision function of the support vector machine (SVM) is calculated as shown in equation (1):

$$f(x) = \operatorname{sign}\left(\sum_{i=1}^{N_s} y_i \alpha_i k(x_i, x) + b\right) \tag{1}$$

x represents the input vector, y_i is the output label of the ith training sample, $k(x_i, x)$ is the kernel function, and α_i and b are the parameters learned by the big data algorithm. The cost coefficient of the regression model is calculated, and the equation is as follows:

$$J(\theta) = \frac{1}{2m}\sum_{i=1}^{m} \left(h_\theta\left(x^{(i)}\right) - y^{(i)}\right)^2 + \frac{\lambda}{2m}\sum_{j=1}^{n} \theta_j^2 \tag{2}$$

m is the number of samples. n is the number of features. $\left(x^{(i)}\right)$ and $y^{(i)}$ are the features and labels of the ith sample. h_θ is the prediction function learned from the regression model. θ_j^2 is the parameter of the regression model, and λ is the regularization parameter. In addition, the calculation equation of classifier function $P(Y = k \mid \mathbf{x})$ of random forest in the process of ensemble learning is as follows:

$$P(Y = k \mid \mathbf{x}) = \frac{1}{N}\sum_{i=1}^{N} P(Y = k \mid \mathbf{x}; \theta_i) \tag{3}$$

k is the category label. $\mathbf{x}$ is the input vector. θ_i is the parameter of different decision trees in a random forest. $P(Y = k \mid \mathbf{x}; \theta_i)$ indicates the probability that the i-th decision tree is classified as k for a given input vector $\mathbf{x}$. The objective function value $J(C, \mu)$ of the big data algorithm and the smart link of the power grid is calculated, and the equation is as follows:

$$J(C, \mu) = \sum_{i=1}^{m} \min_{j=1}^{k} \left\|x^{(i)} - \mu_j\right\|^2 \tag{4}$$

C is the cluster set to which the sample points belong. μ is the centroid of each cluster, k is the number of clusters, and m is the number of samples. The process of smart grid fault identification and intelligent root cause analysis is sorted out, and the structure is shown in Figure 2:

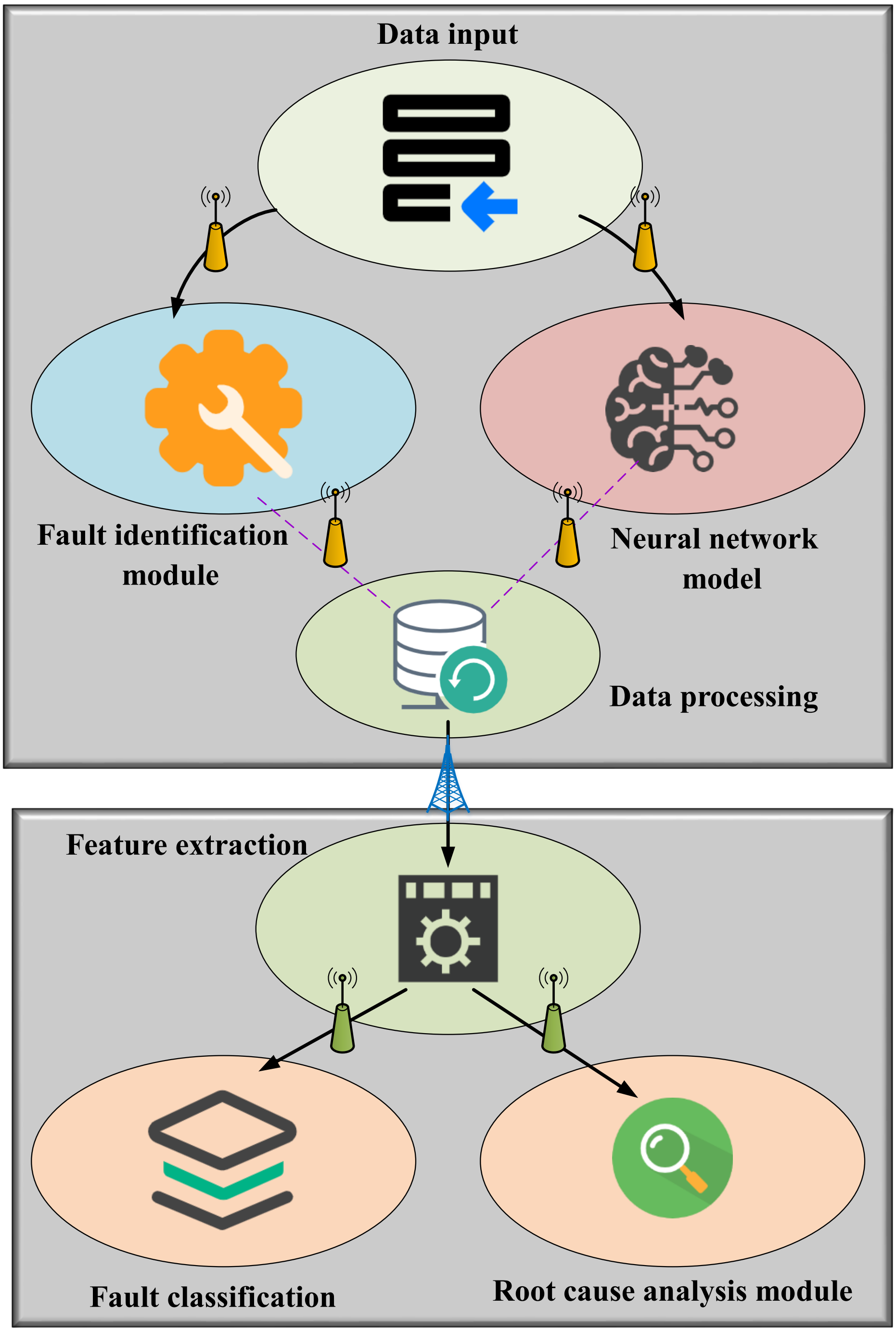

Figure 2 Flow chart of smart grid fault identification and intelligent root cause analysis

3.3 WSN and construction of intelligent monitoring system

This section will focus on the specific construction process of WSN and intelligent monitoring systems. First, appropriate sensor nodes are selected and deployed in key positions in the smart grid system to ensure the comprehensiveness and accuracy of monitoring [45~47]. Next, these sensor nodes are connected to a network through wireless communication technology. To simulate the real intelligent monitoring system, appropriate communication protocols, and data transmission mechanisms are designed to ensure that sensor nodes can stably transmit data to the main control center. To ensure stable data transmission from sensor nodes to the main control center in a power grid system, implementing robust data transmission mechanisms is essential. These mechanisms include reliable protocols, message queuing, quality of service, error detection and correction, redundant paths, real-time monitoring and diagnostics, adaptive transmission control, and security measures. By integrating these mechanisms into the data transmission infrastructure, power grid systems can achieve stable and reliable transmission of sensor data, enabling efficient monitoring, control, and management of the grid infrastructure. Choosing the appropriate protocol is crucial when designing communication protocols and data transmission mechanisms for an intelligent monitoring system. Data volume, latency, reliability, and security should be considered. Network architecture and topology should also be optimized for communication efficiency and scalability. Incorporating error detection and correction mechanisms and security measures such as encryption and authentication are also important. By carefully considering these factors, the system can achieve reliable and secure communication for effective monitoring and control. Meanwhile, efficient data compression and encryption algorithms are adopted to reduce the load of data transmission and protect the security of data. Data compression and encryption algorithms are crucial for reducing transmission load and protecting data in power grid systems. Data compression algorithms reduce the size of data files or streams to minimize bandwidth and storage space. Lossless and lossy compression techniques are used to compress data without losing any information. Data encryption transforms plaintext into ciphertext using cryptographic algorithms to ensure confidentiality and protect sensitive information from unauthorized access or interception. Symmetric and asymmetric encryption techniques are used for encryption. Compression and encryption algorithms are integrated into communication protocols and data transmission frameworks to provide end-to-end data security and efficiency. Data compression reduces the amount of data transmitted by eliminating redundant or unnecessary information. Lossless compression preserves all original data, while lossy compression sacrifices some detail. Data compression improves data transmission efficiency, reduces network congestion, and lowers resource utilization. However, it can increase computational overhead and may result in loss of data quality. The process of establishing WSN and intelligent monitoring system is analyzed, and the structure is shown in Figure 3:

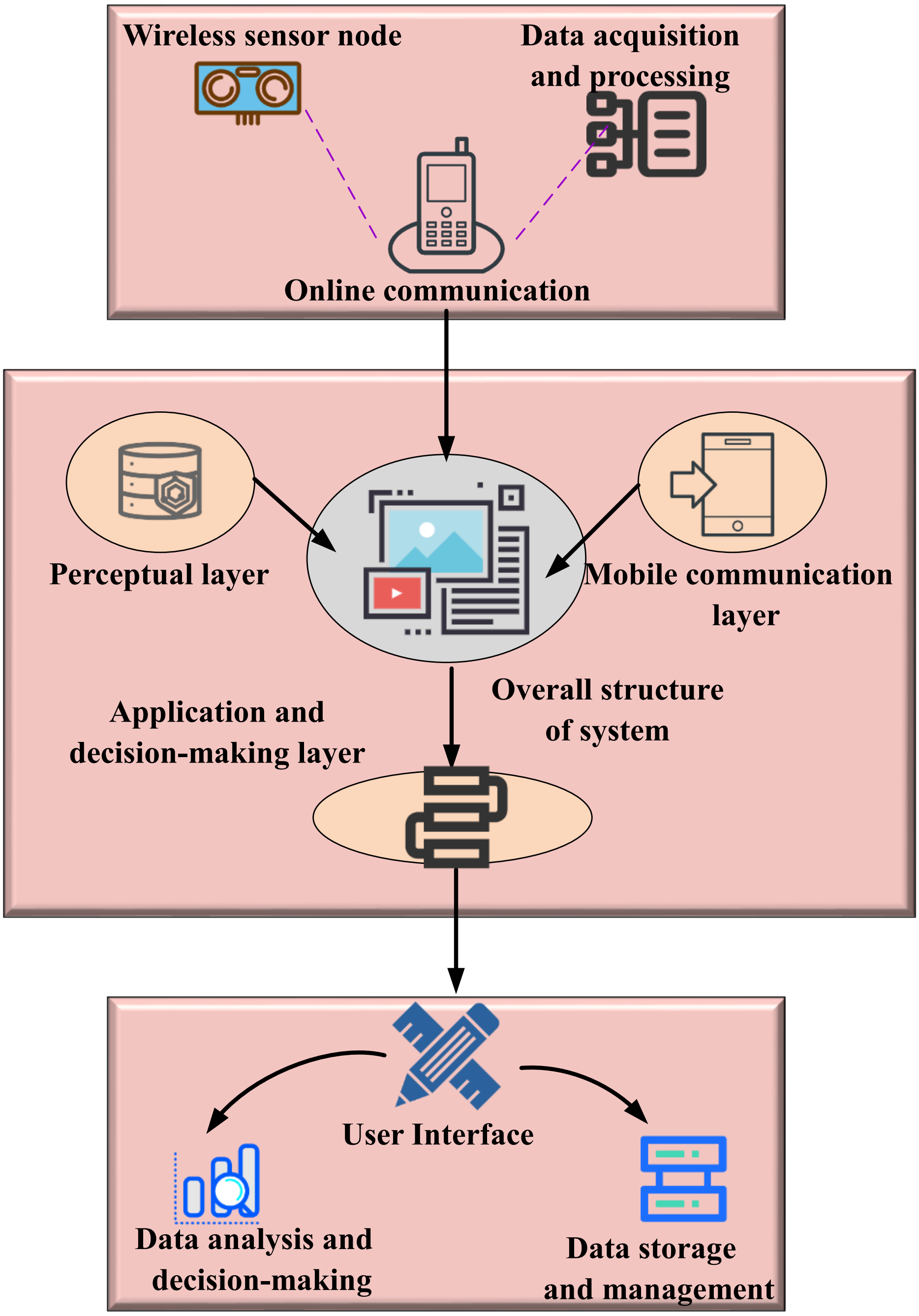

Figure 3 Structure diagram of WSN and intelligent monitoring system

In addition, the proposed model is the intelligent root cause analysis method of the power grid full link based on a big data algorithm (IRCA). To evaluate the performance of the model, the model is compared with the power grid diagnosis method based on machine learning (PGG) and the power grid fault management method based on knowledge map (PGFM). The performance is compared from the aspects of fault location accuracy, fault processing time, monitoring system coverage, root cause analysis efficiency, fault diagnosis accuracy, and system stability evaluation. Fault diagnosis and monitoring systems in power grid environments can be evaluated based on six specific criteria: root cause analysis efficiency, fault diagnosis accuracy, monitoring system coverage, fault location accuracy, fault processing time, and system stability evaluation. These criteria help assess the systems' effectiveness, reliability, and efficiency in identifying, analyzing, and mitigating faults and disturbances in the grid infrastructure.

4. Result and discussion

4.1 Fault handling performance analysis of different power grid link analysis methods

The fault handling performance of different power network link analysis methods is studied and compared, and a detailed analysis is made. To better understand the performance characteristics of these methods, the data change curves from Figure 4 to Figure 6 are provided.

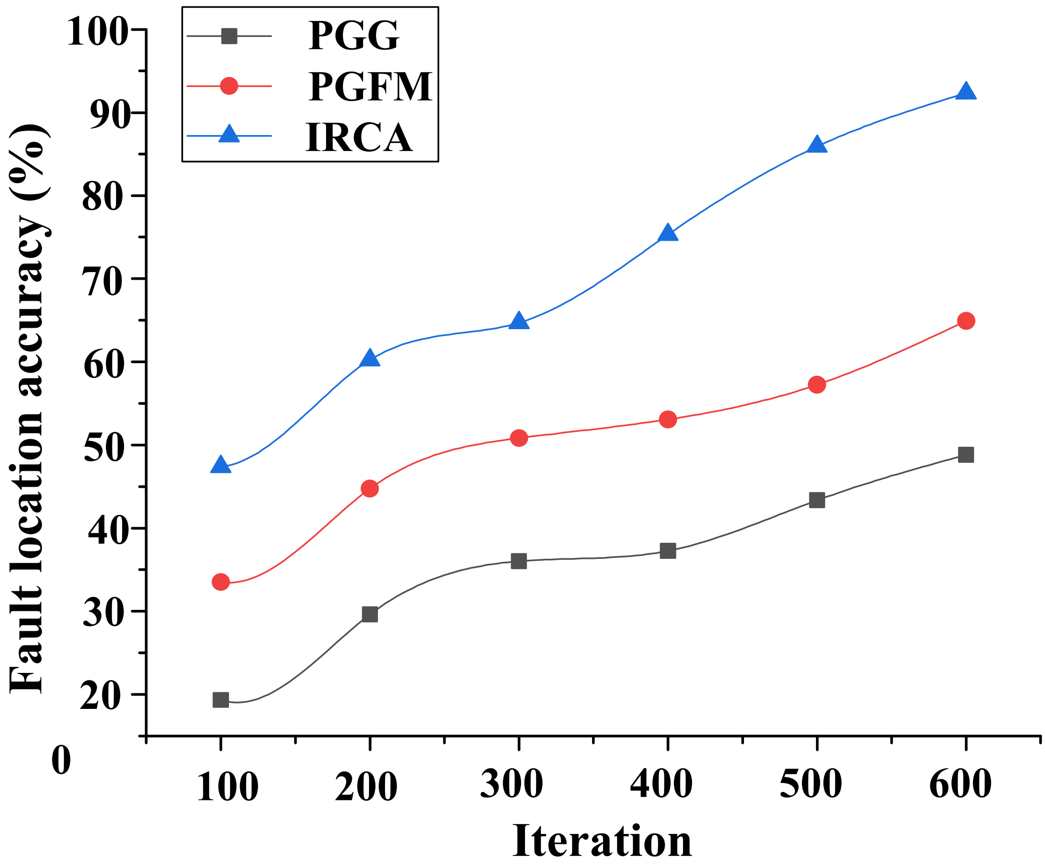

Figure 4 Variation curves of fault location accuracy data of different power grid link analysis methods

It shows that the fault location accuracy of different power grid link analysis methods is gradually improved with the increase of iteration times. At the initial stage, the accuracy of the three methods is low, but with the increase of iteration times, their accuracy is gradually improved. After 600 iterations of the model, the fault location accuracy of the three methods is PGG: 48.81%, PGFM: 64.94%, and IRCA: 92.35% respectively. It shows that the IRCA method shows the highest accuracy in the whole iteration process, and gradually surpasses the other two methods. The Intelligent Root Cause Analysis (IRCA) method uses advanced machine learning techniques to analyze large-scale datasets and extract meaningful insights. IRCA leverages a comprehensive set of features and data sources to inform its analysis and incorporates adaptive learning mechanisms to improve its performance over time. IRCA prioritizes interpretability and explainability in its analysis results, enhancing user understanding and confidence in its findings.

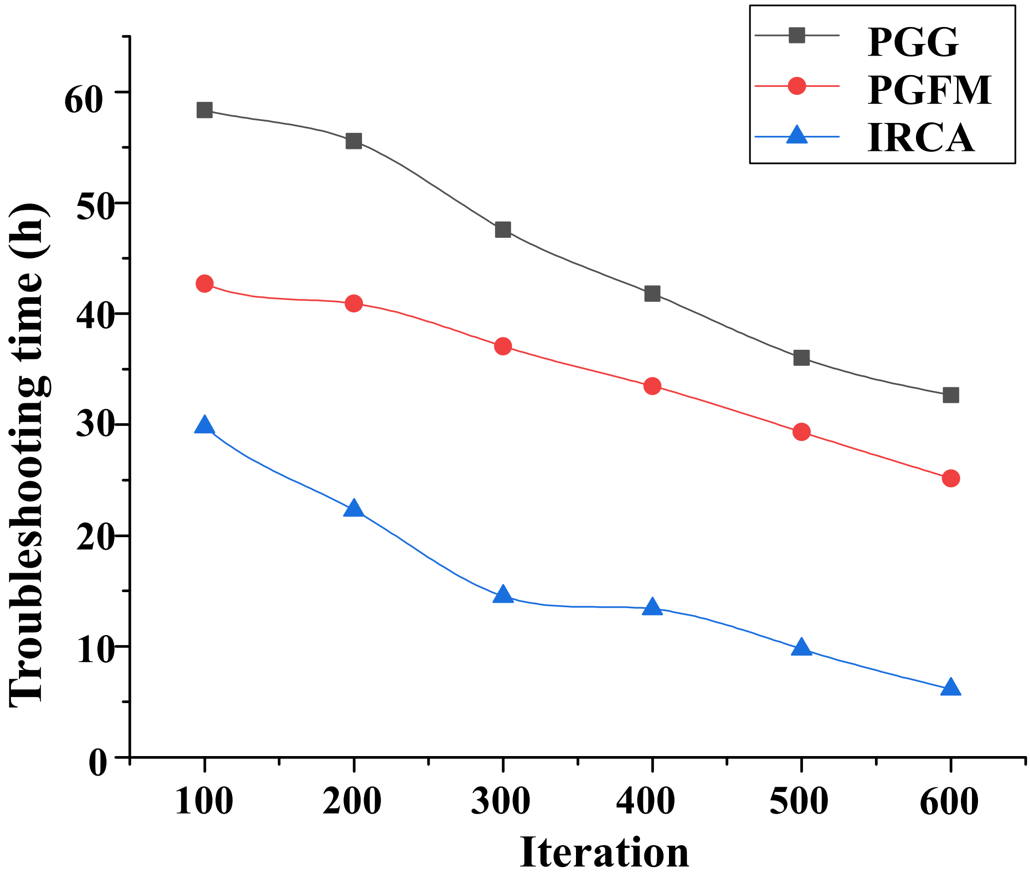

Figure 5 Variation curves of fault handling time data of different power grid link analysis methods

According to the data change trend in Figure 5, it can be observed that the fault handling time of different power grid link analysis methods is gradually decreasing with the increase of iteration times. Algorithm development is an iterative process that involves testing, evaluation, and adjustment. During the initial phase, prolonged iteration time is common as developers experiment with different parameters, algorithms, or optimization techniques. As the algorithm converges towards an optimal solution through successive iterations, processing time decreases across the three methods, improving efficiency and effectiveness. This translates to faster task execution, shorter turnaround times, and enhanced scalability, making the algorithm more suitable for real-time or large-scale applications. In the initial stage, the processing time of the three methods is high, but with the increase in iteration times, their processing time gradually decreases. The high processing time in fault diagnosis methods is due to several reasons, such as extensive search space exploration, computational complexity, data preprocessing, algorithm initialization, and resource constraints. As iteration times increase, processing time typically decreases. Since the 200th iteration, the processing time of the three methods has decreased. Finally, after 600 iterations, the fault handling time of the three methods is PGG: 32.67h, PGFM: 25.16h, and IRCA: 6.16h respectively. It shows that the IRCA method shows the shortest processing time in the whole iteration process, and is gradually superior to the other two methods. While the IRCA method benefits from shorter processing times, it faces potential obstacles that need consideration. These include a trade-off between accuracy and speed, limited exploration of solution space, complexity of data, resource constraints, and dynamic system conditions. Striking the right balance between speed and accuracy is critical to ensure that IRCA delivers reliable and actionable fault diagnosis and mitigation insights.

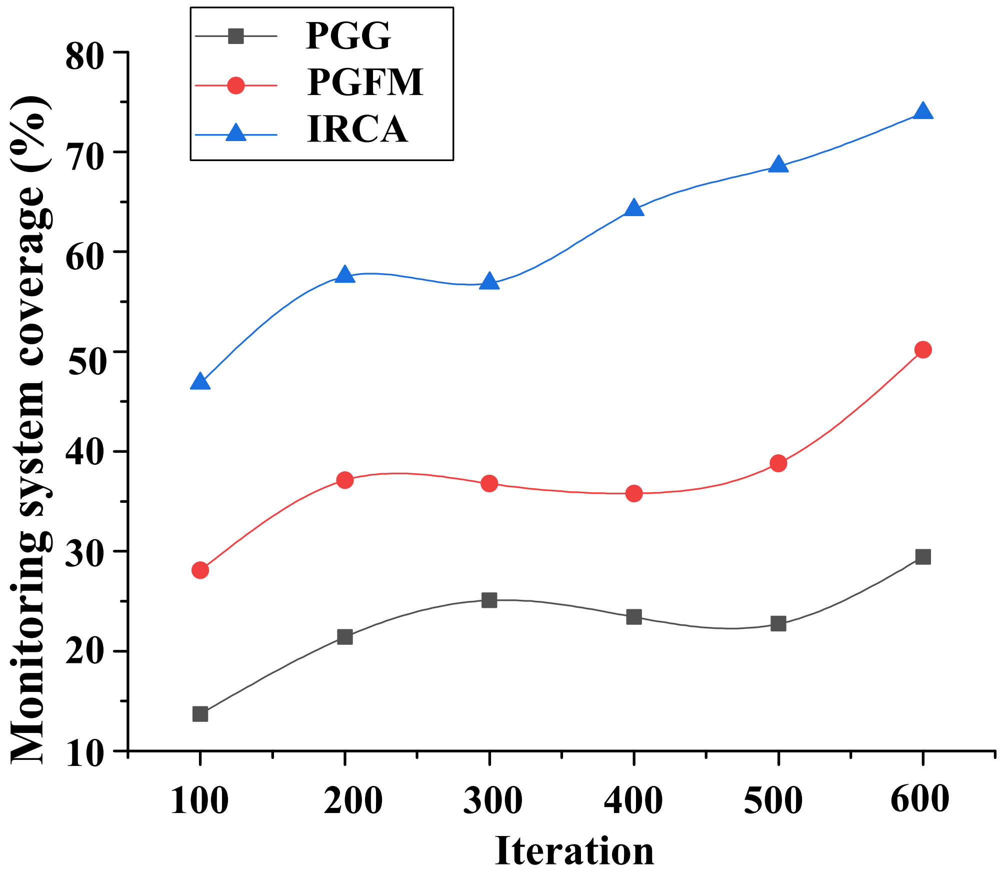

Figure 6 Variation curves of monitoring system coverage data with different power grid link analysis methods

Figure 6 shows the changing trend of fault diagnosis accuracy data of different power grid link analysis methods. According to the data observation, it shows that the accuracy of fault diagnosis of the three methods shows different trends with the increase of iteration times. In the first 200 iterations, the accuracy of the three methods has been significantly improved. The accuracy of the PGG method increased to 21.40%, the PGFM method increased to 37.12%, and the IRCA method reached 57.52%. Finally, after 600 iterations, the fault diagnosis accuracy rates of the three methods are PGG: 29.43%, PGFM: 50.17%, and IRCA: 73.91% respectively. To sum up, according to the data trend analysis in Figure 6, with the increase of iteration times, the fault diagnosis accuracy rate of different power grid link analysis methods shows different trends, and the IRCA method shows a relatively high accuracy rate in the whole iteration process. Data trend analysis involves examining historical data to identify patterns, trends, and anomalies over time. It considers the temporal context of data and uses statistical techniques like time series analysis, regression analysis, and visualization tools like line charts, scatter plots, and heat maps. Data trend analysis examines long-term trends, seasonal variations, cyclic patterns, anomaly detection, forecasting, and continuous monitoring. Organizations can make informed decisions and optimize operations in various domains by understanding the principles and characteristics of data trend analysis.

4.2 Root cause analysis efficiency and monitoring performance analysis of different power grid link analysis methods

In this section, the root cause analysis efficiency and monitoring performance of different power grid link analysis methods are analyzed. The efficiency of Root Cause Analysis (RCA) is analyzed by evaluating the accuracy of root cause identification, time efficiency, effectiveness of analysis techniques, and feedback mechanisms. Power Grid Link Analysis Techniques are monitored using topology, flow, dynamic line rating, and cybersecurity monitoring. These techniques help operators proactively identify and address issues, optimize grid performance, and ensure the reliability and resilience of the power grid infrastructure. To evaluate the performance of these methods in diagnosing and maintaining power grid faults, Figure 7 to Figure 9 are drawn to show the changing trend of relevant data.

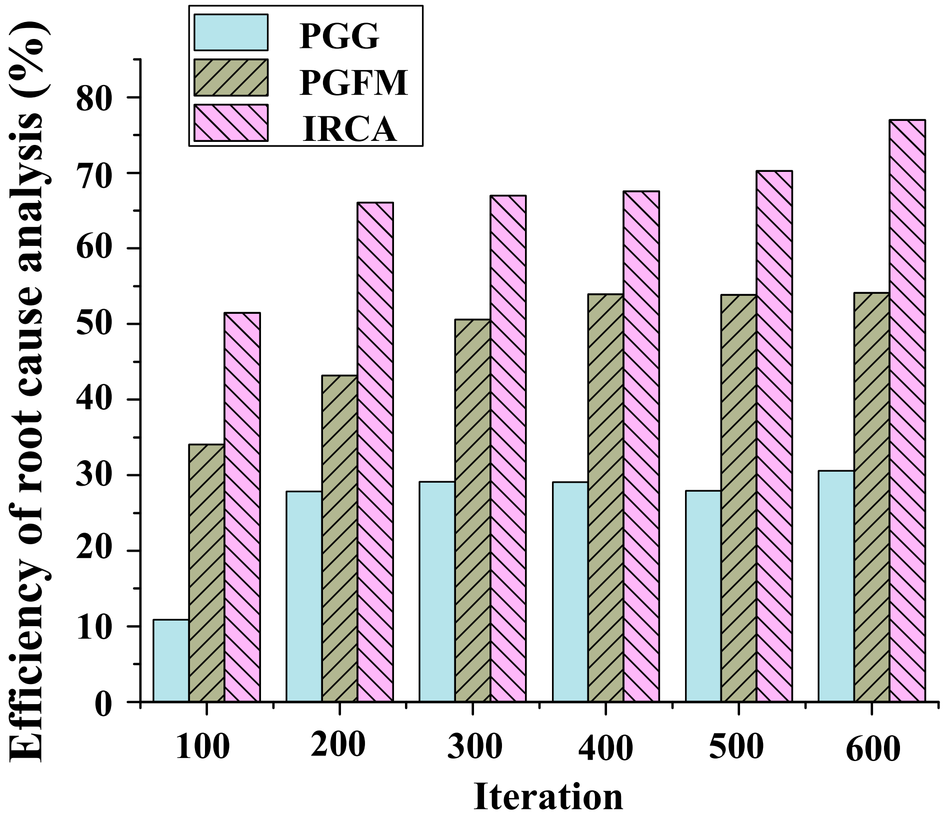

Figure 7 Variation trend of root cause analysis efficiency data of different power grid link analysis methods

According to the data change trend in Figure 7, it can be observed that the root cause analysis efficiency of different power grid link analysis methods presents different change trends with the increase of iteration times. In the initial stage, the root cause analysis efficiency of the PGG method is low, at about 10.88%, the PGFM method is slightly higher, at about 34.07%, and the IRCA method achieves an efficiency of 51.49%. The efficiency of the Parallel Genetic Algorithm (PGG) method for root cause analysis may seem low initially due to several factors. These include the exploratory nature of genetic algorithms, the quality and diversity of the initial population, computational overhead, and the need for careful parameter tuning. However, with sufficient iteration and tuning, PGG can improve its efficiency and converge towards optimal solutions for root cause analysis tasks in big data analytics applications. With the increase in iteration times, the efficiency of the three methods is gradually improved. Finally, after 600 iterations, the root cause analysis efficiencies of the three methods are PGG: 30.58%, PGFM: 54.12%, and IRCA: 76.96% respectively. To sum up, according to the data trend analysis in Figure 7, with the increase of iteration times, the root cause analysis efficiency of different power grid link analysis methods presents different trends, and the IRCA method shows relatively high efficiency in the whole iteration process.

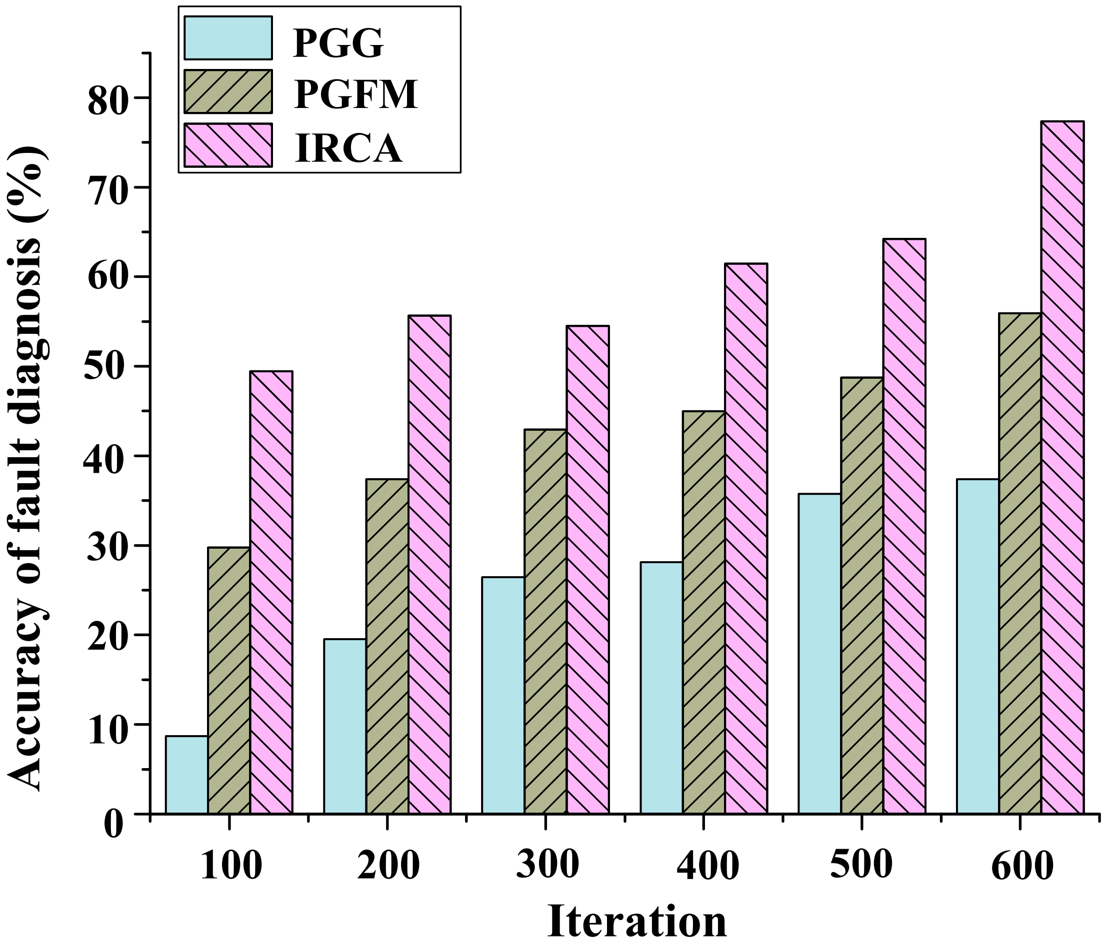

Figure 8 Variation trend of fault diagnosis accuracy data of different power grid link analysis methods

According to the data change trend in Figure 8, it can be observed that the fault diagnosis accuracy rate of different power grid link analysis methods presents different change trends with the increase of iteration times. In the initial stage, the accuracy rate of fault diagnosis of the PGG method is the lowest, about 8.73%, and that of the PGFM method is slightly higher, at about 29.78%, while that of the IRCA method reaches 49.44%. With the increase in iteration times, the accuracy of the three methods is gradually improved. In the first 200 iterations, the accuracy of the three methods has been significantly improved. After 500 iterations, the accuracy of the three methods rises again. The accuracy of the PGG method is 35.76%, the PGFM method is 48.73%, and the IRCA method is 64.20%. With the increase in iteration times, the accuracy of fault diagnosis of different power grid link analysis methods shows different trends, and the IRCA method shows relatively high accuracy in the whole iteration process.

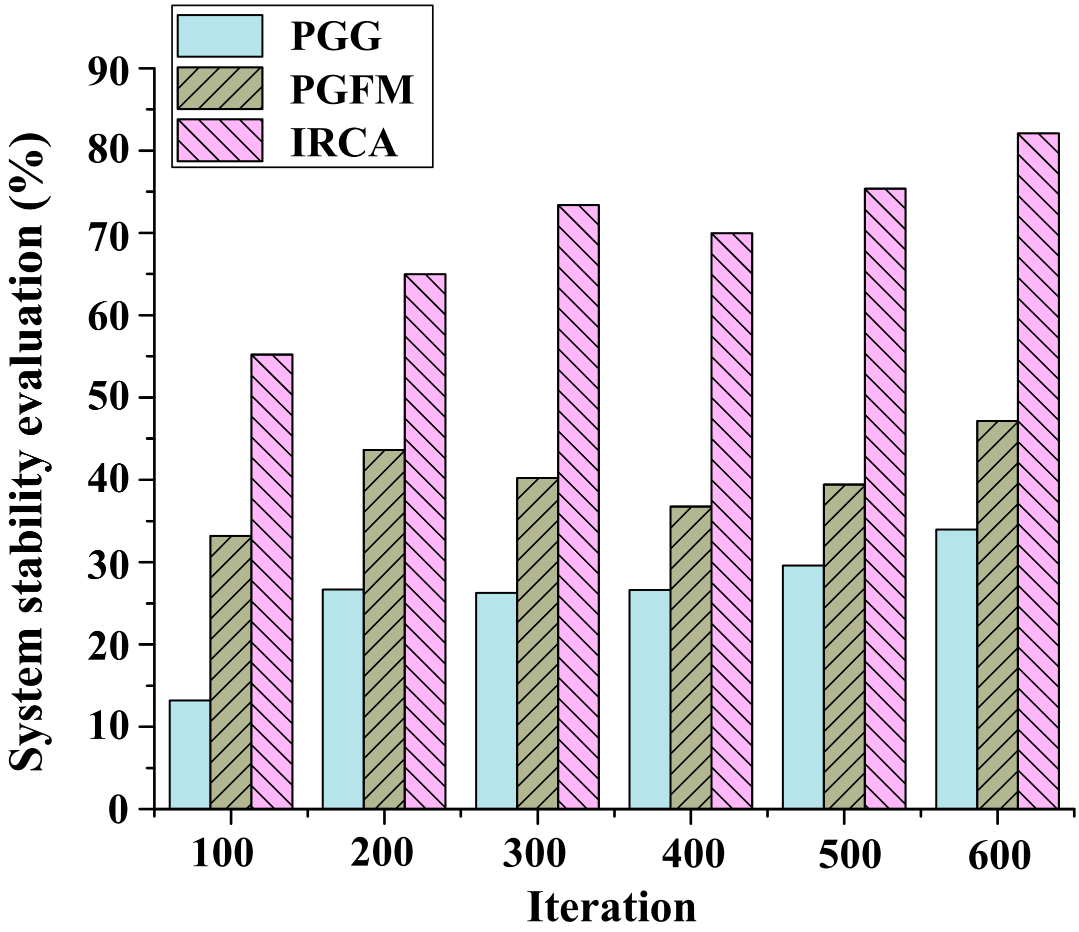

Figure 9 Variation trend of system stability evaluation data of different power grid link analysis methods

Through the data trend analysis in Figure 9, it can be observed that the fault diagnosis accuracy of different power grid link analysis methods shows different changes with the increase of iteration times. In the initial stage, the accuracy of fault diagnosis of the PGG method is about 13.19%, that of the PGFM method is slightly higher than 33.18%, and that of the IRCA method is 55.22%. With the increase in iteration times, the accuracy of the three methods is gradually improved. Finally, after 600 iterations, the fault diagnosis accuracy rates of the three methods are PGG: 33.96%, PGFM: 47.18%, and IRCA: 82.07% respectively. To sum up, according to the data trend analysis in Figure 9, with the increase of iteration times, the fault diagnosis accuracy rate of different power grid link analysis methods shows different trends, and the IRCA method shows a relatively high accuracy rate in the whole iteration process.

4.3 Discussion

In this paper, the fault handling performance of different power grid connection analysis methods is studied, and three analysis methods PGG, PGFM, and IRCA are compared. Firstly, the accuracy, processing time, and data changes of monitoring system coverage of each method are analyzed. The results show that the IRCA method is superior to the other two methods in accuracy and processing time, and its fault location accuracy and fault processing speed increase with the increase of iteration times. It is superior to PGG and PGFM. In addition, the accuracy of fault diagnosis of these three methods also increases with the increase of iteration times, and the IRCA method also shows relatively high diagnostic accuracy in the whole iteration process. Three methods, namely PGG, PGFM, and IRCA, improve fault location accuracy through iterative refinement and learning. These methods explore solution spaces, optimize model parameters, and learn from past data patterns to gradually converge toward more accurate fault localization results. Increasing iteration times enable these methods to accumulate more training data, refine their feature selection process, and optimize their machine learning models, leading to improved fault localization accuracy over time. In addition, by comparing the research results of this paper with previous studies, Vaccaro et al. (2019) [48~49] introduced the advantages and disadvantages, application scenarios, and development trends of various information processing methods from five aspects: data collection, data transmission, data storage, data analysis, and data visualization. It was pointed out that the information processing method was one of the core technologies of the smart grid, which plays an important role in improving the efficiency, reliability, and security of the power grid. Li et al. (2020) [50] used the distributed, tamper-proof, and decentralized characteristics of blockchain to build a power consumption data-sharing platform, which realized the safe transmission and storage of power consumption data. In this paper, an anomaly detection algorithm based on deep learning was used to analyze and identify the anomalies in electricity consumption data. The effectiveness and superiority of this method were verified by experiments, which shows that this method can improve the ability of anomaly detection and attack resistance of smart grids. Kimani (2019) [51] analyzed the security threats and vulnerabilities of smart grid networks from four aspects: IoT equipment, communication protocol, network architecture, and network management. The research put forward some network security protection measures and suggestions, including strengthening the identity authentication, encryption, and update of IoT devices, adopting secure communication protocols and standards, designing multi-level network architecture and topology, and implementing effective network monitoring and auditing.

Based on the above research results, this paper compares and analyzes the fault-handling performance of different power grid connection analysis methods. IRCA method shows better performance in accuracy, processing time, and fault diagnosis accuracy, which is better than PGG and PGFM methods. Compared with other references, this paper focuses on the analysis of fault handling efficiency and monitoring performance of different methods and provides experimental data and trend analysis. These research results have guiding significance for optimizing power grid fault handling and improving power grid reliability and security.

5. Conclusion

In this work, a comprehensive method is adopted to realize the intelligent root cause analysis and monitoring of the whole link of the power grid. Firstly, the data model of the whole link of the smart grid is established, the data from all aspects are integrated, and the data mining technology is used for preprocessing. Then, neural networks, genetic algorithms, WSNs, and association rule mining are comprehensively used to realize intelligent identification, classification, and root cause analysis of power grid faults. During the experiment, the actual cases are collected, processed, and analyzed, and the effect of the proposed system is verified. The research results show that the proposed intelligent root cause analysis and monitoring system of the smart grid can effectively improve the accuracy and efficiency of smart grid fault diagnosis. The system achieves more than 90% positioning accuracy and reduces the fault handling time by more than 50%. In addition, with the increase of iteration times, the accuracy of fault diagnosis of different power network link analysis methods shows different trends, and the IRCA method shows relatively high accuracy in the whole iteration process. However, there are some shortcomings in the research. Although many algorithms are comprehensively used in this work, there may still be other more effective algorithms that have not been considered. Future research can further optimize the algorithm and technology, expand the application fields, meet the challenges of power grid scale expansion and complexity increase, and promote the sustainable development of smart grid.

Declarations

Funding

No funds or grants were received by any of the authors.

Conflict of interest

There is no conflict of interest among the authors.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Code Availability

Not applicable.

Author’s contributions

All Authors contributed to the design and methodology of this study, the assessment of the outcomes, and the writing of the manuscript.

References

[1] Butt O M, Zulqarnain M, Butt T M. Recent advancement in smart grid technology: Prospects in the

electrical power network. Ain Shams Engineering Journal, 2021, 12(1), pp. 687-695.

[2] Omitaomu O A, Niu H. Artificial intelligence techniques in smart grid: A survey. Smart Cities, 2021, 4(2), pp.548-568.

[3] Papadimitrakis M, Giamarelos N, Stogiannos M, et al. Metaheuristic search in smart grid: A review with emphasis on planning, scheduling, and power flow optimization applications. Renewable and Sustainable Energy Reviews, 2021, 145, pp.111072.

[4] Nguyen T N, Liu B H, Nguyen N P, et al. Smart grid vulnerability and defense analysis under cascading failure attacks. IEEE Transactions on Power Delivery, 2021, 36(4), pp.2264-2273.

[5] Abrahamsen F E, Ai Y, Cheffena M. Communication technologies for smart grid: A comprehensive survey. Sensors, 2021, 21(23), pp.8087.

[6] Khalid H, Shobole A. Existing developments in adaptive smart grid protection: A review. Electric Power Systems Research, 2021, 191, pp.106901.

[7] Avancini D B, Rodrigues J J P C, Rabêlo R A L, et al. A new IoT-based smart energy meter for smart grids. International Journal of Energy Research, 2021, 45(1), pp. 189-202.

[8] Bandeiras F, Gomes Á, Gomes M, et al. Exploring Energy Trading Markets in Smart Grid and Microgrid Systems and Their Implications for Sustainability in Smart Cities. Energies, 2023, 16(2), pp.801.

[9] Ghiasi M, Wang Z, Mehrandezh M, et al. Evolution of smart grids towards the Internet of energy: Concept and essential components for deep decarbonization. IET Smart Grid, 2023, 6(1), pp. 86-102.

[10] Talaat M, Alsayyari A S, Alblawi A, et al. Hybrid-cloud-based data processing for power system monitoring in smart grids. Sustainable Cities and Society, 2020, 55, pp.102049.

[11] Wang H, Huang Z, Zhang X, et al. Intelligent power grid monitoring and management strategy using 3D model visual computation with deep learning. Energy Reports, 2022, 8, pp.3636-3648.

[12] Zhu Y, Dong Z, Cheng Z, et al. The neural network extended state-observer for energy system monitoring. Energy, 2023, 263, pp. 125736.

[13] Abujubbeh M, Al-Turjman F, Fahrioglu M. Software-defined wireless sensor networks in smart grids: An overview. Sustainable Cities and Society, 2019, 51, pp.101754.

[14] Dragičević T, Siano P, Prabaharan S R S. Future generation 5G wireless networks for smart grid: A comprehensive review. Energies, 2019, 12(11), pp. 2140.

[15] Manoharan H, Teekaraman Y, Kirpichnikova I, et al. Smart grid monitoring by wireless sensors using binary logistic regression. Energies, 2020, 13(15), pp. 3974.

[16] Bagdadee A H, Hoque M Z, Zhang L. IoT-based wireless sensor network for power quality control in smart grid. Procedia computer science, 2020, 167, pp.1148-1160.

[17] Kamruzzaman M M, Alruwaili O. Energy efficient sustainable wireless body area network design using network optimization with smart grid and renewable energy systems. Energy Reports, 2022, 8, pp.3780-3788.

[18] Bolurian A, Akbari H, Mousavi S, et al. Bi-level energy management model for the smart grid considering customer behavior in the wireless sensor network platform. Sustainable Cities and

Society, 2023, 88, pp.104281.
[19] Atitallah S B, Driss M, Boulila W, et al. Leveraging Deep Learning and IoT big data analytics to support the smart cities development: Review and future directions. Computer Science Review, 2020, 38, pp.100303.
[20] Kaffash S, Nguyen A T, Zhu J. Big data algorithms and applications in intelligent transportation system: A review and bibliometric analysis. International Journal of Production Economics, 2021, 231, pp.107868.
[21] Wang J, Xu C, Zhang J, et al. Big data analytics for intelligent manufacturing systems: A review. Journal of Manufacturing Systems, 2022, 62, pp.738-752.
[22] Li X, Liu H, Wang W, et al. Big data analysis of the Internet of Things in the digital twins of smart cities based on deep learning. Future Generation Computer Systems, 2022, 128, pp.167-177.
[23] Venkatachalam K, Prabu P, Alluhaidan A S, et al. Deep belief neural network for 5G diabetes monitoring in big data on edge IoT. Mobile Networks and Applications, 2022, 27(3), pp. 1060-1069.
[24] Garrett K A, Bebber D P, Etherton B A, et al. Climate change effects on pathogen emergence: Artificial intelligence to translate big data for mitigation. Annual Review of Phytopathology, 2022, 60, pp.357-378.
[25] Zhang Y, Shi X, Zhang H, et al. Review on deep learning applications in frequency analysis and control of modern power systems. International Journal of Electrical Power & Energy Systems, 2022, 136, pp.107744.
[26] Ammari C, Belatrache D, Touhami B, et al. Sizing, optimization, control and energy management of hybrid renewable energy system—A review. Energy and Built Environment, 2022, 3(4), pp. 399-411.
[27] Hoang A T, Nguyen X P. Integrating renewable sources into energy system for the smart city as a sagacious strategy towards the clean and sustainable process. Journal of Cleaner Production, 2021, 305, pp. 127161.
[28] Mohamad F, Teh J, Lai C M. Optimum allocation of battery energy storage systems for power grid enhanced with solar energy. Energy, 2021, 223, pp. 120105.
[29] Shadfar H, Ghorbani Pashakolaei M, Akbari Foroud A. Solid-state transformers: An overview of the concept, topology, and its applications in the smart grid. International Transactions on Electrical Energy Systems, 2021, 31(9), pp. e12996.
[30] Mohapatra H, Rath A K. A fault-tolerant routing scheme for advanced metering infrastructure: an approach towards smart grid. Cluster computing, 2021, 24(3), pp.2193-2211.
[31] Mostafa N, Ramadan H S M, Elfarouk O. Renewable energy management in smart grids by using big data analytics and machine learning. Machine Learning with Applications, 2022, 9, pp.100363.
[32] Bhattacharya S, Chengoden R, Srivastava G, et al. Incentive mechanisms for smart grid: state of the art, challenges, open issues, future directions. Big Data and Cognitive Computing, 2022, 6(2), pp.47.
[33] Zidi S, Mihoub A, Qaisar S M, et al. Theft detection dataset for benchmarking and machine

learning-based classification in a smart grid environment. Journal of King Saud University-Computer and Information Sciences, 2023, 35(1), pp.13-25.

[34] Alsokhiry F, Annuk A, Mohamed M A, et al. An innovative cloud-fog-based smart grid scheme for efficient resource utilization. Sensors, 2023, 23(4), pp.1752.

[35] Parvin K, Hannan M A, Mun L H, et al. The future energy internet for utility energy service and demand-side management in smart grid: Current practices, challenges, and future directions. Sustainable Energy Technologies and Assessments, 2022, 53, pp.102648.

[36] Tiwari S, Jain A, Ahmed N M O S, et al. Machine learning-based model for prediction of power consumption in the smart grid-smart way towards a smart city. Expert Systems, 2022, 39(5), pp.e12832.

[37] Laayati O, El Hadraoui H, Guennoui N, et al. Smart Energy Management System: Design of a Smart Grid Test Bench for Educational Purposes. Energies, 2022, 15(7), pp.2702.

[38] Adnan M, Srivastava G, et al. Instability detection and prevention in smart grids under asymmetric faults. IEEE Transactions on Industry Applications, 2020, 56(4), pp. 4510-4520.

[39] Wilson A J, Reising D R, Hay R W, et al. Automated identification of electrical disturbance waveforms within an operational smart power grid. IEEE Transactions on Smart Grid, 2020, 11(5), pp. 4380-4389.

[40] Belhadi A, Djenouri Y, Srivastava G, et al. Privacy reinforcement learning for fault detection in the smart grid. Ad Hoc Networks, 2021, 119, pp. 102541.

[41] Elbouchikhi E, Zia M F, Benbouzid M, et al. Overview of signal processing and machine learning for smart grid condition monitoring. Electronics, 2021, 10(21), pp. 2725.

[42] Kong Z, Tang B, Deng L, et al. Condition monitoring of wind turbines based on spatio-temporal fusion of SCADA data by convolutional neural networks and gated recurrent units. Renewable Energy, 2020, 146, pp. 760-768.

[43] Huang Z, Zhu J, Lei J, et al. Tool wear predicting based on multi-domain feature fusion by deep convolutional neural network in milling operations. Journal of Intelligent Manufacturing, 2020, 31, pp. 953-966.

[44] Azamfar M, Singh J, Bravo-Imaz I, et al. Multisensor data fusion for gearbox fault diagnosis using 2-D convolutional neural network and motor current signature analysis. Mechanical Systems and Signal Processing, 2020, 144, pp. 106861.

[45] Tang S, Yuan S, Zhu Y. Convolutional neural network in intelligent fault diagnosis toward rotatory machinery. IEEE Access, 2020, 8, pp. 86510-86519.

[46] Chou J S, Hsu S C, Ngo N T, et al. Hybrid machine learning system to forecast electricity consumption of smart grid-based air conditioners. IEEE Systems Journal, 2019, 13(3), pp. 3120-3128.

[47] Alladi T, Chamola V, Rodrigues J J P C, et al. Blockchain in smart grids: A review on different use cases. Sensors, 2019, 19(22), pp. 4862.

[48] Moreno E J J, Morales M O, Tejeida P R, et al. A comprehensive review on smart grids: Challenges and opportunities. Sensors, 2021, 21(21), pp. 6978.

[49] Vaccaro A, Pisica I, Lai L L, et al. A review of enabling methodologies for information processing in smart grids. International Journal of Electrical Power & Energy Systems, 2019, 107, pp. 516-522.
[50] Li M, Zhang K, Liu J, et al. Blockchain-based anomaly detection of electricity consumption in smart grids. Pattern Recognition Letters, 2020, 138, pp. 476-482.
[51] Kimani K, Oduol V, Langat K. Cyber security challenges for IoT-based smart grid networks. International Journal of Critical Infrastructure Protection, 2019, 25, pp. 36-49.