NEW SOLUTIONS FOR SMART GRIDS WITH HIGH-PENETRATION DISTRIBUTED ENERGY RESOURCES

EDITED BY: Yang Li, Liang Chen and Chao Long
PUBLISHED IN: Frontiers in Energy Research

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88974-411-4
DOI 10.3389/978-2-88974-411-4

About Frontiers

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

Frontiers Journal Series

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

Dedication to Quality

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

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

NEW SOLUTIONS FOR SMART GRIDS WITH HIGH-PENETRATION DISTRIBUTED ENERGY RESOURCES

Topic Editors:
Yang Li, Northeast Electric Power University, China
Liang Chen, Nanjing University of Information Science and Technology, China
Chao Long, Cranfield University, United Kingdom

Citation: Li, Y., Chen, L., Long, C., eds. (2022). New Solutions for Smart Grids With High-Penetration Distributed Energy Resources. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88974-411-4

Table of Contents

frontiers in Energy Research

EDITORIAL
published: 10 January 2022
doi: 10.3389/fenrg.2021.739253

Editorial: New Solutions for Smart Grids With High-Penetration Distributed Energy Resources

Liang Chen[1], *Chao Long*[2] *and Yang Li*[3]*

[1]School of Automation, Nanjing University of Information Science and Technology, Nanjing, China, [2]School of Water, Energy and Environment, Cranfield University, Cranfield, United Kingdom, [3]School of Electrical Engineering, Northeast Electric Power University, Jilin, China

Keywords: distributed energy, smart grid, energy storage, cyber security, DC power system

Editorial on the Research Topic

New Solutions for Smart Grids With High-Penetration Distributed Energy Resources

OPEN ACCESS

Edited by:
S. M. Muyeen,
Qatar University, Qatar

Reviewed by:
Ioulia Papaioannou,
European Commission, Italy

***Correspondence:**
Yang Li
liyang@neepu.edu.cn

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

Received: *10 July 2021*
Accepted: *17 December 2021*
Published: *10 January 2022*

Citation:
Chen L, Long C and Li Y (2022) Editorial: New Solutions for Smart Grids With High-Penetration Distributed Energy Resources. Front. Energy Res. 9:739253. doi: 10.3389/fenrg.2021.739253

INTRODUCTION

The application of distributed energy resources (DERs) is an effective way to solve the environmental pollution problem, and ensure a sustainable energy supply just like wind and solar energy. However, with the rapid developments of DERs, the high penetration of renewable energy will bring new challenges to power system operation due to the random fluctuation of power generation. Nowadays, new emerging techniques in smart grid, such as power storages, cyber-physical systems and DC transmission systems, provide more regulatory methods for enabling higher penetration of DERs.

This special issue is organized to introduce the latest research in the field of high penetration of DERs. At present, 18 papers have been accepted for this special issue. These papers can be sorted into the following 7 categories including: 1) Cyber security; 2) DC transmission system; 3) Energy storage; 4) Short-circuit current calculation of wind turbines; 5) Energy trading; 6) Parameter and topology identification; 7) other topics. The sections below introduce the major topics of the papers covered in each category.

CYBER SECURITY TECHNIQUES

The security and economic operation of power system with high penetration of DERs relies on the accurate operating information which can be obtained by measuring equipment and communication systems. However, with the rapid developments of information technologies, the cyber-attacks on smart grids cause blackouts and disrupt power systems. Aiming at this problem, the detection methods of cyber-attacks have been developed to ensure that the accurate information can be obtained. Qu et al. propose a novel detection method of the false data injection attack based on cyber-physical genes. Wang et al. propose a model for cyber and physical data fusion using data link

technologies for detecting attacks on a cyber-physical power system. Wang et al. propose a reconstruction method for missing data in power system measurements based on least squares generative adversarial networks, which is able to reconstruct the missing data in the case of high latitude and high loss rate.

FAULT ANALYSIS AND CONTROL OF DC TRANSMISSION SYSTEMS

High-voltage direct current (HVDC) transmission has already been applied in power systems. The HVDC transmission is more flexible than AC transmission system, which can enhance the penetrations of DERs. On the other hand, the DC fault will seriously affect the operation safety of power systems. Sun et al. establish the mathematical model of modular multilevel converter considering AC feed, and an improved analysis method of single-pole grounding fault of flexible DC network grounded by metal loop is proposed. Li et al. propose a region boundaries determination method for critical commutation failures in multi-infeed HVDC systems under unbalanced short circuit faults. Mao et al. propose a rapid recovery strategy of hybrid Multi-infeed HVDC after AC failure, which can effectively suppress continuous commutation failure and improve the recovery speed of AC voltage and DC power.

ENERGY STORAGE IN SMART GRIDS

Large scale penetration of renewable energy is a challenge for power systems due to the reason that the AC power networks are lack of sufficient flexibilities to deal with the random fluctuations of power sources. Energy storage is an effective way to deal with the penetration problem of DERs. Yang et al. decouple the insufficient flexibility and wind power curtailment caused by transmission congestion, and then a joint planning model of energy storage system and transmission congestion is established. Li et al. establish an optimization method of energy storage system (ESS) to optimize the capacities and locations of distributed ESS to support the voltage regulation of distribution networks.

SHORT-CIRCUIT CURRENT CALCULATION OF WIND TURBINES

Doubly-fed induction generators (DFIGs) are one of the most common kinds of wind generators used in power systems. It is necessary to analyze the influences of short-circuit currents on a power system. Yin proposed a calculation method of the short-circuit current of DFIGs considering the influence of rotor current dynamics. The author also studies the short-circuit current calculation method of a DFIG at the initial fault time with uninterrupted excitation based on the characteristics of non-abrupt changes of flux linkage.

ELECTRICITY PRICE AND TRADING

The demand response is an effective way to enhance the power system flexibility and promote the penetration of DERs. The electricity price is a main regulation method of power loads, which can respond to the power balance of smart grid. Li et al. study a real-time price decision model which adjusts the price in an acceptable range of heating users. Li et al. review the recent studies and projects of energy trading based on blockchain.

PARAMETER AND TOPOLOGY IDENTIFICATION

The random fluctuations of DERs make power system dynamics more frequent than before. The accurate parameters and topology of power grids are vital for the control and management of power dynamics. Qu et al. propose a dynamic optimization method of transmission line parameters based on grey support vector regression. Wang et al. propose a method to identify the topology of a power network based on a knowledge graph and the graph neural network, which can adapt to errors and informational conflicts in the graph data.

OTHER TOPICS

Aiming at the allocation problem of service resources, Jin et al. propose a data association fusion method based on an improved simhash algorithm for allocating the address of power customers according to the short text and unstructured characteristics of the address. Huang et al. propose a calculation method for the temperature rise of oil-immersed transformers considering the vibration of insulating oil. Fu et al. analyze the driving forces of CO_2 emissions change of manufacturing industry of China, and conclude that it is necessary to develop a smart grid to reduce the proportion of coal in the energy structure and increase the proportion of renewable energy. Zhang et al. investigate the economic feasibility of grid-balancing plants via a systematic overall decomposition-based methodology for real geographical zones and flexibility-need scenarios.

CONCLUSION

The papers in this special issue cover different techniques in the smart grid with high penetration of DERs. The outcomes of these research works indicate that the emerging techniques are essential and helpful to promote the penetration of DERs.

AUTHOR CONTRIBUTIONS

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

FUNDING

This work is partly supported by the Natural Science Foundation of Jilin Province, China under Grant No. YDZJ202101ZYTS149.

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

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

frontiers in Energy Research

BRIEF RESEARCH REPORT
published: 17 February 2021
doi: 10.3389/fenrg.2021.634207

Dynamic Optimization Method of Transmission Line Parameters Based on Grey Support Vector Regression

Zhaoyang Qu[1,2], *Miao Li*[3], *Zhenming Zhang*[1,2]*, *Mingshi Cui*[4] *and Yuguang Zhou*[3]

[1]School of Computer Science, Northeast Electric Power University, Jilin, China, [2]Jilin Engineering Technology Research Center of Intelligent Electric Power Big Data Processing, Jilin, China, [3]State Grid Jilin Electric Power Company Limited, Changchun, China, [4]State Grid Inner Mongolia Eastern Electric Power Company, Hohhot, China

Aiming at the problem of insufficient accuracy and timeliness of transmission line parameters in the grid energy management system (EMS) parameter library, a dynamic optimization method of transmission line parameters based on grey support vector regression is proposed. Firstly, the influence of operating conditions and meteorological factors on the changes of parameters is analyzed. Based on this, the correlation quantification method of transmission line parameters is designed based on Pearson coefficient, and the influence coefficient value is obtained. Then, with the influence coefficient as the constraint condition, a method for selecting strong influence characteristics of line parameters based on improved Elastic Net is proposed. Finally, based on the grey prediction theory, a grey support vector regression (GM-SVR) parameter optimization model is constructed to realize the dynamic optimization of line parameter values under the power grid operation state. The effectiveness and feasibility of the proposed method is verified through the commissioning of the reactance parameters of the actual local loop network transmission line.

Keywords: transmission line parameters, strong influence feature selection, parameter correction, grey support vector regression, elastic net algorithm

OPEN ACCESS

Edited by:
Yang Li,
Northeast Electric Power University, China

Reviewed by:
Chen Liang,
Nanjing University of Information Science and Technology, China
Shaoyan Li,
North China Electric Power University, China

***Correspondence:**
Zhenming Zhang
15143289522@163.com

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

Received: *27 November 2020*
Accepted: *13 January 2021*
Published: *17 February 2021*

Citation:
Qu Z, Li M, Zhang Z, Cui M and Zhou Y (2021) Dynamic Optimization Method of Transmission Line Parameters Based on Grey Support Vector Regression. Front. Energy Res. 9:634207. doi: 10.3389/fenrg.2021.634207

INTRODUCTION

Overhead transmission lines are the main components of the power grid. Various advanced calculations of the power system, such as grid modeling, state estimation, power flow calculation, and relay protection settings, require accurate transmission line parameters (Bendjabeur et al., 2019). In the calculation of the actual transmission line parameters, most of the power system operators use electrical equipment to artificially apply a certain voltage after the transmission line is erected and not put into operation, and calculate the double-ended electrical parameters of the line through precision measuring instruments (Xiao et al., 2016). However, when the transmission line is energized and operated, affected by the operating conditions of the transmission network, line environment, seasonal changes and other factors, there are certain differences between the actual line parameters and the parameter values originally calculated (Asprou et al., 2018; Yu et al., 2018). The transmission line parameter error will seriously reduce the calculation accuracy of the state estimation in the local grid area, then affect the accuracy of the calculation results of various advanced applications based on the state estimation, and seriously interfere with the analysis and control of the power system by dispatchers (Liu et al., 2020).

Therefore, studying the accuracy and dynamic modification of transmission line parameters has important theoretical and practical significance.

For a long time, in order to improve the accuracy of transmission line parameters, domestic and foreign experts have carried out many related researches on transmission line parameter estimation and parameter optimization, which can be divided into two main categories:

1) Based on power grid state estimation, expand calculations from different angles, such as augmented state estimation method (Xue et al., 2014), normal equation estimation method (Tang et al., 2018), Kalman filter estimation method (Wang et al., 2019c) and residual sensitivity analysis method (Su et al., 2019). Among them, the literature (Wang et al., 2019a; Kong et al., 2020) proposed a method to generate suspicious lines by calculating measurement deviation and line balance state, and used variable step integration method to estimate the parameter correction value of suspicious lines. Literature (Xue et al., 2019) proposed a comprehensive method based on PQ decoupling combined with augmented state estimation and sensitivity methods to achieve line parameter error correction. Literature (Viafora et al., 2019) proposes a two-step state estimation algorithm that takes into account line temperature. However, these methods generally face the problem of ill-conditioned coefficient matrix. In particular, a large range of line parameters can be extended to state values to achieve the solution of parameter values, which results in too high matrix dimensions, seriously affected state estimation results, and the calculated parameter values cannot meet the basic accuracy requirements of various advanced applications of power grid energy management system.
2) Based on the measured value of the transmission line section, the parameter optimization based on artificial intelligence algorithm includes particle swarm optimization method (Qu et al., 2018; Zhu et al., 2020), genetic algorithm (Shi et al., 2009; Li et al., 2020) and Tabu search method (Dai, 2020). Among them, the literature (Huang et al., 2019) proposed a method for estimating line parameters of distribution network based on radial basis function neural network, which can obtain accurate line parameters by using the measured values at both ends of the line. In literature (Ren et al., 2019; Dutta et al., 2020), multi-time SCADA and PMU measurements of a single line are respectively used for parameter estimation. Simulation results can effectively estimate the impedance parameters of the line. Literature (Ghiasi et al., 2019) trained and learned a large number of multi-section active power data, the sparsity of the line parameter difference and the practical lasso algorithm are used to correct and solve the suspicious line. This type of method has become the mainstream research idea, which provides important reference and theoretical guidance for subsequent research.

Based on the above research on various optimization methods, the current transmission line parameter estimation or optimization method has two shortcomings in the solution calculation: ① The effective influence of the line operation mode on the parameters is not considered and calculated, resulting in the lack of precision of parameter values. ②The calculation of transmission line parameters in the revised scheme cannot be combined with the state data of the latest line operation mode. Only historical data is used to calculate line parameter values, and the generated parameter values are not real-time, and it is not appropriate to be applied to advanced real-time calculations of various power grids.

Transmission line parameters such as resistance (R), reactance (X) and susceptance (B) are the prerequisites for modern power system operation, control and planning research (Salam, 2020). In this paper, by analyzing the internal and external factors that interfere with the line parameters during the operation of the transmission line, the relevant influencing factors reflecting the operating conditions of the line and the natural environmental parameters are initially determined, and the degree of influence is quantified. Then, a strong-influence feature selection model based on the constraint of the influence coefficient is constructed to break the limitation of the specific input quantity of the traditional electrical calculation model. More comprehensive line parameter influence characteristics are considered as parameter optimization considerations. Based on grey prediction theory, a grey support vector regression (GM-SVR) parameter optimization model is constructed to achieve dynamic adjustment of line parameter values under power grid operation.

RELATED INFLUENCING FACTORS AND QUANTIFICATION OF TRANSMISSION LINE PARAMETERS

Analysis of Related Influencing Factors of Transmission Line Parameters

The transmission grid will generate heat during long-distance transmission (Zhang et al., 2019). As the demand on the load side of the transmission grid changes, the line operation mode must be adjusted as needed, which will cause the heat generation of the line to change. At the same time, the air convection heat flow and solar radiation in the natural environment around the line will also cause the line to generate heat. Since the transmission line is a metal conductor, which is a good conductor of heat, there is obvious thermal expansion and contraction. With the temperature change caused by internal or external factors of the line, the resistivity and sag of the line will change, which will cause the line length and cross-sectional area to be different from the initial parameter values (Wang et al., 2019b; Beňa et al., 2020). In addition, the transmission lines are affected by severe weather (Sony and Chettiar, 2020), the geometric uniform distance of the three-phase conductors and the geomagnetic field in special areas are also different from the original parameter calculation scenarios (Zhu et al., 2018; Qu et al., 2019).

Therefore, for different transmission lines erected in different scenarios, the leading factors that affect line parameter values are different, but through in-depth impact analysis of the line's

multi-period historical operation data and meteorological environment data, there are still traces to follow. Based on the measurement of the supervisory control and data acquisition system (SCADA) of the grid, the paper extracts the internal operating mode factors that affect the parameter changes such as active power, reactive power, current, and voltage amplitude at both ends of the line. Based on the monitoring of the National Meteorological Data Center, the external natural environmental factors that affect the change of parameters such as temperature and wind speed in the area where the line is located are extracted.

Quantification of Related Influencing Factors Based on Pearson Coefficient

This paper analyzes the correlation between measured data and transmission line parameters, and designs a method for quantifying the correlation of parameter influencing factors based on Pearson coefficients to quantify the degree of influence between each influencing factor and line parameters.

The specific quantitative method steps are as follows:

Step1: Extract the SCADA measurement characteristic data at both ends of the line through the transmission line ID. At the same time, in the process of considering the operation mode of the line, the influence of the natural environment of the line on the parameters itself is further considered, and meteorological environmental data in the area where the line is located are extracted.

Step2: Fuse the extracted line operation mode and meteorological environment data based on time series to construct a complete dataset of influencing factors:

$$\text{DataSet} = \begin{cases} P_{ic} = \{x_1, x_2, \ldots, x_n\} \\ P_v = \{y_1, y_2, \ldots, y_n\} \end{cases}. \tag{1}$$

In the **Eq. 1**: P_{ic} and P_v respectively represent the set of influencing factors of line parameters and the set of line sequence parameters.

Step3: In order to accurately measure the degree of influence of transmission line parameter influence characteristics on each parameter value, a correlation quantification function based on Pearson correlation coefficient is designed:

$$\begin{aligned} r_{ab} = \text{cor}(x, y) &= \frac{\text{cov}(x, y)}{\eta_x \eta_y} \\ &= \frac{\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})/n - 1}{\sqrt{\sum (x_i - \overline{x})/n - 1} \cdot \sqrt{\sum (y_i - \overline{y})/n - 1}}. \end{aligned} \tag{2}$$

In the **Eq. 2**: η_x, η_y and $\text{cov}(x, y)$ respectively represent the standard deviation of sample x, the standard deviation of sample y, and the covariance between samples. ab represents P_{ic}, P_v the number of the correlation calculation vector in the set. r_{ab} represents the quantized influence coefficient value, $r_{ab} \in [-1, 1]$.

Step4: Impact trend analysis. Analyze the closeness and influence trend between variables according to the magnitude and the sign of r_{ab}.

STRONG INFLUENCE FEATURE SELECTION METHOD BASED ON IMPROVED ELASTIC NET

The purpose of strong influence feature selection is to lay the foundation for establishing an optimal route parameter optimization model. Excluding redundant influence features, only the independent variables that have a strong influence on the parameter value are included in the training data set, which effectively alleviates the computational cost caused by high-dimensional training features.

The principle of the Elastic Net (Zhao et al., 2020) algorithm is to add two penalty items on the basis of least squares. It is often used to process sample data with multi-collinearity, especially when there is a high degree of correlation between data features. However, the double shrinkage of traditional Elastic Net leads to low efficiency and high bias in the selection of strong influence features. In order to correct this effect, the quantized value Pearson influence coefficient is eliminated as the constraint condition, and the quantized value r_{ab} between each feature and the line parameter is weighted to l_1 and l_2 to realize the re-calibration of the secondary partial penalty coefficient. The line parameter strong influence feature selection model based on ElasticNet-r is constructed to screen out the features that have a strong influence on the line parameters, and provide an effective set of influence characteristics for the optimization of transmission line parameters.

The solution process of Elastic Net influence characteristic coefficients is divided into three stages:

First, the Ridge regression coefficient is determined, and then the lasso shrinkage coefficient is used. $\lambda_1 \|\overline{\omega}\|_1 + \lambda_2 \|\overline{\omega}\|_2^2$ is the penalty term of Elastic Net algorithm, which is a convex combination of Lasso and Ridge penalty terms.

Then, let $\theta = \lambda_1 + \lambda_2, \lambda = \lambda_1/\lambda_1 + \lambda_2$, then the specific principle of strong influence feature selection is:

$$\begin{aligned} EN\_r(r, \lambda) = \arg\min_{\overline{\omega}} \Bigg(& \frac{1}{2n} \sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{p} x_{ij} \overline{\omega}_j \right)^2 + \theta\lambda \sum_{j}^{p} r_j |\overline{\omega}_j| \\ & + \frac{\theta(1 - \lambda)}{2} \sum_{j}^{p} r_j \overline{\omega}_j^2 \Bigg). \end{aligned} \tag{3}$$

In the **Eq. 3**: r is the influence coefficient of parameter influence characteristic x_a and this kind of parameter y_b. The value of θ determines the proportional relationship between lasso regression and ridge regression, which is determined by iterative calculation during the solution process. λ is a non-negative regularization parameter. n represents the number of data samples in the constructed transmission line section dataset. $\overline{\omega}$ represents the influence characteristic coefficient.

Finally, after establishing each influencing feature coefficient, the strong influencing feature selection with line parameters can be obtained by **Eq. 4**:

$$\text{features} = \{F_i | \overline{\omega} \neq 0, 1 < i < p\}. \tag{4}$$

DYNAMIC OPTIMIZATION OF TRANSMISSION LINE PARAMETERS

The entire optimization process in this paper consists of two parallel stages. The first stage uses the multi-dimensional feature time series data in the transmission line parameter strong influence characteristic dataset to construct a single characteristic grey model to dynamically predict the next time section strong influence characteristic value. The second stage uses the strong adaptability and fault tolerance of support vector regression (Cao et al., 2017), and trains the support vector regression (SVR) parameter optimization model based on the complete and strong impact feature dataset, and sets the parameter deviation coefficient and deviation floating factor constrained optimization relationship to ensure that the optimization result is within the acceptable accuracy range. At the same time, the real-time strong influence characteristic value of grey prediction is used as the input of the parameter optimization model, and the iterative calculation completes the solution of the optimization model line parameters.

Prediction Model of Strong Influence Eigenvalues Based on GM (1,1)

As a forecasting method that adapts to systems with uncertain factors, the grey forecasting method is predicted by observing the future value of historical data series (Hu et al., 2020). Aiming at the characteristics of weak randomness on the load side during the operation of the power system. Use discrete random numbers to evolve time series data with strong influence characteristics into generated numbers. The randomness of the generated number sequence is significantly weakened and contains characteristic operation laws. Furthermore, a dynamic model in the form of grey differential equations is established to complete the accurate prediction of the eigenvalues of strong influence in real-time conditions.

It is known that a certain transmission line parameter strongly influences the characteristic element sequence data:

$$f_i^{(0)} = \left(f_i^{(0)}(1), f_i^{(0)}(2), \ldots, f_i^{(0)}(n)\right). \tag{5}$$

Where f_i represents a certain feature in the strong influence feature dataset. $f_i^{(0)}$ represents the original data sequence of the characteristic. Although the data sequence has a complex appearance, it reflects the operating state of the power system as a whole, so it must contain some inherent law.

Construct the GM (1,1) transmission line parameter strong influence characteristic grey prediction model, and dynamically calculate the predicted value.

$$f_i^{(0)}(k) + \alpha z_i^{(1)}(k) = \delta, \tag{6}$$

Where $f_i^{(0)}(k)$ is the model grey derivative, which α represents the trend coefficient of the strong influence feature. $z_i^{(0)}(k)$ is the model whitening background value, which δ represents the amount of grey effect.

The time response sequence of the grey prediction model with strong influence characteristics of transmission line parameters is:

$$\widehat{f}_i^{(1)}(k+1) = \left[f_i^{(1)}(0) - \frac{\delta}{\alpha}\right]e^{-\alpha t} + \frac{\delta}{\alpha} \tag{7}$$

According to the originally selected method of generating the grey sequence of the strong-influence feature value, it is restored to the original feature sequence, so as to obtain the predicted value of the next time section during the operation of the transmission line accordingly.

$$\widehat{f}_i^{(0)}(k+1) = \widehat{f}_i^{(1)}(k+1) - \widehat{f}_i^{(1)}(k). \tag{8}$$

Dynamic Optimization Algorithm of Transmission Line Parameters Based on GM-SVR

The strong-influence eigenvalue prediction model realizes the prediction of the operating conditions and meteorological conditions of the transmission line, and integrates it into the transmission line parameter SVR predictor, designs the dynamic optimization algorithm of the transmission line parameters based on GM-SVR, and dynamically solves the parameter optimization The specific execution steps are as follows:

Step 1: Organize the training dataset of strong influence features $T = \{(\overrightarrow{f}_1, y_1), (\overrightarrow{f}_2, y_2), \ldots, (\overrightarrow{f}_t, y_t)\}$. Among them $T_F = \{\overrightarrow{f}_1, \overrightarrow{f}_2, \ldots, \overrightarrow{f}_t\}^T$ represents the set of strong influence characteristic sequences of massive historical sections, $\overrightarrow{f}_i$ including the strong influence characteristic value under the current time section, the time section index $i = 1, 2, \ldots, t$.

$$T_F = \begin{Bmatrix} f_1^{(1)}, f_1^{(2)}, \ldots, f_1^{(n)} \\ f_2^{(1)}, f_2^{(2)}, \ldots, f_2^{(n)} \\ \ldots \\ f_t^{(1)}, f_t^{(2)}, \ldots, f_t^{(n)} \end{Bmatrix}. \tag{9}$$

The transmission line parameters of the corresponding time section during the steady state operation of the power system $T_Y = \{y_1, y_2, \ldots, y_t\}^T$ are **Eq. 10**, y_i including the resistance (R), reactance (X), and susceptance(B) parameter values under the current time section.

$$T_Y = \begin{Bmatrix} R_1, X_1, B_1 \\ R_2, X_2, B_2 \\ \ldots \\ R_t, X_t, B_t \end{Bmatrix}. \tag{10}$$

Step 2: The kernel function of the optimization model is determined. From the input space to the feature space, the traditional mapping method will lead to an explosive increase in dimensionality. The radial basis kernel function has the characteristics of strong locality and stability. It is widely used in the regression analysis process of massive samples and is effective Reduce the calculation amount and storage space of the parameter optimization model.

$$K\left(\overrightarrow{f}_i, y_i\right) = \exp\left(-\left(\left\|\overrightarrow{f}_i - y_i\right\|^2/\beta^2\right)\right). \tag{11}$$

In the **Eq. 11**, $K(\overrightarrow{f}_i, y_i)$ return the calculated result value of the sample after dimension transformation. $\overrightarrow{f}_i, y_i$ represents the data sample of transmission line parameter section. $\left\|\overrightarrow{f}_i - y_i\right\|$ represents the norm of the vector. β is the only hyperparameter in the kernel function, which realizes the search for an optimal balance point between the accuracy of the parameter correction model and the generalization ability.

Step 3: Set the parameter optimization coefficients of the transmission line, and establish the cost function.

$$\mathrm{fun}_{\cos t} = \min \frac{1}{2}\|\xi\|^2 + C\sum_{i=1}^{l}(\rho_i + \rho_i^*). \tag{12}$$

Most of the parameter deviations are within the scope of the grid standards, it will not affect the follow-up power system state estimation and other related advanced applications, More attention should be paid to the excessive deviation of the parameters caused by the real-time changes of the transmission line operating conditions and meteorological conditions, so the optimization coefficient value is to ensure that the parameter value is constrained within an acceptable accuracy range.

$$s.t.\begin{cases} y_i - \left(\xi^T\overrightarrow{f}_i + b\right) < \varepsilon + \rho_i, \\ \left(\xi^T\overrightarrow{f} + b\right) - y_i < \varepsilon + \rho_i^*, \\ \rho_i, \rho_i^* \geq 0, \end{cases} \tag{13}$$

Where ε is the optimization coefficient of transmission line parameters. ρ indicates the deviation factor, for the excessive deviation parameters, the soft boundary method is adopted, and the upper and lower deviation floating factors are adopted, ρ_i, ρ_i^* will effectively improve the generalization ability of the parameter optimization model.

Step 4: Establish the dynamic optimization relationship of line parameters, namely SVR regression model, to fit the relationship between the strong influencing characteristic variables and the transmission line parameter variables.

$$Model(T) = \sum_{i=1}^{m}\left(\widehat{\mu}_i - \mu_i\right)\overrightarrow{f}_i^T T + b. \tag{14}$$

The constraints $\mathrm{Model}(\overrightarrow{f}_i) - y_i - \varepsilon - \rho_i = 0$ and $y_i - \mathrm{Model}(\overrightarrow{f}_i) - \varepsilon - \widehat{\rho}_i = 0$ cannot be established at the same time, the deviation of the transmission line parameters will only be larger or smaller than the current actual parameter values. Therefore, at least one of μ_i and $\widehat{\mu}_i$ is zero.

Step 5: In order to realize the dynamic prediction of the optimization model and further improve the accuracy of the tuning results, the strong influence characteristic value $f_{k+1}^{(1)}, f_{k+1}^{(2)}, \ldots, f_{k+1}^{(n)}$ of the transmission line parameters predicted by GM (1,1) is introduced as the input of the correction model. Calculate the corresponding transmission line parameters $\mathrm{Model}(\overrightarrow{f}_{k+1})$, and further determine the optimized value. When $\left|\mathrm{Model}(\overrightarrow{f}_{k+1}) - y_{k+1}\right| > \varepsilon$, return to step 3 to recalculate the cost function value. When $\left|\mathrm{Model}(\overrightarrow{f}_{k+1}) - y_{k+1}\right| \leq \varepsilon$, the parameter optimization value is considered accurate and the optimization algorithm ends.

EXPERIMENTAL ANALYSIS

In this paper, a local loop transmission network in a province is taken as an example, and the method proposed in the article is used to optimize and debug the loop network reactance parameters. **Figure 1** shows the simulation topology connection of the ring-type power grid system based on Matlab/Simlink (In the **Figure 1**, the red component represents the power node, the blue component represents the 220 KV plant node, and the green component represents the transmission line).

The historical measurement data of the ring network area within 1 year were extracted from the SCADA database of the Provincial Dispatching Center and the database of the National Meteorological Center. By weighing the transmission line parameter optimization accuracy and the computational cost caused by the model learning long-time section dataset, the training dataset is diluted on the premise that the parameter optimization accuracy meets the actual application. The time section interval is set to 3 h, including Measurement of 2,920 moments. According to the analysis of electrical mechanism, after the characteristics irrelevant to parameter correction of line cross-section features are preliminarily screened out, there are still as many as 10 associated attributes in the section dataset extracted and constructed.

Solving the Influence Coefficient of Transmission Line Parameters

In order to reduce the interference of low-impact features in the historical section dataset to the later model training, the correlation between transmission line parameters and each impact feature was analyzed and quantified. In this stage, considering that there is little difference in significant influence rules between different circuits in the same local ring network area, and in order to avoid weakening the influence relationship between measurement features and parameters over a long time span. Statistical experimental analysis was carried out around the concentration time of the Zhenxin line in May. Quantify the influence of the measurement characteristics of the first and last ends of the transmission line and the meteorological measurement characteristics on the reactance value.

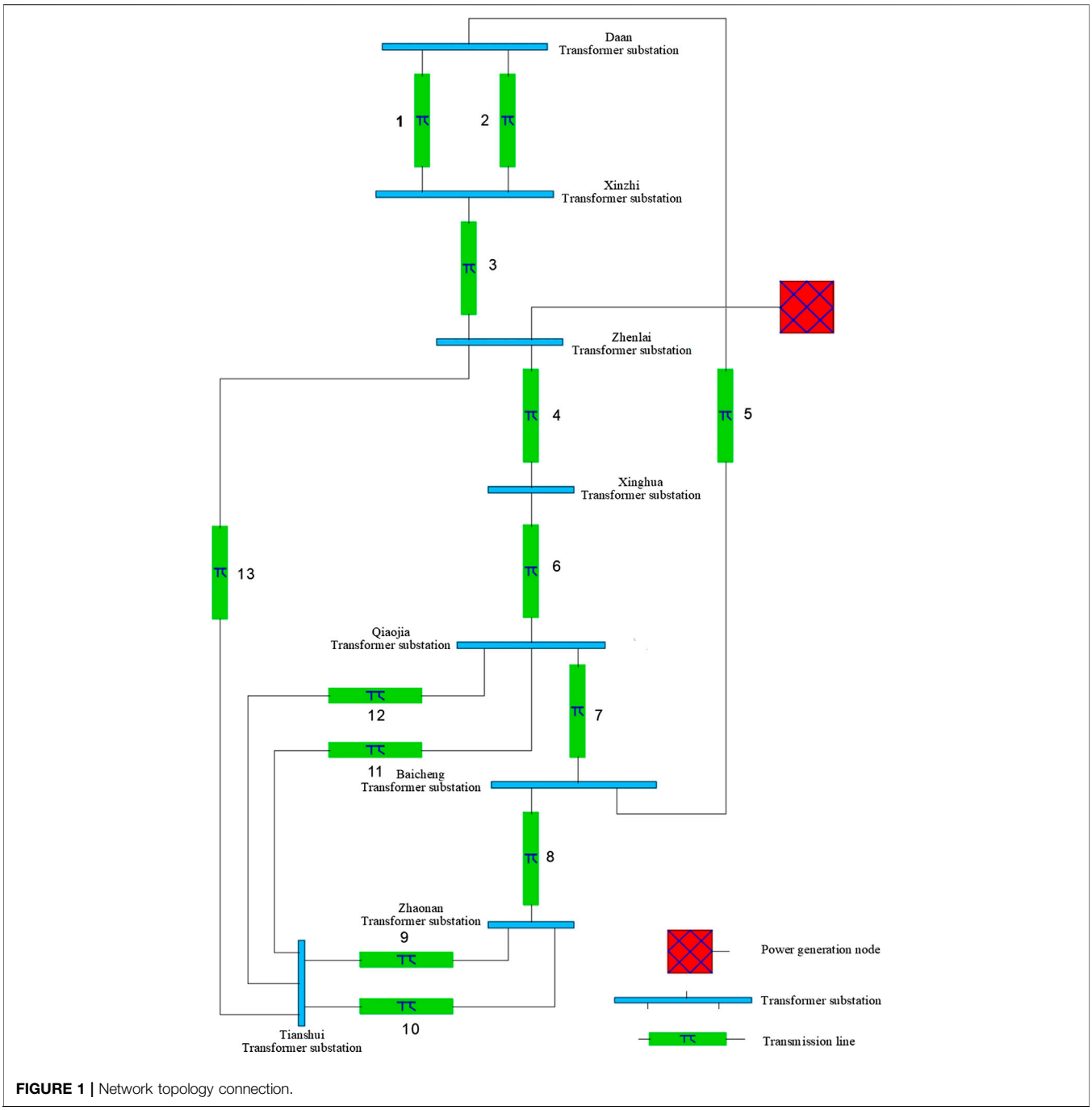


FIGURE 1 | Network topology connection.

Through correlation analysis and quantitative results, it can be found that the absolute value of the Pearson correlation coefficient of the transmission line's first-end active power (P), first-end current (I) and ambient temperature and reactance parameter (X) are greater than 0.7. It shows that the five transmission line measurement characteristics have a strong influence on the parameter values under the time sequence state, and they are not negligible characteristics when training the reactance value correction model.

Strong Influence Feature Weight Calculation

Using the complete dataset of influence factors of transmission line parameters constructed to verify the advantages of ElasticNet-r based strong influence feature selection method in this paper. The Ridge algorithm, Lasso algorithm, Elastic Net algorithm and this algorithm are used to compare the selection simulation results and model performance. Specific experiments are as follows:

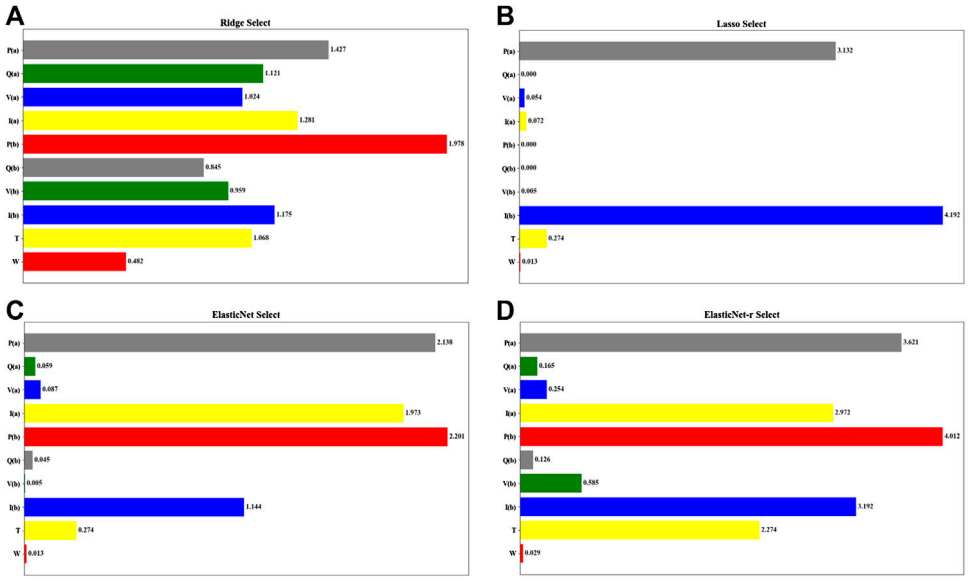


FIGURE 2 | Comparison of the selection results of different algorithms with strong influence.

TABLE 1 | Dynamic optimization of transmission line parameters.

Line number	Line name	Original reactance parameters	Manual tuning	Algorithm optimization	
				Parameter value optimization	Remarks
1	NewDa Jia Line	1.0473	1.04	1.3563	Small change
2	NewDa Yi Line	1.0473	2.09	1.2943	Small change
3	ZhenXin Line	0.5484	2.31	2.0013	Increase
4	ZhenXing Line	1.0478	1.04	1.0478	Constant
5	BaiDa Line	2.8128	2.81	2.8128	Constant
6	Xingqiao Line	1.0600	1.06	1.0600	Constant
7	QiaoBai Line	2.1831	1.27	1.3498	Decrease
8	BaiTao Line	3.1040	2.86	2.4350	Decrease
9	TianTao Jia Line	0.4483	0.44	0.4659	Small change
10	TianTao Yi Line	0.4483	0.44	0.4659	Small change
11	TianQiao Jia Line	1.3901	1.01	0.6996	Decrease
12	TianQiao Yi Line	0.6996	1.01	0.6996	Constant
13	TianZhen Line	2.5043	2.50	2.4424	Small change

Verify the filtering effect of different algorithms in strongly influencing feature selection. The initial alpha of each algorithm is set to 0.1, and max_iter is set to 10,000. Elastic Net algorithm and ElasticNet-r algorithm's L1_ratio is set to 0.5. In the process of model training, the value of the super-parameter is iteratively adjusted, and the feature selection result under the optimal state of different algorithms is strongly influenced, as shown in **Figure 2**.

From **Figures 2A–D**, it can be seen that each algorithm uses model training on a complete dataset of transmission line parameters to determine the corresponding weights for different features, and then achieves the retention and elimination of influencing features.

Circuit Reactance Parameter Tuning

Taking the selected line "Head end active power," "Head end current," "Terminal active power," "Terminal current," and "Environment temperature" as input characteristics, and the optimized value of transmission line parameters as output. Establish a dynamic optimization model for GM-SVR transmission line parameters, the result of tuning the reactance parameters of the experimental loop network is shown in **Table 1**, and it is compared with the tuning result of the engineer combining the least square method and his own experience.

Table 1 shows that the method in this paper can be adapted to different transmission lines and has a certain generalization ability. At the same time, comparing the least square method of each line and the results of manual tuning, the accuracy value cannot meet the accuracy requirements of the reactance parameter value of the transmission line, and it is highly dependent on the experience of the engineer, resulting in the timeliness of the optimized parameter result value, which

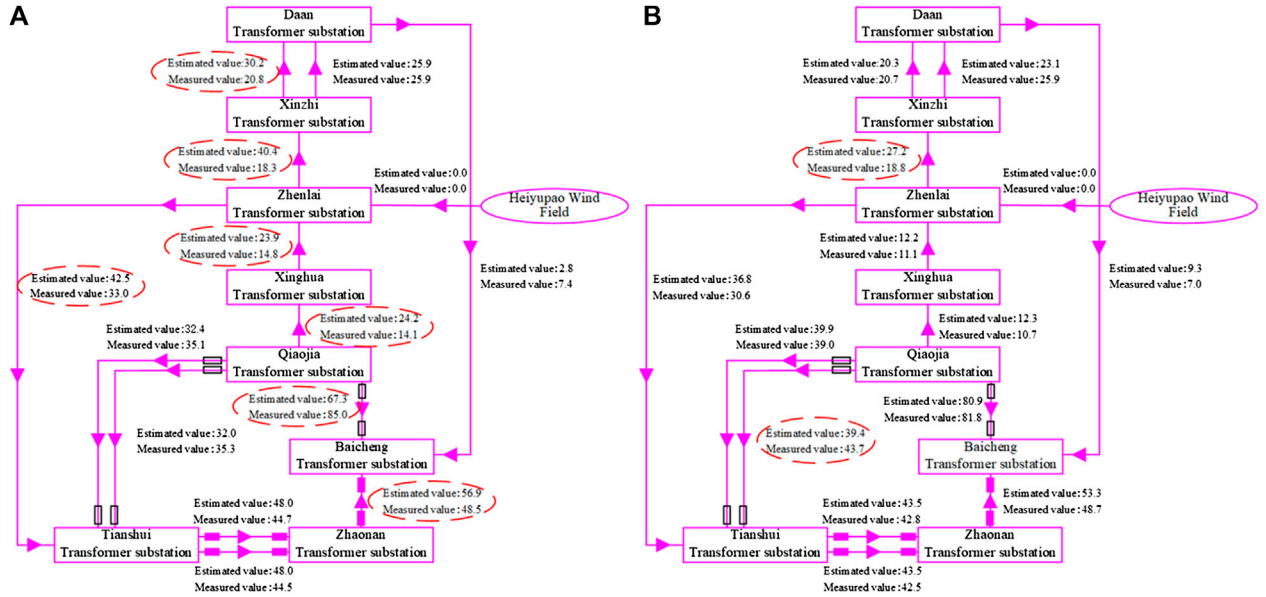


FIGURE 3 | Using original parameters **(A)** and optimized parameters **(B)** for state estimation.

cannot be long-term Maintain the accuracy of the power system's advanced calculation results such as state estimation. The method in this paper only parameter optimization models trained and learned through historical datasets, without manual experience. The parameter optimization results can meet the accuracy requirements of reactance parameters, and as long as the training sample coverage is long enough Time section, the parameter tuning results are sufficient to have good timeliness.

Validation Verification of Transmission Line Parameter Optimization Method

In order to verify the effectiveness of the algorithm optimization parameter values, the transmission line parameter values optimized by the algorithm are filled into the state estimation program simulation module of a provincial smart grid dispatching technical support system, and the results of the transmission line state estimation are observed. **Figure 3A** is the state estimation result using the original parameters:

China power grid stipulates that the state estimation qualification threshold for the active power measurement of 220 kV grid lines is 6.1 MW. **Figure 3A** uses the original parameters for state estimation. The results show that the residual error of seven lines exceeds the threshold, which exceeds 50% of the total number of lines in the local ring network.

It can be seen from **Figure 3B** that only two lines (ZhenXin Line and TianZhen Line) have residuals exceeding the qualified threshold, and the state estimation active residuals of the remaining lines are all small, and the overall residuals are significantly smaller than when the original parameters are used. The pass rate of state estimation has been significantly improved.

CONCLUSION

This paper proposes a new method for dynamic optimization of transmission line parameters based on grey support vector regression (GM-SVR), which effectively improves the accuracy and timeliness of transmission line parameters in the grid energy management system (EMS) parameter library. The time series changes of transmission line parameters are affected by the operating conditions of the power grid and the complex meteorological environment. The strong influence feature selection model of the proposed method can accurately screen out features that have a strong impact on transmission line parameters. The limitation of specific input quantity is broken when solving line parameter values based on electrical mechanism modeling.

The analysis and verification results of actual engineering examples show that the optimized transmission line parameters of this method can significantly improve the pass rate of power grid state estimation. It is of reference significance for the accurate monitoring of power grid operation status and the safe dispatch of power system. However, this method still has shortcomings in considering the influence of parameters such as the geomagnetic field and the geometric distance between lines in special areas. This will be a subject to be studied in depth in the next stage.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

ZQ: designed this study. ML: contributed to the strong influence feature selection method based on improved Elastic Net. ZZ: contributed to transmission line parameter optimization algorithm and article experimental analysis. MC: collected and cleansed the data. YZ: performed the related influencing factors and quantitative analysis of transmission line parameters. All authors contributed to the writing of the article and all agreed to the submitted version of the article.

FUNDING

This paper was supported in part by science and technology innovation development plan project of Jilin (201830817), National Natural Science Foundation of China (Grant No. 51437003).

REFERENCES

Asprou, M., Kyriakides, E., and Albu, M. M. (2018). Uncertainty bounds of transmission line parameters estimated from synchronized measurements. *IEEE Trans. Instrumen. Measure.* 68 (8), 2808–2818. doi:10.1109/TIM.2018.2867966

Bendjabeur, A., Kouadri, A., and Mekhilef, S. (2019). Novel technique for transmission line parameters estimation using synchronised sampled data. *IET Gener., Transm. Distrib.* 14 (3), 506–515. doi:10.1049/iet-gtd.2019.0702

Beňa, Ľ., Gáll, V., Kanálik, M., Kolcun, M., Margitová, A., Mészáros, A., et al. (2020). Calculation of the overhead transmission line conductor temperature in real operating conditions. *Electr. Eng.* 9, 1–12. doi:10.1007/s00202-020-01107-2

Cao, J., Fang, Z., Qu, G., Sun, H., and Zhang, D. (2017). An accurate traffic classification model based on support vector machines. *Int. J. Network Mgmt* 27 (1), e1962. doi:10.1002/nem.1962

Dai, Y. (2020). Real-time dynamic line rating of transmission lines using live simulation model and Tabu search. *IEEE Trans. Pow. Deliv.* 8, 1. doi:10.1109/TPWRD.2020.3014911

Dutta, R., Patel, V., Chakrabarti, S., Sharma, A., Das, R. K., and Mondal, S. (2020). Parameter estimation of distribution lines using SCADA measurements. *IEEE Trans. Instrument. Measure.* 70, 1–11. doi:10.1109/TIM.2020.3026116

Ghiasi, S. M. S., Abedi, M., and Hosseinian, S. H. (2019). Mutually coupled transmission line parameter estimation and voltage profile calculation using one terminal data sampling and virtual black-box. *IEEE Access* 7, 106805–106812. doi:10.1109/access.2019.2901813

Hu, Z., Xu, X., Su, Q., Zhu, H., and Guo, J. (2020). Grey prediction evolution algorithm for global optimization. *Appl. Math. Model.* 79, 145–160. doi:10.1016/j.apm.2019.10.026

Huang, R., Guo, M. F., and Chen, Y. W. (2019). Parameter estimation of distribution system based on radial basis neural network. *Electr. Eng.* 20 (4), 42–46.

Kong, X. L., Zhang, X. Y., Wang, C. S., Li, P., Yu, L., and Jiang, X. D. (2020). "Adaptive self-optimizing state estimation method of distribution network in complex condition," in Proceedings of the CSEE, Lisbon, Portugal, October 18–20, 2020. Available at: http://kns.cnki.net/kcms/detail/11.2107.TM.20200828.1626. 005.html.

Li, Z. X., Zhang, J. T., Chen, Y. X., Weng, H. L., and Lv, L. P. (2020). A new fault location algorithm of line voltage cross-correction for three-terminal transmission lines based on dynamic calculation of real-time parameters. *Power Syst. Technol.* 44 (2), 761–768. doi:10.13335/j.1000-3673.pst.2019.0690

Liu, Y., Wang, B., Zheng, X., Lu, D., Fu, M., and NengLing, T. (2020). Fault location algorithm for non-homogeneous transmission lines considering line asymmetry. *IEEE Trans. Power Deliv.* 35 (5), 2425–24378. doi:10.1109/TPWRD.2020.2968191

Qu, Z., Dong, Y., Qu, N., Wang, L., Li, Y., Zhang, Y., et al. (2019). "Survivability evaluation method for cascading failure of electric cyber physical system considering load optimal allocation. *Math. Prob. Eng.* 2019 (2), 1–15. doi:10.1155/2019/2817586

Qu, Z., Zhang, Y., Qu, N., Wang, L., Li, Y., and Dong, Y. (2018). Method for quantitative estimation of the risk propagation threshold in electric power CPS based on seepage probability. *IEEE Access* 6, 68813–68823. doi:10.1109/access.2018.2879488

Ren, P., Abur, A., and Lev-Ari, H. (2019). "Tracking transmission line parameters in power grids observed by PMUs," in 2019 IEEE milan power tech, Milan, Italy, June 23–27, 2019 (Piscataway, NJ: IEEE), 1–6.

Salam, M. A. (2020). *Transmission line parameters and analysis. Fundamentals of electrical power systems analysis*. Singapore: Springer.

Shi, Z. B., Li, Y., and Yu, T. (2009). "Short-term load forecasting based on LS-SVM optimized by bacterial colony chemotaxis algorithm," in 2009 international conference on information and multimedia technology, Jeju Island, South Korea, December 16–18, 2009 (Piscataway, NJ: IEEE), 306–309.

Sony, M., and Chettiar, M. V. (2020). Stochastic modelling of weather-related transmission line outages. *Int. J. Oper. Res. Inf. Syst.* 11 (1), 66–82. doi:10.4018/ijoris.2020010103

Su, H., Li, P., Fu, X., Yu, L., and Wang, C. (2019). Augmented sensitivity estimation based voltage control strategy of active distribution networks with pmu measurement. *IEEE Access* 7, 44987–44997. doi:10.1109/ACCESS.2019.2908183

Tang, K. J., Dong, S. F., and Song, Y. H. (2018). A real-time N-1 AC power flow calculation method based on FPU-CPU heterogeneous computing framework. *Proceed. CSEE* 38 (15), 4329–4338+4633. doi:10.13334/j.0258-8013.pcsee.172148

Viafora, N., Morozovska, K., Kazmi, S. H. H., Laneryd, T., Hilber, P., and Holbøll, J. (2019). Day-ahead dispatch optimization with dynamic thermal rating of transformers and overhead lines. *Elec. Power Syst. Res.* 171, 194–208. doi:10.1016/j.epsr.2019.02.026

Wang, C., Centeno, V. A., Jones, K. D., and Yang, D. (2019a). Transmission lines positive sequence parameters estimation and instrument transformers calibration based on PMU measurement error model. *IEEE Access* 7, 145104–145117. doi:10.1109/access.2019.2944818

Wang, R. H., Lu, Z. W., Wang, Y. L., Yang, C., and Li, C. Y. (2019b). Research on influence of line arrester on lightning protection performance of 66kV double-circuit transmission lines on the same tower. *J. Northeast Elec. Power Univ.* 39 (5), 17–22.

Wang, Y., Sun, Y., Dinavahi, V., Cao, S., and Hou, D. (2019c). Adaptive robust cubature Kalman filter for power system dynamic state estimation against outliers. *IEEE Access* 7, 105872–105881. doi:10.1109/access.2019.2932261

Xiao, Y., Fan, Y., Cheng, L., and Deng, J. (2016). A parameter measurement theory of single and double circuit AC transmission lines. *Proceed. CSEE* 36 (20), 5515–5522+5727. doi:10.13334/j.0258-8013.pcsee.152393

Xue, A. C., Zhang, Z. Y., Zhang, J. M., Chang, N. C., and Bi, T. S. (2014). An augmented state estimation method for transmission line parameters based on maximum normal measurement rate. *Autom. Electr. Power Syst.* 38 (10), 61–65. doi:10.1109/PESGM.2015.7285847

Xue, A., Xu, F., Xu, J., Chow, J. H., You, H., and Bi, T. (2019). Correction of phasor measurements independent of transmission line parameters. *IEEE Transa. Smart Grid* 11 (1), 346–356. doi:10.1109/TSG.2019.2921819

Yu, X. L., Zhang, H. X., and Wang, M, X. (2018). An algorithm for power system state estimation considering line temperature. *Proceed. CSEE* 38 (9), 2561–2570+2824. doi:10.13334/j.0258-8013.pcsee.171161

Zhang, S., Dong, X., Xing, Y., and Wang, Y. (2019). "Analysis of influencing factors of transmission line loss based on GBDT algorithm," in In 2019 international conference on communications, information system and computer engineering (CISCE), Haikou, China, July 5–7, 2019 (Piscataway, NJ: IEEE), 179–182.

Zhao, G., Hu, J., He, J., and Wang, S. X. (2020). A novel current reconstruction method based on elastic Net regularization. *IEEE Trans. Instrument. Measure.* 69 (10), 7484–7493. doi:10.1109/TIM.2020.2984819

Zhu, H., Yu, Z. X., and Yan, J. T. (2018). UHV line tree barrier to predict the rsk of transmission and simulation analysis. *J. Northeast Elec. Power Univ.* 38 (2), 21–27. doi:10.19718/j.issn.1005-2992.2018.02.004

Zhu, T. C., Nie, Y. X., Li, Z., and Li, X. Y. (2020). Parameter estimation method of stray capacitance of capacitive voltage transformer based on improved particle swarm optimization algorithm. *Autom. Electr. Power Syst.* 44 (4), 178–192. doi:10.7500/AEPS20190430012

Conflict of Interest: Author ML was employed by the State Grid Jilin Electric Power Company Limited, China. Author MC was employed by the State Grid Inner Mongolia Eastern Electric Power Company, China. Author YZ was employed by the State Grid Jilin Electric Power Company Limited, China.

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

frontiers in Energy Research

METHODS
published: 17 February 2021
doi: 10.3389/fenrg.2020.613331

Power System Network Topology Identification Based on Knowledge Graph and Graph Neural Network

Changgang Wang[1], *Jun An*[1,2]* *and Gang Mu*[2]

[1]Key Laboratory of Modern Power System Simulation and Control and Renewable Energy Technology, Ministry of Education (Northeast Electric Power University), Jilin, China, [2]School of Electrical Engineering, Northeast Electric Power University, Jilin, China

The automatic identification of the topology of power networks is important for the data-driven and situation-aware operation of power grids. Traditional methods of topology identification lack a data-tolerant mechanism, and the accuracy of their performance in terms of identification is thus affected by the quality of data. Topology identification is related to the link prediction problem. The graph neural network can be used to predict the state of unlabeled nodes (lines) through training on features of labeled nodes (lines) with fault tolerance. Inspired by the characteristics of the graph neural network, we applied it to topology identification in this study. We propose a method to identify the topology of a power network based on a knowledge graph and the graph neural network. Traditional knowledge graphs can quickly mine possible connections between entities and generate graph structure data, but in the case of errors or informational conflicts in the data, they cannot accurately determine whether the relationships between the entities really exist. The graph neural network can use data mining to determine whether a connection obtained between entities is based on their eigenvalues, and has a fault tolerance mechanism to adapt to errors and informational conflicts in the graph data, but needs the graph data as database. The combination of the knowledge graph and the graph neural network can compensate for the deficiency of the single knowledge graph method. We tested the proposed method by using the IEEE 118-bus system and a provincial network system. The results showed that our approach is feasible and highly fault tolerant. It can accurately identify network topology even in the presence of conflicting and missing measurement-related information.

Keywords: knowledge graph, power network, topology identification, graph neural network, knowledge inference

OPEN ACCESS

Edited by:
Chao Long,
Cranfield University, United Kingdom

Reviewed by:
Canbing Li,
Shanghai Jiao Tong University, China
Liang Chen,
Nanjing University of Information Science and Technology, China

***Correspondence:**
Jun An
anhuianjun@163.com

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

Received: *02 October 2020*
Accepted: *21 December 2020*
Published: *17 February 2021*

Citation:
Wang C, An J and Mu G (2021) Power System Network Topology Identification Based on Knowledge Graph and Graph Neural Network. Front. Energy Res. 8:613331. doi: 10.3389/fenrg.2020.613331

INTRODUCTION

With access to new energy sources continuously increasing and the scale of power grids growing, the variability and complexity of the operation of power grids has increased drastically (Li et al., 2018). In recent years, several blackouts in areas across the world have caused substantial economic losses and led to adverse social impacts. The prevalent online security defense systems for power grids, which focus on modeling, simulation, and fault prediction, have been severely challenged. With the rapid development of power grid measurement systems and the increasing maturity of big data technology, the state cognition and control of grid operations based on operational information on power grids has emerged as the new model for their secure operation (Liu et al., 2018). In addition, the increasing

uncertainties surrounding the use of renewable power has provided unprecedented challenges for ensuring the secure and economic operation of the current power systems (Li et al., 2019).

Understanding the topological structure of a power grid and laws of changes within it in a timely manner based on measurements is the basis for realizing the data-driven operation and control of power grids. Research on identifying the topology of a power network can be divided into several classes. Several studies have used the incidence matrix and the adjacency matrix constructed by using the switching state of the system to determine its connectivity and track topological changes (Zhu et al., 2011; Ma et al., 2014; Lourenco et al., 2015). Such methods are less tolerant of faults and conflicting telemetry data, and their effect depends entirely on the quality of remote signaling data. Other attempts have been made to identify the network topology based on graph theory and object-oriented technology (Song et al., 2005; Nian et al., 2008; Wang et al., 2009; Li et al., 2011; Yansheng et al., 2017). Such approaches involve using technology for the matrix-based identification of topology (Wang et al., 2019). They can compensate for the lower computational efficiency of the original process in terms of timeliness by optimizing the method, but are still reliant on remote signaling data for topology identification and have poor fault tolerance. Recently, the topology identification problem has been transformed into an optimal combination of measurement data problem. (Liu et al., 2019). It is a new attempt to identify network topology based entirely on data. The feasibility of this approach has been verified. Still, this method is not fault tolerant, and requires complete information concerning the active power. The determined optimal combination of data is directly affected when data are missing or when significant errors occur. Furthermore, this algorithm requires a large amount of redundant information as a supporting condition. As Power Management Units are not fully operational, owing to the massive investment required, the requisite supplementary data are difficult to obtain. In summary, current methods of topology identification depend on the quality of data. We thus need a fault-tolerant method of topology identification.

The use of knowledge graph technology has emerged in recent years in several fields, and has attracted widespread attention from both industry and academia (Liu and Wang, 2018). The knowledge graph is a mesh knowledge base of entities with attributes linked through relationships. Its value lies in organizing related information at minimal cost and generating useful knowledge. The topology of a power network is graph data reflecting node-to-node relationships. So the knowledge graph is an excellent vehicle for describing topological information (Park et al., 2019). However, information conflicts may occur in the knowledge graph due to errors or missing information. We thus need a data mining method to identify relationships that cannot be confirmed owing to the informational conflicts.

The graph neural network (GNN) is applicable mainly to non-Euclidean spatial data (Park et al., 2019). It has been applied to a variety of fields (Wu et al., 2018; Jing et al., 2020; Li et al., 2020; Xie et al., 2020; Yu et al., 2020). Multiple types of graph neural networks have been derived for different problems (Zhang et al., 2019; Nikolentzos et al., 2020; Su et al., 2020; Wu et al., 2020). Because the knowledge graph is a kind of graph data, a combination of the GNN and the knowledge graph can be used to solve all types of knowledge graph-related problems. For example, some researchers have modeled the knowledge graph with the GNN (Nathani et al., 2019). The GNN-based approach can adequately capture information on complex and hidden patterns in ternary neighborhoods. It can also achieve the relational complementation of the knowledge graph, in contrast with knowledge-based reasoning that uses individual ternary relationships. A GNN-based entity alignment scheme has been proposed (Cao et al., 2019), and it has been experimentally shown to yield highly consistent data that improve the quality of the knowledge graph on multiple datasets.

In the context of power network topology, a knowledge graph can indicate relationships between entities in the network. The graph neural network can serve as a technical tool for analyzing graph data, mining, and reasoning for relationships among entities in the knowledge graph. It also allows us to consider missing information and determine the correctness of the data, significantly improving the quality of the knowledge graph. Thus, the characteristics of the knowledge graph and the graph neural network are fully compatible with the requirements of topological recognition.

This paper proposes a method of topological identification based on the knowledge graph and the graph neural network. First, we set-up an entity–relationship–entity and entity–attribute/attribute value information triad based on remote signaling data, telemetry data, and a database of component information to construct an informational knowledge base. Second, based on the graph neural network, we determine the relationships that cannot be established owing to conflicting information. Finally, based on knowledge inference, we generate the topology of the power network and track changes in it. The results of tests show that the proposed method is fault tolerant and can accurately determine the network topology of a simulated network as well as a real power grid.

DEFINITION AND STRUCTURE OF KNOWLEDGE GRAPH

Definition of Knowledge Graph

A knowledge graph is a structured semantic knowledge base used to describe concepts and their interrelationships in the physical world in symbolic form. The basic unit is the "entity–relation–entity" triad, and the entity and its associated attribute–value pairs are interconnected through relationships to form a networked knowledge structure.

Based on the above definitions, we can draw the following conclusion:

(1) Essentially, the knowledge graph is a semantic network that reveals the relationships between entities.
(2) The research value of the knowledge graph lies in the fact that it is a reticular knowledge base built based on data and

information. It organizes the information into usable knowledge at minimum cost.

(3) The applicative value of the knowledge graph is that it can change the existing method of information retrieval. On the one hand, retrieval is achieved by reasoning. On the other hand, it can graphically display the structured knowledge after classification.

Architecture of the Knowledge Graph

The architecture of the knowledge graph consists of a data layer and a schema layer, as shown in **Figure 1**. The data layer is toward the logical architecture of the knowledge graph and the schema layer toward its technical architecture. The primary function of the data layer is the extraction of data and information to generate relationships inherent within the latter. The schema layer builds on top of the data layer and is an essential part of processing and analyzing data to generate knowledge, including such means as information fusion and processing, and creating knowledge based on inference.

METHODS OF CONSTRUCTING POWER GRID TOPOLOGY-ORIENTED KNOWLEDGE GRAPHS

The power information system records a large number of structured or semi-structured data. These data are descriptions of the measured attributes and results, and the determined entity attributes. All these data provide the necessary information units to construct knowledge graphs of the topology of the power network.

Methods of Entity–Property Construction

A large number of fields in databases recorded by the power management information system are deconstructed. The measured set of items of entity information O is thus obtained. It contains all information on nodal entities and line entities, given as:

$$O = \{N_i, N_j, L_{ij}\} \tag{1}$$

where N_i and N_j are node information units and L_{ij} is a line information unit.

The information provided by each entity corresponds to its entity attribute and attribute value. For example, an entity node (N_i) section corresponds to the description node inject telemetry (K_{ni}) and node inject telemetry (P_{ni}). Remember that an entity node (N_i) has the following attributes:

$$v_i = (K_{ni}, P_{ni}) \tag{2}$$

The resulting set of properties corresponding to the entity is as follows:

$$V = \{v_i\} \tag{3}$$

Entity O and attribute V correspond in a one-to-one manner. The expression for entity–attribute Q is as follows:

$$Q = (O, V) \tag{4}$$

Methods of Constructing Entity Relationships

Relationships are key to the knowledge graph because they require its use to solve a number of problems. In general, they make it possible to further parse the entity correspondences in the entity information set O, and to organize the entity–relationship information. The expression for the relationship is given as:

$$R_{ni,lij} = \left(N_i, L_{ij}\right) \tag{5}$$

where N_i represents node i, L_{ij} represents the line ij directly connected to node i, and $R_{ni,lij}$ is the value of the relationship between them. 1 indicates that they are directly connected and 0 indicates that they are not.

Methods for Constructing Knowledge Graph Triangles

In this paper, entity–relation–entity and entity–property–attribute (value) are transformed into an integrated entity (attribute–attribute

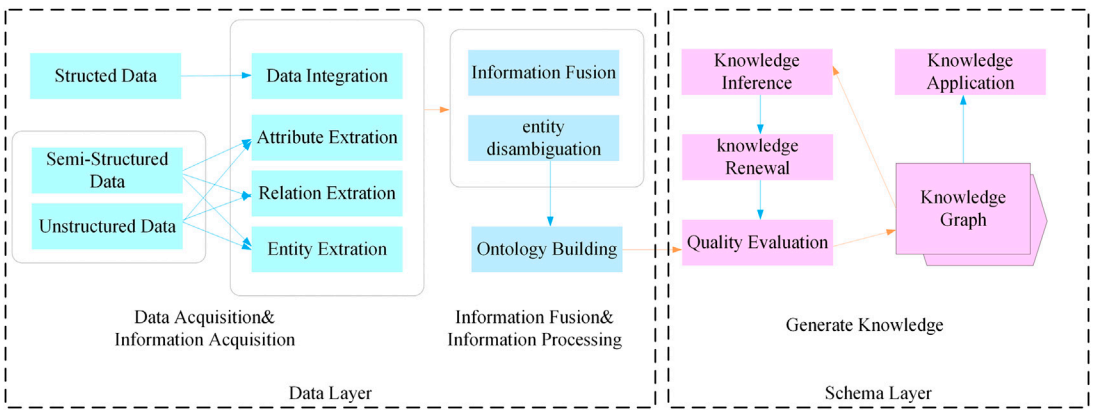


FIGURE 1 | The architecture of the knowledge graph.

value)–relation–entity (attribute–attribute value) triad to form a knowledge base for generating topologies. The expression is given as:

$$G = (O, V, R) \tag{6}$$

where G denotes the set of ternary groups, O the set of entities, and each object is unique. V denotes the set of attributes, including the number of nodes and states of line switches, the magnitude of node-injected power, the first-end power of the line, and the line first-end power difference. R denotes the set of relationships of direct correlations between entities.

GRAPHICAL NEURAL NETWORK

Traditional neural networks have had great success in extracting features from spatial Euclidean data, but are deficient in processing non-Euclidean spatial data. In recent years, researchers have learned from the ideas of convolutional networks, cyclic networks, and deep autoencoders. They have used this knowledge to design a neural network structure, called the graph neural network, for processing graph data.

A graph is composed of many vertices and edges, usually represented and stored using an adjacency matrix. The spatial features in graph data have the following characteristics:

(1) Node features: Each node has certain features.
(2) Structural features: Each node in the graph data has certain structural features.

Thus, there is a connection between pairs of nodes. In general, graph data should consider both node and structural information, and a graph neural network should be able to automatically learn both the characteristics of a node and node-to-node correlation-related information.

GCNs are the first graph-based networks to apply the simple convolution operation to image processing for graph data processing. This paper explains the principles of graph convolutional neural networks mainly from the perspective of spatial construction.

A grid can be substituted with a graph $g = (\Omega, A)$, where Ω is defined as a discrete set of nodes $\Omega \in \mathbb{R}^{N\times 1}$, and W denotes the set of edge weights $W \in \mathbb{R}^{N\times N}$. A straightforward way to represent edge weights on W is to set a threshold $\delta > 0$. Nodes adjacent to Node j can be represented as:

$$N_i = \left\{ i \in \Omega : W_{ij} = \delta \right\} \tag{7}$$

As shown in **Figure 2**, we want to convolve node 6, and adjacent to it are nodes 1 and 5 (including node 6 itself); that is $N[6]$. Therefore, a convolution of node 6 can be expressed as $W_{1,6}x_1 + W_{5,6}x_5 + W_{6,6}x_6$, and x denotes the features of every node, and w denotes the weight of the convolution.

o_v, as the convolution of the node v, can be expressed as:

$$o_v = \sum_{u \in N[v]} W_{u,v}x_u \tag{8}$$

The input node feature may be a vector with dimensions D. A single convolution operation may contain multiple convolution kernels with dimensions d. We convolve every dimension of the input feature and sum it to obtain the convolution of node v. The formula is as follows:

$$o_{v,j} = \sum_{i=1}^{D} \sum_{u \in N[v]} w_{i,j,u,v}x_{u,i}, j = 1\cdots d \tag{9}$$

The basic idea of the GCN is to reduce the dimensionality of a node's high-latitude adjacency information in the graph to a low-dimensional vector representation by aggregating feature-related information from adjacent nodes. It can aggregate the global information of the graph to represent features of the nodes. The process of updating the feature-related information of the node is shown in **Figure 3**.

The convolution operator can be defined as (Kipf and Welling, 2017):

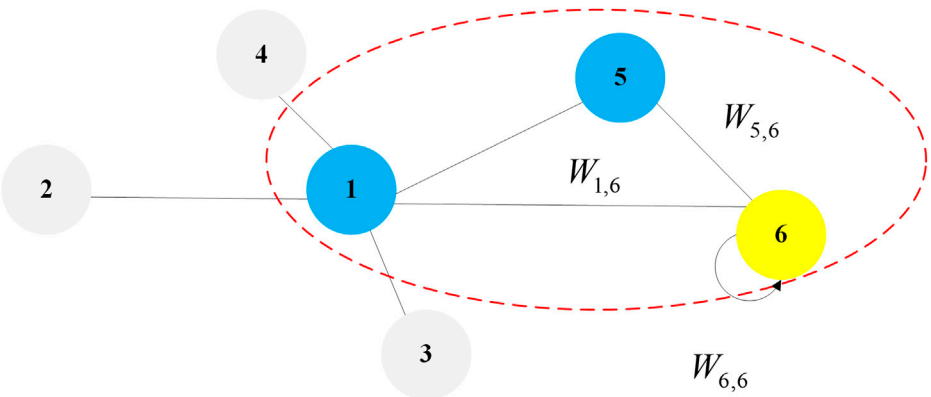


FIGURE 2 | Schematic diagram of spatial convolution for node 6.

$$h_v = f\left(\frac{1}{|n_{(v)}|}\sum_{u \in N[v]} WX_u + b\right), \forall v \in \Omega \quad (10)$$

where h_v is the linearly transformed feature value of node v after aggregating features of neighboring nodes, u is the set of nodes adjacent to v, $n_{(v)}$ is the number of nodes neighboring v, W is the weight, and b is the bias. The formula for passing features through multiple hidden layers is given as:

$$h_v^{k+1} = f\left(\frac{1}{|n_{(v)}|}\sum_{u \in N[v]} W^k h_u^k + b^k\right), \forall v \in \Omega \quad (11)$$

The matrix form of the above equation is as follows:

$$H^{(k+1)} = \sigma\left(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}H^{(k)}W^{(k)}\right) \quad (12)$$

The adjacency matrix A does not contain information on the nodes themselves. To solve this problem, given $\tilde{A} = A + I$, I is a unit array. Define the degree matrix D, $D \in \mathbb{R}^{N\times N}$, that is a diagonal matrix characterizing the degree of connectivity of a node to other nodes $\overline{D_{ii}} = \sum \overline{A_{ij}}$. Thus, the normalized $\tilde{D}^{-1/2}\tilde{A}\tilde{D}^{-1/2}$ is used as a medium to pass information in the topological space to solve the diagonal problem so that the training covers information on the nodes. $H^{(k)} \in \mathbb{R}^{N\times F}$ is the feature matrix after k-layer activation, $H^{(0)} = X$, and $W^{(k)}$ is the learnable parameter matrix of the kth layer.

In short, compared with the basic structure of a neural network, the MLP, the feature matrix multiplied by the weight matrix $H = \sigma(XW)$, adds an adjacency matrix to the graph neural network $H = \sigma(AXW)$.

The core framework of a graph neural network consists of two matrices: 1) the adjacency matrix A, $A \in \mathbb{R}^{N\times N}$, where N is the number of nodes, and 2) the feature matrix X, $X \in \mathbb{R}^{N\times F}$, where N represents the number of nodes in the graph. F is the number of dimensions of the feature of each node. Moreover, the output matrix is Z, $Z \in \mathbb{R}^{N\times C}$, and C represents the number of convolutional kernels used to implement the classification. The graph neural network is designed to maximally match the input and the output, and continually train itself to classify the input data. **Figure 4** shows the flow of the graph neural network.

PROPOSED METHOD FOR IDENTIFYING POWER NETWORK TOPOLOGY

Constructing Overall Topology of Power Grid

Through information extraction, elements of knowledge such as entities, relations, and properties can be extracted from the original corpus to obtain a series of basic factual expressions.

The relationship between entities can be extended according to knowledge inference. For example, the node Na in the entity information set O directly connects to line L_{ab}, and node N_b in the set is also directly connected to L_{ab}. The relationship can be extended to N_a-L_{ab}-N_b. Multiple entity–relation unit elements can form a topology containing all entity–relation information.

Identifying Power Network Topology

The above is not identical to knowledge. To obtain a structured and networked knowledge system, it is necessary to carry out knowledge processing and reasoning. Knowledge processing and knowledge reasoning are the essential means and critical parts of the construction of a knowledge graph. In practice, remote signaling data may be mixed with incorrect information owing to the possible existence of the blade gate relation. Telemetry data may also be combined with errors, such that the resulting data do not comply with the operating rules of the power system. Ambiguity may thus persist within the entity–relation–attribute triad, which is required for the most critical step in topological identification based on the knowledge graph. The authenticity of the entity–relation units is thus verified through the data mining method of graph neural networks.

The main problem in determining power network topology is to identify the connectivity of the lines based on their properties (values). In applying the graph neural network, we use the relationship between entities in the knowledge graph (lines) as "vertices" and the entities (nodes) as "edges." In this way, the connectivity of lines is directly generated by line training, which helps to avoid the conventional approach of first determining the nodes and then identifying connectivity. The follow chart of the proposed method is shown in **Figure 5**.

The specific steps are as follows:

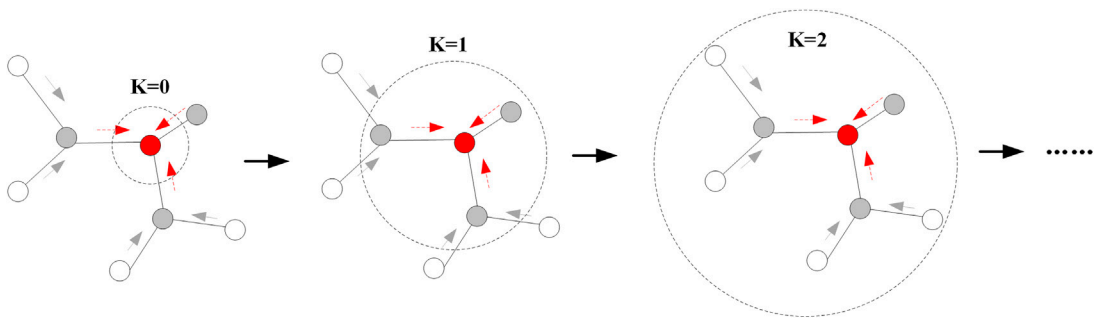


FIGURE 3 | Schematic diagram of node feature update based on graph convolutional neural network.

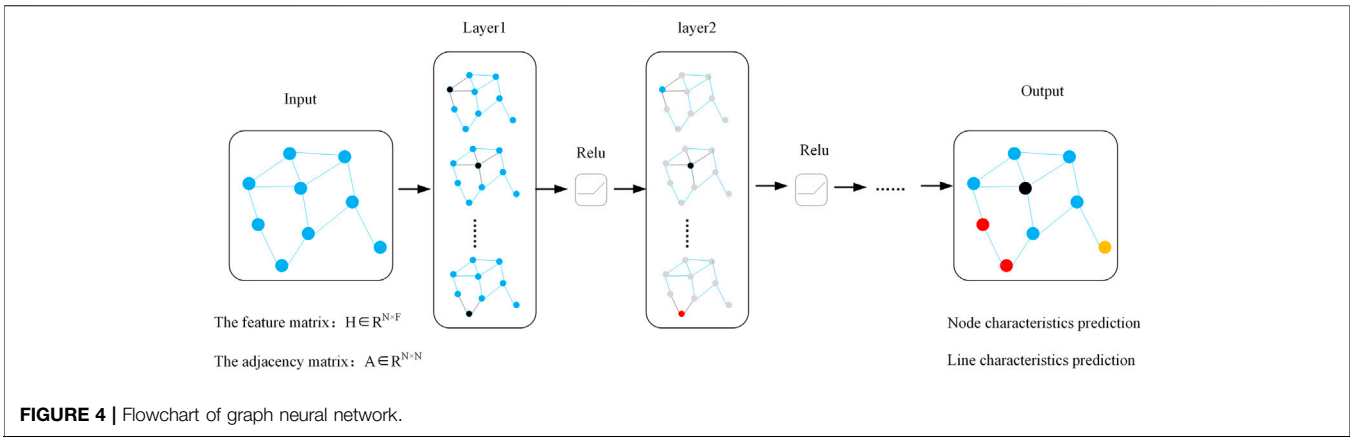


FIGURE 4 | Flowchart of graph neural network.

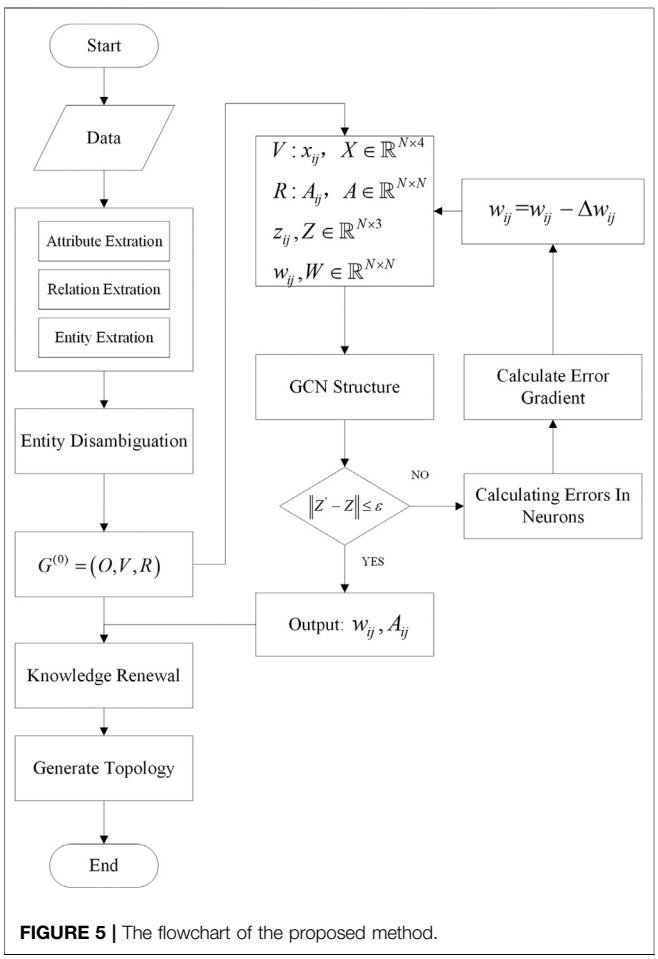


FIGURE 5 | The flowchart of the proposed method.

Step 1. Extract the attributes and attribute values of edges in the knowledge graph that contain the line input power P_{ij} and the line output power P_{ji}, and the head switch statuses of the line S_{ij} and S_{ji} to form a description of the characteristic matrix of the line:. P_{ij} or P_{ji} is the measured power, and S_{ij} and S_{ji} are the descriptions of the on and off states of the line: on is 1 and off is 0.

TABLE 1 | comparison of variables and parameters.

Node i	Node j	Line input power/ MW	Line output power/ MW	Line head switching status	Line end switching status
8	30	0	0	0	0
48	49	0	0	0	0
17	18	0	0	1	1
21	22	−42.837	43.256	0	0
34	36	0	−30.162	1	1
76	77	−61.150	0	1	1
84	85	0	36.791	1	1
92	100	31.479	0	1	1
105	107	26.755	−26.347	0	1
114	115	1.358	−1.358	1	0

Step 2. P_{ji} and P_{ij} in x_{ij} are normalized to form $\hat{x}_{ij}$ to ensure that the eigenvalues $\hat{x}_{ij}$ are in the interval [-1,1], such that all $\hat{x}_{ij}$ are aggregated into X, $X \in \mathbb{R}^{N\times4}$, and N is the total number of lines.

Step 3. Define z_{ij} as the line category identifier, and make all aggregates into Z, which contains all line categories for the graph neural network training comparison.

Step 4. Extract all entity relations of the knowledge graph $G = (O, V, R)$, that is, the node–line relationship, to form an adjacency matrix A, with a line as the vertex and a node as the edge, to describe the direct connection between lines. A_{ij} represents whether there is a common node between lines i and j.

Step 5. Set $IT_{\max}$ to the maximum number of iterations. Substitute the adjacency matrix A and the feature matrix X into the feature transfer formula, calculate Z'. $\|Z' - Z\|$ as a loss function, and modify the weight matrix continuously. After iterative computation, use the weight matrix W of each hidden layer, on behalf of the graph neural network after self-training rules of decision, directly to determine line connectivity.

Step 6. After connectivity determination, update the adjacency matrix A'. Finally, generate the topology by deducing the correlation between the internal line and the line according to depth search to analyze the generated topology and determine whether isolated nodes and islands are in operation.

Algortihm 1 | Overall training and testing process
Input:
Graph $G(V,\varepsilon)$; Input features $\{X_v, \forall v \in V\}$; randomly initialized network weight W; non-linearity σ; the maximum number of iterations IT_{max}; GCN structure;
Output:
Network model , the weight value W, Generate topology
1.**Function:**
2.Calculating the graph structure G;
3.Normalize input features;
4.Integrate features to get GCN input through G;
5.Integrate features to get corresponding labels.
6.**for** range of IT_{max} **do**
7.Calculate network through input;
8.Calculate loss through network output and labels;
9.Do backpropagation algorithm through optimizer, update network weight W
10.**end**
11.Calculate network output;
12.Do anti-normalization for the network output
13.generate topology;
14.Save network model;

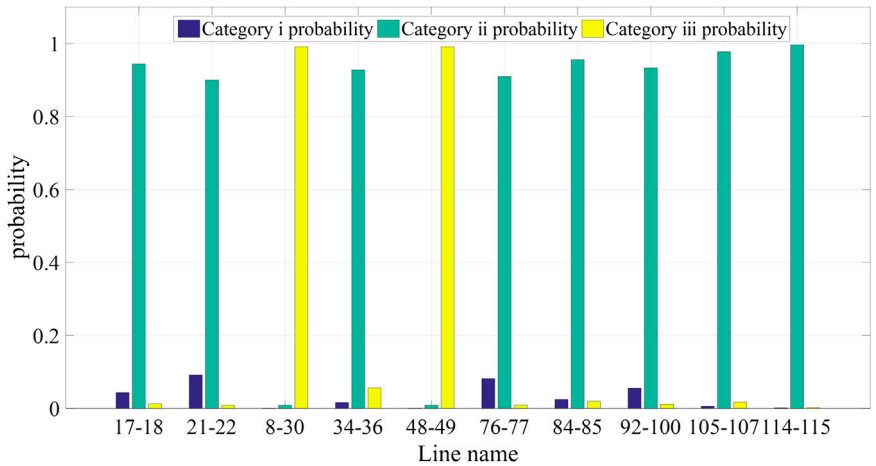


FIGURE 6 | Histogram of probability of line determination for a disconnected line or missing information.

EXPERIMENT AND RESULTS

Experimental Results on a Test Sample

The IEEE 118-node system containing 186 lines was used for our experiment. The initial active power of the nodes was obtained by the MATPOWER software simulation. We set-up two disconnected lines: 8–30 and 48–49. The lines 17–18, 21–22, 34–36, 76–77, 84–85, 92–100, 105–107, and 114–115 contained conflicting information. The remaining lines were normal. Details of the information in the lines are shown in **Table 1**. Algorithm 1 shows the overall training and testing processes.

The above sample was selected to demonstrate the effectiveness of the proposed method. We followed the flow of Algorithm 1 in calculating the sample data, and then used the result of the judgment of this sample with the original features of the line. The probability of categorization and the number of line characteristics that determine the presence of conflicting information and "off" lines are shown in **Figure 6** and **Table 2**.

It is clear that only lines 8–30 and 48–49 had a high probability of falling into category iii in the above diagram. According to the results of training of the graph neural network, these two lines were judged to be disconnected, consistently with the original perturbation.

According to the probability distribution, lines 27–115, 114–115, 68–116, 12, 117, 54–55, 80–81, 75–118, and 76–118 were all judged to be connected but had conflicting information. The results were consistent with the original settings. We updated the knowledge graph based on the above results of identification. A comparison of the results of topological identification is shown in **Figure 7**.

Experimental Results on the IEEE 118-Node System

We used an IEEE 118-node system containing 186 lines. The initial active power of the nodes was obtained by MATPOWER software simulation. A 1% random error obeying a Gaussian

TABLE 2 | Probability distribution of the presence of conflicting information and disconnected lines in various classifications.

Node *i*	Node *j*	Category i probability	Category ii probability	Category iii probability
8	30	0.0051	0.0163	**0.9786**
48	49	0.0051	0.0163	**0.9786**
17	18	0.0012	**0.8854**	0.1134
21	22	0.0009	**0.9987**	0.0004
34	36	0.0001	**0.9996**	0.0004
76	77	0.0003	**0.9969**	0.0028
84	85	0.0055	**0.9772**	0.0172
92	100	0.0016	**0.9963**	0.0021
105	107	0.0019	**0.9979**	0.0002
114	115	0.0014	**0.9984**	0.0002

The number in bold is the value with the largest probability distribution for each line. It indicates that the lines are grouped into the corresponding categories.

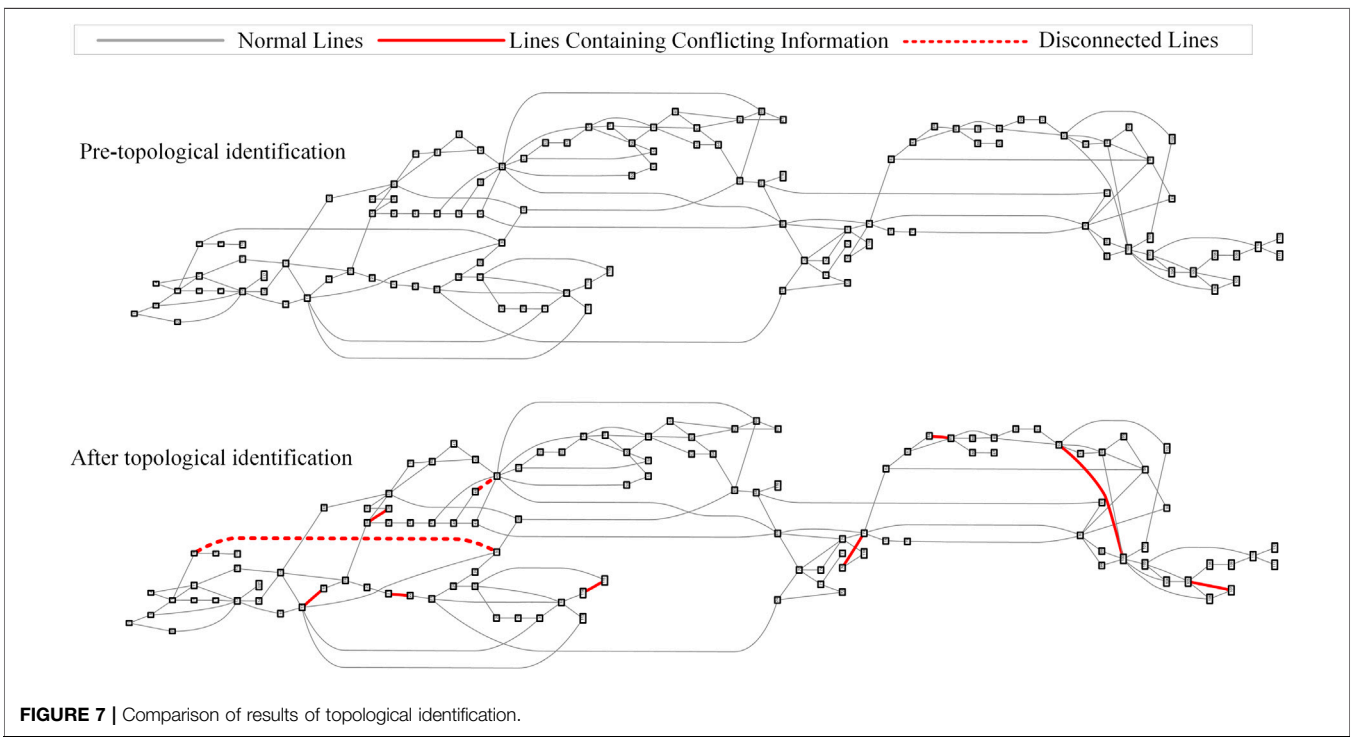


FIGURE 7 | Comparison of results of topological identification.

distribution for active node power was mixed in to generate 1,000 training samples. For each sample, two lines were randomly disconnected, and eight lines were made to randomly miss some attribute values. We formed an information base for the topology of the power grid according to the construction of the knowledge graph, and organized the data and relationships in the node, line feature matrix, and adjacency matrix. The feature matrices of 900 training samples were labeled as follows: lines in label category i represented connected lines, those in label category ii represented lines with conflicting information in the feature matrix despite being related lines, and lines in label category iii represented disconnected lines. One hundred samples were retained as unlabeled samples.

Seven hundred labeled samples were used as input to the graph neural network for training and another 200 were used to compare the ingested topology with the generated topology through the rules learned from the graph neural network.

For each sample, the results of identification of line connectivity could be organized into a 186 × 1 column vector, and the number of dimensions of the matrix of results for 200 samples were 186 × 200. By marking different categories with different colors, a map of the results of connectivity was obtained for all samples. We statistically determined the determinations of the remaining 199 samples. The results are shown in **Figure 8**.

In the figure above, the abscissa represents the number of test samples and the ordinate the line number. The blue blocks represent lines judged to belong to category ii and the red blocks represent those judged to be in category iii. Each sample contained two category iii lines and eight category ii lines. Compared with the initial topology, the graph neural network method accurately identified the disconnected lines.

By using the rules learned from the graph neural networks, we identified the topologies of the remaining 100 unlabeled samples. The results are shown in **Figure 9**.

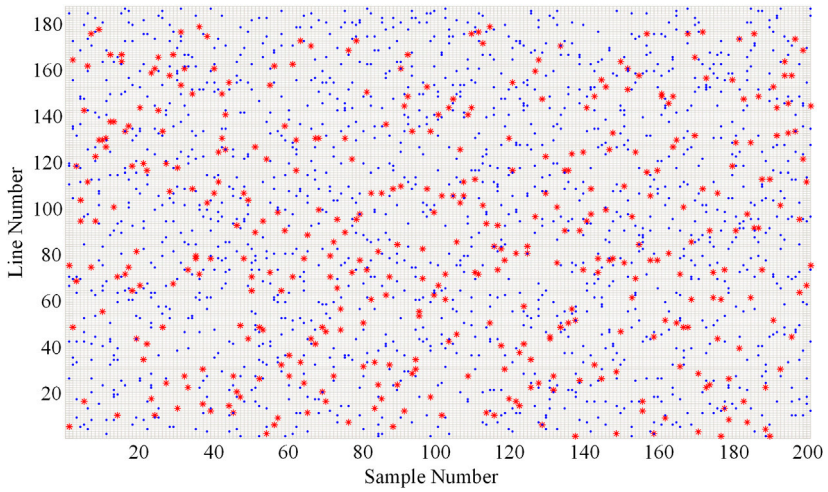


FIGURE 8 | Scatter plot of results of identification of test samples.

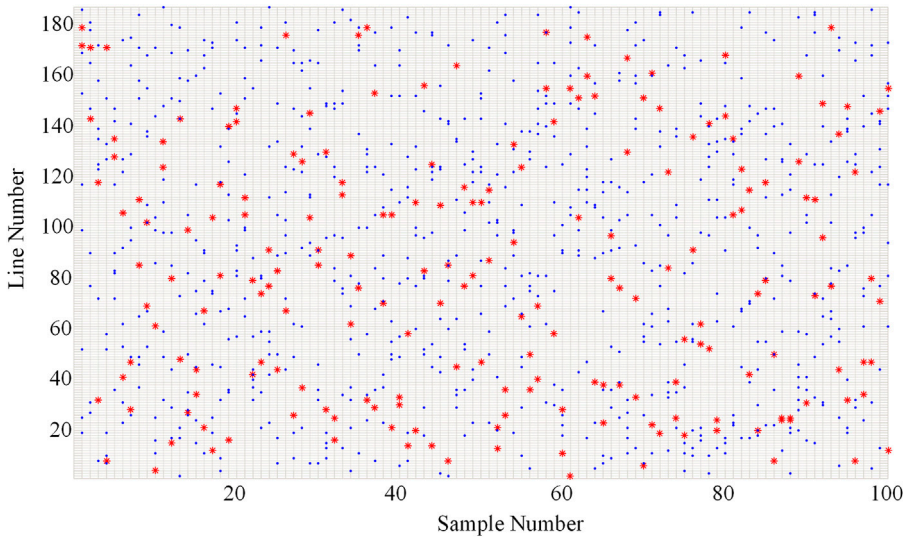


FIGURE 9 | Scatter plot of results of identification of unlabeled samples.

As shown in the figure above, we were able to identify the operating status of each line. Each unlabeled sample contained two Category iii lines and eight Category ii lines, which is consistent with the original assumption.

For the samples remaining as labels, the right decision is made to compare the decision results with their original feature quantities, and the results are consistent.

To further verify the proposed method, we selected the above 100 unlabeled samples as test samples. The method used for comparison with the proposed one was a topology identification method based only on knowledge graphs. Both methods are applied to the same data sample.

Each sample contained 184 connected lines and two disconnected lines. Eight of the connected lines were partially missing information. As is shown in **Figure 8**, the proposed method accurately identified the line connections and determined the classification of each line. We thus determined that there were 18,400 connected lines and 200 disconnected lines in the 100 test samples; 800 of the connected lines had missing partial information, which was consistent with the original settings. The comparison of the results with those of the method for topology identification based only on knowledge mapping is shown in **Figure 10**.

The method for topology identification based only on knowledge mapping accurately identified disconnected lines, but its accuracy of identifying connected lines was 95.65%. This is mainly because knowledge-based topology identification was less fault tolerant than the proposed method and struggled to deal with missing information.

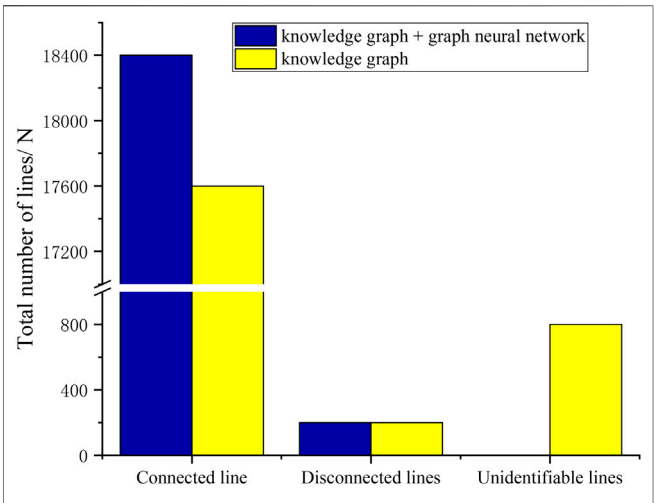


FIGURE 10 | Results of topological identification of the two methods.

In summary, the proposed method for topology identification based on a combination of the knowledge graph and the graph neural network can determine the state of line connectivity, identify topological changes, and has strong fault tolerance. Even in the case of conflicting information, it can identify the topology of the power network, and is thus superior to the traditional method of recognition based only on the knowledge graph.

Experimental Results on a Provincial Network

The proposed method was applied to a provincial network consisting of 132 nodes and 181 lines. The data were recorded at intervals of 1 min. We combined the knowledge graphs and the graph neural networks to identify the topology of the power network.

Historical operational data provided a sufficient number of samples to train the graph neural network, thus enabling it to make more accurate decisions. Nevertheless, with increasing numbers of training samples, the number of dimensions of the

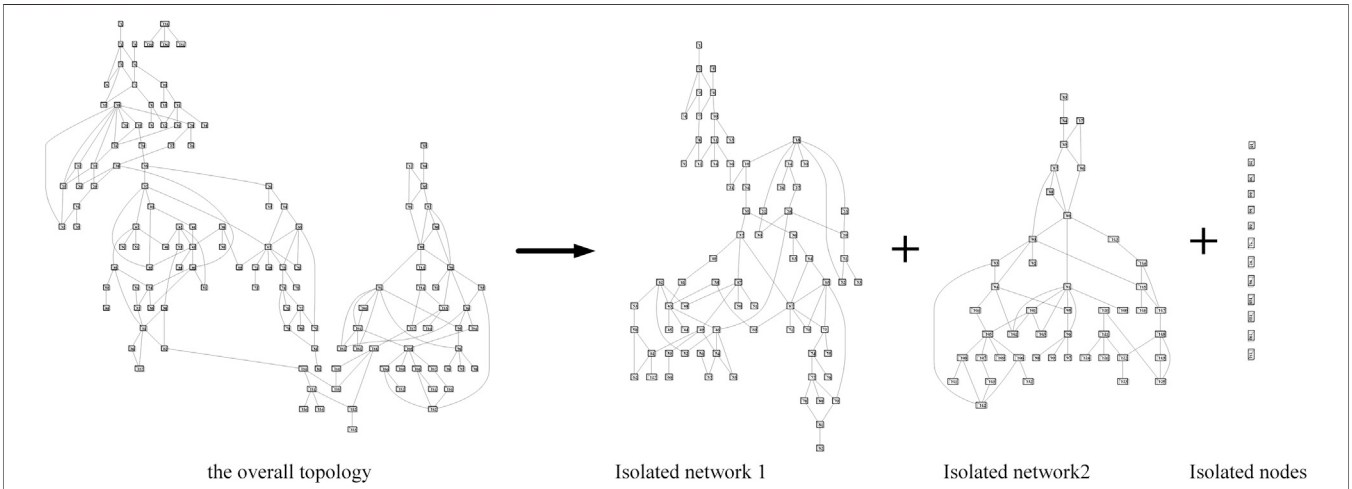


FIGURE 11 | Network structure after topology identification.

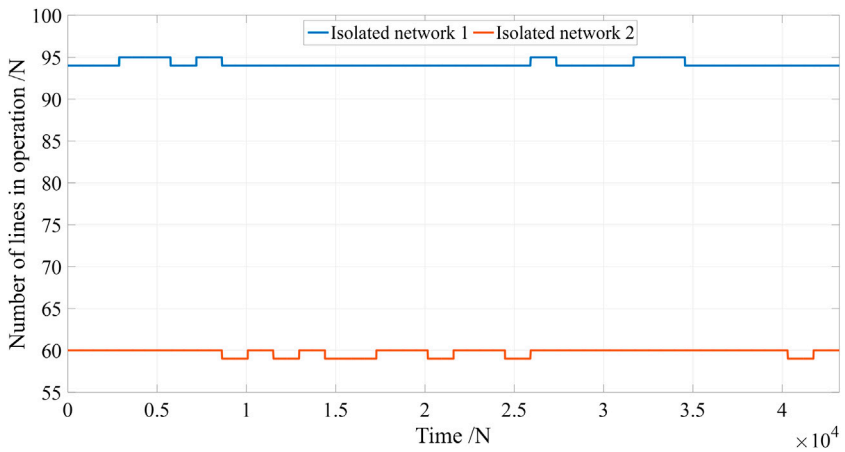


FIGURE 12 | Changes in the number of lines in operation on the isolated networks in the time series.

TABLE 3 | Information on the composition of the network structure.

Network structure	Node name	Line name
Isolated nodes	'15','25','39','60','66','68','72','97','98','128','129','130','131'	
Isolated network 1	'1','2','3','4','5','6','7','8','9','10','11','12','13','14','16','18','19','20','21','22','23','24','26','27','28','29','30','31','32','33','34','35','36','37','38','40','41','42','43','44','45','46','47','48','49','50','51','52','53','54','55','56','57','58','59','61','62','63','64','65','67','69','70','71','73','74','75','76','77','78','79','80','81','82','127'	'1-2','2-3','2-4','2-6','3-4','3-7','5-6','6-7','6-10','7-8','8-9','8-12','10-11','10-13','11-12','11-14','11-16','16-21','18-19','18-20','18-22','18-23','18-24','18-32','19-21','19-34','22-30','23-29','24-26','24-27','27-28','28-29','28-30','28-46','29-31','31-32','31-33','34-35','35-36','35-37','36-63','36-64','37-40','37-47','37-67','38-43','38-69','40-41','41-43','42-48','42-52','42-53','43-44','43-45','43-46','46-52','46-54','46-55','47-48','47-49','47-50','47-51','49-59','49-61','53-58','54-55','54-57','56-57','58-61','58-62','61-62','61-127','64-65','64-67','65-67','65-70','65-73','65-79','67-69','67-70','67-71','67-73','73-74','73-75','74-76','74-77','77-78','77-79','77-80','78-81','79-81','81-82'
Isolated network 2	'17','83','84','85','86','87','88','89','90','91','92','93','94','95','96','99','100','101','102','103','104','105','106','107','108','109','110','111','112','113','114','115','116','117','118','119','120','121','122','123','124','125','126','132'	'17-85','17-86','83-84','84-85','85-86','85-87','86-89','87-88','87-89','87-90','88-89','89-90','89-91','89-113','90-92','90-93','90-94','90-115','91-95','91-96','91-100','91-101','91-102','91-103','93-94','93-112','94-95','94-102','94-104','95-96','96-99','101-102','101-105','104-105','105-106','105-107','105-108','105-109','106-112','106-132','107-110','108-111','108-112','110-112','113-114','114-115','114-117','115-116','115-117','117-118','118-119','118-122','118-125','119-125','120-121','120-125','121-122','121-124','121-126','122-123'

adjacency matrix A increased from the original $N \times N$ to the original $\tilde{N} \times \tilde{N}$, $\tilde{N} = N*(t+1)$. Considering that the training time increased with the amount of input data, we trained the samples with a window length of 10 min and a step length of 10 min, used the number of line features in period t as input, and determined the network topology in period t+1 based on the results of training.

Based on one-month data from the power grid, we found that the provincial network comprised two orphaned network topologies and several isolated nodes. The details are shown in **Figure 11**.

As shown in **Figure 6**, the full-time network consisted of 13 remote nodes and two isolated networks, with a total of 119 nodes in continuous operation and 153–155 lines involved. The structural information of the network is shown in **Table 3**.

We verified that 13 of the isolated nodes were planned nodes that had not yet been put into use in the given month, which is consistent with the results of topology identification. Isolated network 1 contained 75 nodes with 94–95 lines in operation, of which lines 28–46, 42–52, and 65–67 had pitching and cutting variations. The remaining lines in isolated network 1 were in stable operation. Isolated network 2 contained 44 nodes and 59–60 lines in operation, of which lines 91–96 and 118–125 had pitching and cutting variations. The number of lines in operation in chronological order in two isolated networks are shown in **Figure 12**.

The results show that the proposed method is feasible. It can accurately identify the state of line switching, generate network topology, update the knowledge graph, and improve its quality. It also provides a good basis for analyzing the islanding operation, demodulation control, and the operating state and control of power grids.

CONCLUSION

The main contribution of this paper is the proposal of a method based on a knowledge graph and a graph neural network. This method transforms the traditional problem of topology identification into one of inferring the state of line connections based on graphical data for topology identification.

The proposed method is distinct from methods that used knowledge graphs only for topology identification because it contains an additional process for inferring conflicting information. After obtaining the overall topology covering all entities in the network, we performed knowledge inference on contradictory information based on graph neural networks. We then determined the line connectivity and updated the entity–relation information in the graph. In this way, we compensated for the deficiency of topological identification based only on the knowledge graph while making substantial gains in identifying topology and tracking changes in it.

We experimentally demonstrated that our proposed method is fault tolerant, unlike traditional methods, and correctly identifies line connectivity even in the case of informational conflicts.

We tested the proposed approach on an operational power grid to show that it can identify network topology, isolated island operations in the network, and can track changes in topology. This can yield valuable information for analyzing power system operations.

The graph neural network expands the training samples using historical data. The dimensionality of the adjacency matrix increases exponentially, which inevitably reduces training efficiency. Optimizing the graph neural network to improve efficiency should be investigated in future work.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

CW: Data curation, Writing–Original draft preparation and Writing–Reviewing JA: Conceptualization and Methodology GM: Supervision and Editing.

FUNDING

This work was supported by the National Natural Science Foundation of China (No. 51877034).

REFERENCES

Cao, Y. X., Liu, Z. Y., Li, C. J., Liu, Z. Y., Li, J. Z., and Chua, T. S. (2019). "Multi-channel graph neural network for entity alignment," in 57th annual meeting of the association for computational linguistics, Florence, Italy, August 26, 2019 (Stroudsburg, PA: Acl 2019), 1452–1461.

Jing, Y., Wang, J., Wang, W., Wang, L., and Tan, T. (2020). Relational graph neural network for situation recognition. *Pattern Recogn.* 108, 107544. doi:10.1016/j.patcog.2020.107544

Kipf, T. N., and Welling, M. (2017). "Semi-supervised classification with graph convolutional networks," in 5th International Conference on Learning Representations, Toulon, France, April 24, 2017–April 26, 2017.

Li, F., Liu, T., and Jiang, D. (2011). Distribution network reconfiguration with multi-objective based on improved immune algorithm. *Dianwang Jishu/Power Syst. Technol.* 35 (7), 134–138. doi:10.13335/j.1000-3673.pst.2011.07.028

Li, M., Ma, Z., Wang, Y. G., and Zhuang, X. (2020). Fast Haar transforms for graph neural networks. *Neural Network.* 128, 188–198. doi:10.1016/j.neunet.2020.04.028

Li, Y., Wang, J., Zhao, D., Li, G., and Chen, C. (2018). A two-stage approach for combined heat and power economic emission dispatch: combining multi-objective optimization with integrated decision making. *Energy* 162, 237–254. doi:10.1016/j.energy.2018.07.200

Li, Y., Yang, Z., Li, G., Zhao, D., and Tian, W. (2019). Optimal scheduling of an isolated microgrid with battery storage considering load and renewable generation uncertainties. *IEEE Trans. Ind. Electron.* 66 (2), 1565–1575. doi:10.1109/tie.2018.2840498

Liu, D., Zhang, R., Li, B., Yun, Z., Yang, H., and Li, Z. (2019). Research on reverse recognition method of power system topological structure based on state measurement. *Zhongguo Dianji Gongcheng Xuebao/Proc. Chin. Soc. Electr. Eng.* 39 (24), 7107–7117. doi:10.13334/j.0258-8013.pcsee.182454

Liu, Y., Zhang, N., and Kang, C. (2018). A review on data-driven analysis and optimization of power grid. *Dianli Xitong Zidonghua/Autom. Electr. Power Syst.* 42 (6), 157–167. doi:10.7500/AEPS20170922003

Liu, Z., and Wang, H. (2018). Retrieval method for defect records of power equipment based on knowledge graph technology. *Dianli Xitong Zidonghua/Autom. Electr. Power Syst.* 42 (14), 158–164. doi:10.7500/AEPS20180103007

Lourenco, E. M., Coelho, E. P. R., and Pal, B. C. (2015). Topology error and bad data processing in generalized state estimation. *IEEE Trans. Power Syst.* 30 (6), 3190–3200. doi:10.1109/Tpwrs.2014.2379512

Ma, J., Zhang, Y., Ma, W., and Wang, Z. (2014). Power network topological analysis based on incidence matrix notation method and loop matrix. *Dianli Xitong Zidonghua/Autom. Electr. Power Syst.* (12), 74–80. doi:10.7500/AEPS20131212003

Nathani, D., Chauhan, J., Sharma, C., and Kaul, M. (2019). "Learning attention-based embeddings for relation prediction in knowledge graphs," in 57th Annual Meeting of the Association for Computational Linguistics, Florence, Italy, ACL, July 28, 2019–August 2, 2019, (Stroudsburg, PA: Association for Computational Linguistics (ACL)), 4710–4723.

Nian, M., Dong-Yuan, S., and Xian-Zhong, D. J. P. S. T. (2008). A novel method for fast power network topology formation and partial revision based on graph theory. *Power Syst. Technol.* 32 (13), 35–39.

Nikolentzos, G., Dasoulas, G., and Vazirgiannis, M. (2020). k-hop graph neural networks. *Neural Networks* 130, 195–205. doi:10.1016/j.neunet.2020.07.008

Park, N., Kan, A., Dong, X. L., Zhao, T., and Faloutsos, C. (2019). Estimating node importance in knowledge graphs using graph neural networks. Proceedings of the 25th Acm Sigkdd International Conferencce on Knowledge Discovery and Data Mining Kdd'19, Anchorage, AK, August 4–8, 2019, (New York, NY: ACM), 596–606.

Song, S.-Q., Zhu, Y.-L., and Yu, H. (2005). A power network topology tracking method based on graph theory and artificial intelligence search technique. *Power Syst. Technol.* 29 (19), 45–49.

Su, C., Gao, S., and Li, S. (2020). GATE: graph-attention augmented temporal neural network for medication recommendation. *IEEE Access* 8, 125447–125458. doi:10.1109/Access.2020.3007835

Wang, S., Xu, Q., Zhang, G., and Yu, L. (2009). Modeling of wind speed uncertainty and interval power flow analysis for wind farms. *Dianli Xitong Zidonghua/Autom. Electr. Power Syst.* 33 (21), 82–86. doi:10.1016/j.apm.2007.10.019

Wang, T., Liu, Y., Gu, X., and Qin, X. (2019). Vulnerable lines identification of power grid based on cascading fault space-time graph. *Zhongguo Dianji Gongcheng Xuebao/Proc. Chin. Soc. Electr. Eng.* 39 (20), 5962–5972. doi:10.13334/j.0258-8013.pcsee.181730

Wu, B., Liu, Y., Lang, B., and Huang, L. (2018). DGCNN: disordered graph convolutional neural network based on the Gaussian mixture model. *Neurocomputing* 321, 346–356. doi:10.1016/j.neucom.2018.09.008

Wu, Z., Pi, D., Chen, J., Xie, M., and Cao, J. (2020). Rumor detection based on propagation graph neural network with attention mechanism. *Expert Syst. Appl.* 158, 113595. doi:10.1016/j.eswa.2020.113595

Xie, Y., Xu, H., Li, J., Yang, C., and Gao, K. (2020). Heterogeneous graph neural networks for noisy few-shot relation classification. *Knowl. Base Syst.* 194, 105548. doi:10.1016/j.knosys.2020.105548

Yansheng, L., Jing, L. I., Yadi, L., Lingyun, W. U., Qiang, L. I., Jun, Z., et al. (2017). Large power grid topology analysis based on graph partitioning. *Power Syst. Prot. Control* 45 (23), 108–115. doi:10.7667/PSPC171060

Yu, M., Zhang, Z., Li, X., Yu, J., Gao, J., Liu, Z., et al. (2020). Superposition graph neural network for offshore wind power prediction. *Future Generat. Comput. Syst.* 113, 145–157. doi:10.1016/j.future.2020.06.024

Zhang, C., Yu, J. J. Q., and Liu, Y. (2019). Spatial-temporal graph attention networks: a deep learning approach for traffic forecasting. *IEEE Access* 7, 166246–166256. doi:10.1109/Access.2019.2953888

Zhu, W., Zhou, X., and Tang, R. (2011). Elasticity network topology mapping for power grid. *Zhongguo Dianji Gongcheng Xuebao/Proc. Chin. Soc. Electr. Eng.* 31 (31), 109–117. doi:10.1016/S1872-5805(11)60064-4

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

frontiers in Energy Research

ORIGINAL RESEARCH
published: 26 February 2021
doi: 10.3389/fenrg.2021.634603

A Real-Time Electricity Price Decision Model for Demand Side Management in Wind Power Heating Mode

Qiang Li[1], Jian Li[1], Zhengyong Huang[1], Fulin Fan[2] and Weijun Teng[3]*

[1]School of Electrical Engineering, Chongqing University, Chongqing, China, [2]Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, United Kingdom, [3]State Grid Henan Electric Power Research Institute, Zhengzhou, China

The problem of wind power curtailment (WPC) during winter heating periods in China's "Three-North regions" is becoming worse. Wind power heating, though being an effective way to increase wind power consumptions, is constrained by high electric heating costs under a peak-to-valley electricity price pattern. This study develops a real-time price (RTP) decision model which adjusts the time-varying RTPs within an acceptable range of heating users based on the WPC distribution over a particular dispatch day. The lower RTPs accompanying the higher WPC can guide the electric heating user side equipped with regenerative electric boilers (REBs) to actively increase REB imports to absorb additional wind generation. Then, the demand side response using REBs under the RTP scheme is optimized to minimize the total heating cost met by electric heating users while assisting in the large-scale wind generation accommodation. The total heating costs and WPC reductions under different heating scenarios are compared and discussed alongside the effectiveness of the RTP-based demand side management in terms of reducing the WPC and heating costs and increasing the feasibility of wind power heating during winter heating periods.

Keywords: demand side management, electric heating, real-time electricity price, regenerative electric boiler, wind power heating

OPEN ACCESS

Edited by:
Yang Li,
Northeast Electric Power University, China

Reviewed by:
Chen Liang,
Nanjing University of Information Science and Technology, China
Shaoyan Li,
North China Electric Power University, China

****Correspondence:***
Qiang Li
419737100@qq.com

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

Received: *28 November 2020*
Accepted: *12 January 2021*
Published: *26 February 2021*

Citation:
Li Q, Li J, Huang Z, Fan F and Teng W (2021) A Real-Time Electricity Price Decision Model for Demand Side Management in Wind Power Heating Mode. Front. Energy Res. 9:634603. doi: 10.3389/fenrg.2021.634603

INTRODUCTION

By the end of 2019, the cumulative installed capacity of wind power in China reached 210 GW, and the annual national wind generation was about 405.7 TWh (National Energy Administration, 2020). Though the total installed capacity ranked first in the world, the severe issue of wind power curtailment (WPC) restricts China from becoming the largest wind power utilization country. The "Three-North regions" of China having abundant and concentrated wind energy resources are suffering from the most significant WPC, accounting for 98% of the national WPC of 41.9 TWh in 2017 (National Energy Administration, 2020). This is mainly because the installed capacity of renewable energy in the "Three-North regions" is inversely correlated with the local electricity demand. As of the end of 2017, grid-connected wind farms in the "Three-North regions" contributed to about 74.23% of the national total installed capacity and had a wind power penetration rate as high as 70%, while the local electricity consumption reached 36% of the national volume only.

In particular, the electricity generation mix of the "Three-North regions" is dominated by thermal power, most of which is supplied by combined heat and power (CHP) units. In winter heating periods, the mode of "ordering power by heat" reduces the ability of thermoelectric units to adjust their electrical power outputs, resulting in a more significant WPC phenomenon (Liu, 2013; Lü et al.,

2013; Yao et al., 2018). The CHP units providing 92% of the total installed capacity of thermal power in Jilin Province must increase their minimum export limits to meet the heat demand required during winter heating periods. In addition, heat demand peaks generally coinciding with electricity demand troughs further reduce the feasible region of wind power acceptance and compress the space for wind power to access the local grids.

Given actual characteristics of the "Three-North regions," improving the flexibility of CHP units is expected to significantly benefit for the wind power integration (Li et al., 2015; Cui et al., 2016; Xie et al., 2016; Li et al., 2018a; Li et al., 2020). A chance-constrained programming-based scheduling model was developed in (Li et al., 2020) to deal with the insufficient flexibility of an integrated energy system with CHP units. To address the problem of the CHP economic emission dispatch, a two-stage approach combining multi-objective optimization with integrated decision-making was proposed in (Li et al., 2018a). With the deployment of electric boiler and thermal energy storage, a coordinated wind power accommodating dispatch model was developed in (Cui et al., 2016) to improve the operational flexibility of CHP units. In addition, heating with wind power, as an effective solution to the WPC problem, has been significantly promoted in China in recent years (Comprehensive Department of National Energy Administration, 2015). The principle of the wind power heating mode and its role and prospects in improving the grid's capability of wind power accommodation during off-peak periods were investigated in (Wang et al., 2013).

Energy storage systems (ESSs) can convert energy between multiple forms and enable the time shift of a particular energy form, permitting an effective coordination between power supply and demand (Chen et al., 2015; Gifford et al., 2020; Fan et al., 2019). Nowadays, regenerative electric boiler (REB) as one of the heat storage technologies has been deployed to alleviate the WPC (Xu and Jiang, 2015; Cui et al., 2016). An integrated energy system including wind farms, REBs, and ESS was modeled in (Xu and Jiang, 2015) where the WPC was minimized taking into account the ESS scale and the operating conditions of REBs. The China's energy "13th Five-Year Plan" anticipated expanding the scale of the wind power heating mode and promoting the integration of REBs with electric heating (Energy Administration of Jilin Province, 2017). In order to maximize wind power consumptions, the import of heat storage is expected to exactly track the WPC. However, the peak-to-valley pattern of electricity price currently adopted in the "Three-North regions" affects the user's choices of winter heating methods and thus the extent to which the heat storage can assist in the wind power accommodation.

The electricity users in a long-standing consumption mode are reluctant to adjust their consumption behavior to match the load schedule. In order to promote their active participation in the load shaping, an attractive price signal needs to be offered as an incentive (Ketterer, 2014; Shen et al., 2016). Given a reasonable price elasticity of demand, the demand side response (DSR) to short-term variations of electricity prices plays an increasingly important role in balancing short-term supply and demand, especially during peak periods, and in dealing with fluctuations of renewable power outputs (Khodayar and Shahidehpour, 2013; De Jonghe et al., 2014; Sánchez de la Nieta et al., 2014; Shafie-khah et al., 2015). The demand side self-adjustment is more economical and practical than investing in the power supply side to meet the peak demand. Therefore, encouraging users to actively switch to the wind power heating mode is expected to be a practical and direct way to increase wind power consumptions during winter heating periods. This can be realized by reducing real-time electricity prices (RTPs) in response to WPC signals, which guides the users to actively increase the electric heating demands to consume the additional wind generation. The introduction of the DSR under the RTP scheme not only permits an active demand side management (DSM) for the WPC alleviation (Sioshansi and Short, 2009; Ai et al., 2016; Shin et al., 2017; Schmidt and Linder, 2020) but can also reduce heating costs of users, increasing their attractiveness of the electric heating in cost compared to the conventional heating. An RTP decision method has been proposed in (Li et al., 2018b) which constructed a bi-level optimal scheduling model to maximize the system profit.

Most research related to RTPs only sets a single electricity price signal in response to WPC. From the perspective of supply and demand in electricity markets, the article develops an RTP decision model to adjust the time-varying RTPs based on the scale and temporal distribution of WPC in a dispatch day. Then, the responses of electric heating with REBs under the RTP scheme are optimized to minimize the total electric heating cost while realizing the large-scale accommodation of wind generation. The daily heating costs and WPC reductions under four particular heating scenarios are evaluated to verify the effectiveness of the proposed RTP decision model for DSM in terms of reducing the electric heating cost and increasing wind power consumptions.

This article is structured as follows: *Theoretical Basis of RTP Decision Model for DSM* section introduces the theoretical basis of the RTP decision model for DSM in the wind power heating mode; *RTP Decision Model Based on WPC Distribution* section describes the RTP adjustment based on WPC distributions; *DSR Model Under RTP Scheme for Wind Power Heating* section develops a DSR model using REBs under the RTP scheme; *Results and Model Validation* section evaluates the performance of the proposed models in reducing the WPC and heating costs; and *Conclusion and Future Work* section presents conclusions and recommendations for further work.

THEORETICAL BASIS OF RTP DECISION MODEL FOR DSM

Wind Power Accommodation by Electric Heating With Heat Storage

The daily curves of total available wind power outputs and electrical and heat demands during a typical high-wind winter heating period of 2015–2016 in a particular province in the "Three-North regions" are shown in **Figure 1**. The feasible

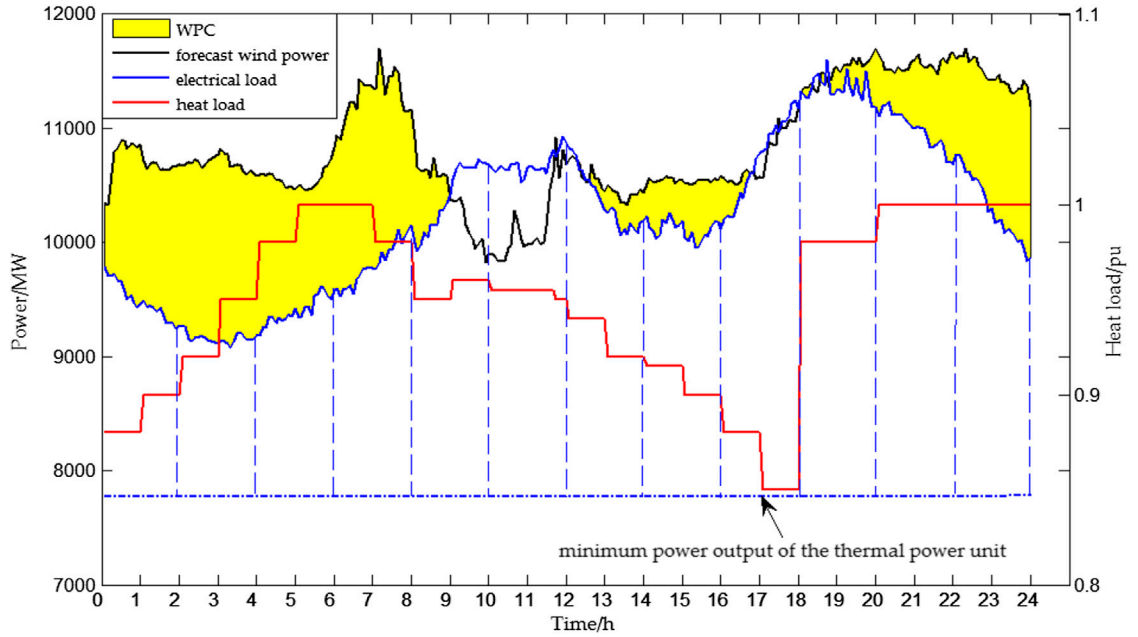


FIGURE 1 | Daily profiles of total wind power (MW), electrical load (MW), and heat load (p.u.) in a particular province during a typical high-wind winter heating period of 2015–2016.

region of wind power acceptance denoted by the area between blue dashed lines is determined by the rise of the electrical demand above the minimum export limit of the thermal power unit. The exceedance of available wind power outputs over the feasible region, that is, denoted by the yellow area, must be curtailed so as to avoid violating the minimum export limit of the thermal power unit. It can be seen that the large-scale WPC mostly coincides with electrical demand troughs at night, while there still exists a small amount of WPC in some particular time slots in the daytime. The high WPC is mainly caused by the mismatch in daily profiles between available wind power outputs and electrical demands.

It is noted that heat demands and available wind power outputs are shown to have similar diurnal trends where heat demand peaks mostly coincide with high wind power outputs. This permits the use of the electric heating with heat storage to meet heat demands while increasing wind power consumptions. The otherwise curtailed wind power can supply the electric heating demand and/or charge the heat storage during electrical off-peak periods; then, the stored heat is released to the users when needed, which alleviates the constraint of "ordering power by heat" and greatly improves the system flexibility.

REBs are widely used for regenerative electric heating due to their advantages of long service life (20–25 years), simple structure, high heating efficiency, intelligent control, and zero labor cost. There are two main types of REBs, that is, a solid REB and a water REB. The former is studied in this article due to its small footprint, large heat storage capacity per unit volume, and no additional ancillary equipment.

Electric Heating Users With Wind Power Heating

Heating as a kind of livelihood guarantee project requires a high reliability of heat supply. With expanding the scale of the wind power heating mode, the intermittency of wind generation requires the introduction of high-cost conventional generation to ensure the supply reliability. However, the current low prices of fossil fuels make the conventional heating method more attractive to heat users, which slow-paced the development of the electric heating. The contradiction between the electric heating demands and the cost of wind power heating has become the focus of market feasibility studies. To deal with this issue, the study develops an RTP decision model where the time-varying RTPs are adjusted to reflect the scale of predicted WPC during a particular dispatch day and guide electric heating users to actively increase their needs to consume the additional wind generation.

The demand change of a single user has little impact on an entire system. In order to facilitate the DSM, the study integrates the electric heating users within a particular area(s) into a single user set where each user is considered to share the same heating behavior. In addition, it is presumed here that the daily WPC will fully meet the heating demand with the operation of REBs.

RTP DECISION MODEL BASED ON WPC DISTRIBUTION

Before a particular dispatch day, wind farms will provide a system operator (SO) with their day-ahead wind power forecasts. Then,

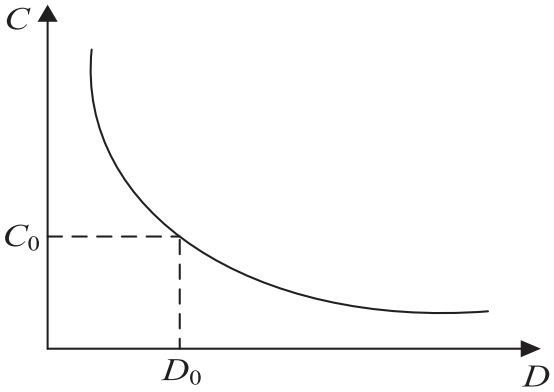


FIGURE 2 | Typical demand curve with quantity D_0 and price C_0 being transacted at the equilibrium point.

the SO combines available wind power forecasts with the electrical demand forecast curve to predict the hourly WPC in the dispatch day, based on which hourly RTPs are determined subject to a set of constraints and issued prior to the dispatch day. The relatively lower RTPs accompanying higher WPC forecasts can encourage a user set to choose the time of heating supplied by the otherwise curtailed wind power and conduct day-ahead transactions according to its heat demand.

Given day-ahead predictions of available wind power outputs $P_{wind.f}(t)$ (MW) and electrical demands $P_{load}(t)$ (MW), the WPC forecast $P_{qf}(t)$ (MW) at each 1-hour time step t within the dispatch day is determined by **Eq. 1**, avoiding the violation of the minimum export limit $P_{h.min}$ (MW) of the thermal power unit.

$$P_{qf}(t) = P_{wind.f}(t) + P_{h.\min}(t) - P_{load}(t). \tag{1}$$

From a range of factors such as electricity prices, weather, and national economy, electricity prices are the dominant factor affecting electricity consumptions of users. The electricity has the common attribute of general goods, that is, a contrary relationship between demand and price, as shown in **Figure 2**, where D_0 and C_0 represent the quantity and the price transacted at an equilibrium point, respectively (the supply curve that determines the equilibrium point is not plotted for brevity).

The demand curve in **Figure 2** can also describe the negative correlation between RTPs and WPC forecasts within a dispatch day, which is approximated here by **Eq. 2**.

$$\begin{aligned} C_i &= C_i^* \times C_B \\ C_i^* &= \frac{\frac{1}{D_i}}{\sum_{i=1}^{n}\frac{1}{D_i}} \end{aligned} \tag{2}$$

where C_i (\$/kWh) denotes the RTP at the *ith* ($i = 1, \ldots, n$) time step where wind farms are forecast to be curtailed by D_i (kWh); C_B (\$/kWh) is the benchmark value of RTPs equaling the sum of C_i; and C_i^* denotes the per unit RTP relative to C_B which inversely depends on D_i. Furthermore, C_i should be maintained within a reasonable range below $C_{i\max}$ (\$/kWh):

$$C_i \le C_{i\max} \tag{3}$$

In addition, it must be profitable for a wind farm to inject the otherwise curtailed wind power onto the grid. This means that C_i paid by the user set should be greater than the incremental operating cost c_w of the wind farm (i.e., 0.0087 \$/kWh in this study):

$$C_i > c_w \tag{4}$$

When there is no WPC at a particular time step, the RTP will not be issued and no transactions will occur at that time step.

DSR MODEL UNDER RTP SCHEME FOR WIND POWER HEATING

Given the REB installation providing a user set with a certain amount of controllable demand, the user set will actively shape its demand curve to match available wind power or WPC forecasts so as to exploit the price opportunity under the RTP scheme described in *RTP Decision Model Based on WPC Distribution* section. The user set can respond to RTP signals by increasing electricity consumptions at times of low RTPs which result from high WPC forecasts, and vice versa. The DSR using REBs under the RTP scheme not only helps to put additional wind generation onto the grid but also reduces heating costs met by users. From the perspective of the user set, this section will develop a DSR model using REBs to minimize the total heating cost by optimizing the heating behavior under the RTP scheme.

Objective Function

The objective function of the DSR model is the total electric heating cost of a user set in a particular dispatch day:

$$\min F = \min \sum_{i=1}^{n} C_i \Delta D_i \tag{5}$$

where ΔD_i (kWh) denotes the WPC reduction consumed by the REBs at the *ith* time step. The imports of the REBs supplied by otherwise curtailed wind power at each time step in the dispatch day are determined to minimize the objective function F. It is noted that the DSR model developed here deals with the optimization of the REB imports only, while the REB exports are not simulated in this work.

Constraint Condition

The daily electricity consumption needed for electric heating, denoted by E_h (kWh), limits the total volume of WPC that can be alleviated on a particular dispatch day:

$$E_h = \frac{WS_{hot}T_h}{\eta} \ge \sum_{i=1}^{n} \Delta D_i \tag{6}$$

where W is the heat load index (W/m^2), S_{hot} is the heating area (m^2), T_h is the daily heating duration (h), and η is the

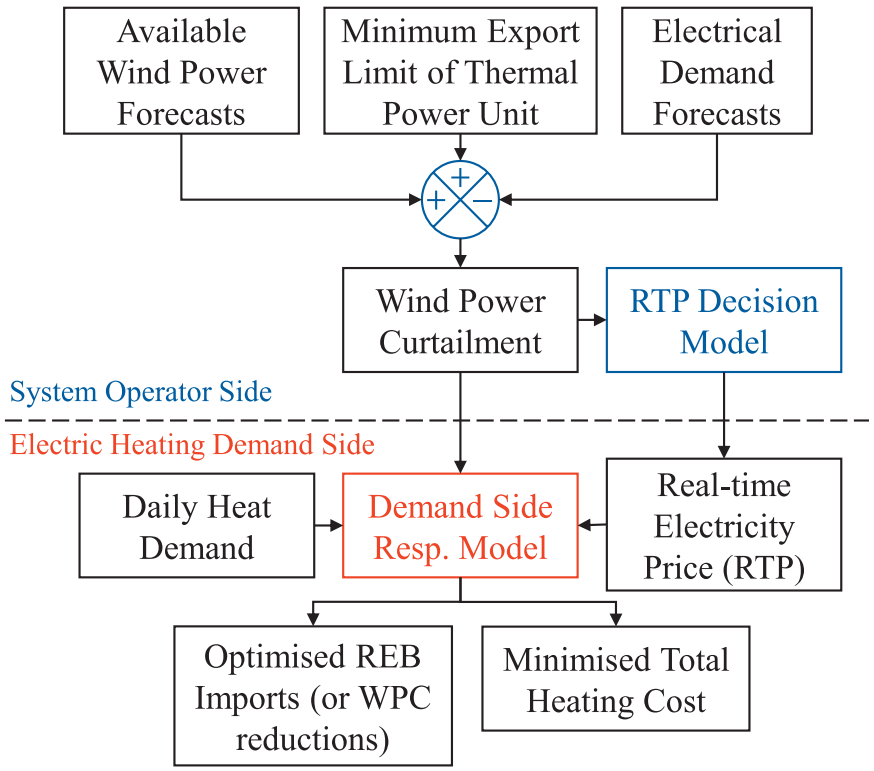


FIGURE 3 | Process of the DSR model under the RTP scheme.

electrothermal conversion efficiency. In addition, ΔD_i is smaller than the WPC volume D_i at the *ith* time step:

$$0 \le \Delta D_i \le D_i \tag{7}$$

.

From the perspective of the user set, the implementation of RTPs should not increase the average electricity price met by the user set during the day. This is formulated by **Eq. 8** as

$$0 \le \frac{\sum_{i=1}^{n} C_i \Delta D_i}{\sum_{i=1}^{n} \Delta D_i} \le C_{av} \tag{8}$$

,

where C_{av} represents the average electricity price (\$/kWh) before the price adjustment. In addition, the import rate $P_{en.in}(t)$ (MW) of an REB at time step t must not exceed its maximum import limit $P_{in.N}$ (MW):

$$0 \le P_{eb.in}(t) \le P_{in.N}. \tag{9}$$

A flowchart describing the optimization of the REB imports under the RTP scheme is shown in **Figure 3**.

As a comparison to the wind power heating cost, the conventional heating cost for the heating area S_{hot} (m^2) is also estimated here based on the conventional heating price C_r (\$/m^2) for an entire heating season and then evenly spread to each day within the season.

TABLE 1 | Current peak-to-valley electricity prices for electric heating.

	Valley section	Flat section	Peak section
Period	21: 00-24: 00 0: 00-7: 00	7: 00-7: 30 11: 30-17: 00	7: 30-11: 30 17: 00-21: 00
Price (\$/kWh)	0.04142	0.07865	0.11588

RESULTS AND MODEL VALIDATION

The developed DSR model using REBs under the RTP scheme for the wind power heating is validated here based on four particular heating days with various WPC distributions (2015.12.18, 2016.01.21, 2016.02.22, and 2016.04.03). The following four different heating scenarios are simulated for each day to evaluate the performance of the proposed approach in putting additional wind generation onto the grid and reducing heating costs met by the user set.

Scenario 1: Conventional heating with a constant conventional heating price;

Scenario 2: Wind power heating under peak-to-valley electricity price scheme with REB importing at off-peak time;

Scenario 3: Wind power heating under peak-to-valley electricity price scheme with REB tracking WPC;

Scenario 4: Wind power heating under the RTP scheme with REB tracking WPC.

TABLE 2 | Technical parameters of a single REB.

Boiler type	Footprint/m^2	Maximum import limit/MW	Total heat storage /MWh	Maximum heat storage temperature/°C
Solid REB	48	3.5	28	850

TABLE 3 | Residential district heating parameters.

Parameter	Value
η	98%
Total heating area, S_{hot}	15*200,000 m^2
Daily heating duration, T_h	24 h
Heat load index, W	58 W/m^2 Beijing Gas and Heating Engineering Design Institute (2010)
Conventional heating price, C_r	3.92 $/m^2
$C_{i\ max}$	0.58 $/kWh

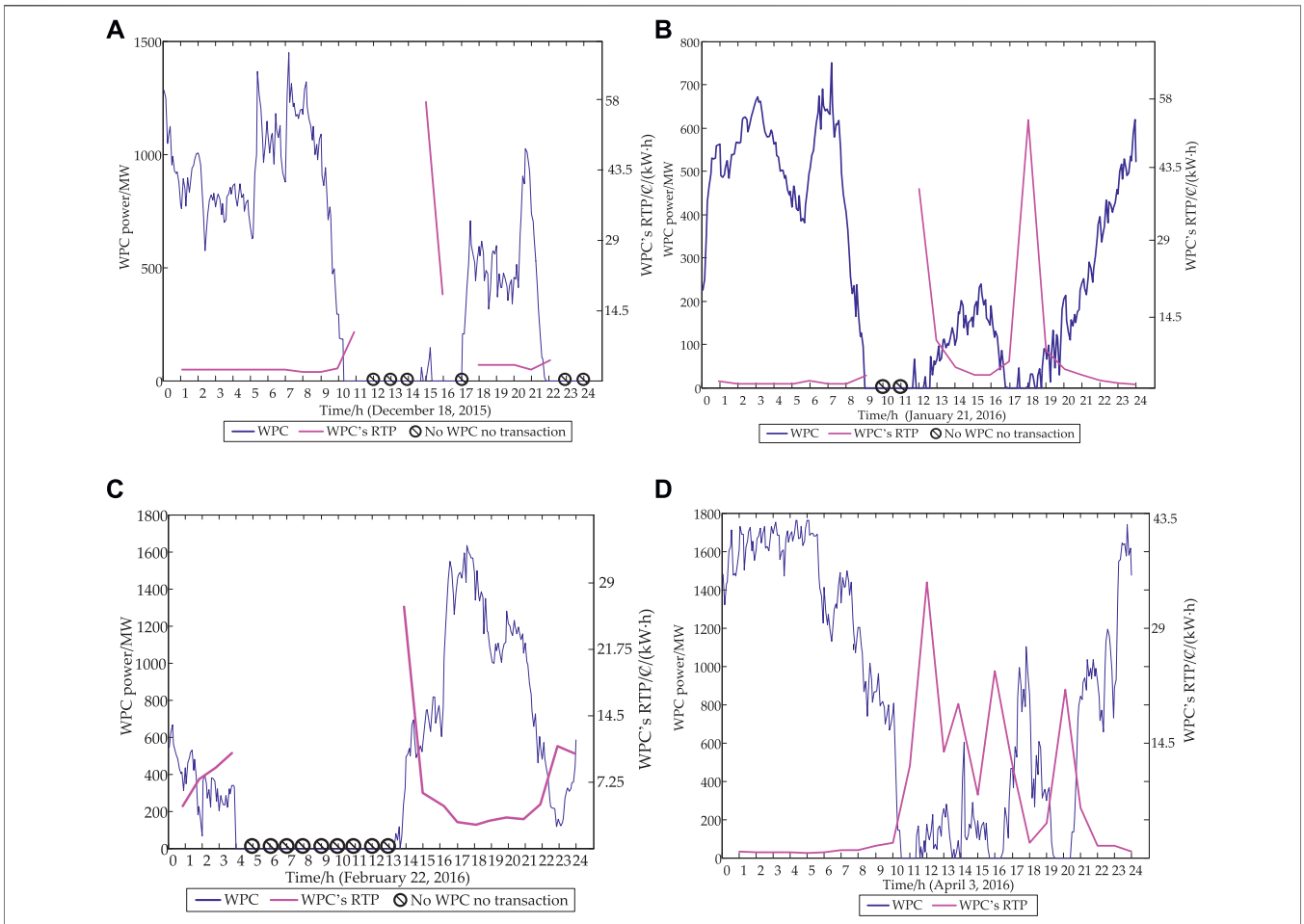


FIGURE 4 | Daily curves of WPC (MW) and corresponding RTPs (¢/kWh) on **(A)** 2015.12.28, **(B)** 2016.01.21, **(C)** 2016.02.22, and **(D)** 2016.04.03.

The 5-minute average available wind power output forecasts and actual demands at a provincial level in the "Three-North regions" over each day are converted into 1-hour average data in line with the adopted time step length. The province studied here currently employs peak-to-valley electricity prices for electric heating, as shown in **Table 1** (Jilin Province Price Bureau, 2018). The total area of the 15 residential districts with the heat demand across the province is about 3 million m^2. Given technical parameters of a solid REB and the residential district heating as listed in **Tables 2**, **3**, respectively, it is evaluated by **Eq. 6** that each 0.2 million m^2 residential district needs to

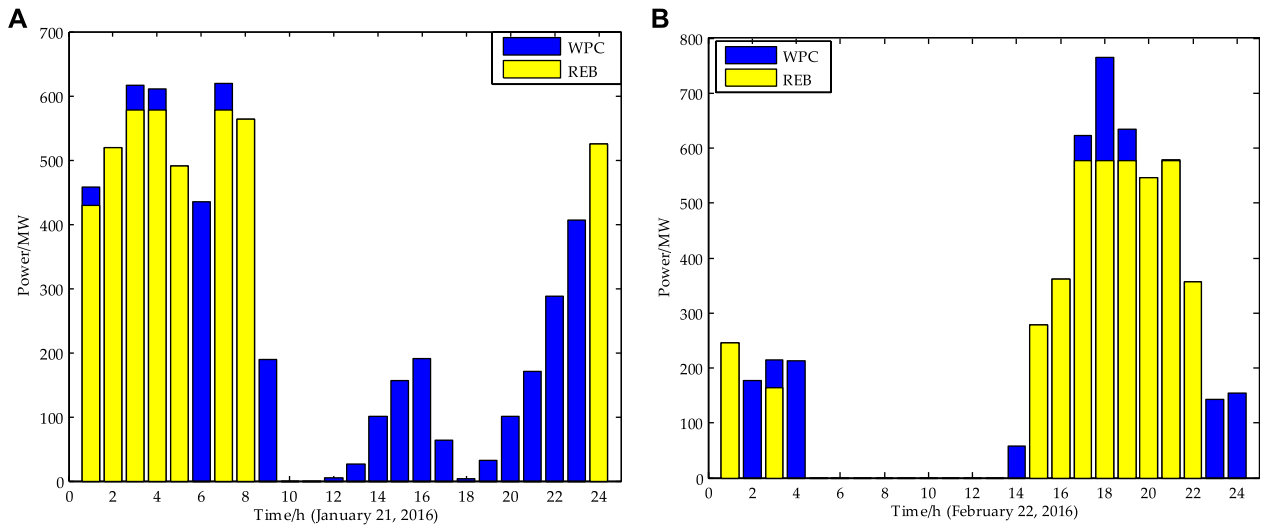


FIGURE 5 | Aggregation of optimized REB imports (MW) and remaining WPC (MW) under Scenario 4 on **(A)** 2016.01.21 and **(B)** 2016.02.22.

be equipped with 11 REBs. Therefore, it is presumed here that a total of 165 REBs are installed within the whole province.

Simulation Results of the RTP Decision Model

Figure 4 shows the daily curves of WPC forecasts on the four heating days and their corresponding RTPs determined by the RTP decision model (for the exact RTP values, the reader is referred to **Table A1** in Appendix). The WPC is shown to not only happen in off-peak periods but may also occur at times of normal or even peak electrical demands. This means that REBs simply importing during electrical demand troughs cannot ensure the WPC to be minimized. Therefore, it is necessary to implement the RTP scheme that guides the user set to charge REBs in response to the WPC, that is, tracking the WPC.

Evaluation on WPC Reductions and Heating Costs

In Scenario 4, the time and rates at which the REBs import are optimized to minimize the total heating cost. **Figure 5** shows the aggregation of the optimized imports of REBs along with the remaining WPC in Scenario 4 on 2016.01.21 and 2016.02.22 where the REB imports greatly increase wind power consumptions which would otherwise be curtailed during off-peak and peak periods, respectively. To illustrate the advantages of the DSR model under the RTP scheme adopted in Scenario 4, total WPC reductions and total heating costs over each day estimated under the four heating scenarios are compared in **Figure 6** and **Table 4**.

The wind power heating mode employed in Scenarios 2–4 is mostly shown to achieve a maximum of 4261.2 MWh WPC reduction with the help of REBs. However, simply charging REBs at off-peaks in Scenario 2 may decrease the REB performance in the WPC reduction depending on the temporal distribution of WPC. For example, only 2080.2 MWh otherwise curtailed wind generation is absorbed by REBs on 2016.02.22 where the WPC mostly occurs during peak demands (see **Figure 4C**).

Compared to the conventional heating method in Scenario 1, the peak-to-valley electricity price pattern adopted in Scenarios 2 and 3 leads to significant electric heating costs. With the implementation of the RTP scheme in Scenario 4, total electric heating costs are largely reduced to the levels close to or even smaller than the total conventional heating cost. The significantly decreased electric heating costs will make the electric heating with REBs more attractive, expanding the scale of the wind power heating mode. In addition, the total heating cost on 2016.02.22 is higher than the total cost on 2016.01.21 under Scenario 4. This is because the WPC occurs at such fewer time steps that higher RTPs are issued on 2016.02.22. Among all heating scenarios, Scenario 4 using REBs to track the WPC under the RTP scheme achieves the greatest WPC reduction and an acceptable total heating cost.

The additional wind generation put onto the grid can reduce fossil fuel consumptions of conventional thermal power plants and thus increase the environmental benefit. The revenue associated with the environmental benefit received by wind farms can contribute to the use-of-system charge paid to the SO who is responsible for the reliable system operation. The win-win situation of wind farms, network owners, and electric heating users illustrates the feasibility of implementing the DSR with REBs under the RTP scheme as proposed in this work.

CONCLUSION AND FUTURE WORK

The electric heating demand management permitted by the heat storage combined with a suitable electricity price incentive is expected to assist in the development of the wind power heating mode during winter heating periods. This study has developed a real-time price (RTP) decision model to introduce an RTP scheme

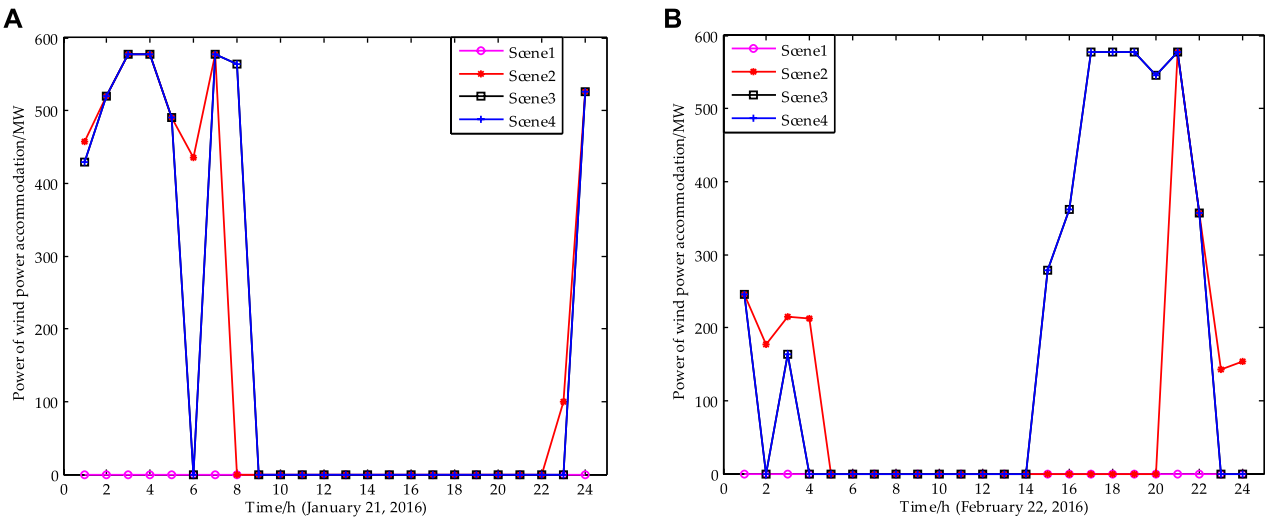


FIGURE 6 | WPC reductions (MW) under the four heating scenarios on **(A)** 2016.01.21 and **(B)** 2016.02.22.

TABLE 4 | Total WPC reductions (MWh) and heating costs (10 k$) under four heating scenarios on two particular days.

	WPC reduction/MWh		Heating cost/10 k$	
	2016.01.21	2016.02.22	2016.01.21	2016.02.22
Scenario 1	0	0	6.992	6.992
Scenario 2	4261.2	2080.2	17.65	17.65
Scenario 3	4261.2	4261.2	19.796	39.148
Scenario 4	4261.2	4261.2	4.542	7.906

taking into account a negative correlation between the RTP and the wind power curtailment (WPC) within a particular dispatch day. The time-varying RTPs not only met the acceptable range of electric heating users but also covered the cost of a wind farm to put otherwise curtailed wind power onto the grid. Given electric heating users being equipped with regenerative electric boilers (REBs), a demand side response (DSR) model using REBs under the RTP scheme was developed to minimize the total electric heating cost by charging REBs at lower RTPs so as to alleviate the higher WPC. The developed RTP decision and DSR models have been tested on four particular heating days at a provincial level in the "Three-North regions" in China. Compared to the currently adopted peak-to-valley price pattern, the RTPs reflecting the WPC distribution in a dispatch day permitted the electric heating users with REBs to exploit the price opportunity while helping to increase wind power consumptions. Furthermore, the implementation of the RTP scheme largely reduced the total electric heating cost met by users to the level close to or even smaller than the conventional heating cost, which may increase the attractiveness of the electric heating with REBs in cost, and in turn, expand the scale of the wind power heating mode. In addition, the DSR model optimizing the REB imports to track the lower RTPs has maximized the WPC reduction in this work, while simply charging the REBs during electrical demand troughs may degrade the WPC alleviation depending on the temporal distribution of WPC.

Building on the present work, the integrated energy system modeled here should be extended to incorporate other renewable energy resources such as photovoltaic energy and reflect transmission constraints on both heat and electricity. This will permit to estimate the most suitable REB locations within the system that maximize the use of renewable energy. Furthermore, capital and operating expenditures of REBs will be compared with the heating cost saving of users relative to the conventional heating. The cost–benefit analysis can indicate the investment return and payback period, assessing the economic feasibility of deploying REBs from the perspective of heating users. In addition, the proposed RTP-based demand side management for the wind or renewable power heating will be further tested on other cities or provinces, for example, in the "Three-North regions," which are suffering from the WPC problem in heating seasons.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

QL performed the experiment and simulation; JL contributed to the conception of the work; ZH performed analysis and prepared the manuscript; FF performed analysis and revised the manuscript; and WT assisted in analysis with constructive discussions.

FUNDING

This research was supported by National Natural Science Foundation of China (NSFC) (51707016).

REFERENCES

Ai, X., Zhou, S., and Zhao, Y. (2016). Study on time of use pricing of user side considering wind power uncertainty. *Power Syst. Technol.* 40 (5), 1529–1535. doi:10.13335/j.1000-3673.pst.2016.05.035 [in Chinese, with English summary]

Chen, L., Xu, F., Wang, X., Min, Y., Ding, M., and Huang, P. (2015). Implementation and effect of thermal storage in improving wind power accommodation. *Proc. CSEE* 35 (17), 4283–4290. doi:10.13334/j.0258-8013.pcsee.2015.17.001 [in Chinese, with English summary]

Comprehensive Department of National Energy Administration (2015). Notice on the work of clean heating using wind power (No. 306, 2015). Available at: http://zfxxgk.nea.gov.cn/auto87/201506/t20150615_1938.htm (Accessed June 5, 2015) [in Chinese, with English summary]. [Online]

Cui, Y., Chen, Z., Yan, G., and Tang, Y. (2016). Coordinated wind power accommodating dispatch model based on electric boiler and CHP with thermal energy storage. *Proc. CSEE* 36 (15), 4072–4081. [in Chinese, with English summary]. doi:10.13334/j.0258-8013.pcsee.152406

De Jonghe, C., Hobbs, B. F., and Belmans, R. (2014). Value of price responsive load for wind integration in unit commitment. *IEEE Trans. Power Syst.* 29 (2), 675–685. doi:10.1109/tpwrs.2013.2283516

Energy Administration of Jilin Province (2017). Notice on issuing the "13th Five-Year plan" for energy development in Jilin. Available at: http://nyj.jl.gov.cn/ghjh/20170522/852.html(Accessed May 22, 2017) [in Chinese, with English summary]. [Online]

Fan, F., Xu, H., and Kockar, I. (2019). "Utilisation of energy storage to improve distributed generation connections and network operation on Shetland Islands," in Proceedings of 25th international conference on electricity distribution, Madrid, Spain, June 3, 2019 (CIRED 2019 Conference), 1764, 1–5.

Beijing Gas; Heating Engineering Design Institute (2010). *City thermal net design criterion CJJ34-2010.* Beijing, China: China Architecture & Building Press, 21–30.

Gifford, J., Ma, Z., and Davenport, P. (2020). Thermal analysis of insulation design for a thermal energy storage silo containment for long-duration electricity storage. *Front. Energy Res.* 8, 99. doi:10.3389/fenrg.2020.00099

Jilin Province Price Bureau (2018). Notice on further clarifying the relevant issues concerning the clean heating price policy of the province. Available at: http://wjj.jl.gov.cn/scjgjg/zyynyjg_67230/dj/201802/t20180223_3716147html(Accessed February 23, 2018) [in Chinese, with English summary]. [Online]

Ketterer, J. C. (2014). The impact of wind power generation on the electricity price in Germany. *Energy Econ.* 44, 270–280. doi:10.1016/j.eneco.2014.04.003

Khodayar, M. E., and Shahidehpour, M. (2013). Stochastic price-based coordination of intrahour wind energy and storage in a generation company. *IEEE Trans. Sustain. Energy* 4 (3), 554–562. doi:10.1109/tste.2012.2228284

Li, Y., Wang, C., Li, G., Wang, J., Zhao, D., and Chen, C. (2020). Improving operational flexibility of integrated energy system with uncertain renewable generations considering thermal inertia of buildings. *Energy Convers. Manag.* 207, 112526. doi:10.1016/j.enconman.2020.112526

Li, Y., Wang, J., Zhao, D., Li, G., and Chen, C. (2018a). A two-stage approach for combined heat and power economic emission dispatch: combining multi-objective optimization with integrated decision making. *Energy* 162, 237–254. doi:10.1016/j.energy.2018.07.200

Li, Y., Yang, Z., Li, G., Mu, Y., Zhao, D., Chen, C., et al. (2018b). Optimal scheduling of isolated microgrid with an electric vehicle battery swapping station in multi-stakeholder scenarios: a bi-level programming approach via real-time pricing. *Appl. Energy* 232, 54–68. doi:10.1016/j.apenergy.2018.09.211

Li, Z., Wu, W., Shahidehpour, M., Wang, J., and Zhang, B. (2015). Combined heat and power dispatch considering pipeline energy storage of district heating network. *IEEE Trans. Sustainable Energy* 7 (1), 12–22. doi:10.1109/TSTE.2015.2467383

Liu, J. (2013). Basic issues of the utilization of large-scale renewable power with high security and efficiency. *Proc. Chin. Soc. Electr. Eng.* 33 (16), 2–8. doi:10.3969/j.issn.2095-0802.2014.11.031 [in Chinese, with English summary].

Lü, Q., Wang, W., Han, S., Yuan, S., Zhang, J., and Li, W. (2013). A new evaluation method for wind power curtailment based on analysis of system regulation capability. *Power Syst. Technol.* 37 (7), 1887–1894. [in Chinese, with English summary]. doi:10.1016/S0924-8579(07)70847-6

National Energy Administration (2020). Wind power grid operation in the first half of 2018. Available at: http://www.nea.gov.cn/201802/01/c_136942234.htm (Accessed February 1, 2020) [in Chinese, with English summary]. [Online]

Sánchez de la Nieta, A. A. S., Contreras, J., Muñoz, J. I., and O'Malley, M. (2014). Modeling the impact of a wind power producer as a price-maker. *IEEE Trans. Power Syst.* 29 (6), 2723–2732. doi:10.1109/tpwrs.2014.2313960

Schmidt, M., and Linder, M. (2020). A novel thermochemical long term storage concept: balance of renewable electricity and heat demand in buildings. *Front. Energy Res.* 8, 137. doi:10.3389/fenrg.2020.00137

Shafie-khah, M., Heydarian-Forushani, E., Golshan, M. E. H., Moghaddam, M. P., Sheikh-El-Eslami, M. K., and Catalão, J. P. S. (2015). Strategic offering for a price-maker wind power producer in oligopoly markets considering demand response exchange. *IEEE Trans. Ind. Inf.* 11 (6), 1542–1553. doi:10.1109/tii.2015.2472339

Shen, J., Jiang, C., Liu, Y., and Wang, X. (2016). A microgrid energy management system and risk management under an electricity market environment. *IEEE Access* 4, 2349–2356. doi:10.1109/access.2016.2555926

Shin, H., Lee, D., and Baldick, R. (2017). An offer strategy for wind power producers that considers the correlation between wind power and real-time electricity prices. *IEEE Trans. Sustainable Energy* 9 (2), 695–706. doi:10.1109/TSTE.2017.2757501

Sioshansi, R., and Short, W. (2009). Evaluating the impacts of real-time pricing on the usage of wind generation. *IEEE Trans. Power Syst.* 24 (2), 516–524. doi:10.1109/tpwrs.2008.2012184

Wang, C., Li, Q., and Xie, G. (2013). Evaluation of wind power heating in facilitating wind power integration capability during valley load period. *Electr. power* 46 (12), 100–106. [in Chinese, with English summary]. doi:10.4028/www.scientific.net/amr.953-954.501

Xie, D., Lu, Y., Sun, J., Gu, C., and Li, G. (2016). Optimal operation of a combined heat and power system considering real-time energy prices. *IEEE Access* 4, 3005–3015. doi:10.1109/access.2016.2580918

Xu, M., and Jiang, D. (2015). Research on energy performance and economic analysis of wind power heating system. *Energy China* 37 (8), 42–47. doi:10.3969/j.issn.1003-2355.2015.08.008 [in Chinese, with English summary]

Yao, S., Gu, W., Zhou, S., Lu, S., Wu, C., and Pan, G. (2018). Hybrid timescale dispatch hierarchy for combined heat and power system considering the thermal inertia of heat sector. *IEEE Access* 6, 63033–63044. doi:10.1109/access.2018.2876718

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

APPENDIX

TABLE A1 | Exact RTP values (¢/kWh) on (a) 2015.12.18, (b) 2016.01.21, (c) 2016.02.22, and (d) 2016.04.03.

Day(a)	Hour	1	2	3	4	5	6	7	8	9	10	11	12
	RTP	1.475	1.524	1.579	1.578	1.578	1.470	1.470	1.410	1.439	1.685	9.257	—
	Hour	13	14	15	16	17	18	19	20	21	22	23	24
	RTP	—	—	58.000	16.364	—	1.907	1.840	2.005	1.598	2.616	—	—
Day(b)	Hour	1	2	3	4	5	6	7	8	9	10	11	12
	RTP	1.146	1.089	1.022	1.027	1.114	1.170	1.021	1.056	1.821	—	—	40.206
	Hour	13	14	15	16	17	18	19	20	21	22	23	24
	RTP	9.188	2.835	2.063	1.818	4.130	54.782	7.595	2.842	1.955	1.427	1.206	1.085
Day(c)	Hour	1	2	3	4	5	6	7	8	9	10	11	12
	RTP	3.448	4.7746	3.940	3.979	—	—	—	—	—	—	—	—
	Hour	13	14	15	16	17	18	19	20	21	22	23	24
	RTP	—	14.551	3.039	2.339	1.357	1.105	1.333	1.549	1.460	2.368	5.937	5.497
Day(d)	Hour	1	2	3	4	5	6	7	8	9	10	11	12
	RTP	1.220	1.122	1.104	1.119	1.088	1.182	1.436	1.395	1.969	2.358	12.079	35.121
	Hour	13	14	15	16	17	18	19	20	21	22	23	24
	RTP	13.779	21.103	8.387	23.961	11.911	2.320	4.305	22.441	6.776	1.950	1.984	1.209

frontiers in Energy Research

ORIGINAL RESEARCH
published: 12 March 2021
doi: 10.3389/fenrg.2020.612909

Storage-Transmission Joint Planning Method to Deal with Insufficient Flexibility and Transmission Congestion

Xiuyu Yang[1,2], Guofeng Chai[2], Xueyuan Liu[2], Minghong Xu[2] and Qi Guo[2]*

[1]School of Electrical and Electronic Engineering, North China Electric Power University, Beijing, China, [2]Key Laboratory of Modern Power System Simulation and Control & Renewable Energy Technology (Northeast Electric Power University), Jilin, China

The insufficient power system flexibility and transmission congestion are two fundamental reasons for wind power curtailment. As the scale of the wind power in the power system is growing rapidly, the two factors of wind power curtailment events coexist and have a certain coupling relationship. The configuration of the energy storage system can not only increase the flexibility of the system but also alleviate transmission congestion. Therefore, the joint planning of energy storage and transmission grid that takes into account the flexibility of the system and the transmission congestion is of great significance to solve the wind curtailment. Hence, this paper first decouples the insufficient flexibility and transmission congestion wind power curtailment, and quantitatively analyzes the impact of transmission capacity on the coupling relationship between the two; second, reveals the principle of joint planning of energy storage system and transmission congestion, and constructs an optimization model, and proposes to set up the capacity of the wind power which connects the power network. The solution procedure, which deals with grid planning and the energy storage system optimization in turn, not only ensures the accuracy of the model, but also significantly reduces the calculation cost. Finally, a case study of a wind power base in Northeast China and an improved Garver-6 system are carried out to verify the effectiveness of the proposed method.

Keywords: transmission congestion, energy storage system, wind power curtailment, transmission grid planning, power system flexibility

OPEN ACCESS

Edited by:
Chao Long,
Cranfield University, United Kingdom

Reviewed by:
Kenneth Okedu,
Caledonian College of Engineering, Oman
Chen Liang,
Nanjing University of Information Science and Technology, China

****Correspondence:***
Xiuyu Yang
yangxiuyu2011@163.com

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

Received: *01 October 2020*
Accepted: *21 December 2020*
Published: *12 March 2021*

Citation:
Yang X, Chai G, Liu X, Xu M and Guo Q (2021) Storage-Transmission Joint Planning Method to Deal with Insufficient Flexibility and Transmission Congestion. Front. Energy Res. 8:612909. doi: 10.3389/fenrg.2020.612909

INTRODUCTION

Vigorously developing wind power is an important measure to promote the low-carbon transformation of energy and respond to the energy crisis and climate change. The development of wind power has increased rapidly in the past 10 years, from the installed capacity of 159.21 million kW in 2009 to 651 million kW in 2019, an increase of more than five times (GWEC, 2019). And countries all over the world have formulated a vision for the development of renewable energy in 2050 (Han et al., 2014; Price water house Coopers LLP et al., 2014). However, in the rapid development of wind power, the issue of wind power curtailment has not been effectively solved. For example, in China, the installed wind power capacity which was 210.05 million kW in 2019 is 8.1 times that of 25.805 million kW in 2009. During this period, the wind curtailment rate

remained high, and the curtailment rate in 2015 and 2016 was as high as 15% and 17% respectively (NEA, 2016). Although the curtailment rate in 2019 reached 4%, there are still some areas with the curtailment rate as high as 14% (National Energy Administration, 2020). Therefore, analyzing the causes of wind curtailment and taking measures scientifically to alleviate or eliminate wind power curtailment is of vital importance to the healthy development of wind power in the future.

The main challenges for solving the wind power curtailment in the power system mainly includes two aspects: one is that the thermal power system lacks sufficiently available downward capacity due to its limitation of minimum technological output, which cannot guarantee effective wind power accommodation (Li et al., 2020; Wang et al., 2020). Another one is the wind power transmission congestion which is due to insufficient available transmission capacity. This normally results from that the construction of the power grid lags behind the development of wind power or the transmission capacity of the planned construction is inadequate (Jorgenson et al., 2017). In addition, in the system where the wind power base is far away from the load center and the proportion of thermal power units is large, the lack of flexibility and transmission congestion exist at the same time. If only considering the power supply side, the flexibility transformation of thermal power units (Zhou and Wang, 2017; Wang et al., 2019) and configuration of energy storage can increase the flexibility of the system. Although the problem of wind abandonment has been alleviated to a certain extent, there will still be wind curtailment caused by transmission congestion. Similarly, if only the transmission network expansion planning has carried out (Correa et al., 2016; Gan et al., 2016; Zhang et al., 2019), although the transmission congestion problem is solved, the wind curtailment caused by the lack of system flexibility still exists. Moreover, the downward regulation capacity of the system can be increased through the configuration of energy storage or flexible transformation of thermal power units at the power supply side. Simultaneously, the problem of transmission congestion can be solved by constructing more transmission lines from the power grid side. Although the above solutions can effectively overcome the wind power curtailment, they will result in the investment overlap, following the huge cost investment and low asset utilization rate.

The energy storage system is a high-quality and flexible resource with time migration ability of power and energy and flexible installation location. Large-scale energy storage system technology is considered as one of the key technical supports for the popularization of renewable energy (Li et al., 2019). The configuration of the energy storage system can not only increase the flexibility of the power supply side but also reduce the peak wind power to alleviate transmission congestion. Therefore, in the context of large-scale wind power grid integration, research on joint planning of energy storage and transmission grids to deal with insufficient flexibility and transmission congestion is of great significance to the overall economy of the system and reducing wind curtailment.

In recent years, many scholars have carried out a series of fruitful work to alleviate the wind curtailment problem, such as energy storage optimization configuration (Dicorato et al., 2012; Del Rosso and Eckroad, 2014; Zheng et al., 2015b; Lu et al., 2018; Masoumzadeh et al., 2018; Al Ahmad and Sirjani, 2020; Padhee et al., 2020), and joint planning of energy storage and transmission grids (Zheng et al., 2015a; Qiu et al., 2017; Bustos et al., 2018; Dvorkin et al., 2018; Jorgenson et al., 2018; Yacar et al., 2018; Zhang and Conejo, 2018; Nikoobakht and Aghaei, 2019; Sima et al., 2019; Wu and Jiang, 2019). In terms of energy storage optimization configuration, Padhee et al., (2020) proposed a hybrid energy storage optimization configuration method for the intermittency and uncertainty of renewable energy, while Al Ahmad and Sirjani (2020) combined the discretization method, multi-objective hybrid particle swarm algorithm and non-dominated sorting genetic algorithm. Zheng et al., (2015a) proposed a two-level planning model that considers the layout and operation of the energy storage system, and it has reached the goal of increasing the penetration rate of wind power. The poor predictability of wind power results in a large change in the actual power generation and the forecasted power in the power market. Dicorato et al., (2012) to meet the hourly electricity curve required by the day-ahead market, a joint planning and operation method for wind power and energy storage system is proposed. The fluctuation of renewable energy output brings about the problem of fluctuations in electricity prices. Masoumzadeh et al., (2018) using energy storage makes the electricity market price fluctuate at a certain level. Energy storage configuration can not only increase the flexibility of the system (Lu et al., 2018), but also delay the construction of transmission lines (Del Rosso and Eckroad, 2014), so the joint planning of energy storage and transmission grids has attracted extensive research interest. Yacar et al., (2018) explores the impact of transmission congestion, loss and different intermittent degree of wind power on energy storage layout and investment (Jorgenson et al., 2018). Comparatively analyzed the respective effects of the deployment of the energy storage system and the expansion plan of the transmission grid in reducing wind power curtailment and solar curtailment and gave the effect of mitigating wind power curtailment that the coordinated planning of the two is superior to their respective plans. Bustos et al., (2018) use of hourly energy requirements for coordinated planning of energy storage and transmission grid verified that the energy storage system is both a supplement to and a substitute for transmission lines. Alleviating transmission congestion can benefit the Energy storage situation. If the expansion of transmission lines will reduce or eliminate such benefits (Dvorkin et al., 2018), fully considered the benefits and return rates of the entire life cycle of the energy storage system, and constructed joint planning of energy storage system layout and transmission grid frame. The capacity of energy storage will decrease over time (Qiu et al., 2017), constructed a joint optimization plan for transmission grid and energy storage system that considers changes in energy storage capacity. Because of the uncertainty of renewable energy (Zhang and Conejo, 2018), from the perspective of system planners proposed an optimization model for the coordinated expansion planning

of transmission and energy storage system that considers long-term and short-term uncertainties. Nikoobakht and Aghaei (2019) proposed a continuous-time model based on stochastic robust optimization technology to coordinate the planning model of transmission expansion and energy storage system. Zheng et al., (2015a), Sima et al., (2019), Wu and Jiang (2019) incorporated system operating and wind power curtailment costs into the joint planning of the energy storage system and transmission grid. Although it can improve the planning economy and reduce wind power curtailment, there is a problem of the excessive computational burden based on timing simulation.

The above references have carried out fruitful research on the optimal configuration of energy storage and the joint planning of energy storage and transmission network from different perspectives. However, there are few studies on the coupling relationship between wind curtailment caused by different reasons. During the joint planning of energy storage and transmission network, only the wind curtailment cost and energy storage operation cost are added into the planning model, and planning boundaries of wind power grid connection points are still calculated based on installed capacity. The joint planning does not make full use of analyzing the coupling relationship between energy storage efficiency and wind abandonment for different reasons and the principle of joint planning of transmission grid to reduce wind abandonment, and there is a problem that the calculation amount of the existing model solving method based on time series simulation is too large. To this end, this paper considers both system flexibility and transmission congestion. It constructs a joint planning method for the energy storage system and transmission grid to deal with insufficient system flexibility and transmission congestion. Firstly, decoupled the insufficient flexibility and the transmission congestion wind power curtailment, and analyzed the time-series coupling relationship between the two under different planning boundaries of the wind farm grid connection point; secondly, build a joint planning model of the energy storage system and transmission grid to deal with insufficient flexibility and transmission congestion on this basis; Finally, designed an iterative solution algorithm to optimize the configuration of energy storage after grid planning, that is, first set the grid-connected point capacity of the dynamic wind power to obtain the grid planning scheme, and obtain the total abandonment of the system under the grid planning topology, and then optimize between the energy storage system and the total system wind power curtailment, in the end, iterate the process repeatedly to arrive at an optimal plan for energy storage and transmission grid.

The main contributions of this paper are as follows:

(1) By assuming sufficient power transmission capacity, the wind power curtailment caused by insufficient flexibility can be obtained, while assuming sufficient system flexibility, the wind power curtailment by transmission congestion can be obtained, and the decoupling is realized;
(2) This article plans the boundary by changing the wind power grid connection point and the obstructed wind power curtailment is obtained at grid points under different planned boundary capacities and established the relationship between total curtailed wind power curtailment and grid-connected point capacity. Optimizing the configuration of energy storage system based on the total wind power curtailment can vastly reduce the difficulty of solving the joint planning model of the energy storage system and transmission grid, and also solve the double dilemma of the excessive shortage of capacity settings at wind power grid connection point;
(3) The joint planning method of energy storage and transmission network is constructed to deal with the lack of flexibility and transmission congestion, which can fundamentally alleviate and solve the problem of wind power curtailment, and improve the utilization rate of energy storage and transmission network.

ANALYSIS OF THE COUPLING RELATIONSHIP BETWEEN INSUFFICIENT FLEXIBILITY AND TRANSMISSION CONGESTION

Large-scale grid integration of wind power puts higher requirements on the system's flexibility capacity and power transmission capacity. The occurrence of wind power curtailment is the main result of the combined effect of insufficient system flexibility and power transmission capacity. However, there is a certain coupling relationship between the abandonment wind of transmission congestion and the lack of flexibility in the unified system.

Quantitative Analysis of Wind Power Curtailment Caused by Insufficient System Flexibility

Assuming that the transmission capacity of the system is fully sufficient, analyze the relationship between the system flexibility and the wind power curtailment. Wind power consumption of the system is determined by flexibility (peak regulation depth) provided by conventional thermal power units. The "wind power feasible region" is defined as the daily dispatch time scale, and the difference between the load and the minimum technical output of the conventional units is the maximum wind power allowed by the system, and the formula is expressed as **(1)**

$$P_{\mathrm{FR}-i}(t) = P_{\mathrm{D}-i}(t) - P_{\mathrm{G.\,min}-i} \tag{1}$$

In **formula (1)**, $P_{\mathrm{FR\text{-}i}}(t)$ is the maximum wind power that the system can accept on dispatching day i; $P_{\mathrm{D}\text{-}i}(t)$ is the load value at time t on dispatching day i; $P_{\mathrm{G.\ min}\text{-}i}$ is the minimum technology for conventional units on dispatching day i contribute.

However, Windpower integration reduces the start-up capacity of conventional units. To ensure the reliability of power supply and consider the uncertainty of wind power

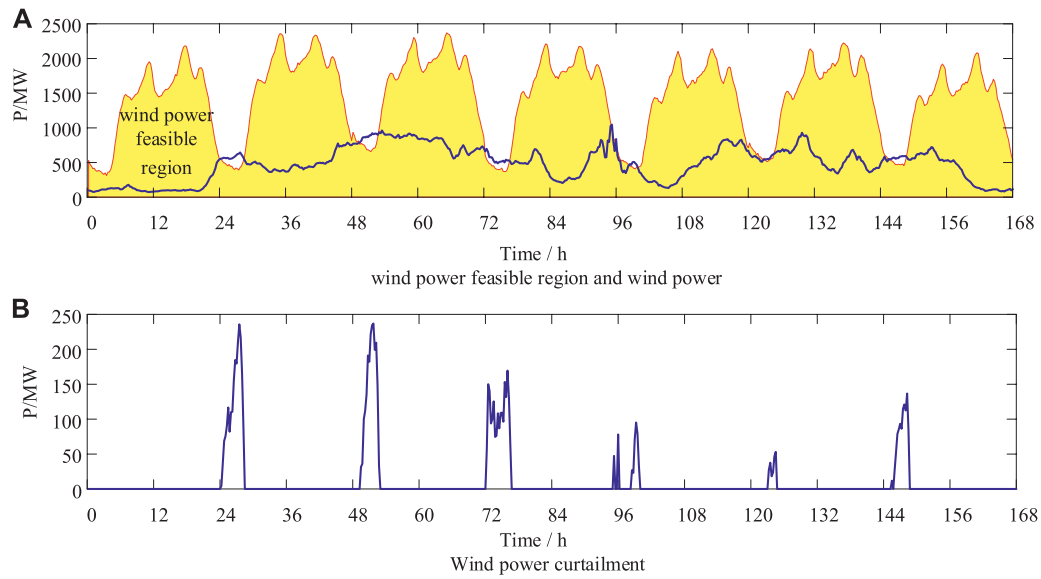


FIGURE 1 | Wind power consumption space and wind power curtailment.

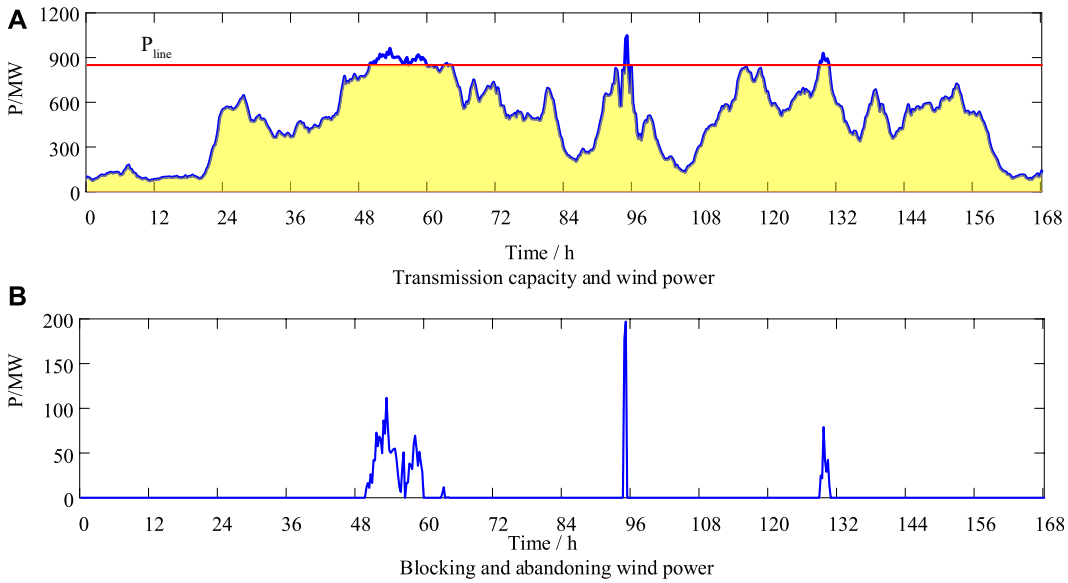


FIGURE 2 | Transmission capacity and blocked wind power curtailment.

output, the start-up capacity of conventional units after wind power is connected is **(2)**

$$P_{G-i} = (P_{D.max-i} - P_{W.min-i})(1 + k) \quad (2)$$

In **formula (2)**, $P_{G\text{-}i}$, $P_{D.\ max\text{-}i}$, $P_{W.\ min\text{-}i}$, and k are the starting capacity, maximum load, minimum wind power, and standby coefficient of conventional units on dispatch day i, respectively.

After knowing the available wind power on the dispatch day i, the wind power beyond the feasible range of wind power will not be absorbed. The expression of wind power curtailment caused by insufficient flexibility system is **(3)**

$$P_{LG-i}(t) = \begin{cases} 0, & P_{W-i}(t) < P_{FR-i}(t) \\ P_{W-i}(t) - P_{FR-i}(t), & P_{W-i}(t) \geq P_{FR-i}(t) \end{cases} \quad (3)$$

In **formula (3)**, $P_{LG\text{-}i}(t)$ is the wind power curtailment caused by insufficient flexibility at time t within the scheduling day i. **Figure 1** shows the feasible region and wind power curtailment in seven dispatch days.

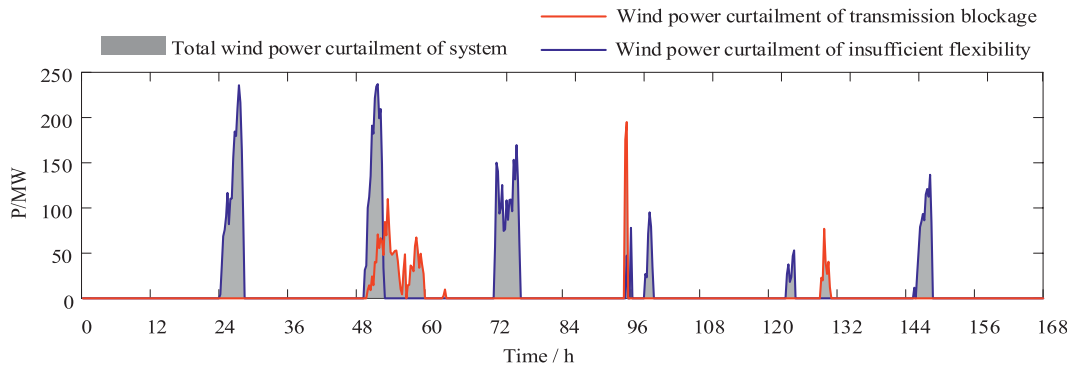


FIGURE 3 | The coupling of insufficient flexibility and blocked wind power curtailment.

Quantitative Analysis of Blockage and Abandonment Caused by Insufficient Transmission Capacity

Assuming that the system is sufficiently flexible, the relationship between transmission capacity and wind power curtailment is analyzed, that is the relationship between the planning of boundary wind power grid connection point and the abandoned wind congestion power. When the wind power $P_{W\text{-}i}$ is greater than the transmission capacity P_{line} of the transmission grid, the phenomenon of wind-blocking and curtailment power occurs. The expression of blocked wind power is **(4)**

$$P_{LT-i}(t) = \begin{cases} 0, & P_{W-i}(t) < P_{line} \\ P_{W-i}(t) - P_L, & P_{W-i}(t) \geq P_{line} \end{cases} \quad (4)$$

In **formula (4)**, $P_{LT\text{-}i}(t)$ is the blocked wind power curtailment power generated at time t within the scheduling day i. **Figure 2** shows the blocked abandonment power during seven dispatch days.

Coupling Relationship Between Two Factors Causing by Wind Power Curtailment

The total abandonment power is obtained by coupling the abandonment power caused by the above two factors. The formula is as **(5)**.

$$P_{LW-i}(t) = P_{LG-i}(t) \cup P_{LT-i}(t) \quad (5)$$

In **formula (5)**, $P_{LW\text{-}i}(t)$ is the total abandonment power at time t within the dispatch day i. **Figure 3** shows the total abandonment, which is a result of the coupling of insufficient flexibility abandonment and blocking wind power curtailment in seven dispatch days.

Because the load is relatively stable with conventional units, the system's wind power absorption space is relatively stable, and the wind power curtailment caused by insufficient flexibility is relatively stable. However, when planning for the expansion of the transmission grid, the transmission capacity of wind power is different, and the generated wind power is also different. If the power transmission capacity of wind power curtailment power is sufficient, no obstruction or wind power curtailment will occur; if the transmission capacity of wind power is insufficient, wind power obstruction and curtailment will occur.

Taking the networking system of a wind power base in Northeast China as the analysis object, the installed capacity of wind power is 1900 MW, the maximum output per unit value is 0.77, and the system load is 3769 MW. Set up different transmission capacity P_{line}, then analyze the relationship between blocked wind power curtailment, insufficient flexibility curtailment, total wind power curtailment, and curtailment duration. Among them, ΔE_1 and T_1 are the power and duration caused by the transmission congestion throughout the year; ΔE_2 and T_2 are the power and duration of wind power curtailment guided by insufficient flexibility; ΔE and T are the total power and duration of wind power curtailment throughout the year, respectively.

It can be seen from **Table 1**:

(1) When the transmission capacity is greater than the maximum wind power, the system wind power curtailment is determined by the power flexibility capacity, and the transmission capacity which is greater than the maximum wind power will not be used, resulting in excessive transmission investment;
(2) When the transmission capacity is less than the maximum wind power, the blocked wind will be curtailed; when the transmission capacity decreases in steps of 0.1, the blocked wind power ΔE_1 and T_1 will increase accordingly, and the total power and duration of the system will be less than the sum of flexibility and insufficient wind power curtailment. Obstruction wind power curtailment shows that there is a certain overlap between the two wind power curtailment, as shown in **Figure 3**.

PRINCIPLES OF COORDINATED PLANNING BETWEEN THE ENERGY STORAGE SYSTEM AND TRANSMISSION

In the power system with a rapidly growing scale of wind power, insufficient flexibility wind power curtailment coexists. To completely solve or improve wind power curtailment, ensure

TABLE 1 | ΔP, ΔE and duration T under different Pline.

Transmission capacity P_{line}/p.u	Blocking wind power curtailment		Insufficient peak shaving and wind power curtailment		Total wind power curtailment	
	ΔE_1/p.u.	T_1/h	ΔE_2/p.u.	T_2/h	ΔE/p.u.	T/h
1.0	0	0	0.079	1,138.50	0.079	1,138.50
0.8	0	0	0.079	1,138.50	0.079	1,138.50
0.7	0.001	80	0.079	1,138.50	0.0794	1,174.75
0.6	0.01	312.75	0.079	1,138.50	0.086	1,312
0.5	0.038	780	0.079	1,138.50	0.105	1,620.25
0.4	0.097	1,537.75	0.079	1,138.50	0.149	2,126.25

that the system has sufficient flexibility and transmission capacity is necessary.

Based on the above analysis, there is a certain overlap between the two wind power curtailment events, and wind power generation has low energy density characteristics. The configuration of the energy storage system can increase the flexibility of the system, and it can also reduce the transmission demand for wind power at the same time. Therefore, in planning, the coordination and optimization of the energy storage system and transmission lines can significantly reduce wind power curtailment and increase the utilization rate of energy storage and transmission grid. On the one hand, it can increase system flexibility. On the other hand, it can reduce the nature of the storage-transmission joint planning to cope with insufficient flexibility and transmission congestion. The essence of the joint planning is to reduce wind power curtailment while achieving coordinated configuration between the energy storage system and transmission lines, and to minimize the comprehensive cost including transmission investment, energy storage system investment, power generation cost and wind power curtailment cost.

The specific planning principle is: the first step is to input wind power data and load data for the planning target year and give the initial grid topology and conventional power supply structure. In the second step, assuming sufficient power transmission capacity, a time-series simulation is performed to calculate the flexibility and wind power curtailment through the feasible region of wind power. The third step assumes that the flexibility capacity is sufficient, and the power flow is calculated based on a given power topology to get the blocking wind power curtailment. The fourth step is to couple the insufficient flexibility and the blocking wind power curtailment to obtain the system wind power curtailment. The fifth step is to configure energy storage system, and calculate the comprehensive costs including transmission investment, energy storage system investment, power generation cost, and wind power curtailment cost, and then modify the energy storage system configuration to repeatedly calculate the optimal configuration of energy storage system under the current grid topology. Revise the capacity of the wind power grid-connected point (that is, modify the planning boundary of the wind power grid-connected point), obtain a new grid topology through grid planning, and repeat the first step to the fifth step until the optimal plan for the coordinated planning of energy storage system and transmission lines is obtained. The specific principles are shown in **Figure 4**.

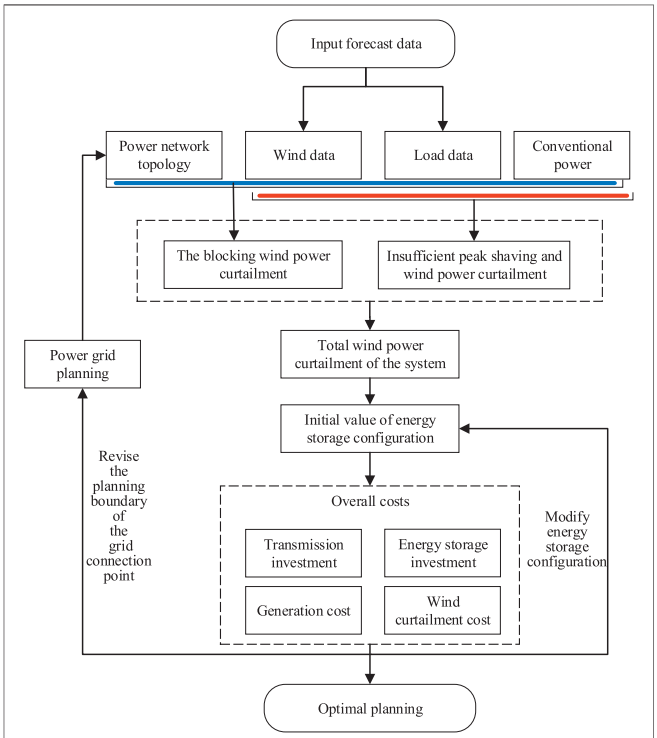


FIGURE 4 | Principles of storage-transmission joint planning.

THE STORAGE-TRANSMISSION JOINT OPTIMIZATION PLANNING MODEL

Planning Model

To achieve the goal of minimizing comprehensive costs such as transmission investment, energy storage investment and wind curtailment cost, this article comprehensively considers the coupling relationship between insufficient flexibility system and wind power curtailment and transmission congestion, and builds a storage-transmission joint optimization model to deal with insufficient flexibility and transmission congestion, and realizes coordination and optimization between energy storage system configuration and expansion of transmission lines. The objective function formula is as **(6)** (Wang et al., 2019):

$$F = \min\left[I\left(P_{\text{line}}\right) + I\left(P_{\text{ess}}, E_{\text{ess}}\right) + L\left(P_{\text{line}}, E_{\text{ess}}\right)\right] \tag{6}$$

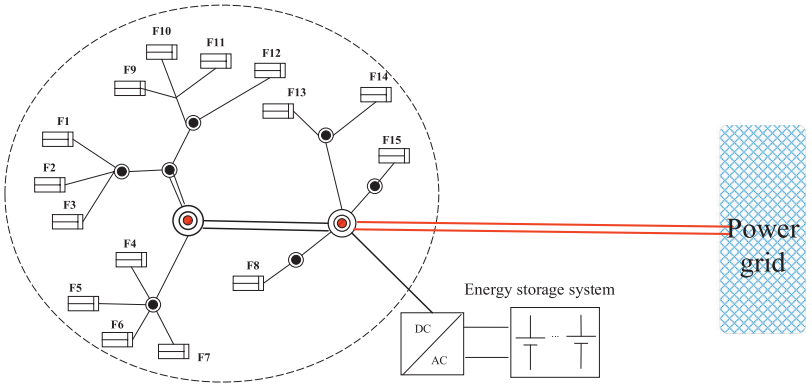


FIGURE 5 | Wind farm cluster centralized grid.

In **formula (6)**, I (P_{line}) is the cost of transmission lines expansion; I (P_{ess}, E_{ess}) is the investment cost of configuring energy storage system; L (P_{line}, E_{ess}) is the cost of wind power curtailment.

(1) Expansion cost of transmission lines: Transmission line expansion cost is the sum of all expansion line costs.

$$I(P_{\text{line}}) = \sum_{i \in \Omega} n_i C_i L_i Z_i, \quad C_i = K_l P_{l-i} \tag{7}$$

In **formula (7)**, n_i is the number of construction times of the ith candidate line; C_i is the unit price per unit length of the ith candidate line, K_l is the unit power cost per unit length of the line, L_i is the length of the ith candidate line, and Z_i is the ith of the 0–1 decision variable for the investment of the line to be selected, $P_{l\text{-}i}$ is the transmission capacity of the ith line to be selected; Ω is the set of lines to be selected.

(2) Energy storage system investment cost: Energy storage investment cost is the sum of energy storage power cost and energy storage capacity cost (Al Ahmad and Sirjani, 2020).

$$I(P_{\text{ess}}, E_{\text{ess}}) = P_{\text{ess}} \times K_{\text{Pess}} + E_{\text{ess}} \times K_{\text{E}} \tag{8}$$

In **formula (8)**, K_{pess} is the purchase cost per unit power converter; K_E is the purchase cost per unit capacity of the energy storage system.

(3) Loss cost of abandoned wind power

$$L(P_{\text{line}}, E_{\text{ess}}) = K_{\text{W}} \times T_{\text{S}} \times \sum_{i=1}^{365} \left(\int_0^{24} P_{LW-i}(t)\mathrm{d}t - \Delta W_{\text{ess}-i} \right) \tag{9}$$

In **formula (9)**, K_w is the penalty value of wind power curtailment per unit, T_s is the planning period, and $\Delta W_{ess\text{-}i}$ is the wind power curtailment absorbed by the energy storage on the ith dispatch day.

TABLE 2 | Coordination and optimization results of energy storage and transmission channels (continuous value).

Name	Configuration	Cost
Transmission capacity	1253 MW	2,305 million RMB
Energy storage capacity	358 MW h	1,070 million RMB
Converter power	178 MW	178 million RMB
Annual wind power curtailment	224,000 MW h	3,712 million RMB
Total costs	7,265 million RMB	

Constraints

(1) Transmission limit constraint of transmission lines: The power flow of the ith transmission line shall be less than its maximum power flow (Qiu et al., 2017).

$$P_i^{\max} < n_i P_{l-i} Z_i \tag{10}$$

P_{imax} is the maximum flow of the i transmission line.

(2) Power balance constraint: The sum of thermal power units, wind power, energy storage and abandon wind power needs to keep balance with the load (Qiu et al., 2017; Zhang and Conejo, 2018).

$$\sum_k P_{\text{gk}}(t) + P_{\text{w}}(t) + P_{\text{ess}}(t) + \Delta P_{\text{w}}(t) = P_{\text{l}}(t) \tag{11}$$

Where is the output power of all thermal power units at time t; $P_w(t)$ is the output power of the wind farm group at time t; $P_{ess}(t)$ is the output power at time t of the energy storage system; $\Delta P_w(t)$ is the abandoned wind power at time t; $P_l(t)$ is the total system load at time t.

(3) The state of charge of the energy storage system and Nissin constraints: The soc of the energy storage system is limited between the minimum SOC and the maximum SOC, and the SOC at the beginning and end of each day is the same (Nikoobakht and Aghaei, 2019; Wu and Jiang, 2019).

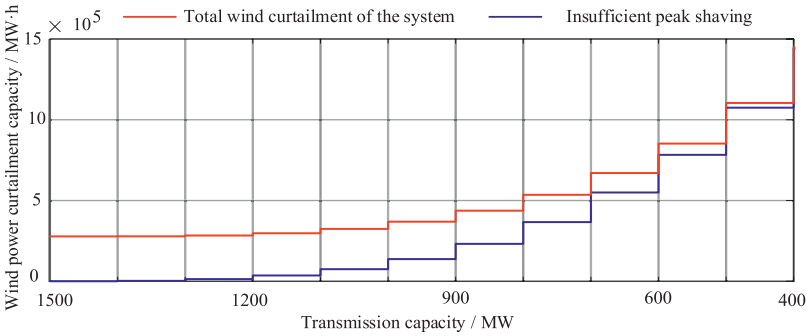


FIGURE 6 | Wind farm cluster centralized grid.

$$\begin{cases} \mathrm{SOC}_{\min} \le \mathrm{SOC}(t) \le \mathrm{SOC}_{\max} \\ \mathrm{SOC}(0) = \mathrm{SOC}(24) \end{cases} \tag{12}$$

(4) The upper and lower limits of the output of thermal power units: The output limit of the thermal power unit is limited to the minimum output limit and rated capacity (Zheng et al., 2015a).

$$P_{gk}^{\min} \le P_{gk}(t) \le P_{gk}^{\max} \tag{13}$$

(5) System and branch flow constraint (Zheng et al., 2015a)

$$\mathrm{B}\theta = P_{\mathrm{G}} + P_{\mathrm{w}} + P_{\mathrm{ess}} - P_{\mathrm{L}} \tag{14}$$

In **formula(13)**, B is the node admittance matrix; θ is the node voltage phase angle vector; $\mathrm{P_G}$ is the output power vector of the thermal power unit; P_w is the output power vector of the wind farm; P_{ess} is the output power vector of the energy storage system device; P_L is the load power vector.

(6) Electricity balance constraint: The sum of thermal power generation and wind power generation is not less than the sum of abandoned wind power and load power (Sima et al., 2019).

$$\mathrm{W_g} + \mathrm{W_w} - \Delta\mathrm{W_w} \ge \mathrm{W}_L \tag{15}$$

In **formula (12)**, $\mathrm{W_g}$, $\mathrm{W_w}$, $\Delta\mathrm{W_w}$, $\mathrm{W_L}$ are the power of thermal power units, wind power generation, wind power generation, and load power respectively.

CASE STUDIES

In order to verify the feasibility and effectivity of the proposed method, this paper analyzes a case of a wind power base in Northeast China and a modified Garver-6 node system, the relevant parameters of wind turbines are shown in **Appendix Table A1**. At the same time, the method in this paper and the traditional energy storage and transmission in literature (Sima et al., 2019), grid planning method without energy storage are contracted. Convenient for expression, the above method is named, and abbreviations:

(1) STJP, the method in this paper named: Storage-transmission joint planning method to deal with insufficient flexibility and transmission congestion, abbreviation: STJP
(2) STP, the traditional storage and coordinate grid planning method in literature (Sima et al., 2019), namely boundary according to the installed capacity for wind power grid planning, abbreviation: STP
(3) NSTP, without energy storage in the transmission planning method, abbreviation: NSTP

The given basic calculation conditions are:

(1) Unit comprehensive cost of 220 kV transmission project K_{l} = 10 thousand yuan/(km MW);
(2) Unit capacity cost of energy storage system K_{E} = 1.5 million yuan/MW h;
(3) The unit cost of converter K_{Pess} = 500,000 yuan/MW;
(4) Unit wind power curtailment penalty Kw = 0.083 million/MW h;
(5) Energy storage system life of 10 years;
(6) Planning period T_{s} = 20 years;

Testing System 1-Centralized Outward Transmission and Transmission Project of the Wind Power Base

The installed capacity of wind power in the centralized transmission system of a wind power base in Northeast China is 1900 MW, the maximum wind power is 1463 MW, the grid side load is 3769 MW, the total length of the transmission channel is 184 km, and the wind power base network topology is shown in **Figure 5**.

To deeply analyze the interaction among transmission capacity, wind abandonment and energy storage in the planning process, based on the transmission capacity of the single-circuit transmission line, the outgoing transmission capacity of the wind power base is set as a continuous value, that is, the transmission capacity can change in 1 MW steps. Therefore, two examples of transmission capacity with continuous value and engineering value are shown.

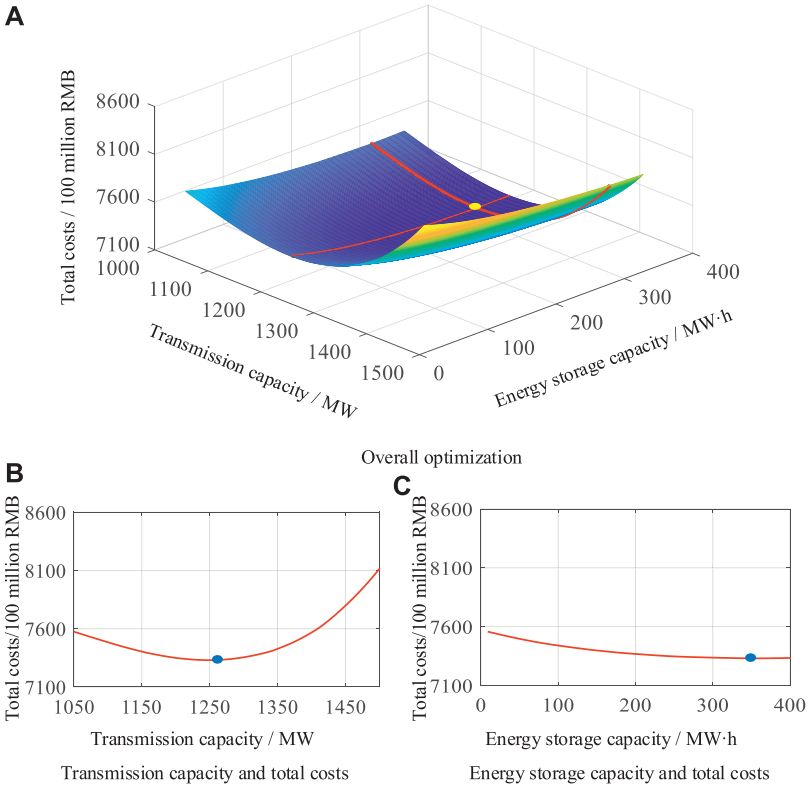


FIGURE 7 | Windfarm cluster centralized grid.

TABLE 3 | Comparison of planning schemes (engineering value).

	STJP		STP		NSTP	
	Configuration	**Cost/million RMB**	**Configuration**	**Cost/million RMB**	**Configuration**	**Cost/million RMB**
Transmission capacity	1200 MW	2,208	1800 MW	3,312	1800 MW	3,312
Energy storage capacity	390 MW h	1,170	296 MW h	888	0 MW h	0
Converter power	188 MW	188	148 MW	148	0 MW	0
Annual wind power curtailment	2.25×10^5MW h	3,732	2.27×10^5MW h	3,774	2.82×10^5MW h	48.5
Total costs	7,298 million RMB		8,122 million RMB		8,162 million RMB	

(1) The transmission capacity as a continuous value

Use the storage-transmission joint optimization model proposed in this paper to deal with insufficient flexibility and transmission congestion. The optimization scheme obtained is shown in **Table 2**. When the wind power transmission channel capacity is 1253 MW, the energy storage system charge/discharge power is 178 MW, and the capacity is 358 MW h, the optimal total cost is 7.265 billion yuan.

The centralized transmission capacity of wind power is reduced in steps of 100 MW. The relationship between the wind power congestion and the total wind power curtailment of the system is analyzed, as shown in **Figure 6**. It can be seen from the figure that when the transmission capacity is 1500 MW, the wind power curtailment of the system is insufficient for flexibility; when the transmission capacity continues to decrease, the blocked wind power curtailment starts to increase, and the total system wind power curtailment also increases, but the blocked wind power curtailment increases more than the total Wind power curtailment grows at a rapid rate. Eventually, wind power curtailment includes insufficient flexibility and wind power curtailment. The total wind power curtailment of the system coincides with the blocked.

The changing trend of total costs in the process of coordination and optimization of the energy storage system and transmission capacity is shown in **Figure 7A**. The relationship between transmission capacity and total costs present a downward concave surface. In the process of increasing transmission capacity, the initial total costs are discarded. The total costs of wind are dominated, and the total costs of power transmission are dominated by the

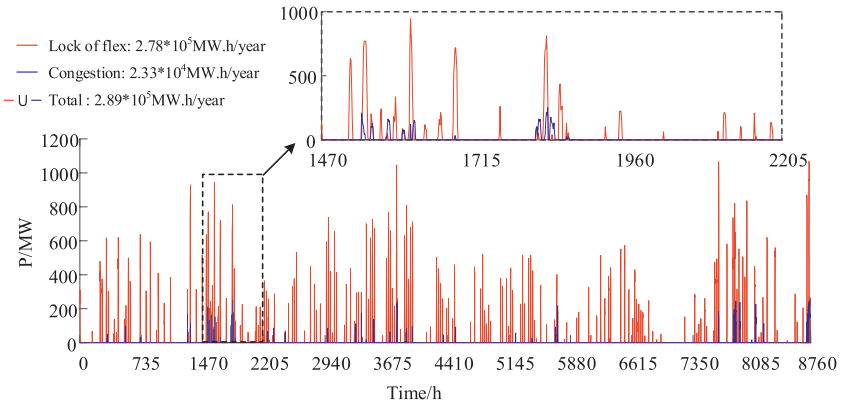


FIGURE 8 | Coupling relationship of wind curtailment with an engineering value of transmission capacity.

transmission investment cost when the optimal solution is reached. The optimal solution is marked with a yellow origin, and the horizontal and vertical directions of the optimal solution are the relationship between transmission capacity, energy storage system capacity and total costs, as shown in **Figures 7B,C**.

(2) Transmission lines capacity is engineering value

After consulting relevant information, the transmission limit of a 220 kV transmission line of 184 km is 300 MW. Under these conditions, this paper compares the various performance of system parameters amount the proposed method STJP, STP (Sima et al., 2019) and NSTP. The comparison results of the above methods are shown in **Table 3**.

As can be seen from **Table 3**, the proposed STJP can effectively reduce the annual wind power curtailment. To be more exact, it is reduced by 5×10^3 MW h and 57×10^3 MW h compared with the STJ method and NSTP method, respectively. In addition, the transmission capacity has dropped from 1800 to 1200 MW when using the proposed STJP. This improves the percentage of utilization of transmission assets. Although the energy storage capacity has increased by 31.7% compared with the STP, the total costs when using STJP method are much lower than the other two costs with STP and NSTP.

As is shown in **Figure 8**, the red line demonstrates the generated wind curtailment due to system's lack of flexibility, whereas the blue one refers to the generated wind curtailment due to transmission congestion. In order the analyze the results of wind curtailment in different condition, the comparison has been made as below. Without consideration of transmission congestion, the summation of wind curtailment in each time period within one year is nearly 2.78×10^5 MW h. However, the summation of wind curtailment in each time period within one year is nearly2.89×10^5 MW h when the consideration is given to the lack of flexibility and transmission congestion. In addition, the sub-figure mainly illustrates the coupling relationship between the lack of flexibility and transmission congestion that lead to wind curtailment. During the period from 1,470 to 1715, there is an overlap of wind curtailment, which results from the above two reason. Therefore, when catcalling the summation of wind curtailment, the overlap part should be omitted. Further, the summation of wind curtailment should be less than the wind curtailment caused by lack of flexibility and transmission congestion respectively.

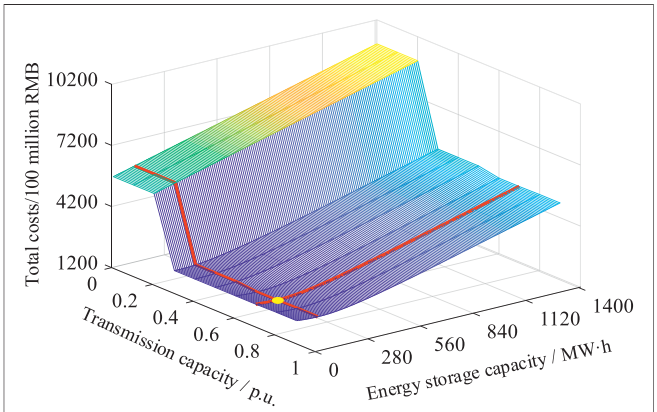


FIGURE 9 | Windfarm cluster centralized grid.

In summary, the proposed method STJP can make full use of the coupling relationship between insufficient flexibility and wind power transmission congestion, which improves the transmission asset utilization and energy storage efficiency. It also achieves the most optimally economic planning schemes among the three methods.

Testing System 2-Garver6 System

In order to identify the universality of the proposed algorithm, the STJP algorithm is also applied in the cast study 5.2 of Garver6 System.

According to the original calculation example of the Garver-6 node system, the load of each node in the system matches the measured load and wind power data of a certain regional power grid in Northeast China. The total load is 763 MW, 6-node wind power installed capacity is 300 MW, and the maximum wind

TABLE 4 | Different planning methods of Garver6 system.

Planning method	STJP		STP		NSTP	
	Configuration	Cost/ million RMB	Configuration	Cost/ million RMB	Configuration	Cost/ million RMB
Wind power grid-connected planning boundary	0.7		1.0		1.0	
Extension line	1-4(2), 1-5(1), 2-3(3), 3-5(3), 2-6(1), 4-6(2)	311	1-2(1), 1-4(3), 1-5(2), 2-3(4), 2-4(1), 3-5(4), 2-6(2), 4-6(2)	338	1-2(1), 1-4(3), 1-5(2), 2-3(4), 2-4(1), 3-5(4), 2-6(2), 4-6(2)	338
Energy storage capacity	129 MW h	387	105	315	0	0
Converter power	49 MW	49	39	19.5	0	0
Annual wind power curtailment	49,500 MW h	821	55,830 MW h	956.028	77,920 MW h	1,340.2
Total costs	1,568 million RMB		1,629 million RMB		1,649 million RMB	

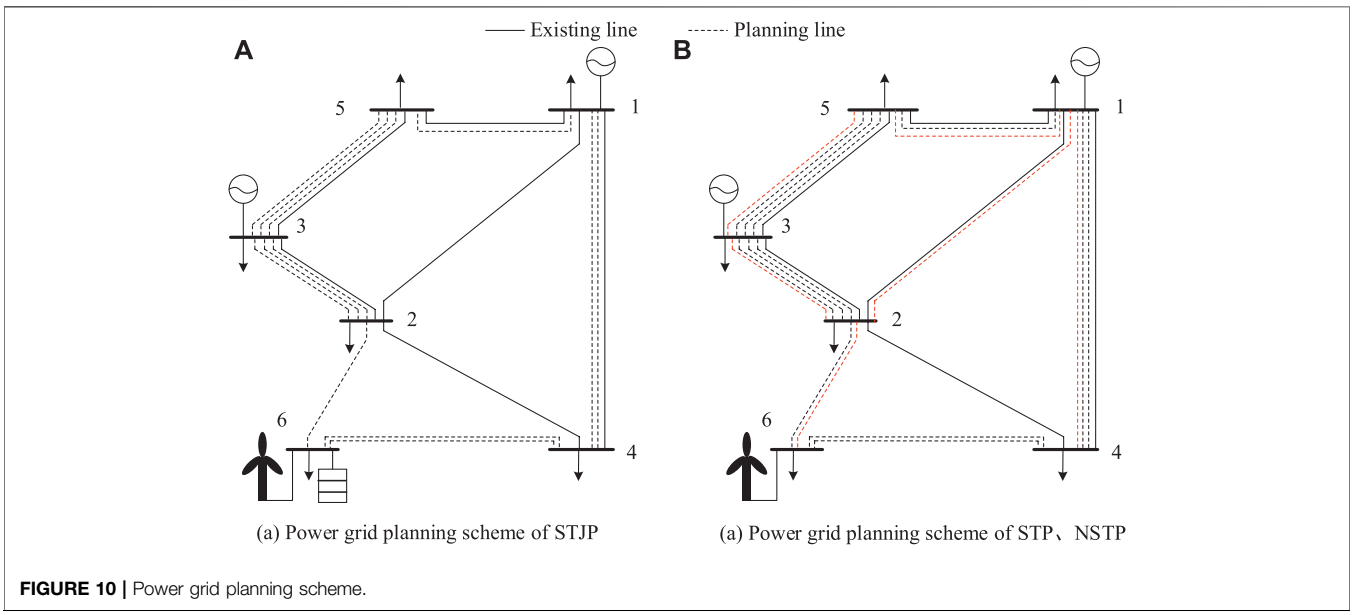


FIGURE 10 | Power grid planning scheme.

power is 261 MW. In the storage-transmission joint optimization using STJP of this article on this example system, the most important point here is to change the boundary of the 6-node wind power grid-connected planning during the optimization process (the rule is: the installed capacity is the benchmark, and the 0.1 installed capacity is used as the step size). Then get blocked wind curtailment, and coupled with insufficient flexibility to obtain the total wind curtailment of the system, and finally optimize the energy storage configuration and the wind curtailment loss. Finally, get the plan of the Garver6 system storage-transmission joint planning, the power grid planning scheme is shown in **Figure 9**.

In **Figure 9** of the overall optimization process, the yellow dot is the optimal solution position of the storage-storage joint planning. As the transmission lines increase or decrease according to the number of lines, the total cost of the wind power grid connection plan increases from small to small, and the total cost first decreases and then increases in a stepwise manner, and several grid-connected planning boundaries are corresponding to the same plan. In the process of energy storage configuration, the total cost is first reduced and then increased during the process of energy storage capacity reduction. The overall optimization process can reflect that the storage-transmission joint planning is effective to deal with insufficient flexibility and transmission congestion and abandonment of wind, which promotes wind power consumption and reduces transmission line investment.

At the same time, STP and NSTP are used for optimization planning. The planning scheme of each planning method is shown in **Table 4**, grid planning topology is shown in **Figure 10**, (a) Plan is for STJP, (b) Plan is for STP and NSTP.

Figure 11 is similar to **Figure 8**. In order the analyze the results of wind curtailment in different condition, the comparison has been made as below. Without consideration of transmission congestion, the summation of wind curtailment in each time period within one year is nearly 7.29×10^4 MW h. However, the

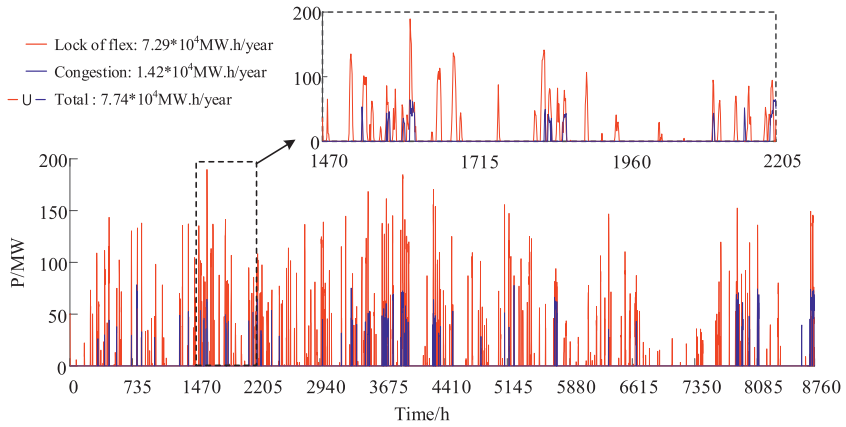


FIGURE 11 | Coupling relationship of wind curtailment in Garver6 System.

summation of wind curtailment in each time period within one year is nearly 7.74×10^4 MW h when the consideration is given to the lack of flexibility and transmission congestion. In addition, the sub-figure mainly illustrates the coupling relationship between the lack of flexibility and transmission congestion that lead to wind curtailment. During the period from 1,470 to 1715, there is an overlap of wind curtailment, which results from the above two reason. Therefore, when catcalling the summation of wind curtailment, the overlap part should be omitted.

As can be seen from **Table 4**, the proposed STJP can effectively reduce the annual wind power curtailment. To be more exact, it is reduced by 6330 MW h and 28,420 MW h compared with the STJ method and NSTP method, respectively. In addition, compared with STJ method and NSTP method, STJP reduces six transmission lines and saves 27 million dollars. In general, although the cost increase of STJP energy storage configuration increases by 72 million dollars, it brings about the decrease of annual wind power curtailment and transmission line investment, thus minimizing the total cost among the three methods.

CONCLUSION

To solve the wind curtailment caused by large-scale wind power grid integration, this paper first analyzes the coupling relationship between insufficient flexibility and wind abandonment events with power transmission congestion, then a joint planning model is constructed for energy storage and transmission grid to deal with insufficient flexibility and transmission congestion. Finally, a case study has been performed to identify the effectiveness of the proposed method by comparison with.

STP and NSTP methods. The main conclusions are as follows:

(1) By applying the planning scheme STJP, the overall system achieve a much lower annual wind power curtailment compared with the other two traditional planning schemes, although it brings a relatively higher energy storage investment, cost of its total investment can be optimized to be far lower than any other planning performance. Taking a wind power base in Northeast China as an example, compared with the STP and NSTP methods, the annual curtailment of wind power is reduced by 2000 MW h and 57000 MW h, respectively, and the total system investment cost has been reduced by 11.3% and 11.8%.

(2) During the optimization configuration of STJP, it gradually reduces the planning boundary of the wind power base grid connection point (rather than the installed capacity of the wind power base), and makes full use of the interaction between energy storage and transmission lines, the problem of low asset utilization of wind power transmission and transmission has been better solved. It is also the same case that the transmission capacity has dropped from 1800 to 1200 MW when applying the STJP with joint optimization of energy storage configuration and transmission capacity.

In summary, the method in this paper makes full use of the coupling relationship between insufficient flexibility and transmission congestion. The planning boundary of the grid connection point of the wind power base (below its installed capacity) was reduced, and the utilization rate of energy storage and transmission lines was greatly increased. Meanwhile, the proposed STJP overrides the STP and NSTP in terms of algorithm efficiency and complexity. With the continuous improvement of energy storage technology and the economy, an effective planning method is provided for the realization of a high proportion of renewable energy power system planning.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material,

further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

XY contributed to the conception of the study; GC performed the experiment; XL contributed significantly to analysis and manuscript preparation; MX performed the data analyses and wrote the manuscript; QG helped perform the analysis with constructive discussions.

FUNDING

This study was supported by the National Key Research and Development Program (2016YFB0900100).

REFERENCES

Al Ahmad, A. K., and Sirjani, R. (2020). "Optimal allocation of energy storage system in transmission system considering wind power", in 7th international conference on electrical and electronics engineering (New York: Institute of Electrical and Electronics Engineers Inc.), 181–187. doi:10.1109/ICEEE49618.2020.9102603

Bustos, C., Sauma, E., De La Torre, S., Aguado, J. A., Contreras, J., and Pozo, D. (2018). Energy storage and transmission expansion planning: substitutes or complements?. *IET Gener. Transm. Distrib.* 12, 1738–1746. doi:10.1049/iet-gtd.2017.0997

Correa, C., Sanchez, A., and Marulanda, G. (2016). Expansion of transmission networks considering large wind power penetration and demand uncertainty. *IEEE Latin Am. Trans.* 14, 1235–1244. doi:10.1109/TLA.2016.7459604

Del Rosso, A. D., and Eckroad, S. W. (2014). Energy storage for relief of transmission congestion. *IEEE Trans. Smart Grid.* 5, 1138–1146. doi:10.1109/TSG.2013.2277411

Dicorato, M., Forte, G., Pisani, M., and Trovato, M. (2012). Planning and operating combined wind-storage system in electricity market. *IEEE Trans. Sustain. Energy.* 3, 209–217. doi:10.1109/TSTE.2011.2179953

Dvorkin, Y., Ricardo, F.-B., Wang, Y., Xu, B., Kirschen, D. S., Pandi, H., et al. (2018). Co-planning of investments in transmission and merchant energy storage. *IEEE Trans. Power Syst.* 33, 245–256. doi:10.1109/TPWRS.2017.2705187

Gan, L., Li, G. Y., and Zhou, M. (2016). Coordinated planning of large-scale wind farm integration system and regional transmission network considering static voltage stability constraints. *Electr. Power Syst. Res.* 136, 298–308. doi:10.1016/j.epsr.2016.03.002

GWEC (2019). *Global Wind Report 2019 [Online]*. Available online at: https://gwec.net/docs/global-wind-report-2019/.

Hand, M. M., Baldwin, S., and De Meo, E. (2014). *Renewable electricity futures study*. Colorado: National Renewable Energy Laboratory.

Jorgenson, J., Denholm, P., and Mai, T. (2018). Analyzing storage for wind integration in a transmission-constrained power system. *Appl. Energy*. 228, 122–129. doi:10.1016/j.apenergy.2018.06.046

Jorgenson, J., Mai, T., and Brinkman, G. (2017). *Reducing wind curtailment through transmission expansion in a wind vision future*. NREL/TP-6A20-67240

Li, Y., Wang, C., Li, G., Wang, J., Zhao, D., and Chen, C. (2020). Improving operational flexibility of integrated energy system with uncertain renewable generations considering thermal inertia of buildings. *Energy Convers. Manag.* 207, 2526. doi:10.1016/j.enconman.2020.112526

Li, Y., Yang, Z., Li, G. Q., Zhao, D. B., and Tian, W. (2019). Optimal scheduling of an isolated microgrid with battery storage considering load and renewable generation uncertainties. *IEEE Trans. Ind. Electron.* 66, 1565–1575. doi:10.1109/TIE.2018.2840498

Lu, Z., Li, H., and Qiao, Y. (2018). Probabilistic flexibility evaluation for power system planning considering its association with renewable power curtailment. *IEEE Trans. Power Syst.* 33, 3285–3295. doi:10.1109/TPWRS.2018.2810091

Masoumzadeh, A., Nekouei, E., Alpcan, T., and Chattopadhyay, D. (2018). Impact of optimal storage allocation on price volatility in energy-only electricity markets. *IEEE Trans. Power Syst.* 33, 1903–1914. doi:10.1109/TPWRS.2017.2727075

National Energy Administration (2020). *Grid connected operation of wind power in 2019 (in Chinese) [Online]*. Available online at: http://www.nea.gov.cn/2020-02/28/c_138827910.htm.

NEA (2016). *Chinese wind curtailment hit 15% in 2015 on 33GW of new capacity [Online]*. Available online at: http://www.rechargenews.com/wind/867893/chinese-wind-curtailment-hit-15-percent-in-2015-on-33gw-of-new-capacity-nea.

Nikoobakht, A., and Aghaei, J. (2019). Integrated transmission and storage systems investment planning hosting wind power generation: continuous-time hybrid stochastic/robust optimisation. *IET Gener. Transm. Distrib.* 13, 4870–4879. doi:10.1049/iet-gtd.2019.0257

Padhee, M., Pal, A., Mishra, C., and Vance, K. A. (2020). A fixed-flexible BESS allocation scheme for transmission networks considering uncertainties. *IEEE Trans. Sustain. Energy.* 11, 1883–1897. doi:10.1109/TSTE.2019.2946133

Price Water House Coopers LLP (PwC) (2014). *Potsdam institute for climate impact research (PIK), and international institute for applied systems analysis (IIASA)% renewable electricity: a roadmap to 2050 for europe and North Africa*.

Qiu, T., Xu, B., Wang, Y., Dvorkin, Y., and Kirschen, D. S. (2017). Stochastic multistage coplanning of transmission expansion and energy storage. *IEEE Trans. Power Syst.* 32, 643–651. doi:10.1109/TPWRS.2016.2553678

Sima, C. A., Popescu, M. O., Popescu, C. L., and Lazaroiu, G. (2019). "Integrating energy storage systems and transmission expansion planning in renewable energy sources power systems", in 54th international universities power engineering conference (New York: Institute of Electrical and Electronics Engineers Inc.). doi:10.1109/UPEC.2019.889348

Wang, Y. C., Lou, S. H., Wu, Y. W., and Wang, S. R. (2020). Flexible operation of retrofitted coal-fired power plants to reduce wind curtailment considering thermal energy storage. *IEEE Trans. Power Syst.* 35, 1178–1187. doi:10.1109/TPWRS.2019.2940725

Wang, Y., Lou, S., Wu, Y., Lv, M., and Wang, S. (2019). Coordinated planning of transmission expansion and coal-fired power plants flexibility retrofits to accommodate the high penetration of wind power. *IET Gener. Transm. Distrib.* 13, 4702–4711. doi:10.1049/iet-gtd.2018.5182

Wu, X., and Jiang, Y. (2019). Source-network-storage joint planning considering energy storage systems and wind power integration. *IEEE Access.* 7, 137330–137343. doi:10.1109/ACCESS.2019.2942134

Yacar, D., Tejada-Arango, D. A., and Wogrin, S. (2018). Storage allocation and investment optimisation for transmission-constrained networks considering losses and high renewable penetration. *IET Renew. Power Gener.* 12, 1949–1956. doi:10.1049/iet-rpg.2018.5472

Zhang, H., Zhang, S., Cheng, H., Wang, Z., and Zhang, J. (2019). "Probabilistic transmission network expansion planning considering integration of wind power", in *IEEE PES innovative Smart grid technologies Asia.* (New York: ISGTInstitute of Electrical and Electronics Engineers Inc.), 3194–3198. doi:10.1109/ISGT-Asia.2019.8881768

Zhang, X., and Conejo, A. J. (2018). Coordinated investment in transmission and storage systems representing long- and short-term uncertainty. *IEEE Trans. Power Syst.* 33, 7143–7151. doi:10.1109/TPWRS.2018.2842045

Zheng, J., Wen, F., Zhou, M., Hu, L., Xu, Q., and Lan, Z. (2015a). "Transmission planning with renewable generation and energy storage", in *10th international conference on Advances in power system control, operation and management.* (New York: Institution of Engineering and Technology). doi:10.1049/ic.2015.0224

Zheng, L., Hu, W., Lu, Q., and Min, Y. (2015b). Optimal energy storage system allocation and operation for improving wind power penetration. *IET Gener. Transm. Distrib.* 9, 2672–2678. doi:10.1049/iet-gtd.2014.1168

Zhou, Y. L., and Wang, D. (2017). An improved coordinated control technology for coal-fired boiler-turbine plant based on flexible steam extraction system. *Appl. Therm. Eng.* 125, 1047–1060. doi:10.1016/j.applthermaleng.2017.07.077

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

NOMENCLATURE

B The node admittance matrix

C_i The unit price per unit length of the *i*th candidate line

F Comprehensive cost

$I(P_{ess}, E_{ess})$ The investment cost of configuring energy storage system

$I(P_{line})$ The cost of transmission lines expansion

K The standby coefficient of conventional units on dispatch day *i*, respectively

K_E The purchase cost per unit capacity of the energy storage system

K_l The unit power cost per unit length of the line

K_{pess} The purchase cost per unit power converter

K_w The penalty value of wind power curtailment per unit

$L(P_{line}, E_{ess})$ The cost of wind power curtailment

L_i The length of the *i*th candidate line

n_i The number of construction times of the *i*th candidate line

NSTP The no-storage transmission planning method

$P_{D.max-i}$ The maximum load of conventional units on dispatch day *i*, respectively

$P_{D-i}(t)$ The load value at time *t* in dispatching day *i*

P_{ess} The output power vector of the energy storage system device

$P_{ess}(t)$ The output power at time *t* of the energy storage system

$P_{FR-i}(t)$ The maximum wind power that the system can accept on dispatching day *i*

P_G The output power vector of the thermal power unit

$P_{G.min-i}$ The minimum technology for conventional units in dispatching day *i* contribute

P_{G-i} The starting capacity of conventional units on dispatch day *i*, respectively

$P_{gk}(t)$ The output power of the thermal power unit at t time

P_{gk}^{max} The maximum output of thermal power units

P_{gk}^{min} The minimum output of thermal power units

P_i^{max} The maximum flow of the *i* transmission line

P_{line} Transmission capacity

P_L The load power vector

$P_L(t)$ The total system load at time *t*.

$P_{LG-i}(t)$ The wind power curtailment caused by insufficient flexibility at time *t* within the scheduling day *i*

P_{l-i} The transmission capacity of the *i*th line to be selected

$P_{LT-i}(t)$ The blocked wind power curtailment power generated at time *t* within the scheduling day *i*.

$P_{LW-i}(t)$ The total abandonment power at time *t* within the dispatch day *i*

P_w The output power vector of the wind farm

$P_w(t)$ The output power of the wind farm group at time *t*

$P_{W.min-i}$ The minimum wind power of conventional units on dispatch day *i*, respectively

SJP The traditional storage joint planning method

SOC(t) The state of charge of the energy storage system

SOC_{max} The maximum charge of energy storage system

SOC_{min} The minimum charge of energy storage system

STJP The method in this passage: Storage-transmission joint planning method

T The total duration of wind power curtailment throughout the year

T_1 The duration caused by the transmission congestion throughout the year

T_2 The duration of wind power curtailment guided by insufficient flexibility

T_s The planning period

W_g The power of thermal power units

W_L Load power respectively

W_w Wind power generation

Z_i The *i*th of the 0–1 decision variable for the investment of the line to be selected

ΔE The total power of wind power curtailment throughout the year

$ΔE_1$ The power caused by the transmission congestion throughout the year

$ΔE_2$ The power of wind power curtailment guided by insufficient flexibility

$ΔP_w(t)$ The abandoned wind power at time *t*

$ΔW_{ess-i}$ The wind power curtailment absorbed by the energy storage on the *i*th dispatch day

$ΔW_w$ Wind power generation

θ The node voltage phase angle vector

Ω The set of lines to be selected

APPENDIX

TABLE A1 | 3 MW double-fed induction wind generator technical parameter.

Operational parameters		Rotor	
Nominal power	3000 kW	Number of blades	3
Power control method	Variable pitch and speed	Rotor diameter	100 m
Nominal wind speed	13 m/s	The altitude of wheel hub center	84.71 m
Cut-in wind speed	3.5 m/s	Rotor speed	9.2–16.4 rpm
Cut-out wind speed	25 m/s	Rotor swept area	7854 m^2
Operating temperature	−30°C ~ +40°C	Rotor angle	2.5°
Longevity	20 years	Angle of mainshaft	5°
Unit security grade	IEC||A	Airfoil section	NACA63, DU
		Material	GFRP
		Weight	67400 Kg

frontiers in Energy Research

TECHNOLOGY AND CODE
published: 17 March 2021
doi: 10.3389/fenrg.2021.644489

False Data Injection Attack Detection in Power Systems Based on Cyber-Physical Attack Genes

Zhaoyang Qu[1,2], Yunchang Dong[1,2], Nan Qu[3], Huashun Li[4], Mingshi Cui[5], Xiaoyong Bo[1,2], Yun Wu[6] and Sylvère Mugemanyi[1]*

[1]School of Electrical Engineering, Northeast Electric Power University, Jilin, China, [2]Jilin Engineering Technology Research Center of Intelligent Electric Power Big Data Processing, Jilin, China, [3]State Grid Jiangsu Electric Power Co., Ltd., Nanjing, China, [4]State Grid Jilin Electric Power Co., Ltd., Jilin, China, [5]State Grid Inner Mongolia Eastern Electric Power Co., Ltd., Hohhot, China, [6]Zhejiang Windey Co., Ltd., Hangzhou, China

In the process of the detection of a false data injection attack (FDIA) in power systems, there are problems of complex data features and low detection accuracy. From the perspective of the correlation and redundancy of the essential characteristics of the attack data, a detection method of the FDIA in smart grids based on cyber-physical genes is proposed. Firstly, the principle and characteristics of the FDIA are analyzed, and the concept of the cyber-physical FDIA gene is defined. Considering the non-functional dependency and nonlinear correlation of cyber-physical data in power systems, the optimal attack gene feature set of the maximum mutual information coefficient is selected. Secondly, an unsupervised pre-training encoder is set to extract the cyber-physical attack gene. Combined with the supervised fine-tuning classifier to train and update the network parameters, the FDIA detection model with stacked autoencoder network is constructed. Finally, a self-adaptive cuckoo search algorithm is designed to optimize the model parameters, and a novel attack detection method is proposed. The analysis of case studies shows that the proposed method can effectively improve the detection accuracy and effect of the FDIA on cyber-physical power systems.

Keywords: cyber-physical power system (CPPS), false data injection attacks (FDIAs), cyber-physical attack genes, stacked autoencoder network, cuckoo search algorithm (CS), attack detection

OPEN ACCESS

Edited by:
Liang Chen,
Nanjing University of Information Science and Technology, China

Reviewed by:
Jun Yin,
North China University of Water Conservancy and Electric Power, China
Shaoyan Li,
North China Electric Power University, China

****Correspondence:***
Yunchang Dong
595245700@qq.com

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

Received: *21 December 2020*
Accepted: *01 February 2021*
Published: *17 March 2021*

Citation:
Qu Z, Dong Y, Qu N, Li H, Cui M, Bo X, Wu Y and Mugemanyi S (2021) False Data Injection Attack Detection in Power Systems Based on Cyber-Physical Attack Genes. Front. Energy Res. 9:644489. doi: 10.3389/fenrg.2021.644489

INTRODUCTION

With the continuous development of information technology, the interaction between information flow and energy flow in power systems is becoming more and more frequent (Yu and Xue, 2016; Xu et al., 2018; Qu et al., 2019). By integrating computing systems, communication networks and physical environments, traditional power systems have evolved into cyber-physical power systems (CPPS) (Yang et al., 2019; He et al., 2020). In the process of production management and dispatch control in smart grids, cyber systems are increasingly inseparable. The integration of renewables is threatening the secure operation of today's CPPS because of inherent uncertainties of renewable power generations (Li et al., 2018; Li et al., 2020). At the same time, however, some vulnerabilities in the cyber system may be exploited by attackers, posing serious threats to the physical system across cyber-physical spaces (Li et al., 2019). They can even cause the temporary paralysis of important infrastructure (Adhikari et al., 2017). FDIAs are advanced and sustainable data integrity attacks in the CPPS. By modifying the collected measurement data, the system state estimation will be

biased and the power grid will be damaged by the incorrect action of switches (Liang et al., 2016). The cyber security of CPPS has gradually attracted people's attention. Furthermore, how to effectively identify the FDIA has become an urgent problem to be solved in the safe and stable operation of power systems.

In recent years, FDIAs in the CPPS have become a hot topic in power system studies. The FDIA was first proposed in Liu et al. (2011). The relevant principles are explained. It is assumed that the attacker can successfully attack the power system by bypassing the traditional bad data detection method when fully grasping the grid topology information and related parameters. But in reality, it is more difficult for an attacker to obtain this information. From the attacker's perspective, research has shown that FDIA can be launched without fully grasping the topology information (Liu et al., 2015). A sequential pattern mining method was proposed in Pan et al. (2015) to accurately extract power system interference and network attack patterns from heterogeneous time synchronization data. This method does not have a suitable division scheme to determine the classification boundary. Through feature engineering to reconstruct new features, a new method of processing abnormal data was proposed (Wang et al., 2019), and then a power system attack detection model based on machine learning was constructed. Existing features were reconstructed in that method, which increased the computational cost. The mechanism and method of FDIA under DC and AC models were studied, and an attack detection method based on the CNN-GRU hybrid model was proposed (Li et al., 2019). The optimal fusion estimation method was designed by adding a compensation factor to study the detection of FDIA attack signal under CPPS (Gao et al., 2019). However, there is no explanation on how to choose the appropriate compensation factor.

At present, the CPPS sensing equipment is gradually increasing, and the amount of data is increasing. The recognition accuracy of traditional methods can no longer meet the increasing actual demand. At the same time, new types of FDIAs are emerging endlessly. Even if the topology information is not fully mastered, it can bypass the traditional detection mechanism to launch attacks.

In summary, the existing research on the detection of FDIAs based on machine learning has the following limitations: 1) less collaborative consideration of cyber and physical data characteristics, unilateral attack detection methods find it difficult to identify complex cyber-physical attacks. 2) The features of the initial CPPS data are complex, and feature selection and transformation directly affect the validity and accuracy of the detection results, and effective feature mining techniques are required. 3) Attack detection speed is an important factor in actual engineering, and it is necessary to increase the model calculation speed.

According to the consideration of the complex characteristics of the CPPS data, this paper proposes a method for identifying the FDIA in the CPPS based on cyber-physical genes. The main contributions of this paper are as follows:

(1) From the perspective of cyber and physical integration of power systems, the concept of "FDIA genes" is proposed. Through the melting processing of information features, the largest information coefficient is introduced for feature selection, and a subset of gene features that are useful for attack detection are screened, and the high-dimensional problem of data features is solved.
(2) An FDIA gene extraction model is constructed based on a stack auto-encoding network (SAE). An unsupervised pre-training encoder is set to extract attack genes, and a supervised fine-tuning classifier is set to classify attacks. Deep learning is used to automatically learn and extract the abstract FDIA gene, which can better represent the essential attributes of the FDIA.
(3) A self-adaptive cuckoo search algorithm (SACS) is proposed to solve and optimize the initialization parameters of the detection model. Historical data are trained and network parameters are fine-tuned to accurately identify the FDIA.

The rest of this paper is organized as follows: according to the principle and characteristics of false data injection attacks, cyber-physical FDIA genes are defined, and the optimal attack gene feature set is selected with the maximum information coefficient in *FDIA Gene Definition and Feature Selection*. The SAE is constructed to extract FDIA genes in *.FDIA Detection Model* A self-adaptive cuckoo algorithm is proposed to solve the model parameter optimization problem, the FDIA detection method based on SACS-SAE is designed to identify attacks in *Attack Detection Method Based on SACS-SAE. Case Studies* demonstrates simulation results on case studies, and finally, the conclusions are drawn in *Conclusion*.

FDIA GENE DEFINITION AND FEATURE SELECTION

FDIAs are a new type of attack against the CPPS. The attacker uses the measurement data of the power system as the target of the attack, and builds the attack vector based on the state estimation equation. Vulnerabilities in the detection mechanism of poor state estimation data are used to tamper with the measurement data at will, prompting the power grid to incorrectly estimate the current operating state. In turn, it will affect the power generation control and dispatch operation of the CPPS, causing system malfunctions. This attack is highly concealed and difficult to be detected. We analyze the principles and characteristics of the FDIA in the CPPS, use the bio-inspired concept of genes, and define the cyber-physical FDIA gene.

Principles of the FDIA

Assuming that the attacker has obtained the complete information of the CPPS, then an FDIA can be constructed. Moreover, it is not discovered by the traditional bad data

detection mechanism. As a result, the state estimation result of the power system is wrong. The process under the direct current system can be described in **Eq. (1)**:

$$Z_a = Z + a = \boldsymbol{H}\theta + a + e \tag{1}$$

Where Z_a is the tampered measurement data, Z is the actual measurement, $\boldsymbol{H}$ is the Jacobian matrix of the power system, θ is the state of the power system, a is the false value injected by the attacker, and e is the amount of error generated in the measurement process.

If c is the false data, it should be satisfied in **Eq. (2)**:

$$\theta_a = \left(\boldsymbol{H}^T W^{-1}\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T W^{-1} Z_a = \theta + c \tag{2}$$

Where W is the diagonal matrix of the measurement error, and the residual expression is shown in **Eq. (3)**:

$$\begin{aligned} e_a &= Z_a - \boldsymbol{H}\theta_a = Z + a - \boldsymbol{H}(\theta + c) \\ &= Z - \boldsymbol{H}\theta + a - \boldsymbol{H}c \end{aligned} \tag{3}$$

From the above equation, when the attacker masters the Jacobian matrix, as long as the injected false data satisfy $a = \boldsymbol{H}c$, the measured data can be tampered with and the traditional bad data detection method can be bypassed to realize the FDIA.

The features of the FDIA in the CPPS are summarized as follows:

(1) The FDIA has a cross-domain interaction in the cyber-physical space. The attacker comprehensively considers the characteristics of cyber-physical coupling, and utilizes cyber-attacks technology to detect, invade, elevate, and control the cyber system. Then the attacker will launch physical attacks such as disconnecting the knife switch to undermine the security of the power system.
(2) The evolution mechanism of FDIA behavior is complicated. A coordinated attack is a finely designed multi-step attack, including the evolution of network behavior over time, and the coordination of multiple attackers that change over space.
(3) The FDIA is hereditary and variant. The same type of attack has the common characteristics of the family. The attack changes the form of the attack through the outer event encapsulation, but the nature of the attack has not changed, making it difficult to identify the attack after the variant.

Cyber-Physical FDIA Genes

The phasor measurement unit has been widely used in the CPPS, providing nearly real-time voltage phase, current phase, and other measurement values for power system personnel (Li et al., 2019) as physical data in the attack detection process. Simultaneously, various intrusion detection systems are deployed in the CPPS. Among them, the collected communication flow, network log and other data are regarded as cyber data. The data of the physical system, and the data of the cyber system together form the cyber-physical data.

The biological gene is the most essential factor used to distinguish different species or different individuals of the same species. It is also the most credible basis for judging the same species or the same individual. Genetic recombination and variation control the family heritability and diversity of organisms (Hagai et al., 2018; Qu et al., 2020).

If the potential key data features that control the law of the FDIA behavior can be found from the original cyber-physical data, the types of attacks can be more effectively detected. Drawing lessons from the idea of the biological gene, the cyber-physical FDIA gene is defined as follows: in the CPPS data set containing the attack, select the key samples in the attack activity, and extract the characteristic data pieces that contain the attack behavior information. First remove the redundant and useless characteristic information, and then perform the self-learning of the inherent law of the characteristic combination change through the process of mining, the genes of the nature of the attack are obtained, and the type of attack can be detected through the uniqueness of each attack gene.

Optimal Attack Gene Feature Set Selection

Due to the increasing number and type of controlled components, the number of key features in the CPPS is relatively large. Data parameters usually have complex properties such as nonlinear correlation, non-functional dependency, and redundancy. For example, the fluctuation of voltage, phase, and frequency show similar fluctuations, indicating that there is a certain degree of redundancy among them (Qu et al., 2018). However, traditional feature selection methods find it difficult to capture the non-functional dependency between features.

Considering that the high-dimensional characteristics, nonlinear correlation, and non-functional dependency of CPPS data have caused serious obstacles in the research and application process. Under the condition of supervised learning, this paper proposes an improved maximum correlation minimum redundancy attack gene feature selection method considering the maximum information coefficient (MIC). The maximum correlation minimum redundancy algorithm determines the optimal set of gene features by calculating the mutual information between different features and categories (Dai et al., 2016). Nevertheless, it is difficult to capture the non-functional dependency between features in CPPS data (Yang et al., 2018). The MIC between features is analyzed using the method of grid partition and mutual information. Given feature x with the probability density $p(x)$ and feature y with the probability density $p(y)$, and the joint probability density is $p(x, y)$, the mutual information $I(x, y)$ between them is shown in **Eq. (4)**:

$$I(x,y) = \iint p(x,y)\log\frac{p(x,y)}{p(x)p(y)}dxdy \tag{4}$$

The set F formed by features $< x, y >$ is cut using grid G. By changing the position of the dividing point, the mutual information value in each sub-grid is calculated, and the maximum mutual information value of the entire grid G is obtained. The maximum mutual information of F under grid G is shown in **Eq. (5)**:

$$I^*(F,x,y) = \max I(F|G) \tag{5}$$

Where $F|G$ means that the set F is divided by the grid G, and the maximum normalized $I^*(F, x, y)$ values obtained by changing different cutting points form the characteristic matrix $\mathbf{M}(F)_{x,y}$ in **Eq. (6)**:

$$\mathbf{M}(F)_{x,y} = \frac{I^*(F,x,y)}{\log \min\{x,y\}} \tag{6}$$

The maximum information coefficient is shown in **Eq. (7)**:

$$\mathrm{MIC}(F) = \max_{xy < B(n)} \left\{ M(F)_{x,y} \right\} \tag{7}$$

Where the range of MIC(F) is [0,1], and $B(n)$ is the upper limit of the grid size. If $B(n)$ is too large, the data in set F may be gathered in a small part of the sub-grid, and if $B(n)$ is too small, less data can be searched. Generally, the effect is best when $B(n) = n^{0.6}$ (Law et al., 2019), and this value is also used in this paper. The greater the MIC between features and categories, the stronger the correlation, and the greater the impact on the final classification results. The greater the MIC between features, the stronger the substitutability between features, that is, the stronger the redundancy. The process of quantitative analysis of correlation and redundancy is shown as follows:

$$C(F,c) = \frac{1}{|F|} \sum_{x_i \in F} \mathrm{MIC}(x_i, c) \tag{8}$$

$$R(F) = \frac{1}{|F|^2} \sum_{x_i, x_j \in F} \mathrm{MIC}\left(x_i, x_j\right) \tag{9}$$

Where C represents the correlation between the feature set F and the attack category c, and R represents the redundancy between the features in the set F. F and $|F|$ are the feature set and the number of features, x_i represents the ith feature, and c represents the category label. MIC (x_i, c) represents the maximum information coefficient between feature i and the target category, and MIC (x_i, x_j) represents the maximum information coefficient between feature i and feature j.

The optimal attack gene feature set realizes feature selection from the perspective of feature correlation and redundancy. This requires the selected set to meet the conditions of maximum correlation and minimum redundancy. The following criteria are used for comprehensive consideration in **Eq. (10)**:

$$\max \varnothing_1(C,R), \ \varnothing_1 = C - R \tag{10}$$

Assuming that the original feature set is F, the optimal feature subset $F_{(m\text{-}1)}$ of m-1 features has been obtained. The process of selecting the mth feature from the remaining F-$F_{(m\text{-}1)}$ features should be satisfied in **Eq. (11)**:

$$\max_{x_j \in F - F_{m-1}} \left[\mathrm{MIC}\left(x_j, c\right) - \frac{1}{m-1} \sum_{x_i \in F_{m-1}} \mathrm{MIC}\left(x_j, x_i\right) \right] \tag{11}$$

At the same time, with the update and development of terminal equipment, CPPS feature information may gradually increase, and the optimal feature set can still be supplemented according to **Eq. (8)**.

Gene Feature Set Selection Process

The algorithm flow is as follows:

Input: original attack feature set F, category label c.

Output: the optimal attack gene feature set F'.

(1) Discretize the continuous data in the feature set F, and the initial value of the feature set F' is a null value.
(2) Calculate the maximum information coefficient of each feature and category label c, and remove irrelevant and weakly related features.
(3) Find the feature F_i that maximizes **Eq. (8)** in F, add it to the optimal attack feature set F', and delete the feature F_i from F.
(4) Repeat step 3) and continue to select features from the remaining features in feature set F.
(5) Obtain the optimal FDIA gene feature set F'.

FDIA DETECTION MODEL

On the basis of attack gene feature selection, the self-learning ability of a stacked autoencoder (SAE) network is used to further extract abstract genes.

Extracting Attack Genes with Unsupervised Pre-Training Encoder.

An SAE detection model mainly includes two parts: an unsupervised pre-training encoder and a supervised fine-tuning classifier. The detection model uses the cyber-physical optimal attack gene feature set after the above feature selection method as the input layer, and the type of attack on the system is used as the output layer of the neural network. Where the physical layer includes physical factors such as voltage, current, and phase angle, and the cyber layer includes information factors such as communication traffic and alarm logs. The SAE detection model is shown in **Figure 1**.

An autoencoder neural network is constructed. The input layer and output layer of the network are kept as much the same as possible, so that the low-dimensional data of the hidden layer extracted from the intermediate gene can represent the original data. This process does not require labels for supervised training. The layer-by-layer greedy training method is used to pre-train each layer of the neural network, and the network parameters are initialized layer by layer. The cyber and physical features are abstracted layer by layer and coded into low-dimensional data features to reduce the difficulty of model training.

$$L = S\left(w^1 X + b^1\right) \tag{12}$$

$$Z = S\left(w^2 L + b^2\right) \tag{13}$$

In **Eq. (11)** and **Eq. (12)**, L and Z are encoders and decoders, $X = \{X_1, X_2, \ldots\ldots, X_n\}$ is the feature to be encoded, w refers to the weight parameter, b refers to the deviation value, and S refers to the activation function sigmoid:

$$S(x) = \frac{1}{1 + \exp(-x)} \tag{14}$$

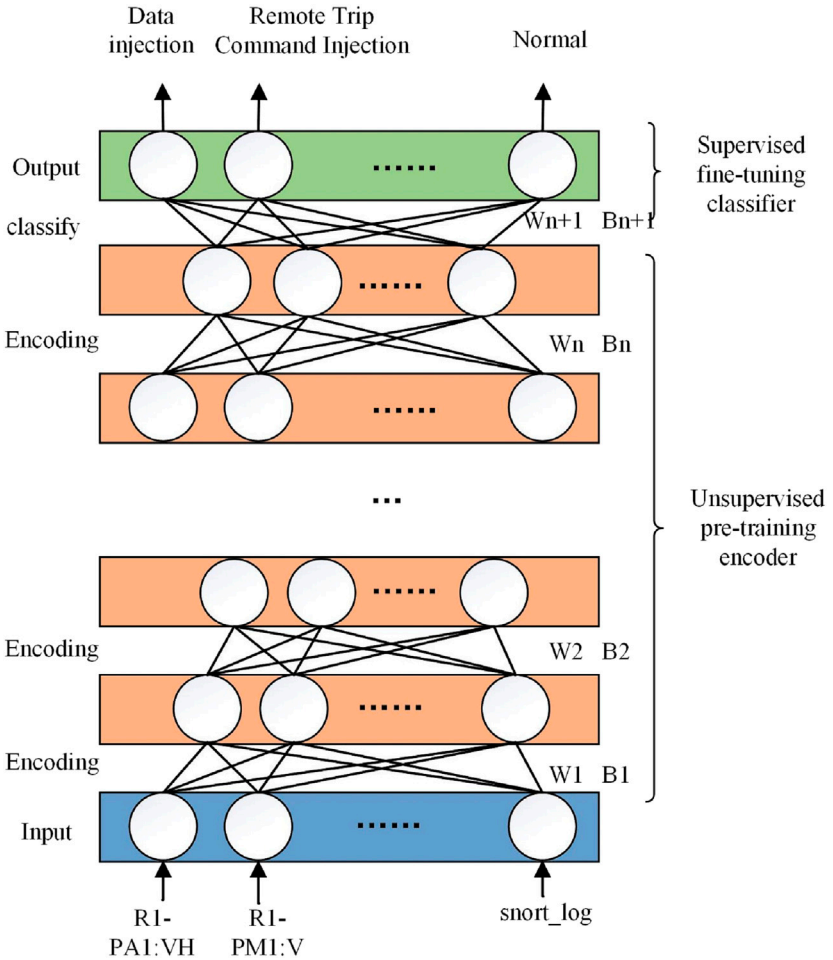


FIGURE 1 | SAE detection model structure.

The loss function is defined as follows:

$$J(w,b) = \frac{1}{n}\left(\sum_{i=1}^{n}(Z - X)^2\right) \tag{15}$$

By minimizing the loss function, the most suitable set of weight parameters (w, b) are obtained to make the neural network optimal. The data are encoded multiple times to obtain the cyber-physical attack gene after dimensionality reduction.

Supervised Fine-Tuning Classifier

The supervised fine-tuning classifier Softmax is constructed for the final attack detection step. The neurons in the output layer are set to N. For N-type CPPS attacks, each neuron represents a type of attack. After the detection is completed, only one neuron of the N output layer neurons has a value of 1, and the remaining N-1 neurons are all 0. The attack mode represented by the neuron with the value of one is the detected attack. The Softmax algorithm is defined in **Eq. (16)**:

$$Y_i = \frac{e^{x_i}}{\sum_{i=1}^{N} e^{x_i}} \tag{16}$$

Where i represents the category index, Y_i represents the probability that the detection result is the ith category, x_i is the input value of the Softmax classifier, and the total number of categories is N.

In order to prevent the parameters from falling into the local optimum, the labeled data set is used for reverse fine-tuning. By comparing the label result with the classification result, the backpropagation algorithm is used to adjust the model parameters. All layer parameters of the entire SAE detection model reach the global optimum as much as possible.

When the SAE detection model adjusts the optimized parameters, the initial parameter setting requirements are higher. The objective function of the initial parameters of the model is expressed in **Eq. (17)**:

$$E = \frac{1}{2}\sum_{k=1}^{n}\left(y'(i) - y(i)\right)^2 \tag{17}$$

Where n is the total number of samples, $y'(i)$ represents the expected output sample, and $y(i)$ represents the actual output sample.

ATTACK DETECTION METHOD BASED ON SACS-SAE.

The initial parameter values of the model are critical to the accuracy of the model (Xing et al., 2019). The initial parameters after unsupervised pre-training have the disadvantages of slow convergence and easily fall into local optimum. Aiming at this problem, a self-adaptive cuckoo search (SACS) algorithm is proposed to optimize the initial parameters of the SAE model. And on this basis, a new method of FDIA detection in the CPPS based on SACS-SAE is obtained.

Model Parameter Optimization

The cuckoo algorithm has the advantages of being a simple model, has few parameters, and strong versatility, which is often used to find the optimal solution of the model (Chitara et al., 2018). However, the step size factor and discovery probability are usually fixed values, which leads to a lack of flexibility and adaptability in the algorithm. A self-adaptive strategy is introduced to optimize this problem.

In the self-adaptive cuckoo algorithm, all bird nests are calculated in each iteration to save the best bird nest storage. The specific algorithm flow is as follows:

Step (1) Randomly generate n initial positions of the bird's nest $[x_1^{(0)}, x_2^{(0)},..., x_n^{(0)}]$. They correspond to the initial weight and threshold parameters of the SAE network model, and the neural network trains the model according to the parameter values. The result is calculated according to **Eqs. (18)Eqs. (19)**.

$$x_i^{(t+1)} = x_i^{(t)} + \alpha \oplus L(\lambda) \tag{18}$$

$$\alpha = \alpha_0\left(x_i^{(t)} - x_b^{(t)}\right) \tag{19}$$

Where x_i is the position of the ith bird's nest in the tth generation, α is the step-length control factor, $\oplus$is the point-to-point multiplication, and $x_b^{(t)}$is the optimal solution of the current generation, α_0 is a fixed value of 0.01, and $L(\lambda)$ is a random search path, obeying Levy distribution:

$$L(\lambda) = \frac{\phi \times \mu}{|\nu|^{1/\beta}} \tag{20}$$

Both μ and ν obey normal distribution; β is 1.5, and the value of ϕ is as follows:

$$\phi = \left\{ \frac{\Gamma(1+\beta) \times \sin\left(\frac{\pi \times \beta}{2}\right)}{\Gamma\left[\left(\frac{1+\beta}{2}\right) \times \beta \times 2^{\frac{\beta-1}{2}}\right]} \right\}^{1/\beta} \tag{21}$$

Step (2) Generally, the larger the step size factor α_0, the stronger the global search ability, but the algorithm convergence accuracy decreases; the smaller the value, the more the optimization accuracy improves, but the convergence speed is slower. In the standard cuckoo algorithm, this value is set to a fixed value, which makes the algorithm's convergence process lack self-adaptability. This paper creates dynamic settings as shown in **Eq. (22)**:

$$\alpha_0 = \exp\left(-\frac{t_i}{t_{\max}}\right) \tag{22}$$

Where t_i represents the current number of iterations, and t_{max} represents the maximum number of iterations. The value of α_0 will gradually decrease with the increase of the number of iterations, to ensure that the step size is gradually reduced, the algorithm satisfies the global search in the early stage, and the optimization accuracy is improved in the later stage.

Step (3) Integrating the above process, the expression of the new individual generated by the self-adaptive cuckoo algorithm is as follows:

$$x_i^{(t+1)} = x_i + \exp\left(-\frac{t_i}{t_{\max}}\right) \times \frac{\phi \times \mu}{|\nu|^{1/\beta}}\left(x_i^{(t)} - x_b^{(t)}\right) \tag{23}$$

Step (4) Calculate all bird's nests in each iteration, and save the best bird's nest position $x_d^{(0)}$.

Step (5) After obtaining a new generation of bird's nest positions, use better bird's nest positions to replace the poorer bird's nest positions of the previous generation to obtain a set of better bird's nest positions $e_k = [x_1^{(k)}, x_2^{(k)},..., x_n(k)]$.

Step (6) Generate a random number *rand* in the range of [0,1]. If *rand* $>p_a$, then discard some of the solutions, and use the preferred random walk to generate the same number of new solutions. As shown in **Eq. (24)**:

$$x_i^{(t+1)} = x_i^{(t)} + rand\left(x_j^{(t)} - x_k^{(t)}\right) \tag{24}$$

Where $x_j^{(t)}$ and $x_k^{(t)}$ represent two random solutions of the t-th generation.

Step (7) Discovery probability p_a is generally a fixed value of 0.25, which determines whether the current solution is retained. In order to prevent the algorithm from falling into the local optimum, the cuckoo algorithm is further improved: the discovery probability p_a is dynamically set, and it gradually increases as the search progresses. In the later stage of the evolution, the global search and local search in the algorithm can be balanced. The convergence accuracy of the algorithm is improved as a whole to avoid falling into the local optimum. As shown in **Eq. (25)**:

$$p_a = p_{a,\max} \times \left(\frac{t_i}{t_{\max}}\right)^3 \tag{25}$$

Where p_a represents the discovery probability of the bird's nest, $p_{a,max}$ represents the maximum discovery probability, t_i represents the current iteration number, and t_{max} represents the maximum iteration number.

Step (8) After obtaining a new set of bird's nest positions, according to the objective function, the bird's nest position with better performance is replaced with the bird's nest position with poor performance in e_k. So far we get the latest set of bird's nest positions $Q_k = [x_1^{(k)}, x_2^{(k)}, \ldots, x_n(k)]$.

Step (9) Find the best bird's nest position $x_d^{(k)}$ in Q_k, if the maximum number of iterations is not reached, return to step 4) to continue searching and optimizing, otherwise output the best position $x_d^{(k)}$.

Step (10) According to the value corresponding to the optimal bird's nest position $x_d^{(k)}$ as the initial parameters of the model, the forward training and reverse adjustment of the model are performed.

SACS-SAE Attack Detection Steps

The SACS is employed for solving the objective function in order to improve the convergence accuracy. So far, the attack detection steps based on the SACS-SAE detection method are as follows.

Step (1)There are non-numerical and infinite values in the data, which seriously affect model training. The entire data containing non-numerical and infinite values are deleted. The detection problem is a multi-classification problem, so the category attribute should be converted to one-hot encoding form. For example, the type of event 1 can be converted to (1, 0, 0,..., 0), and event 41 can be converted to (0, 0,..., 0, 1).

There is a large gap in the value of different features in the original data, which is prone to large errors. The original data are normalized, and the characteristic data are mapped to the range of [−1, 1] according to **Eq. (26)**. Where x_i is the attribute value of a certain column, x_{min} is the minimum value of the column attribute, and x_{max} is the maximum value of the column attribute.

$$x_{new} = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \tag{26}$$

Step (2) After obtaining the standardized data set, use the maximum information coefficient to calculate the correlation between the feature and the category, and the redundancy between the feature and the category, and screen the optimal cyber-physical attack gene feature set based on the maximum mutual information coefficient.

Step (3) Construct an unsupervised pre-training encoder to encode the original data and reduce the dimensionality. The greedy training method is used to train each layer of the network layer by layer, initialize the model parameters, and deeply extract the abstract genes of the input gene feature set to reduce the difficulty of model training.

Step (4) Construct a supervised fine-tuning classifier, followed by a Softmax classifier after the coding layer. The network weights and thresholds initialized by unsupervised pre-training are utilized as the position vector of the bird's nest, and the objective function of the model is utilized as the fitness function of the bird's nest, and the optimal solution is continuously iterated.

Step (5) After initializing the network parameters using the adaptive cuckoo algorithm, on this basis, reverse adjustment and optimization is used to train the neural network parameter weights. CPPS data injection attacks are detected, and the operation and maintenance personnel will take corresponding actions based on the detection results.

CASE STUDIES

The ICS data set is used as a test of the CPPS cyber-attacks detection system in this paper. The ICS data set is the analog CPPS data provided by Mississippi State University (Adhikari et al., 2016). The three-bus and two-wire transmission system used in the data set is modified from IEEE three machines and nine nodes. The system frame configuration is shown in **Figure 2**. Various intelligent electronic devices interact, supplemented by network monitoring equipment such as the SNORT and Syslog systems. G1 and G2 are generators, and PDC refers to the phasor data concentrator, which is responsible for storing and displaying synchronized data and recording historical data of the system. R1 to R4 are intelligent electronic devices (IEDs) that can open or close the circuit breakers (BR1 to BR4), and each IED corresponds to a circuit breaker. The IED is used with a distance protection scheme, which makes the circuit breaker trip automatically when a fault is detected, but it cannot be verified whether it is caused by an actual fault or an information attack. At the same time, the operation and maintenance personnel can also manually issue commands to the IED to trip the circuit breakers BR1 to BR4 during line maintenance.

There are 15 sets of data in the ICS data set, each of which contains about 5000 pieces of information, divided into training set and test set according to the ratio of 8: 2. The measurement data mainly includes PMU, snort log information, control panel, and relay records, etc. It mainly includes six scene types:

(1) Normal: the power system is operating normally and no abnormality occurs.
(2) Short-circuit fault: a short-circuit fault occurs on the line, and the fault may occur in different positions of the line.
(3) Line maintenance: operation and maintenance personnel manually issue a trip command, and usually disable one or more IEDs.
(4) Remote trip command injection attack: command injection attack on single relay (R1-R4) or double relay (R1 and R2, R3 and R4), and send a trip command to the IED.
(5) Relay setting change attack: by disabling the function of IED R1/R2/R3/R4, it cannot be tripped due to faults or commands, causing faults at different locations on the transmission line.
(6) Data injection attack: simulate this situation by modifying the values of parameters such as voltage, current, and sequence components, resulting in different positions of L1 or L2 (10%–19%, 20%–79%, 80%–90%) A failure occurs.

The proportions of data in the six scenarios are 1: 3: 1: 2: 9: 2, which are divided into 37 event types (not including 31–34) according to the location where the fault occurs. The details are shown in **Table 1**.

Evaluation Criteria

In the process of CPPS attack detection, evaluation indicators such as training time, test time, accuracy (*AC*), and precision (*P*) are used to measure the performance of the model. Where the definition of accuracy and precision are shown in **Eqs. (27) Eqs. (28)**.

$$AC = \frac{TP + TN}{TP + TN + FP + FN} \tag{27}$$

$$P = \frac{TP}{TP + FP} \tag{28}$$

Where *TP* (true positive) is the number of records that correctly identify the normal situation, *TN* (true negative) is the number of

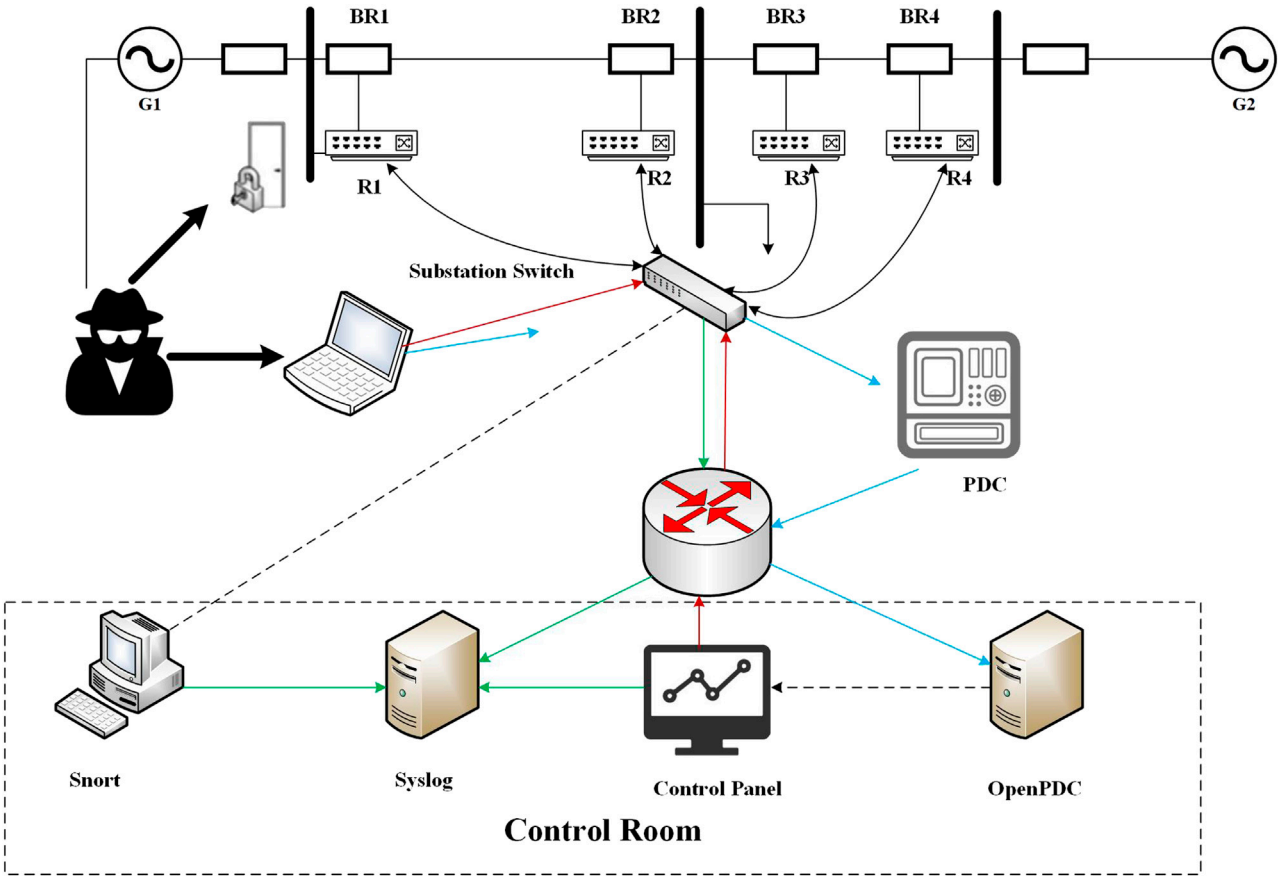


FIGURE 2 | ICS structure.

TABLE 1 | Scene event description.

Scenario	Attack type
41	Normal
1–6	Short circuit fault
13,14	Line maintenance
15–20	Remote trip command injection attack
21–30, 35–40	Relay setting change attack
7–12	Data injection attack

TABLE 2 | SAE parameters.

Parameter	Value
Learning_rate	0.01
Training_epochs	1000
Batch_size	256
Display_step	50
Activation function	Sigmoid
Optimizer	AdamOptimizer

records that correctly identify the attack, *FP* (false positive) is the number of records that incorrectly identify the normal situation, *FN* (false negative) is the number of records that incorrectly identify the attack, training time is represented by T_{train}, and test time is represented by T_{test}.

Simulation Environment

This simulation is carried out in an Intel CPU 2.29 GHz, 128G memory, 64-bit hardware environment and Windows 10 workstation environment, and is implemented using the Python language and TensorFlow framework. Some initial parameter settings of SAE are shown in **Table 2**.

Case Analysis

In order to verify the effectiveness of the proposed method of the FDIA detection in the CPPS, this paper designs two cases:

Case 1: The influence of different network parameters is analyzed on the accuracy of the case.

The key factors such as SACS initialization parameters, the detection effectiveness of the selected optimal gene feature set, and the number of hidden layers of the gene extraction network are discussed. Under the condition of other parameters remaining unchanged, a certain parameter is analyzed.

Case 2: The machine learning algorithm and the SACS-SAE detection method are analyzed for differences in detection

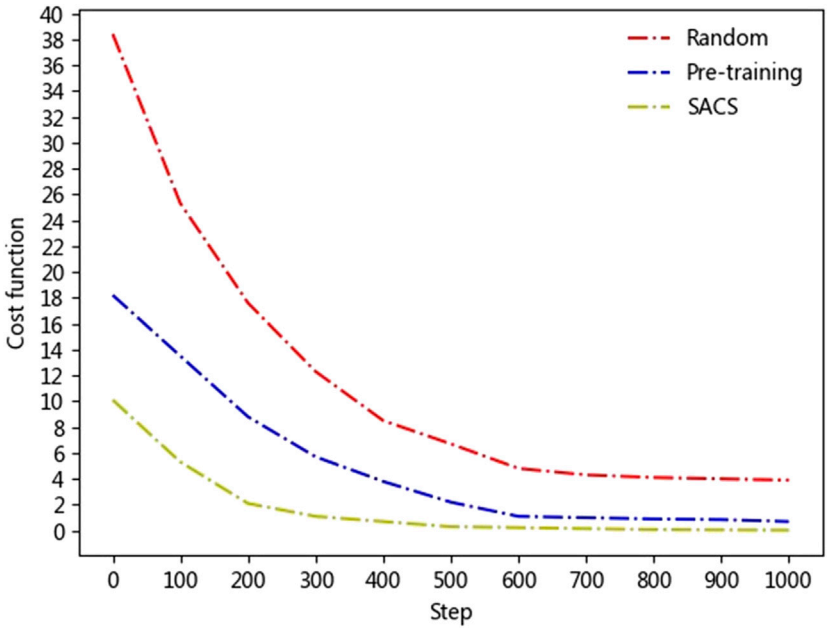


FIGURE 3 | Comparison of different initialization parameters.

accuracy and detection speed. KNN, SVM, random forest, decision tree, XGBOOST and other algorithms, and the SACS-SAE detection method proposed in this paper are used for comparative analysis.

The data is divided into 15 groups. For each group of data, it is divided into a training set and a test set at a ratio of 9: 1. At the same time, a 10-fold cross-validation method is used to optimize the model parameters as much as possible. The simulation is carried out on each data set, and then the detection accuracy of the 15 sets of data is averaged as the final result.

(1) Influence analysis of model parameters

(1) The influence of parameter initialization on detection results

The number of hidden layers of the gene extraction network is set to 3, and all feature parameters are used for training to analyze the degree of influence of different initialization parameters on the model. The result is shown in **Figure 3**.

The results show that the loss function and neural network convergence speed are better than the other two methods after SACS initialization parameters. The main reason is that the FDIA detection process is a non-convex optimization problem, and there are many local optimal solutions. At the same time, in the deep network BP algorithm optimization process, the error gradually decreases, and the problem of "gradient diffusion" is prone to occur. After pre-training initialization, the parameters can be initialized to a suitable place, but it is still difficult to achieve the global optimum. The initial parameters obtained by SACS have been fine-tuned to converge the training process to a more ideal state, and the model training speed has also been improved.

(2) The influence of optimal attack gene feature set selection on detection results

The maximum correlation minimum redundancy algorithm is improved to determine the cyber-physical optimal attack gene feature set, and the feature correlation ranking is shown in **Figure 4**.

In the process of selecting the optimal attack gene feature, about 75% of the 128 gene features can provide high learning value, and the two features with the highest correlation are R1 and R2's A phase voltage phase angle, respectively. The features are selected according to the above content, and the optimal feature subset is determined to include 56 features such as A-C phase voltage, phase angle, phase current, and snort log information. In order to verify the effectiveness of this method, three different model structures are set up: Model 1 (the first 32 features in relevance ranking), Model 2 (the optimal feature selection is 56 features), and model 3 (the first 96 features in relevance ranking). The number of hidden layers of the network is set to 3, and the adaptive cuckoo algorithm is used to initialize the parameters. The analysis result is shown in **Figure 5**.

The results show that the accuracy of the model trained with the first 32 gene features is significantly lower than the other two models. The reason is that there are relatively few selected features and some effective information is missing. The accuracy of Model two is slightly higher than Model 3, and its training time is shorter than Model 3. The reason is that Model 3 has a large gene feature dimension, and there are some redundant features and weakly related features. The model is affected by these features to a certain degree, which increases the complexity of the model training process, resulting in a slight decrease in recognition accuracy. When the optimal attack gene feature set

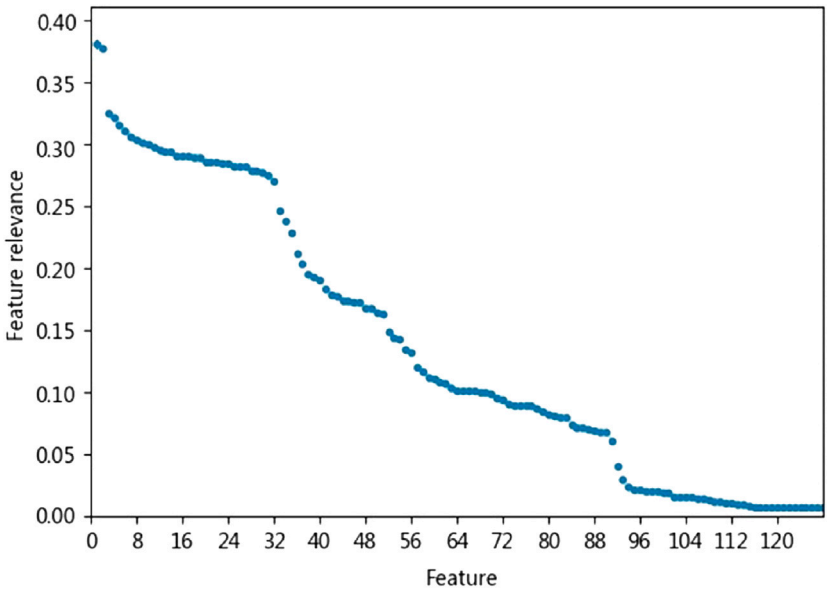


FIGURE 4 | Attack gene feature correlation ranking.

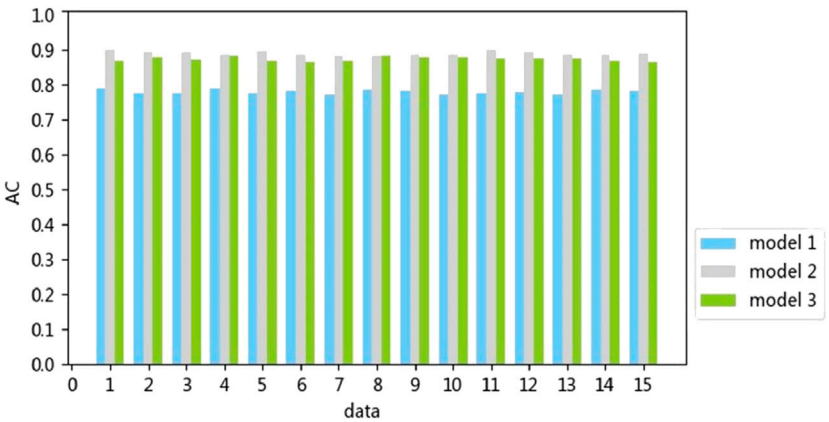


FIGURE 5 | Comparison of different attack gene feature numbers.

proposed in this paper is selected, a relatively balanced state can be achieved between the model detection accuracy and the training time. At the same time, the optimal cyber-physical attack gene feature selection method proposed in this paper has been verified not only to improve the accuracy of model detection, but also to accelerate the effectiveness of model training.

(1) The influence of the number of hidden layers of gene extraction on the detection result

The number of hidden layers of the network in deep learning is critical to the abstract gene extraction results. On the premise that the initialization parameters and the optimal gene feature set are determined, five different hidden layer experiments are set up. The most suitable one is selected as the depth of the gene extraction network for the FDIA detection model. The comparative analysis result is shown in **Figure 6**.

As shown in **Figure 7**, the training time increases greatly as the number of hidden layers increases, while the test time is maintained at about 1 s to complete. As the number of hidden layers increases, the hidden abstract genes in the data are deeply extracted, but at the same time, the training time in the deep learning model increases significantly, which can easily lead to overfitting. When the detection model is set with four hidden layers, the accuracy rate is the largest, and the average accuracy rate reaches 93.7%.

(2) Comparison and analysis with traditional detection methods

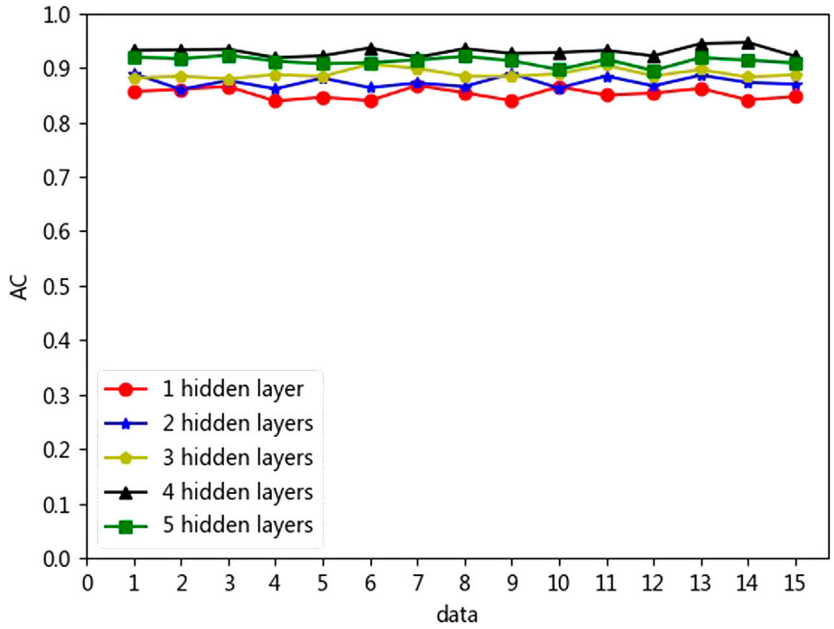


FIGURE 6 | Detection results with different hidden layers of attack gene extraction.

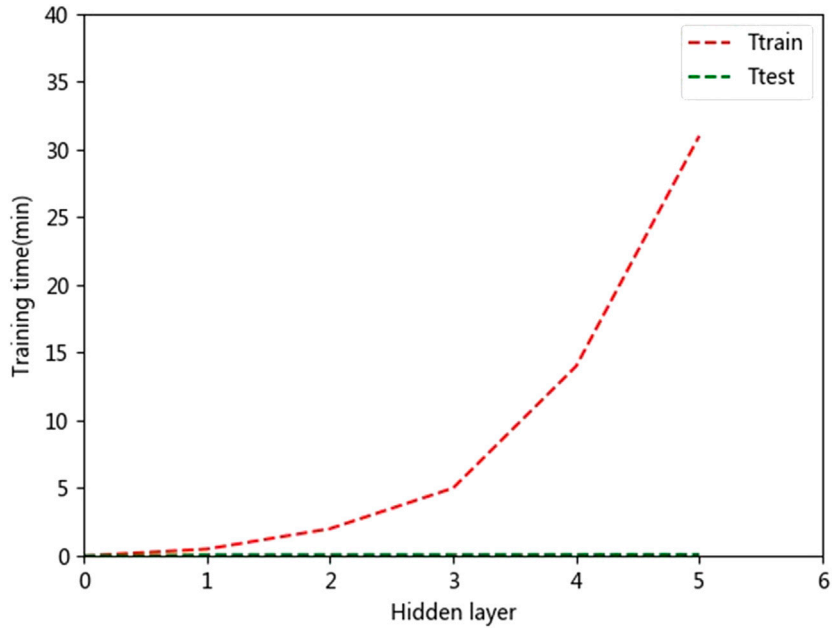


FIGURE 7 | Model training time with different hidden layers of gene extraction.

Six detection methods including KNN, SVM, random forest, decision tree, XGBOOST, and SACS-SAE are selected for comparative analysis. For other algorithms, the optimal results are selected for comparison after multiple parameter adjustments. As shown in **Figure 8**, the results show that in terms of accuracy, the SACS-SAE detection method with four hidden layers has better performance than the other traditional machine learning methods. The reason is that the stack auto-encoding network has a strong expression ability, combined with the global optimization ability of the SACS. The recognition accuracy has been significantly improved compared with other algorithms.

As shown in **Figure 9**, after the algorithm with the maximum information coefficient is improved for genetic feature selection, the recognition accuracy and training time are better than the

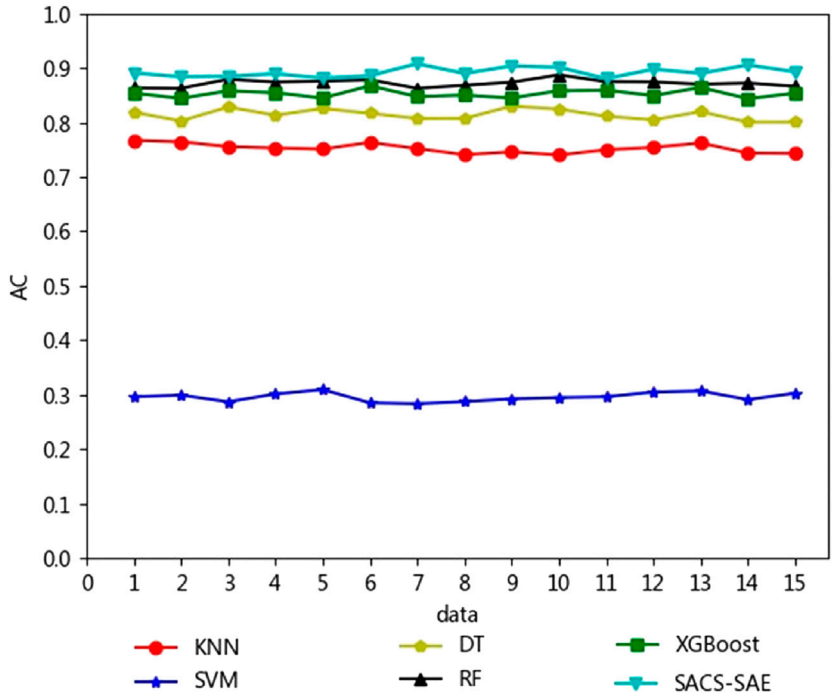


FIGURE 8 | Detection results with different detection methods.

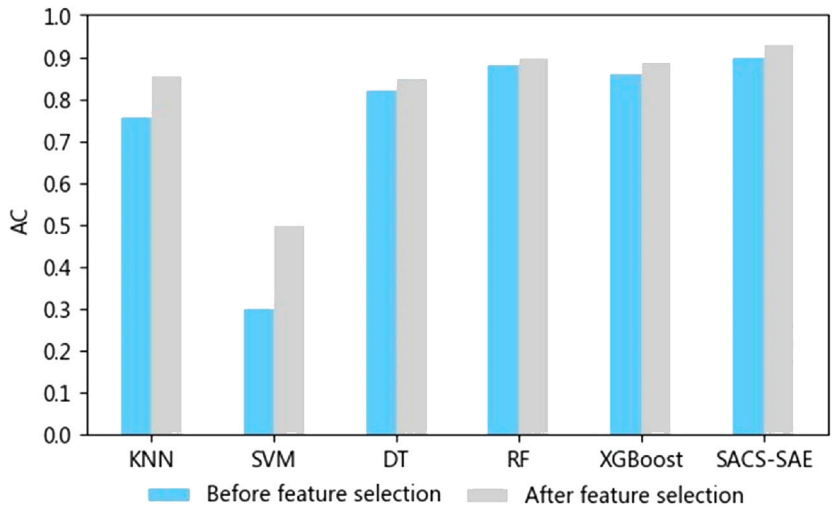


FIGURE 9 | The influence of attack gene feature selection with different methods.

original feature selection to a certain extent, which proves the effectiveness and universality of the method. However, in terms of training model time, the KNN algorithm has the fastest training speed, and the detection model proposed in this paper has a relatively large number of hidden layers and a complex network, and the model training speed is relatively slow. The results of AC, P, $T_{train,}$ and T_{test} of the attack detection algorithm before and after gene feature selection are shown in **Table 3**.

CONCLUSION

From a data-driven perspective, a bio-inspired concept of genes is introduced to provide a new idea for the detection of FDIAs in the CPPS.

(1) The maximum information coefficient is used to analyze the correlation and redundancy in the features to determine the

TABLE 3 | Comparison of performance with different models.

Detection method		*AC* (%)	*P* (%)	T_{train} (/min)	T_{test} (/s)Ttest
KNN	Before	75.8	74.3	0.3	1.2
	After	85.8	86.3	0.1	0.8
SVM	Before	32.6	30.8	17	3.8
	After	51.3	52.4	8	2
DT	Before	82.4	81.1	5	1.5
	After	82.7	81.5	2	1.3
RF	Before	88.1	86.4	7	2.1
	After	89.6	87.2	6	1.7
XGBoost	Before	89.3	87.8	2	1.1
	After	90.4	91.3	0.3	0.8
SACS-SAE	Before	89.3	88.2	18	1.2
	After	93.7	93.2	14	0.8

optimal attack gene feature set. The analysis of case studies shows that the selection of gene features will affect the results of attack detection. The attack gene feature selection method in this paper achieves a good balance in terms of detection accuracy and training time.

(2) A stack auto-encoding network is constructed for the abstract extraction of genes, and a method for detecting FDIAs based on SACS-SAE is proposed. The analysis of case studies shows that the proposed method is suitable for CPPS attack detection in high-dimensional space, and the detection accuracy and precision rate are both over 90%, and the identification accuracy is better than traditional methods. The amount of hidden layers is not as high as it could be. In the attack identification process, a suitable number of hidden layers needs to be selected.

(3) The loss function and the training convergence speed perform better after SACS initialization parameters.

However, the structure of the deep network model is complex and the model training time is long. In order to speed up the model training process, the dynamic optimization method of the learning rate and other parameters in the training process will be considered in future work.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

ZQ: data curation, writing and software. YD: supervision, writing, reviewing and editing. NQ: conceptualization and methodology. HL: project administration. MC: software and formal analysis. XB: reviewing and editing. YW: validation and editing. SM: software and editing. All the co-authors reviewed and contributed to the revision of the manuscript

FUNDING

This work is supported by the State Key Program of National Natural Science Foundation of China (No. 51437003), the Science and Technology Development Plan of Jilin Province (No. 20200401097GX), and the Science and Technology Development Plan of Jilin Province (No. 20180201092GX).

REFERENCES

Adhikari, U., Morris, T. H., and Pan, S. (2017). Applying hoeffding adaptive trees for real-time cyber-power event and intrusion classification. *IEEE Trans. Smart Grid* 9 (5), 4049–4060. doi:10.1109/TSG.2017.2647778

Adhikari, U., Morris, T. H., and Pan, S. (2016). Applying non-nested generalized exemplars classification for cyber-power event and intrusion detection. *IEEE Trans. Smart Grid* 9 (5), 3928–3941. doi:10.1109/TSG.2016.2642787

Chitara, D., Niazi, K. R., Swarnkar, A., and Gupta, N. (2018). Cuckoo search optimization algorithm for designing of a multimachine power system stabilizer. *IEEE Trans. Ind. Applicat.* 54 (4), 3056–3065. doi:10.1109/tia.2018.2811725

Dai, J., Li, Y., He, K., and Sun, J. (2016). R-fcn: Object detection via region-based fully convolutional networks. Pro. NIPS. *arXiv preprint. arXiv:1605.06409*, 379–387.

Gao, L., Chen, B., and Yu, L. (2019). Fusion-based FDI attack detection in cyber-physical systems. *IEEE Trans. Circuits syst. II: express briefs* 67 (8), 1487–1491. doi:10.1109/TCSII.2019.2939276

Hagai, T., Chen, X., Miragaia, R. J., Rostom, R., Gomes, T., Kunowska, N., et al. (2018). Gene expression variability across cells and species shapes innate immunity. *Nature* 563 (7730), 197–202. doi:10.1038/s41586-018-0657-2

He, R., Xie, H., Deng, J., Feng, T., Lai, L. L., and Shahidehpour, M. (2020). Reliability modeling and assessment of cyber space in cyber-physical power systems. *IEEE Trans. Smart Grid* 11, 3763–3773. doi:10.1109/TSG.2020.2982566

Law, R., Li, G., Fong, D. K. C., and Han, X. (2019). Tourism demand forecasting: a deep learning approach. *Ann. Tourism Res.* 75, 410–423. doi:10.1016/j.annals.2019.01.014

Li, Y., and Zeng, J. (2019). Improved convolutional neural network based detection method for false data injection attack on power grid. *Automation Electric Power Syst.* 43 (20), 97–104. doi:10.7500/AEPS20180919001

Li, Y., Li, Z., and Chen, L. (2019). Dynamic state estimation of generators under cyber attacks. *IEEE Access* 7, 125253–125267. doi:10.1109/access.2019.2939055

Li, Y., Wang, C., Li, G., Wang, J., Zhao, D., and Chen, C. (2020). Improving operational flexibility of integrated energy system with uncertain renewable generations considering thermal inertia of buildings. *Energ. Convers. Management* 207, 112526. doi:10.1016/j.enconman.2020.112526

Li, Y., Yang, Z., Li, G., Zhao, D., and Tian, W. (2018). Optimal scheduling of an isolated microgrid with battery storage considering load and renewable generation uncertainties. *IEEE Trans. Ind. Electronics* 66 (2), 1565–1575. doi:10.1109/TIE.2018.2840498

Liang, G., Zhao, J., Luo, F., Weller, S. R., and Dong, Z. Y. (2016). A review of false data injection attacks against modern power systems. *IEEE Trans. Smart Grid* 8 (4), 1630–1638. doi:10.1109/TSG.2015.2495133

Liu, X., Bao, Z., Lu, D., and Li, Z. (2015). Modeling of local false data injection attacks with reduced network information. *IEEE Trans. Smart Grid* 6 (4), 1686–1696. doi:10.1109/tsg.2015.2394358

Liu, Y., Ning, P., and Reiter, M. K. (2011). False data injection attacks against state estimation in electric power grids. *ACM Trans. Inf. Syst. Secur.* 14 (1), 1–33. doi:10.1145/1952982.1952995

Pan, S., Morris, T., and Adhikari, U. (2015). Classification of disturbances and cyber-attacks in power systems using heterogeneous time-synchronized data. *IEEE Trans. Ind. Inf.* 11 (3), 650–662. doi:10.1109/tii.2015.2420951

Qu, Z., Dong, Y., Liu, S., Shen, Z., Yu, J., and Li, Y. (2020). Bioimmunological method based security technology of ubiquitous power internet of things. *Automation Electric Power Syst.* 44 (2), 1–12. doi:10.19718/j.issn.1005-2992.2019-03-0038-06

Qu, Z., Dong, Y., Qu, N., Wang, L., Li, Y., Zhang, Y., et al. (2019). Survivability evaluation method for cascading failure of electric cyber physical system considering load optimal allocation. *Math. Probl. Eng.* 2019, 1–15. doi:10.1155/2019/2817586

Qu, Z., Zhang, Y., Qu, N., Wang, L., Li, Y., and Dong, Y. (2018). Method for quantitative estimation of the risk propagation threshold in electric power CPS based on seepage probability. *IEEE Access* 6, 68813–68823. doi:10.1109/access.2018.2879488

Wang, D., Wang, X., Zhang, Y., and Jin, L. (2019). Detection of power grid disturbances and cyber-attacks based on machine learning. *J. Inf. Security Appl.* 46, 42–52. doi:10.1016/j.jisa.2019.02.008

Xing, X., Sun, Q., Zhang, P., and Li, M. (2019). Research on distribution network fault recovery and reconstruction based on deep-first search and colony algorithms. *J. Northeast Electric Power Univ.* 39 (3), 38–43. doi:10.19718/j.issn.1005-2992.2019-03-0038-06

Xu, L., Guo, Q., Yang, T., and Sun, H. (2018). Robust routing optimization for smart grids considering cyber-physical interdependence. *IEEE Trans. Smart Grid* 10 (5), 5620–5629. doi:10.1109/TSG.2018.2888629

Yang, F., Xu, Y., Qian, Y., Li, Z., Sheng, G., and Jiang, X. (2018). Application of correlation analysis techniques in feature extraction and selection for DC partial discharge signals of XLPE cables. *Power Syst. Technol.* 42 (5), 1653–1660. doi:10.13335/j.1000-3673.pst.2017.1866

Yang, M., Wang, J., and Du, J. (2019). The complement of the missing data based on the extreme learning machine and granger test in wind power. *J. Northeast Electric Power Univ.* 39 (5), 9–16. doi:10.19718/j.issn.1005-2992.2019-05-0009-08

Yu, X., and Xue, Y. (2016). Smart grids: A cyber-physical systems perspective. *Proc. IEEE* 104 (5), 1058–1070. doi:10.1109/jproc.2015.2503119

Conflict of Interest: NQ was employed by the company State Grid Jiangsu Electric Power Company Limited, China. HL was employed by the company State Grid Jilin Electric Power Company Limited, China. MC was employed by the company State Grid Inner Mongolia Eastern Electric Power Company Limited, China. YW was employed by the company Zhejiang Windey Company Limited, China.

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

frontiers in Energy Research

ORIGINAL RESEARCH
published: 29 March 2021
doi: 10.3389/fenrg.2021.651807

A Reconstruction Method for Missing Data in Power System Measurement Based on LSGAN

Changgang Wang[1,2], *Yu Cao*[2]*, *Shi Zhang*[2] *and Tong Ling*[2]

[1]Key Laboratory of Modern Power System Simulation and Control and Renewable Energy Technology, Ministry of Education (Northeast Electric Power University), Jilin, China, [2]School of Electrical Engineering, Northeast Electric Power University, Jilin, China

The integrity of data is an essential basis for analyzing power system operating status based on data. Improper handling of measurement sampling, information transmission, and data storage can lead to data loss, thus destroying the data integrity and hindering data mining. Traditional data imputation methods are suitable for low-latitude, low-missing-rate scenarios. In high-latitude, high-missing-rate scenarios, the applicability of traditional methods is in doubt. This paper proposes a reconstruction method for missing data in power system measurement based on LSGAN (Least Squares Generative Adversarial Networks). The method is designed to train in an unsupervized learning mode, enabling the neural network to automatically learn measurement data, power distribution patterns, and other complex correlations that are difficult to model explicitly. It then optimizes the generator parameters using the constraint relations implied by true sample data, enabling the trained Generator to generate highly accurate data to reconstruct the missing data. The proposed approach is entirely data-driven and does not involve mechanistic modeling. It can still reconstruct the missing data in the case of high latitude and high loss rate. We test the effectiveness of the proposed method by comparing three other GAN derivation methods in our experiments. The experimental results show that the proposed method is feasible and effective, and the accuracy of the reconstructed data is higher while taking into account the computational efficiency.

Keywords: generative adversarial networks, least-squares, measurement data, missing data, reconstruct data

OPEN ACCESS

Edited by:
Liang Chen,
Nanjing University of Information Science and Technology, China

Reviewed by:
Shaoyan Li,
North China Electric Power University, China
Jun Yin,
North China University of Water Conservancy and Electric Power, China

***Correspondence:**
Yu Cao
ycao@neepu.edu.cn

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

Received: *11 January 2021*
Accepted: *19 February 2021*
Published: *29 March 2021*

Citation:
Wang C, Cao Y, Zhang S and Ling T (2021) A Reconstruction Method for Missing Data in Power System Measurement Based on LSGAN. Front. Energy Res. 9:651807. doi: 10.3389/fenrg.2021.651807

INTRODUCTION

As the power grid-scale continues to grow, especially with renewable energy generation's accession, the power system operation's uncertainty has increased dramatically. The above situation brings unprecedented challenges to ensure the power system's security and economic operation (Li et al., 2019). In recent years, with the flourishing development of supervisory control and data acquisition systems, as well as the increasing maturity of technologies such as big data and deep learning, power security situation prediction has gradually formed new security warning modes based on data-driven to grasp, control and predict the operation status of the power system, which is different from the traditional modeling and presupposing working conditions. It shows the significant value of data for secure power system operation.

The reliability of measurement data directly affects the conclusions from the data-based analysis of the power system operation behavior. Only conclusions based on reliable data analysis can reflect

the system operation's true status (Wang et al., 2020). However, in practice, the supervisory and data acquisition system, due to the data acquisition process, measurement process, transmission modes, storage modes, and other segments, may break down or suffer interference, which will lead to lost or missing data (Jing et al., 2018). To grasp, control, and predict the power system's operation status based on data-driven, the primary problem we need to solve is reconstructing the missing data.

State estimation, a fundamental technology for advanced applications in energy management systems, has made a remarkable contribution to grid data estimation (Ho et al., 2020). On the premise that there are a few missing data and the system has complete observability, we can treat the missing data as data to be estimated and then apply state estimation to estimate the missing data's concrete values (Miranda et al., 2012). Nevertheless, applying state estimation has two major prerequisites: the system needs to meet complete observability, full parameter information (network topology and line parameters). In general, to meet the system's complete observability, the measurement system provides many redundant data. In the case of a high rate of missing data, the state estimation cannot accomplish the task of estimating missing data when the system's complete observability requirement cannot be satisfied.

Traditionally, the methods for reconstructing missing data are mainly based on the filling method, which can be subdivided into the data filling method based on statistical analysis and the data filling method based on machine learning from the methodological perspective. The former is more common, such as regression Imputation, mean Imputation, and hot-deck Imputation are widely used in practice. The principle is to give reasonable reconstruction values through statistical analysis to reduce the calculation bias caused by missing data (Kallin Westin, 2004). The latter mostly uses supervised learning, semi-supervized learning, and unsupervized learning to achieve the effective reconstruction of missing values (Comerford et al., 2015; Sun et al., 2018; Li et al., 2020). Data reconstruction methods based on statistical analysis are simple and efficient, but reconstructed data accuracy is weak. Although the data reconstruction method based on machine learning class has high accuracy, it requires corresponding multiple mechanism modeling when dealing with multiple missing data, and its practicality is doubtful in the case of high latitude and high missing rate.

The correlations, data distribution characteristics, and data change patterns existing among the power system measurement data can be used as an auxiliary basis for reconstructing the missing data, which can greatly enrich the data's information. The defect of traditional methods is that they do not rationalize the application of such information. The birth of GAN (Generative Adversarial Network) has solved this problem. Initially, GAN made breakthroughs in image inpainting and high-resolution graphic reconstruction (Wang et al., 2017). Indeed, restoring the missing part of the image and reconstructing the missing data of power system measurements belong to the same problem. Both of them generate the missing part following the objective law considering the assigned partial constraints (Dong et al., 2019). In the image restoration problem, GAN can automatically learn the complex distribution pattern among data through the training of neural network in an unsupervized form, and then generate the data to meet the objective law, solving the problem of high data latitude and complex modeling (Wu et al., 2017).

GAN has attracted scholars from home and abroad, and many studies have been conducted. J. Lan et al. have proposed a CGAN (Conditional Generative Adversarial Networks) model with the inclusion of classification label information to enrich the original true and false binary classification into a multi-type determination. The introduced label information can be used as an additional criterion to verify the generation results and contribute to the correction of the generation results (Lan et al., 2018). A. Borgia et al. have applied GAN to generate pedestrian walking postures and Interpolate the video to enrich the video information, thus improve the accuracy of pedestrian recognition (Borgia et al., 2019). C. Ledig et al. proposed the SRGAN (Super-Resolution Generative Adversarial Networks) model to accomplish the task of improving the image resolution (Ledig et al., 2017). In the domain of missing image restoration, M. Wang et al. applied GAN to reconstruct the obscured part of the face in the image to enrich the training sample, thus improving the accuracy of recognizing facial expressions (Wang et al., 2019a). R. A. Yeh et al. applied a deep generative model to the image reconstruction problem to guarantee that the image realism constraint is satisfied during reconstruction (Yeh et al., 2017). To solve the problem of gradient disappearance and dispersion during GAN training, S. Wang et al. replaced the original objective function to train GAN with minimized Wasserstein distance as the objective function, which improved the stability of training. However, applying WGAN makes computational efficiency significantly decreased (Wang et al., 2019b).

In summary, data restoration methods based on statistical analysis are relatively simple and straightforward but not very practical in the case of high dimensionality and high loss rates. The state-estimation-based data restoration method is limited by the conditions required for mechanism modeling, and the available premise is that absolute preconditions must be provided. The GAN-based data restoration method solves the former deficiency to some extent. It overcomes the limitations of conditions for the method and can still reconstruct data and restore data in high-dimensional and high-lost rate cases. However, the original GAN may suffer from gradient disappearance and dispersion during training due to the loss function's limitation. The improved WGAN, a GAN derivative method, solves gradient disappearance during training by modifying the loss function. Nevertheless, the consequent computational burden makes the training efficiency drop significantly. It is worth investigating how to find a generative adversarial network that can overcome the gradient disappearance and consider computational efficiency.

In this paper, we propose to apply LSGAN (Least Squares Generative Adversarial Networks) to the problem of reconstructing missing data from power system measurements. The proposed

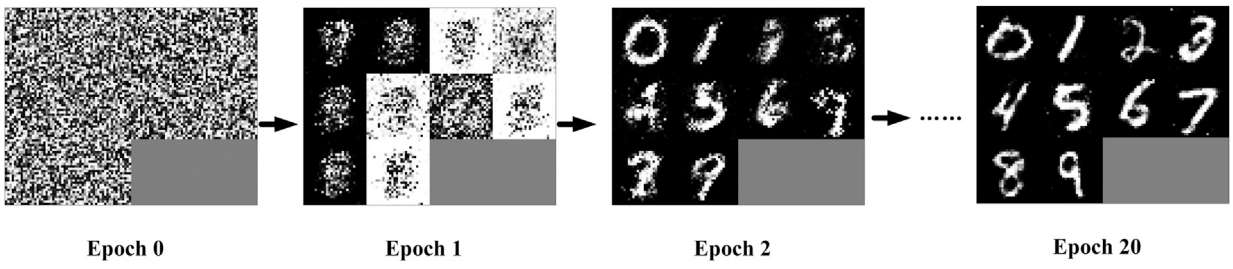


FIGURE 1 | The schematic diagram for the effects of GAN in generating images within the iterative training process.

method learns the data's objective distribution pattern to generate highly accurate reconstructed data that conforms to the inter-data complex pattern. Unlike other GANs, LSGAN replaces the cross-entropy loss function with the least-squares loss function when applying GAN in reconstructing missing data. We use this different distance metric from the traditional one to build an adversarial network with more stable training, faster computational convergence, and higher quality in the generated data. It solves unstable training due to gradient disappearance and diffusion and the low computational efficiency of traditional GANs. The experimental results show that comparing with GAN, CGAN, and WGAN methods, the data generated by LSGAN can still guarantee high accuracy in the case of multiple data missing, which provides a good data basis for applying large volume data to analyze the power system operation behavior.

RELATED WORK

Generating Adversarial Networks

GAN is an unsupervized learning model. It was first proposed by Ian J. Goodfellow and other researchers in 2014 (Goodfellow et al., 2014a). GAN was mainly used to generate images in unsupervized training mode in the beginning. It shows the effect of generating handwritten digital images by GAN training with the MNIST training dataset as the sample in **Figure 1**. We can see that evolution from the initial noise-filled image to a cleaned handwritten digital image.

The model embodies the idea of a "zero-sum game": for the two participants in the game, under tough competition, if one gains, it must mean that the other loses. The sum of gains and losses for both participants is always "zero," and there is no possibility of cooperation between them. With this concept of the non-cooperative game, GAN is composed of Generator and Discriminator.

Generator is a neural network used to learn the distribution pattern of data within a sample and generate new sample data to meet the pattern accordingly. The technical route: the Generator trains an arbitrarily distributed vector z to obtain $x' = G(z)$, where x' is obeying distribution $P_G(x)$. Then we obtain the optimal solution G^* by modifying the Generator's internal weight parameters to obey the real data distribution $P_{\text{data}}(x)$. Then there is the following equivalence relation: $P_G(x) \approx P_{\text{data}}(x)$. According to the above idea, the expression of G^* is as shown in **Eq. 1**:

$$G^* = \arg\left\{\min_G \left(\text{Div}\left(P_{\text{data}}(x), P_G(x)\right)\right)\right\}. \tag{1}$$

Discriminator is also a neural network, a binary classifier mainly used to determine whether the input data is from the sample data or the generated data. Its purpose is to discriminate the disparity between the original and generated samples more precisely. The disparity can be expressed in **Eq. 2** as follows:

$$v(D,G) = \mathrm{E}_{x\sim P_{\text{data}}} \log D(x) + \mathrm{E}_{x\sim P_G} \log(1 - D(x)), \tag{2}$$

Where $\mathrm{E}_{x\sim P_{\text{data}}}$ denotes the expectation of x within the sample distribution $P_{\text{data}}(x)$. $\mathrm{E}_{x\sim P_G}$ denotes the expectation of x within the sample distribution $P_G(x)$.

The Discriminator aims at maximizing $v(D,G)$. $\max_D v(D,G)$ makes Discriminator better distinguish between D and G. The larger the value of $\log D(x)$ and the smaller the value of $\log(1 - D(x))$, the trained Discriminator gets higher the discriminative accuracy.

We denote $\max_D v(D,G) = \text{Div}\left(P_{\text{data}}(x), P_G(x)\right)$, then $G^* = \arg\left\{\min_G \max_D v(D,G)\right\}$. The Discriminator's goal is to maximize the gap between the generated data distribution and the true data distribution. The Generator's goal is to minimize the gap between the generated data distribution and the true data distribution.

In summary, the Discriminator in GAN is trained to maximize the correctness of the labels assigned to the sample data and the "generated data." The Generator in GAN is trained to minimize the correct recognition of the "generated data" by the Discriminator. This adversary training process allows the Discriminator to reach the Nash equilibrium. Meanwhile, the Generator can generate "generated data" similar to the sample data and successfully "trick" the Discriminator.

According to the proof procedure in the companion paper, we can see that **Eq. 3** is the minimized Jensen-Shannon divergence (Goodfellow et al., 2014a):

$$\begin{aligned} v(D,G) &= -2\log 2 + KL\left(P_{\text{data}}(x) \| \frac{1}{2}\left(P_{\text{data}}(x) + P_G(x)\right)\right. \\ &\quad \left. + KL\left(P_G(x) \| \frac{1}{2}\left(P_{\text{data}}(x) + P_G(x)\right)\right)\right) \\ &= -2\log 2 + 2JS\left(P_{\text{data}}(x) \| P_G(x)\right). \end{aligned} \tag{3}$$

The original GAN has one defect: initially, the distribution of the "generated data" obtained by the Generator may not overlap

with the real data distribution. In this case, using the original JS divergence as a measure of the "distance" between the two distributions may fail. It results in the gradients disappearing and diffusing during training, thus failing to generate high-quality data.

CONDITIONAL GENERATIVE ADVERSARIAL NETWORKS

CGAN is a conditional generative adversarial networks model based on the GAN with conditional extensions. Suppose both the Generator and the Discriminator apply to some additional condition y, such as class labels. In that case, the data can be calibrated during the generation process by attaching y to the input layer for input to the Generator and Discriminator.

In the Generator, the noise is input along with the corresponding condition y, and the real data x and condition y are used as the Discriminator's objective function. According to the corresponding literature's derivation process, we can obtain **Eq. 4** (Mirza and Osindero, 2014).

$$\min_G \max_D v(D,G) = \mathrm{E}_{x\sim P_{\text{data}}} \log D(x|y) + \mathrm{E}_{x\sim P_G} \log(1 - D(x|y)). \tag{4}$$

From the above equation, the optimization process of CGAN for the objective function $v(D,G)$ is similar to that of GAN. Where $\mathrm{E}_{x\sim P_{\text{data}}} \log D(x|y)$ denotes the probability that the Discriminator identifies whether the data x is real data under condition y; $\mathrm{E}_{x\sim P_G} \log(1 - D(x|y))$ denotes the probability that the generated data x is judged to be real data by the Discriminator under condition y.

In summary, it can be seen that CGAN is an improvement of the unsupervized GAN model into a supervised model. The added condition helps to improve the accuracy of the generated data. However, since the objective function continues to use the GAN's objective function, there remains a scenario when the data distribution pattern of generated data P_G differs significantly from that of data P_{data}. It can make the case that the Generator's generated data would never pass the Discriminator's validation, and thus the gradient disappears and diffuses.

Wasserstein Distance-Based Generative Adversarial Networks

Both the original GAN and CGAN have the same problem: applying JS divergence as a measure of the "distance" between two distributions leads to the gradients disappearing and dispersion in training. In response to these problems, WGAN uses the Wasserstein distance to measure the difference between the true data distribution and the generated data distribution (Arjovsky et al., 2017).

The Wasserstein distance, also called Earth-Mover distance, is used to measure the distance between two distributions. Its expression is shown in **Eq. 5**:

$$W(p_d, p_g) = \inf_{\lambda \sim \prod(p_d, p_g)} \mathrm{E}_{(x,y)\sim\lambda} \|x - y\|, \tag{5}$$

Where $\Pi(p_d, p_g)$ is the set of all joint distributions after the combination of p_d and p_g. $\mathrm{E}_{(x,y)\sim\lambda}\|x - y\|$ is the expected value of the distance between the true sample x and the generated sample y under compliance with the joint distribution λ. $\inf_{\lambda \sim \prod(p_d, p_g)} \mathrm{E}_{(x,y)\sim\lambda}\|x - y\|$ is the lower bound that we manage to take for this expected value among all possible joint distributions.

For this Earth-Mover distance, we can intuitively interpret it as the "distance" used to move the "mound (p_d)" to the location of the "mound (p_g)" under the "planning path" of λ. $W(p_d, p_g)$ is the minimum distance under the "optimal path."

Wasserstein's advantage over traditional distance measures is that Wasserstein distance can still describe the distance between two distributions even if they do not overlap, overcoming the problem of gradient disappearance and dispersion in training due to no overlap between the two distributions. Although WGAN solves gradient disappearance and dispersion in training, the additional computational load increases the computational cost and reduces the training efficiency.

LEAST SQUARES GENERATIVE ADVERSARIAL NETWORKS

LSGAN is an optimization model of GAN proposed by Mao Xudong and other scholars in 2017 (Mao et al., 2017). It mainly addresses the traditional GAN's two defects: the low quality of the generated images and the training process's instability. The difference is that GAN's objective function is changed from the cross-entropy loss function to the least-squares loss function. Consequently, a more stable and faster converging adversarial network with high generation quality is born.

The objective function of LSGAN is defined as **Eq. 6** (Mao et al., 2017).

$$\begin{cases} \min_D V_{LSGAN}(D) = \frac{1}{2}\mathrm{E}_{x\sim p_{data}(x)}(D(x) - b)^2 + \frac{1}{2}\mathrm{E}_{x\sim p_G(x)}(D(G(x)) - a)^2, \\ \min_G V_{LSGAN}(G) = \frac{1}{2}\mathrm{E}_{x\sim p_G(x)}(D(G(x)) - c)^2, \end{cases} \tag{6}$$

Where a and b denote the labels of the fake and true data, respectively. c denotes the value that the Generator expects the Discriminator to trust for the fake data (Ma et al., 2019).

Two options are given for the values of a, b, c.

1) Add $(1/2)\mathrm{E}_{x\sim p_{\text{data}}(x)}(D(x) - c)^2$ to **Eq. 6**. Since $\min_G V_{LSGAN}(G)$ does not contain the generator parameters, it does not change the optimal solution. Thus we can obtain **Eq. 7**.

$$\begin{cases} \min_D V_{LSGAN}(D) = \frac{1}{2}\mathrm{E}_{x\sim p_{\text{data}}(x)}(D(x) - b)^2 + \frac{1}{2}\mathrm{E}_{x\sim p_G(x)}(D(G(x)) - a)^2, \\ \min_G V_{LSGAN}(G) = \frac{1}{2}\mathrm{E}_{x\sim p_{\text{data}}(x)}(D(x) - c)^2 \frac{1}{2}\mathrm{E}_{x\sim p_G(x)}(D(G(x)) - c)^2, \end{cases} \tag{7}$$

Maintaining the Generator constant, we can obtain the optimal solution of the Discriminator as in **Eq. 8**.

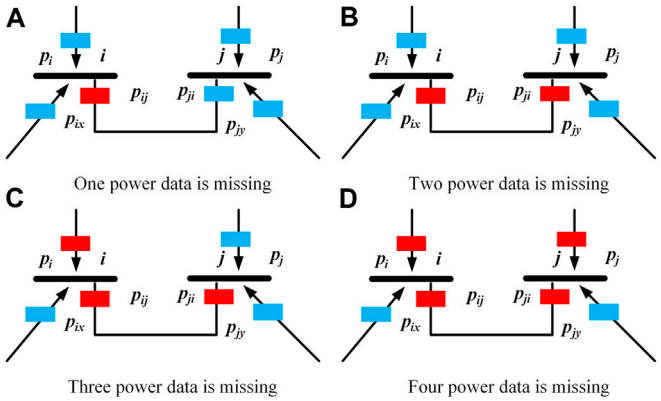


FIGURE 2 | Schematic diagram of the missing active power data in different cases.

$$D^* = \frac{bP_{\text{data}}(x) + aP_G(x)}{P_{\text{data}}(x) + P_G(x)}. \tag{8}$$

From **Eqs. 7**, **8**, we derive **Eq. 9**.

$$v(G) = \int_x \frac{((b-c)(P_{\text{data}}(x) + P_G(x)) - (b-a)P_G(x))^2}{P_{\text{data}}(x) + P_G(x)} dx. \tag{9}$$

If we set b-c = 1 and b-a = 2, we can obtain the following equations.

$$\begin{aligned} v(G) &= \int_x \frac{((P_{\text{data}}(x) + P_G(x)) - 2P_G(x))^2}{P_{\text{data}}(x) + P_G(x)} dx \\ &= \chi^2_{\text{Pearson}}(P_{\text{data}}(x) + P_G(x) \| 2P_G(x)), \end{aligned} \tag{10}$$

where χ^2_{Pearson} is the Pearson chi-square divergence. For example, when we set a = -1, b = 1, and c = 0, the objective function is as follows:

$$\begin{cases} \min\limits_D V_{LSGAN}(D) = \frac{1}{2}\mathrm{E}_{x\sim p_{data}(x)}(D(x)-1)^2 + \frac{1}{2}\mathrm{E}_{x\sim p_G(x)}(D(G(x))+1)^2, \\ \min\limits_G V_{LSGAN}(G) = \frac{1}{2}\mathrm{E}_{x\sim p_G(x)}(D(G(x)))^2. \end{cases} \tag{11}$$

2) By setting b = c, the Generator generates data that is as similar as possible to the true data distribution. For example, if we set a = 0 and b = c = 1, respectively, the objective function is as follows:

$$\begin{cases} \min\limits_D V_{LSGAN}(D) = \frac{1}{2}\mathrm{E}_{x\sim p_{\text{data}}(x)}(D(x)-1)^2 + \frac{1}{2}\mathrm{E}_{x\sim p_G(x)}(D(G(x)))^2, \\ \min\limits_G V_{LSGAN}(G) = \frac{1}{2}\mathrm{E}_{x\sim p_G(x)}(D(G(x))-1)^2. \end{cases} \tag{12}$$

The main idea of LSGAN is to provide a smoothing and non-saturating gradient loss function for the Discriminator. In this way, D "pulls" the data generated by the generator G to the true data distribution $P_{\text{data}}(x)$, so that the distribution of the data generated by G is similar to $P_{\text{data}}(x)$. In this paper, we choose the second scheme as the objective function.

LSGAN-BASED METHOD FOR RECONSTRUCTING MISSING POWER DATA

Data does not exist in isolation. There are often various constraints between data, which describe the relationship between the data. The data must meet this correlation between the data and not be contradicted by each other. There is a constraint relationship between the power data in the power system: during the system's operation, the power balance is satisfied at all times.

The power grid is composed of nodes and lines. The grid's power balance can be divided into two types of balance: node power balance and line power balance. For the node, the power directly related to the node satisfies the principle that the total power injected into the node is equal to the total power out of the node. For the line, the difference of the actual power at both ends is the power loss of branches.

The blue blocks in **Figure 2** represent normal power data, and the red blocks represent missing power data.

When the case in a) occurs, P_{ij} power data is missing. According to the line power balance principle, we can reconstruct the missing data Pij by the power at the other end of the branch, and the branch loss power. We can also reconstruct the missing power P_{ij} by directly related to node i according to the node power balance principle.

When the case in b) occurs, P_{ij} and P_{ji} power data are missing. We can reconstruct the missing power P_{ij} and P_{ji} by directly correlating the power with node i and node j, respectively, according to the node power balance principle.

When the case in c) occurs, the missing power P_{ji} can be reconstructed according to the node power balance principle by directly correlating the power with node j. Then the missing power P_{ij} according to the line power balance principle. Finally, the missing power P_{ii} can be reconstructed according to the node power balance principle by directly correlating the power with node i.

When the case in d) occurs, the system does not meet the observability. We can no longer complete the data reconstruction task by applying the power balance principle alone. The data reconstruction method in this paper can solve this problem.

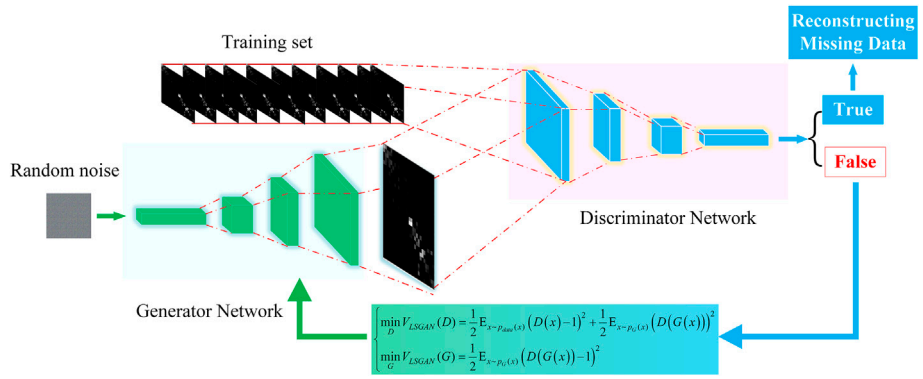


FIGURE 3 | Flow chart of the reconstruction method for missing power data based on LSGAN.

The data we currently acquire are fully structured data recorded at the same time sampling scale with the correlation measurements of different stations. There are topological linkage relationships between the physical objects it represents, so each time section's data are data with topological constraints.

Unlike regular data, power system measurement data embody spatially constrained relationships between each other. Therefore, we can add spatiality to the characteristics of describing data. By organically integrating the grid data's spatial correlations with the grid power data, we can enrich the data's distribution characteristics and enhance the data's representable dimensions. It contributes to improving the learning effect of generative adversarial networks on data characters.

We use the adjacency matrix A to describe the network topology. If a topology consists of n nodes, its adjacency matrix is an $n \times n$ matrix $A \in \mathbb{R}^{n\times n}$. The values of each element a_{ij} inside it are as follows:

$$a_{ij} = \begin{cases} 1, & if\ (v_i, v_j) \in E, \\ 0, & if\ (v_i, v_j) \notin E, \end{cases} \tag{13}$$

Where v is a one-dimensional array storing information about the graph's vertices, v_i and v_j denote node i and node j. E is a two-dimensional array storing information about the edges (nodes directly interconnected) in the graph. $(v_i, v_j) \in E$ means that node i and node j are directly interconnected. Otherwise, the two nodes are not directly connected.

We place the adjacency matrix's non-zero elements to the line active power P_{ij} and place the diagonal elements to the node injected active power value P_{ii}. The node active power correlation matrix $P_{relation}$ is thus generated as shown in **Eq. 14**, $P_{\text{relation}} \in \mathbb{R}^{n\times n}$.

$$P^{ij}_{\text{relation}} = \begin{cases} P_{ij}, & if\ (v_i, v_j) \in E, \\ 0, & if\ (v_i, v_j) \notin E, \\ P_{ii}, & if\ i = j. \end{cases} \tag{14}$$

The node active power correlation matrix P_{relation} forms the database for describing the grayscale map of active power distribution in a single section. The magnitude of the matrix's values determines the corresponding color blocks' lightness and darkness in the grayscale map. We can analogize the process of generating new pictures by unsupervized learning in generative adversarial networks to reconstruct the missing power data. The specific process is shown in **Figure 3**.

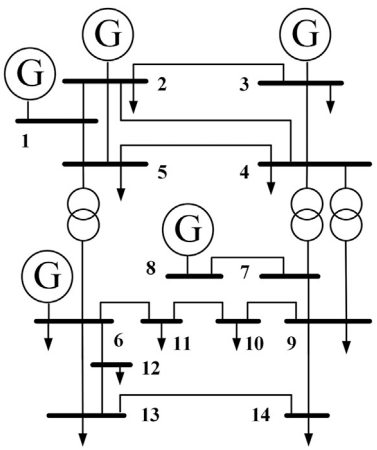


FIGURE 4 | IEEE14-bus system network topology.

EXPERIMENT AND RESULTS

Experimental Results on a Test Sample

In this paper, the IEEE14-bus system shown in **Figure 4** was used to verify the proposed method's effectiveness. The network consisted of 14 nodes and 20 equivalent transmission lines. To make the examples more general, we increased each load in the example system by 1–10% in equal proportions, for a total of 10 growth percentages, while not changing the rated output of the generator nodes. A Gaussian perturbation of 0.01 was added to each growth amount to generate 1,000 data samples for 10,000 data samples.

The nodal active power correlation matrix P_{relation} and the nodal reactive power correlation matrix Q_{relation} were generated

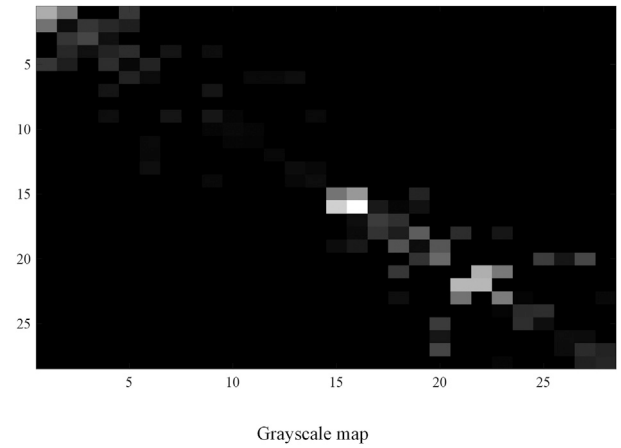


FIGURE 5 | The grayscale map is drawn from the original sample data.

for these data samples according to the nodal active power correlation matrix's composition mode. We stitched the two matrices together along the diagonal to form a new matrix Y, $Y \in \mathbb{R}^{n\times n}$:

$$Y = \begin{bmatrix} P_{\text{relation}} & 0 \\ 0 & Q_{\text{relation}} \end{bmatrix}. \tag{15}$$

We set the base values of active and reactive power as P_{base} = 500 MW and Q_{base} = 50 MW and then took the standardized values for the Y matrix's active and reactive parts, respectively. Thus, the normalized power correlation matrix X was generated.

$$X = \begin{bmatrix} P_{\text{relation}}/P_{\text{base}} & 0 \\ 0 & Q_{\text{relation}}/Q_{\text{base}} \end{bmatrix}. \tag{16}$$

If we mapped the normalized power correlation matrix X as a graph, then the magnitude of the matrix X values determined the color block's lightness or darkness at the corresponding position. The original sample data mapping graph is shown in **Figure 5**.

As seen in **Figure 5**, the data distribution characteristics shown in the graph are that the data are concentrated around the diagonal, and the upper and lower triangles are approximately symmetric. It is consistent with the distribution characteristics of the original data.

We divided the above 10,000 data samples into the training set and test set in the ratio of 8:2. The dimension of the training set X_{train} was (28,28,8000), and the dimension of the test set X_{test} was (28,28,2000). On the test set X_{test}, we set 10 active power missing and 10 reactive power missing in their node power correlation matrix. The specific information was node injected power: $P_{1\text{-}1}$, $P_{3\text{-}3}$, $Q_{1\text{-}1}$, $Q_{2\text{-}2}$, $Q_{6\text{-}6}$, $Q_{8\text{-}8}$, and line power: $P_{1\text{-}2}$, $P_{2\text{-}1}$, $P_{5\text{-}1}$, $P_{3\text{-}2}$, $P_{5\text{-}2}$, $P_{2\text{-}3}$, $P_{5\text{-}4}$, $P_{4\text{-}5}$, $Q_{1\text{-}2}$, $Q_{2\text{-}1}$, $Q_{4\text{-}5}$, $Q_{5\text{-}4}$, $Q_{8\text{-}7}$, $Q_{7\text{-}8}$.

We set the batch training number of GAN, CGAN, WGAN, and LSGAN to 32 and the maximum number of training epochs to 50. Regarding the optimizer, we chose to use Adaptive Moment Estimation (Adam). Based on the comparison of the relevant parameters in the literature (Ruder, 2016), we set the parameters as follows: Image size was 28 × 28, the learning rate of the Discriminator was 3e-4, the learning rate of the Generator was 3e-4, beta1 was 0.5, and beta2 was 0.999.

We got the mapping graphs of the data generated under the same training batch at each method's 50th epoch. Then they were used to compare and analyzed the effectiveness of learning data features by each method. The results were shown in **Figure 6**.

As can be seen from **Figure 6**, the data generated by each method display the main distribution characteristics of "the data are concentrated around the diagonal, and the upper and lower triangles show approximate symmetry." However, judging from the details, the grayscale map of the data generated by LSGAN is most close to the grayscale map of the original sample. The data generated by GAN differs significantly from the original data. Both CGAN and WGAN methods are deficient in the accuracy of generating non-diagonal data.

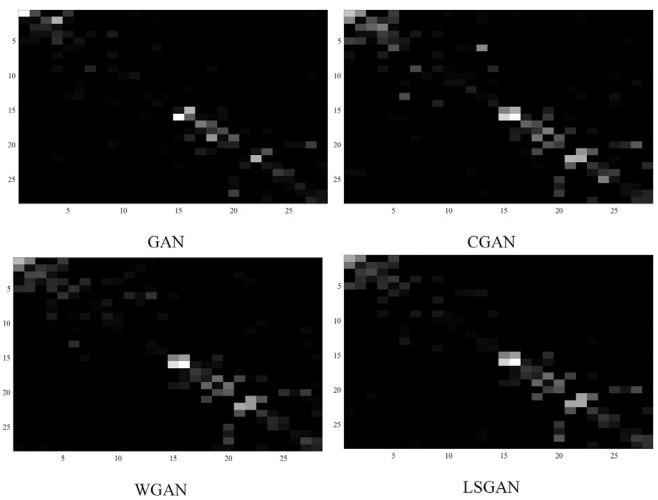


FIGURE 6 | The grayscale map is drawn from the generated data by GAN, CGAN, WGAN, LSGAN.

TABLE 1 | The AE and APE statistics of the generated active data.

	GAN		CGAN		WGAN		LSGAN	
	AE (MW)	**APE (%)**	**AE (MW)**	**APE (%)**	**AE (MW)**	**APE (%)**	**AE (MW)**	**APE (%)**
$P_{1\text{-}1}$	115.3 ± 68.5	49.60 ± 29.44	**7.1** ± 17.1	**3.07** ± 7.36	19.0 ± 14.3	8.18 ± 6.13	9.9 ± **4.4**	4.24 ± **1.9**
$P_{2\text{-}1}$	66.4 ± 48.6	43.46 ± 31.85	11.2 ± 28.1	7.33 ± 18.41	22.4 ± 16.4	14.67 ± 10.75	**8.9** ± **4.9**	**5.8** ± **3.21**
$P_{5\text{-}1}$	42.2 ± 36.1	57.98 ± 49.67	24.4 ± 71.4	33.60 ± 98.12	10.3 ± 7.5	14.14 ± 10.28	**4.2** ± **3.2**	**5.76** ± **4.41**
$P_{1\text{-}2}$	57.0 ± 42.6	36.29 ± 27.11	8.8.0 ± 23.6	5.57 ± 15.02	19.4 ± 13.9	12.37 ± 8.87	**7.5** ± **4.7**	**4.75** ± **2.98**
$P_{3\text{-}2}$	23.4 ± 16.4	32.95 ± 23.09	23.0 ± 64.1	32.41 ± 90.20	13.9 ± 10.7	19.52 ± 15.12	**4.6** ± **3.5**	**6.5** ± **4.88**
$P_{5\text{-}2}$	31.4 ± 23.8	77.09 ± 58.46	33.6 ± 81.9	82.34 ± 200.75	10.1 ± 8.3	24.85 ± 20.37	**2.7** ± **2.1**	**6.62** ± **5.06**
$P_{2\text{-}3}$	43.7 ± 29.3	59.59 ± 40.00	7.1 ± 18.5	9.66 ± 25.18	9.0 ± 6.9	12.28 ± 9.38	**5.7** ± **4.1**	**7.73** ± **5.61**
$P_{3\text{-}3}$	47.6 ± 32.6	50.56 ± 34.66	19.0 ± 44.5	20.20 ± 47.25	13.9 ± 10.3	14.75 ± 10.91	**4.3** ± **3.0**	**4.52** ± **3.23**
$P_{5\text{-}4}$	24.4 ± 21.2	40.53 ± 35.14	17.5 ± 46.7	29.04 ± 77.51	12.8 ± 9.7	21.29 ± 16.12	**5.7** ± **4.2**	**9.39** ± **7.02**
$P_{4\text{-}5}$	33.3 ± 29.5	55.69 ± 49.29	13.9 ± 44.2	23.25 ± 73.83	16.6 ± 10.0	27.81 ± 16.78	**4.6** ± **3.2**	**7.63** ± **5.33**

1 The bold values represent the minimum average error values and percentage error values of the active power obtained by different methods.

TABLE 2 | The AE and APE statistics of the generated reactive data.

	GAN		CGAN		WGAN		LSGAN	
	AE (MW)	**APE (%)**	**AE (MW)**	**APE (%)**	**AE (MW)**	**APE (%)**	**AE (MW)**	**APE (%)**
$Q_{1\text{-}1}$	7.3 ± 5.1	46.69 ± 32.86	0.8 ± 2.2	5.29 ± 14.46	0.7 ± 0.5	4.68 ± 3.38	**0.4** ± **0.3**	**2.78** ± **1.65**
$Q_{1\text{-}2}$	8.3 ± 5.5	40.54 ± 26.99	0.7 ± 2.1	3.36 ± 10.39	0.6 ± 0.4	2.84 ± 2.17	**0.3** ± **0.2**	**1.35** ± **1.04**
$Q_{2\text{-}1}$	7.7 ± 4.7	27.9 ± 17.1	0.6 ± 1.7	2.06 ± 6.05	0.6 ± 0.4	2.07 ± 1.56	**0.4** ± **0.3**	**1.3** ± **1**
$Q_{2\text{-}2}$	8.7 ± 6.2	25.38 ± 18.15	**1.1** ± 2.6	**3.26** ± 7.46	1.7 ± 0.7	4.88 ± 2.12	1.7 ± **0.5**	5.09 ± **1.38**
$Q_{4\text{-}5}$	4.4 ± 3.4	34.72 ± 27.29	**1.3** ± 2.8	**10.24** ± 22.32	1.5 ± 0.7	12.06 ± 5.39	1.6 ± **0.3**	12.4 ± **2.78**
$Q_{5\text{-}4}$	5.9 ± 5.2	53.86 ± 47.43	1.4 ± 3.1	12.37 ± 27.98	1.5 ± 0.8	13.99 ± 7.53	**1.4** ± **0.5**	**12.36** ± **4.78**
$Q_{8\text{-}7}$	10.4 ± 6.6	42.46 ± 26.95	**1.6** ± 2.6	**6.35** ± 10.43	3.3 ± 0.8	13.3 ± 3.2	3.4 ± **0.3**	13.69 ± **1.21**
$Q_{7\text{-}8}$	10.8 ± 6.6	45.86 ± 28.11	**1.4** ± 2.2	**5.93** ± 9.42	3.2 ± 0.7	13.59 ± 3.14	3.1 ± **0.3**	13.19 ± **1.36**
$Q_{8\text{-}8}$	6.9 ± 4.4	28.07 ± 18.05	**1.4** ± 2.1	**5.89** ± 8.73	3.3 ± 0.8	13.52 ± 3.29	3.4 ± **0.4**	13.87 ± **1.45**
$Q_{6\text{-}6}$	6.5 ± 2.6	46.36 ± 18.2	**2.1** ± 3.8	**14.66** ± 27.22	4.2 ± 1.2	30.02 ± 8.79	4.2 ± **0.5**	29.55 ± **3.62**

2 The bold values represent the minimum average error values and percentage error values of the reactive power obtained by different methods.

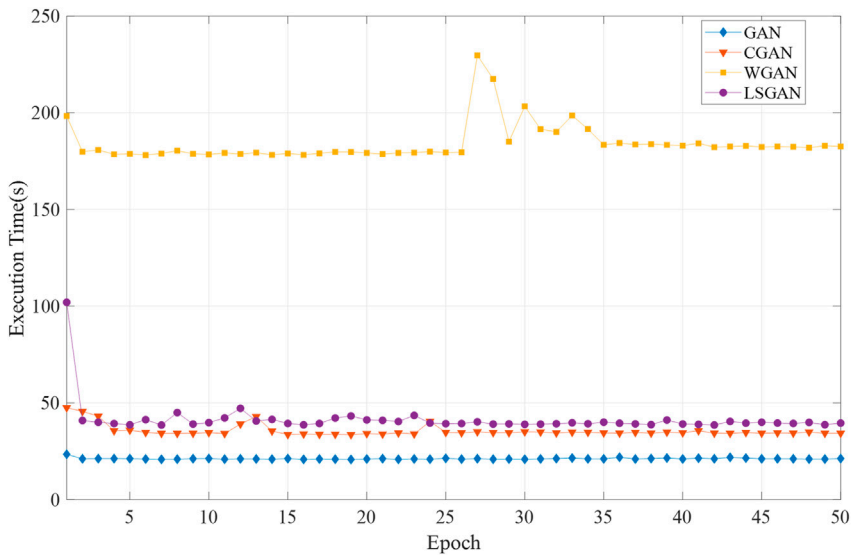


FIGURE 7 | Computational time consumption curves for each training cycle of different methods.

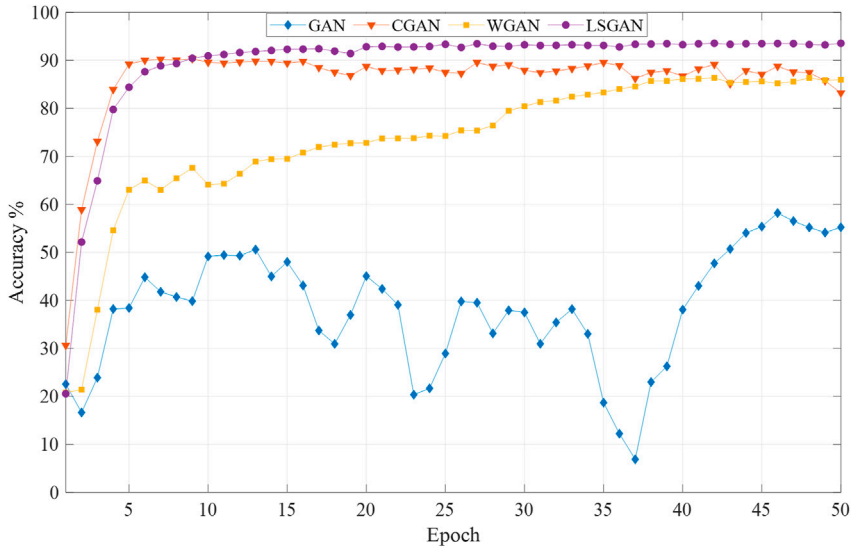


FIGURE 8 | Mean accuracy of reconstructed data by different methods during the training process.

TABLE 3 | The AE and APE statistics of the generated reactive data (IEEE 39-bus system).

	$P_{1\text{-}1}$	$P_{2\text{-}1}$	$P_{1\text{-}2}$	$P_{3\text{-}2}$	$P_{2\text{-}3}$
AE (MW)	7.05 ± 5.20	11.05 ± 8.28	4.81 ± 3.60	10.10 ± 7.46	9.07 ± 5.17
APE (%)	7.22 ± 5.33	6.33 ± 4.74	2.77 ± 2.07	3.17 ± 2.34	2.83 ± 1.62
	$P_{3\text{-}3}$	$P_{4\text{-}5}$	$P_{5\text{-}6}$	$P_{4\text{-}4}$	$P_{5\text{-}4}$
AE (MW)	5.30 ± 4.06	6.62 ± 5.09	6.91 ± 4.87	8.33 ± 6.22	8.60 ± 6.68
APE (%)	1.65 ± 1.26	3.35 ± 2.58	1.29 ± 0.91	1.67 ± 1.24	4.35 ± 3.38

To further measure the accuracy of the data generated by the above four methods, we counted the absolute error (AE), absolute percentage error (APE) of the 20 sets of missing data corresponding to the generated data under the 50th epoch of each method. The mean and standard deviation statistics were shown in **Tables 1**, **2**.

As seen in **Tables 1**, **2**, the LSGAN method generates data with small errors and the highest accuracy in most of these cases. Although the error of the generated data under the CGAN method in some measurements is smaller than the error of the generated data based on the LSGAN method, the LSGAN method is more stable in terms of standard deviation. These indicate that LSGAN outperforms the other three methods in most cases in terms of generating data effects.

To obtain the training effect of the proposed method in the training process, we counted the computation time consumed and the reconstructed data's average accuracy under different epochs. The results were shown in **Figures 7**, **8**.

As seen in **Figure 7**, GAN is the most efficient in terms of computational efficiency. CGAN and LSGAN are approximately more efficient than WGAN. WGAN takes the most time to compute and is the least efficient.

As seen in **Figure 8**, the highest accuracy of the data reconstructed by GAN is only 58.19%, and the training effect is unstable. It is mainly due to the gradient disappearance and dispersion during the training process. CGAN makes the accuracy of the reconstructed data reach 90.24% in the first 7 epochs. It indicates that the method can obtain the reconstructed data with high accuracy in a short training period. However, the accuracy of the reconstructed data by CGAN decreases as the training continues. It indicates that although CGAN improves the GAN-based reconstructed data's accuracy, the training instability still exists. The accuracy of the data reconstructed by WGAN steadily increased during the continuous training process and finally reached 86.13%. In contrast, the accuracy of the data reconstructed by LSGAN is not as high as that of CGAN in a short period. However, the accuracy of the generated data has been steadily improving with increasing training epochs, and the highest accuracy reaches 93.57%, which is significantly better than the other three methods.

Experimental Results on LSGAN

In this section, we applied the IEEE 39-bus system and the IEEE 118-bus system to test the method's effectiveness in this paper. The IEEE 39-bus system had 39 nodes and 46 lines. The dimension of the nodal active power correlation matrix $P_{\text{relation-}39}$, which was composed based on this system's tide data, was 39 × 39. We followed the LSGAN-related

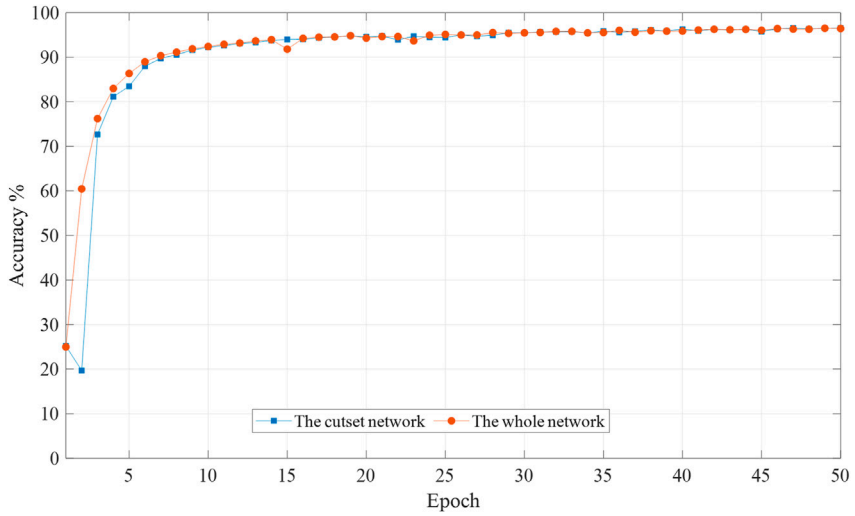


FIGURE 9 | Mean accuracy of reconstructed data by different network forms during the training process (IEEE 39-bus system).

TABLE 4 | The AE and APE statistics of the generated reactive data (IEEE 118-bus system).

	P_{4-5}	P_{5-8}	P_{8-9}	P_{10-9}	P_{15-15}
AE (MW)	5.51 ± 4.09	10.45 ± 7.51	5.85 ± 4.44	6.11 ± 4.55	4.13 ± 3.14
APE (%)	5.34 ± 3.97	3.09 ± 2.22	1.33 ± 1.01	1.36 ± 1.01	4.60 ± 3.50
	P_{15-17}	P_{25-23}	P_{25-25}	P_{26-26}	P_{25-27}
AE (MW)	4.98 ± 3.64	3.13 ± 2.42	1.79 ± 1.39	5.96 ± 4.37	1.68 ± 1.34
APE (%)	4.80 ± 3.51	1.88 ± 1.45	0.82 ± 0.63	1.90 ± 1.39	1.17 ± 0.93

parameters in the previous section and modified the weight coefficient matrix dimension to fit the nodal active power correlation matrix derived from the IEEE 39-bus system.

We added a 1% Gaussian perturbation to each load in the test system without changing the generator nodes' rated output to generate 1,000 data samples as training samples. The proposed method was then trained iteratively for 50 epochs, generating 250 data samples per epoch. We selected some nodes active power and directly associated line power as missing data, as follows: P_{1-1}, P_{2-1}, P_{1-2}, P_{3-2}, P_{2-3}, P_{3-3}, P_{4-5}, P_{5-6}, P_{4-4}, P_{5-4}

We counted the absolute error (AE), absolute percentage error (APE) of the 10 sets of missing data corresponding to the generated data under each method's 50th epoch. The mean and standard deviation statistics were shown in **Table 3**.

As seen from the above table, the generated data have few errors, and the majority of the generated data have an accuracy of over 92%. The difference between data-driven and mechanism-based modeling is that the former is not constrained by the system operating conditions. The above process is executed under whole network conditions. How effective the proposed method is under partial data conditions. We performed the following experiments: we extracted bus 1–14 in the whole network to form a cut-set network, and the cut-set network was made to contain all missing measures. We treat the contact line between the cut-set network and the whole network as a separate line. Then we modify the dimensionality of the relevant parameters within LSGAN to fit the new network. The accuracy of the generated data changes under the two network forms is counted. The average accuracy of the generated data trained by the whole network and the cut-set network under different training cycles is compared as shown in **Figure 9**.

As can be seen from the above figure, the data generated by the whole network (39 buses) is more accurate than that generated by the cut-set network (14 buses) at the beginning of the training process. In the later period, the data accuracy of both generated data is the same. The data-driven data restoration approach does not rely on external conditions such as network parameters and does not require complete network data. The purpose of data restoration can also be achieved with cut-set data.

We made similar experiments as above in the IEEE 118-bus system. The IEEE 118-bus system had 118 nodes and 186 lines. The dimension of the nodal active power correlation matrix $P_{\text{relation-}118}$, which was composed based on this system's tide data, was 118 × 118. We followed the LSGAN-related parameters in the previous section and modify the weight coefficient matrix dimension to fit the nodal active power correlation matrix derived from the IEEE 118-bus system.

We selected some nodes' active power and directly associated line power as missing data, as follows: P_{4-5}, P_{5-8}, P_{8-9}, P_{10-9}, P_{15-15}, P_{15-17}, P_{25-23}, P_{25-25}., P_{26-26}, P_{25-27}. We counted the absolute error (AE), absolute percentage error (APE) of the 10 sets of missing data corresponding to the generated data under each method's 50th epoch. The mean and standard deviation statistics were shown in **Table 4**.

As seen from the above table, the generated data have minor errors, and the majority of the generated data have accuracy above 95%. It indicates that the method in this paper can be extended to apply in larger networks. To further verify whether

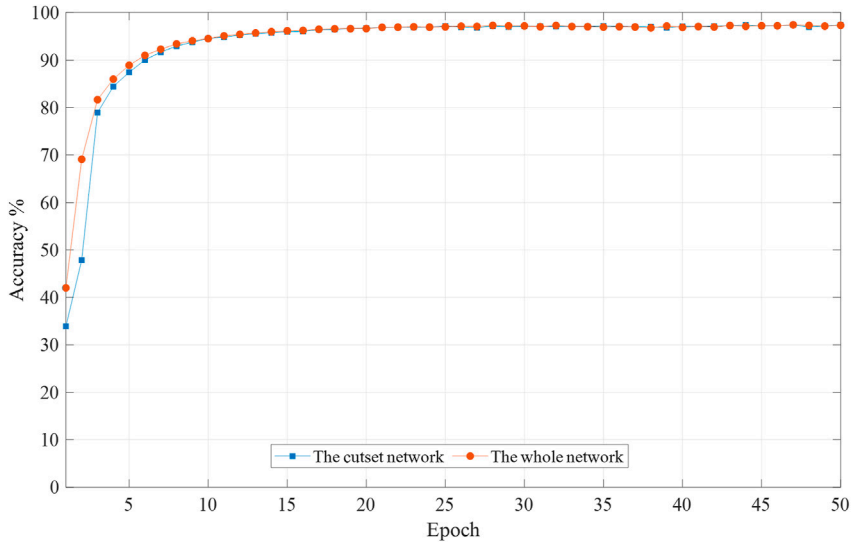


FIGURE 10 | Mean accuracy of reconstructed data by different network forms during the training process (IEEE 118-bus system).

the cut-set network still works in large networks, we select a cut-set network consisting of bus 1–30 for training, containing all the missing data. The average accuracy of the generated data trained by the whole network (118 nodes) and the cut-set network (30 nodes) under different training epochs are compared, as shown in **Figure 10**.

As shown from the above figure, the test results are similar to those of the IEEE 39-bus system test. At the beginning of the training process, the accuracy of the data generated by the whole network training is higher than that of the data generated by the cut-set network training. With continuous iterative training, the two generated data with the same accuracy at the later stage. Although the accuracy is similar between the two networks, the cut-set network is used as the training sample to streamline the data's composition and decrease the non-essential data. The computation time is only 1/10 of the whole network, which greatly improves the computation efficiency.

LSGAN does not depend on external operating conditions' constraints but needs to have sample data as the necessary foundation. With the great development of power system information, the power grid has established many measurement systems reflecting the system operation status, such as the SCADA system, which records active power, reactive power, voltage, and power grid frequency. The sampling interval of this system is 1 min, and each measurement day records 1,440 data. The amount of SCADA data recorded by a provincial power grid in a day can reach several GB, which provides a good training sample base for this paper's method. Grid measurement data has spatial and temporal properties. When there is a high dimensional and high loss rate case in the power grid, the data before and after the time series can be used as training samples to identify the missing data, and the data at the same time every day can be used as supplementary samples to assist in determining the missing data. The actual system is complex and variable. How to construct the internal architecture of LSGAN, select the sample data, and apply this paper to the actual system needs to continue to be studied in-depth, which is not discussed too much in this paper.

CONCLUSION

This paper's main contribution is to propose a method based on LSGAN to reconstruct missing data for missing measurement data in power systems. We transformed the problem of reconstructing missing data into the problem of repairing missing parts in images. The method in this paper provided a new approach to solve the problem of restoring missing data. LSGAN learned correlations and distribution features among data by unsupervized self-gaming training mode. By changing the latent parameters in the Generator, it enabled the Generator to generate data that matched the objective laws of real data. The proposed method was able to cope with data loss in power systems due to improper handling and provide a solid technical basis for ensuring data integrity.

Unlike the traditional GAN model, LSGAN replaced the original objective function from cross-entropy loss function to least-squares loss function. It ensured gradient descent by penalizing those samples far from the decision boundary, solving gradient disappearance and dispersion. Moreover, the least-squares iterative computation was efficient.

It was experimentally verified that the proposed method could still reconstruct the missing data in the case of multiple power data missing. In the comparison experiments with GAN, CGAN, and WGAN models, respectively, the LSGAN-based data reconstruction method could steadily improve the generated data accuracy during the continuous training process with

higher final accuracy than GAN, CGAN, and WGAN models under the same epoch. The computational efficiency was 4.5 times higher than that of WGAN.

The method in this paper was entirely data-driven and did not involve mechanistic modeling. A cut-set network could be constructed on-demand to streamline the composition of the data, thus avoiding non-essential computational burden, and its accuracy was similar to that of data generated by the overall network training. The method was able to generate data with high accuracy for the restoration data problem. It was mainly because the least squares-based loss function imposes a large penalty on the boundary data. Although it enabled to improve the accuracy of the generated data, the method had some limitations for cases where diverse sample data needed to be generated.

In this paper, the proposed method was validated based on the IEEE 14-bus system, IEEE 39-bus system, and IEEE 118-bus system. The feasibility of the method was demonstrated. It should be noted that many issues are worthy of attention and further study for application to actual large-scale power grids. For example, designing LSAGN internal deep neural networks for large-scale power systems and how to select the training set reasonably are all issues that we plan to study in-depth in the future.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

Conception and design of study: CW Acquisition of data: SZ Drafting the manuscript: YC, TL Analysis and interpretation of data: CW, YC Revising the manuscript critically for important intellectual content: YC, CW.

FUNDING

This work was supported by the National Natural Science Foundation of China (No. 51437003).

REFERENCES

Arjovsky, M., Chintala, S., and Bottou, L. (2017). "Wasserstein generative adversarial networks," in Proceedings of the 34th international conference on machine learning-volume 70, Sydney, Australia, August 6–11, 2017. Editor D. Precup and Y. Whye Teh (Sydney, NSW, Australia: JMLR.org), 214–223.

Borgia, A., Hua, Y., Kodirov, E., and Robertson, N. (2019). "GAN-based pose-aware regulation for video-based person Re-identification," in 2019 IEEE winter conference on applications of computer vision (WACV), Waikoloa Village, HI, January 11, 2019 (New York, NY: IEEE), 1175–1184. doi:10.1109/WACV.2019.00130

Comerford, L., Kougioumtzoglou, I. A., and Beer, M. (2015). An artificial neural network approach for stochastic process power spectrum estimation subject to missing data. *Struct. Saf.* 52, 150–160. doi:10.1016/j.strusafe.2014.10.001

Dong, J., Yin, R., Sun, X., Li, Q., Yang, Y., and Qin, X. (2019). Inpainting of remote sensing SST images with deep convolutional generative adversarial network. *IEEE Geosci. Remote Sensing Lett.* 16 (2), 173–177. doi:10.1109/LGRS.2018.2870880

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., et al. (2014a). "Generative adversarial nets," in Advances in neural information processing systems, Montreal, QC, Demember 8–13, 2014 (CanadaCurran Associates, Inc.).

Ho, C. H., Wu, H. C., Chan, S. C., and Hou, Y. (2020). A robust statistical approach to distributed power system state estimation with bad data. *IEEE Trans. Smart Grid* 11 (1), 517–527. doi:10.1109/Tsg.2019.2924496

Jing, L., Wei, L., and Sheng, Y. (2018). Research on abnormal diagnosis for power grid equipment archival data based on machine learning. *Electric Power Inf. Commun. Tech.* 16 (7), 21–27. doi:10.16543/j.2095-641x.electric.power.ict.2018.07.004

Kallin Westin, L. (2004). *Missing data and the preprocessing perceptron*. Umeå, Sweden: Umeå University.

Lan, J., Guo, Q., and Sun, H. (2018). Demand side data generating based on conditional generative adversarial networks. *Energ. Proced.* 152, 1188–1193. doi:10.1016/j.egypro.2018.09.157

Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., et al. (2017). "Photo-realistic single image super-resolution using a generative adversarial network," in 2017 IEEE conference on computer vision and pattern recognition (CVPR), Honolulu, HI, July 21–26, 2017 (New York, NY: IEEE), 105–114. doi:10.1109/CVPR.2017.19

Li, L., Liu, H., Zhou, H., and Zhang, C. (2020). Missing data estimation method for time series data in structure health monitoring systems by probability principal component analysis. *Adv. Eng. Softw.* 149, 102901. doi:10.1016/j.advengsoft.2020.102901

Li, Y., Yang, Z., Li, G., Zhao, D., and Tian, W. (2019). Optimal scheduling of an isolated microgrid with battery storage considering load and renewable generation uncertainties. *IEEE Trans. Ind. Electron.* 66 (2), 1565–1575. doi:10.1109/TIE.2018.2840498

Ma, J., Yu, W., Liang, P., Li, C., and Jiang, J. (2019). FusionGAN: a generative adversarial network for infrared and visible image fusion. *Inf. Fusion* 48, 11–26. doi:10.1016/j.inffus.2018.09.004

Mao, X., Li, Q., Xie, H., Lau, R. Y. K., and Smolley, S. P. (2017). "Least squares generative adversarial networks," in 2017 IEEE international conference on computer vision (ICCV), Venice, Italy, October 22, 2017 (New York, NY: IEEE), 2813–2821. doi:10.1109/ICCV.2017.304

Miranda, V., Krstulovic, J., Keko, H., Moreira, C., and Pereira, J. (2012). Reconstructing missing data in state estimation with autoencoders. *IEEE Trans. Power Syst.* 27 (2), 604–611. doi:10.1109/TPWRS.2011.2174810

Mirza, M., and Osindero, S. (2014). Conditional generative adversarial nets. arXiv: https://arxiv.org/abs/1411.1784.

Ruder, S. (2016). An overview of gradient descent optimization algorithms. arXiv: https://arxiv.org/abs/1609.04747.

Sun, Y.-Y., Jia, R.-S., Sun, H.-M., Zhang, X.-L., Peng, Y.-J., and Lu, X.-M. (2018). Reconstruction of seismic data with missing traces based on optimized Poisson Disk sampling and compressed sensing. *Comput. Geosci.* 117, 32. doi:10.1016/j.cageo.2018.05.005

Wang, C., Mu, G., and Cao, Y. (2020). A method for cleaning power grid operation data based on spatiotemporal correlation constraints. *IEEE Access* 8, 224741–224749. doi:10.1109/ACCESS.2020.3044051

Wang, K., Gou, C., Duan, Y. J., Yilun, L., and Zheng, X. H. (2017). Generative adversarial networks: the state of the art and beyond. *Zidonghua Xuebao/Acta Automatica Sinica* 43, 321–332. doi:10.16383/j.aas.2017.y000003

Wang, M., Wen, X., and Hu, S. (2019a). "Faithful face image completion for HMD occlusion removal," in IEEE international symposium on mixed and augmented reality adjunct, Beijing, China, October 10–18, 2019 (New York, NY: IEEE), 251–256. doi:10.1109/ISMAR-Adjunct.2019.00-36

Wang, S., Chen, H., Pan, Z., and Wang, J. (2019b). A reconstruction method for missing data in power system measurement using an improved generative adversarial network. *Proc. Chin. Soc. Electr. Eng.* 39, 56–64. doi:10.13334/j.0258-8013.pcsee.181282

Wu, X., Xu, K., and Hall, P. (2017). A survey of image synthesis and editing with generative adversarial networks. *Tinshhua Sci. Technol.* 22 (6), 660–674. doi:10.23919/TST.2017.8195348

Yeh, R. A., Chen, C., Lim, T. Y., Schwing, A. G., Hasegawa-Johnson, M., and Do, M. N. (2017). "Semantic image inpainting with deep generative models," in 2017 IEEE conference on computer vision and pattern recognition (CVPR), Honolulu, HI, October 21–26, 2017 (New York, NY: IEEE), 6882–6890. doi:10.1109/CVPR.2017.728

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

frontiers in Energy Research

BRIEF RESEARCH REPORT
published: 06 April 2021
doi: 10.3389/fenrg.2021.644580

Rapid Recovery Control Method Based on Improved VDCOLs for Hybrid Multi-Infeed DC Transmission System After AC Failure

Chizu Mao[1], Xianchao Liu[2], Qing Li[1], Zhihua Xu[1], Yechun Xin[2] and Tuo Wang[2]*

[1] Maintenance and Test Center of CSG EHV Power Transmission Company of China Southern Power Grid Co., Ltd., Guangzhou, China, [2] Key Laboratory of Modern Power System Simulation and Control and Renewable Energy Technology, Ministry of Education, Northeast Electric Power University, Jilin, China

OPEN ACCESS

Edited by:
Liang Chen, Nanjing University of Information Science and Technology, China

Reviewed by:
Yanbo Chen, North China Electric Power University, China
Kenneth E. Okedu, National University of Science and Technology (Muscat), Oman
Shaoyan Li, North China Electric Power University, China

***Correspondence:**
Xianchao Liu
liuxianchao2017@aliyun.com

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

Received: *21 December 2020*
Accepted: *11 March 2021*
Published: *06 April 2021*

Citation:
Mao C, Liu X, Li Q, Xu Z, Xin Y and Wang T (2021) Rapid Recovery Control Method Based on Improved VDCOLs for Hybrid Multi-Infeed DC Transmission System After AC Failure. Front. Energy Res. 9:644580. doi: 10.3389/fenrg.2021.644580

Continuous commutation failure is very likely to occur in the hybrid Multi-infeed high-voltage direct current (HMIDC) after AC failure. In order to improve the recovery quality after HMIDC failure, an AC-DC voltage-dependent current order limiter (VDCOL) based on system strength index is proposed in this article. Firstly, the control mode transition process and system recovery process after DC failure are analyzed based on the hybrid multi-infeed DC transmission port model. Then, considering the impact of AC voltage and DC voltage input signals of VDCOL on AC voltage recovery and DC power recovery, respectively, the interaction factor and strength index of the hybrid multi-infeed system are constructed. Moreover, the weight coefficient of AC and DC voltage is calculated according to the strength of the multi-infeed system. Finally, a three-infeed hybrid DC transmission simulation model is built in the MATLAB/Simulink digital simulation platform. The simulation results demonstrate that the rapid recovery strategy proposed in this article can effectively suppress continuous commutation failure and improve the recovery speed of AC voltage and DC power.

Keywords: HMIDC system, commutation failure, system strength, VDCOL, control method

INTRODUCTION

High-voltage direct current (HVDC) transmission has gradually become the main method of long-distance, large-capacity, and cross-regional transmission with significant advantages of not being restricted by the system stability and lower cost compared with AC power transmission (Lee et al., 2020; Li et al., 2020). HVDC types include line-commutated converter-based high-voltage direct current (LCC-HVDC) and voltage source converter-based high-voltage direct current (VSC-HVDC) transmission according to commutation methods. The Hybrid multi-infeed direct current (HMIDC) transmission system is formed under multiple VSCs and LCCs with similar electrical distances being connected to the receiving grid. Combining the advantages of LCC and VSC, the HMIDC transmission system has a larger transmission capacity and more flexible operation mode (Li et al., 2018; Bakas et al., 2020). However, when the receiving power grid fails, LCC stations absorb a large amount of reactive power at the same time, and the infeed bus voltage with a short

electrical distance will drop quickly. If the AC commutation voltage drops to a certain value, the commutation margin will be insufficient leading to continuous commutation failure (Lin et al., 2020; Song et al., 2020). If the receiving grid can not recover from the commutation failure as soon as possible, it may cause a large area of power supply interruption and greatly affect the power supply reliability of the power grid. Therefore, it is necessary to design a rapid recovery strategy for HMIDC systems.

In order to investigate the recovery behavior of the HMIDC system after failure, the failure recovery process of the first commutation failure by using the steady-state operating curve of the DC system is analyzed in Hong et al. (2020b) and Liu et al. (2020a), where continue decreasing of the extinction angle is validated to be the essential reason for continuous commutation failure. At present, the voltage-dependent current order limiter (VDCOL) is one of the important control units in the HVDC system. Its main function is to limits the reactive power absorbed by the DC system by lowering the DC current order, which is suitable to improve the recovery speed of the feeding system (Nguyen et al., 2017; Huang and Wang, 2018). Zhang et al. (2020) analyzed the control effect of VDCOL under AC voltage and DC voltage drive modes, respectively, and proposed a VDCOL module considering AC voltage input signal conversion, in order to give consideration to the quality of converter bus and DC current recovery. However, the control conversion design of this method is relatively complex, and the improper setting of conversion criteria may be aggravated. Zeng et al. (2017) proposes a coordinated VDCOL control method of multi-infeed DC transmission based on a wide-area measurement system (WAMS). The impact of the fault on each DC transmission electronic system is evaluated based on the measurement results to adjust and control the voltage input signal of the DC subsystem VDCOL. However, this strategy has little consideration of the impact of voltage participating factor on the recovery process of different AC system strengths.

The existing literature mainly focuses on improving the recovery quality of the multi-infeed DC system by improving the input signal of VDCOL. However, the fast recovery control strategy considering the interaction among the infeed systems including VSC stations is urgently needed to be further studied. In order to improve the recovery characteristics of receiving power grids after failure, an AC-DC voltage-dependent current order limiter (VDCOL) based on system strength index in reference (Zeng et al., 2017) is further supplemented for HMIDC. The control mode transition process and recovery process of the HMIDC system after failure are analyzed. Considering the impact of infeed system voltage on AC commutation voltage recovery and DC power recovery, respectively, the relationship between the strength of the infeed system and the AC/DC voltage weight factor is fixed. The VDCOL driven by AC-DC voltage is designed according to the weight factor of infeed system voltage. The effectiveness of the proposed fast recovery strategy is validated by building a triple HMIDC simulation model in MATLAB/Simulink. The results show that the proposed method can avoid the subsequent commutation failure and improve the recovery quality to a certain extent by virtue of the voltage support capacity of the strong power grid and the reactive power regulation capacity of the VSC.

OPERATION PRINCIPLE OF HYBRID MULTI-INFEED DC TRANSMISSION SYSTEM

The HMIDC system refers to the network formed when VSC-HVDC and LCC-HVDC are connected with the same or similar electrical distance to the converter bus (Wei et al., 2020). For any hybrid multi-infeed DC transmission system, it is assumed that a total of N inverter stations are connected to the receiving AC grid, which contains k LCC inverter stations and N-k VSC inverter stations. The model of the hybrid multi-infeed DC system can be established based on the multi-port Davenan equivalent method, as shown in **Figure 1**.

Operating Principle of VDCOL

VDCOL mitigates the reactive power demand of converter bus by reducing the DC current order, The relationship between voltage and current is as follows (Jiang and Chiang, 2013):

$$I_d = \begin{cases} I_L, U_{in} \le U_L \\ \dfrac{I_H - I_L}{U_H - U_L} U_{in} + \dfrac{U_H I_L - U_L I_H}{U_H - U_L} U_{in}, U_L < U_{in} < U_H \\ I_H, U_{in} \ge U_H \end{cases} \tag{1}$$

where I_H, I_L, U_H, and U_L are the upper and lower limits of DC current and input voltage, respectively. The output current instruction value I_d of low-voltage current-limiting control mode depends on the input voltage detection value U_{in}.

Recovery Process HMIDC After Failure

LCC-HVDC inverter adopt DC current control, extinction angle control and current error control (CEC) (Hong et al., 2020a; Xiao and Li, 2020). In normal operation, the LCC inverter is in the state of extinction angle control. When a serious failure occurs, the transient operation process of the HMIDC system is roughly divided into three stages (Liu et al., 2020b), as shown in **Figure 2**.

(1) Start-up stage of CEC. The converter bus voltage of the LCC-HVDC inverter side drops significantly, causing commutation failure. The DC voltage drops and the DC current rises rapidly. The operating point A is shifted to the operating point B. The current deviation control of the LCC inverter is switched to current control, and operating point B moves to point B′.
(2) Start-up stage of VDCOL. VDCOL compulsively reduces the DC current order to control the power transmission, which results in operating point shifts from B to C. The inverter recovers from commutating failure to realize normal commutating. The DC voltage increases, and the operating point moves to point C′. At the same time, the system operates stably under the state of "low voltage and small current" until the fault is removed.

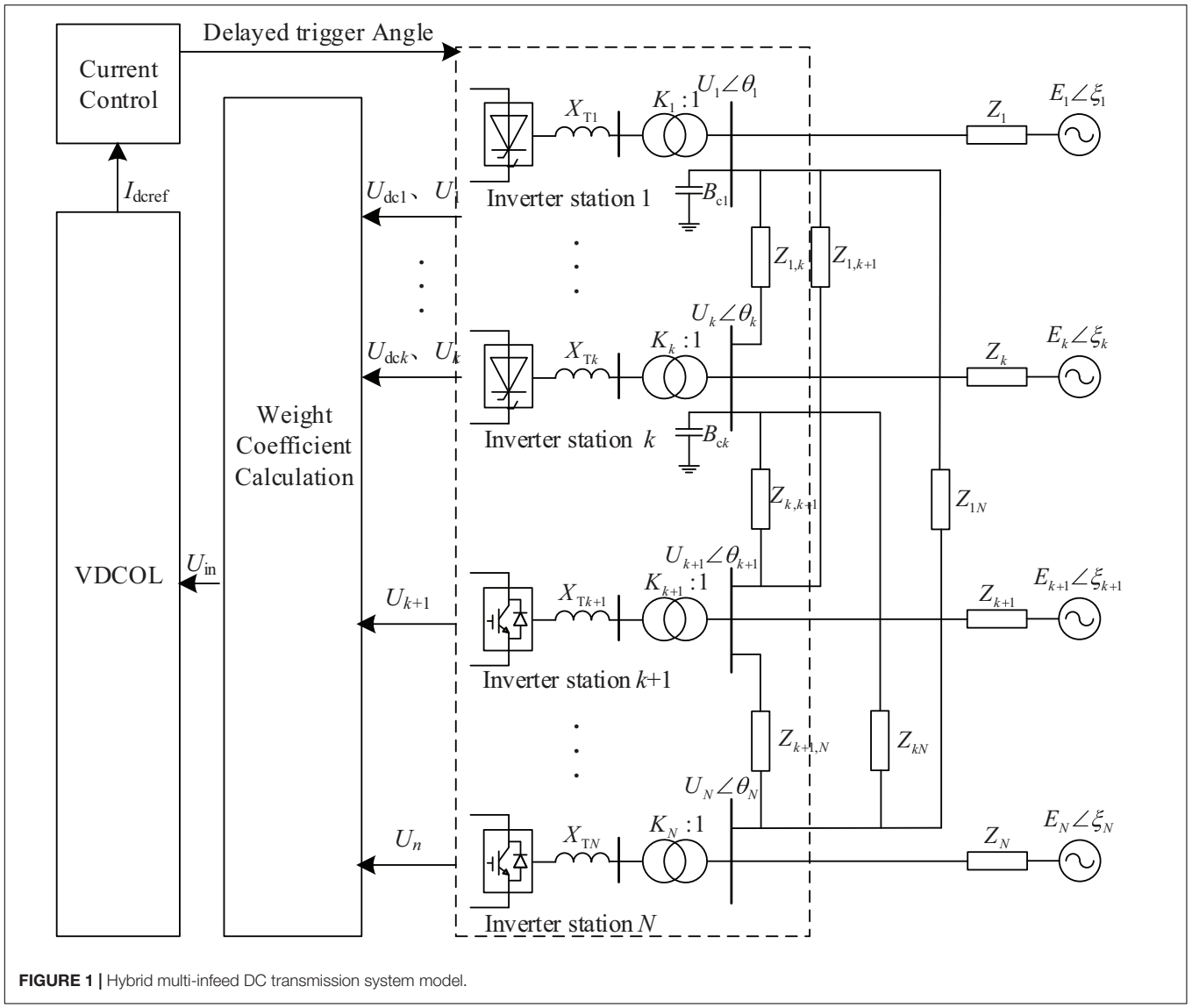


FIGURE 1 | Hybrid multi-infeed DC transmission system model.

(3) CEC conversion stage. When the system operating point C moves so that the current deviation value is less than the set threshold, the current deviation control converted from the current control to extinction angle control, and finally stabilized at operating point D.

IMPROVED VDCOL RAPID RECOVERY CONTROL METHOD

According to the impact of VDCOL on the stability of different infeed systems, VDCOL with mixed input of DC voltage and AC commutation voltage is constructed by collecting voltage data of converter bus. The AC/DC voltage is weighted and used as the input of VDCOL so that DC power recovery not only depends on the DC voltage but also takes into account the voltage recovery of the AC system at the receiving grid. The improved VDCOL control framework is shown in **Figure 1**.

Improved VDCOL Input Signal

For the LCC converter, the collected converter bus voltage and DC voltage are weighted by the signal calculation to obtain the nominal value of the input signal of VDCOL. The calculation formula is as follows:

$$\begin{cases} U_{in\_i} = \xi_{i0} U_{dci} + \sum_{j=1}^{k} \xi_{ij} U_{ac_i} \\ \sum_{j=0}^{k} \xi_{ij} = 1 \end{cases} \tag{2}$$

where U_{in_i} represents the input voltage signal of the VDCOL of the i-th terminal LCC converter; $\underline{U_{dci}}$ represents the DC voltage of the i-th terminal LCC converter; U_{ac_j} represents the AC voltage of j-th terminal LCC converter; ξ_{i0} represents DC voltage weighting coefficient of the i-th terminal LCC converter; ξ_{ij} represents AC voltage weighting

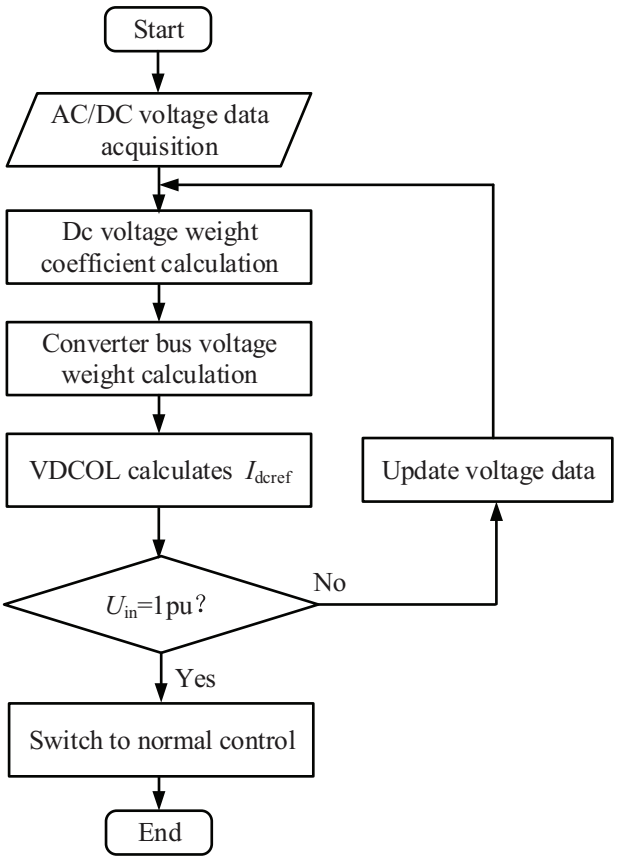


FIGURE 2 | Diagram of VDCOL operation principle.

coefficient of the *j*-th terminal LCC converter; and *k* is the number of LCC.

DC Input Signal Weight Coefficient Calculation

In the recovery process after fault removal, the HMIDC system not only needs rapid recovery of DC power but also needs to consider the degree of disturbance to the voltage stability of the receiving AC grid. The smaller the strength of the receiving grid is, the greater effect of low-voltage current limitation. To this end, the generalized short-circuit ratio (MIESCR) of all hybrid multi-infeed DC systems is established as follows (Liu and Chen, 2013; Xiao et al., 2020):

$$MIESCR_i = \frac{S_{Ni} - Q_{Fi}}{P_{Ni} + \sum_{\substack{j=1 \\ j \neq i}}^{n} MIIF_{ji} P_{Nj}} \tag{3}$$

where S_{Ni} is the short-circuit capacity of the *i*-th terminal receiving grid; Q_{Fi} is the reactive power compensation capacity of the *i*-th terminal receiving grid; and P_{Ni} is the rated power of the *i*-th LCC-HVDC.

$MIIF_{ji}$ is a generalized multi-feed interaction factor, which is used to evaluate the extent of voltage interaction between DC converter stations (Denis and Göran, 2013; Xiao et al., 2017). It is defined as follows: when a symmetric three-phase reactor is put into the converter bus *i*, resulting in a 1% decrease in the voltage on the converter bus *j*, the voltage change rate on the converter bus *j*.

$$\mathbf{M}IIF_{\mathbf{ji}} = \frac{\triangle \mathbf{U_j}}{\mathbf{1\%U_i}} \tag{4}$$

$MIIF_{ji}$ reflects the effect degree of the change of the AC voltage of bus *i* on bus *j*.

The larger *MIESCR* is, the greater receiving AC system strength is, and the change of converter bus voltage or that of the converter bus voltage near DC system has little impact on system stability. Therefore, change of DC voltage can be taken as the main determinant. DC voltage weight coefficient constructed is as follows:

$$\xi_{i0} = \begin{cases} \left| \dfrac{MIESCR_i - CSCR}{CSCR} \right|, 0 < MIESCR_i < 2CSCR \\ 1, MIESCR_i \geq 2CSCR \end{cases} \tag{5}$$

where ξ_{i0} represents DC voltage weighting coefficient of the *i*-th terminal LCC; $CSCR_i$ represents the critical short circuit ratio of the *i*-th terminal LCC.

Combined with Eqs 3, 4, it can be seen that all AC voltage weight coefficients are expressed as follows:

$$\xi_{aci} = 1 - \xi_{i0} \tag{6}$$

where ξ_{aci} represents the weighting coefficients of the *i*-th terminal LCC commutation voltage.

As is known from Eqs 5,6, DC voltage weighting coefficient represents the intensity risk degree of the receiving system. The higher the system's strength security is, the less the impact of the recovery process on the stability of the AC system, and the DC voltage and active power can be the priority recovery targets. On the contrary, the voltage recovery of the AC system should be given priority. The DC voltage weight coefficient of VDCOL is positively correlated with the MIESCR.

AC Voltage Input Signal Weight Coefficient Calculation

In the HMIDC system, DC transmission subsystems with similar electrical distances have strong interaction. In case of failure, the degree of converter bus drop is closely related, which can also affect the fault recovery characteristics of the MIDC system. Therefore, it is necessary to add the participation factor of the commutator bus into the AC input signal to represent the coupling interaction between the multi-infeed DC systems.

The coupling interaction between the DC receiving grids affects the AC voltage recovery speed of the converter bus. If the converter bus *i* are greatly influenced by bus *j*, the strength $MIESCR_j$ of the *j*-th terminal receiving system is small, and the weight coefficient of AC commutation voltage U_j of VDCOL at the *i*-th terminal is larger, which can restore the voltage of the *j* bus stably and quickly restore DC power of the *i*-th transmission, improve the frequency stability of the AC system, and reduce

power failure loss. Therefore, the voltage weight coefficient of the converter bus should be as follows:

$$\frac{\xi_{i1}}{MIESCR_1} = \frac{\xi_{i2}}{MIESCR_2} = \cdots = \frac{\xi_{ik}}{MIESCR_k} \quad (7)$$

HMIDC Fast Recovery Control Method Steps

In conclusion, the framework of the HMIDC quick recovery control method can be established in this work. The proposed coordinated control method of VDCOL in the hybrid multi-infeed DC system is summarized as follows:

(1) Collect the converter bus voltage and DC voltage at LCC-HVDC and MMC-HVDC.
(2) Calculate the DC voltage weight coefficient and AC weight coefficient of each infeed system.
(3) Calculate the AC voltage weight coefficient of each infeed system.
(4) VDCOL works and outputs the DC current reference value I_{dcref}.
(5) Input DC current reference value I_{dcref} into current control, and switch to normal control when $U_{in} = 1$p.u.

CASE STUDY

Simulation Parameter

In order to verify the effectiveness of the proposed HMIDC fast recovery control method, the electromagnetic simulation model of the HMIDC system is built in the MATLAB/Simulink, including two LCC-HVDC systems and one MMC-HVDC system. In stable operation, current control (2 kA) is adopted for the LCC-HVDC rectifier station, and extinction angle control (18°) is adopted for the inverter side. The MMC-HVDC rectifier station adopts active power (1000 MW) and reactive power control (0 MVar), while the inverter adopts DC voltage (500 kV) and reactive power control (0 MVar). The main parameters are shown in **Table 1**.

The two LCC-HVDC and MMC-HVDC are denoted as the first, second and third DC infeed subsystems, respectively, and the equivalent impedance is as follows: $Z_1 = 5.5 + j20.43\Omega$, $Z_2 = 4.7 + j17.6\Omega$, and $Z_3 = 5.08 + j18.92\Omega$; The contact impedance is: $Z_{12} = 3.8 + j47\Omega$, $Z_{23} = 0.25 + j3.14\Omega$, and $Z_{31} = 1.3 + j15.7\Omega$. The effective short-circuit ratio of hybrid multi-feed of the three receiving systems can be calculated based on the above parameters: $HMIESCR_1 = 2.60$, $HMIESCR_2 = 2.32$, and $HMIESCR_3 = 2.16$.

TABLE 1 | Main parameters of HMIDC system.

System parameter	Rectifier side	Inverter side
AC system voltage/kV	525	525
Rated frequency /Hz	50	50
Rated active power/MW	1000	1000
Transformer ratio /kV	450/200/200	450/200/200
Transformer capacity /MVA	1200	1200
Current limiting reactor/H	0.35	0.35
Rated DC voltage/kV	500	
Rated DC current/kA	2	
DC line resistance/Ω	3.5	

According to the above calculation, the input signals of two VDCOL of LCC-HVDC inverter station can be obtained as follows:

$$\begin{aligned} U_{in_1} &= 0.7U_{d1} + 0.1U_{a1} + 0.1U_{a2} + 0.1U_{a3} \\ U_{in_2} &= 0.55U_{d1} + 0.17U_{a1} + 0.15U_{a2} + 0.13U_{a3} \end{aligned} \quad (8)$$

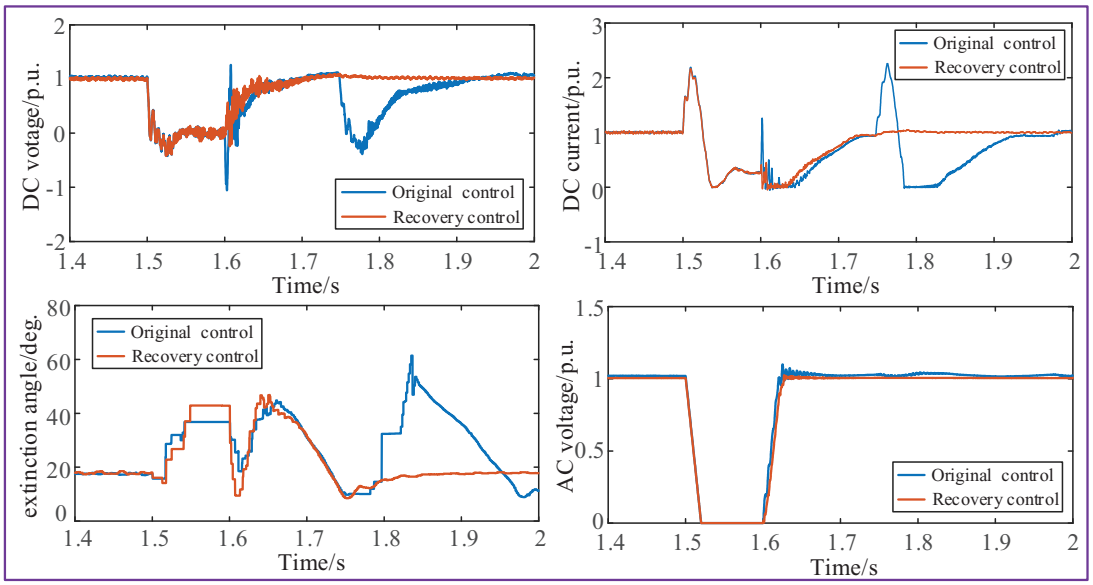
Three-Phase Short Circuit Simulation Verification

When $t = 1.5$ s, the three-phase short circuit fault is set at the first LCC-HVDC converter bus, and the fault lasted for 0.1 s. Through comparative analysis, the simulation results of DC voltage VDCOL and the mixed input VDCOL proposed verified the effectiveness of the rapid recovery control. The simulation results are shown in **Figure 3**.

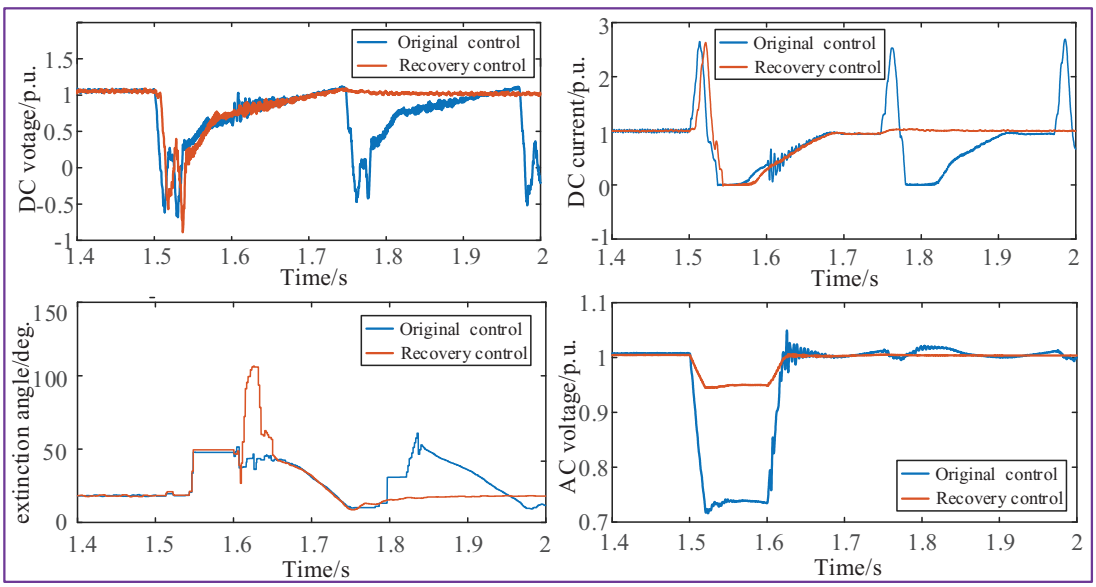
Figure 3A shows the simulation results of DC voltage, DC current, extinction angle, and inverter converter bus voltage of subsystem 1. The converter bus voltage drops to 0 rapidly when three-phase short-circuit fault occurs ($t = 1.5$ s), and the DC current rises extremely fast. At this time, the current deviation control is started and converted from extinction angle control to current control. Under the action of two kinds of VDCOL strategy for current limit, DC current declined at about $t = 1.53$ s, commutation failure has happened once at this point. When the three-phase short circuit fault was cleared at $t = 1.6$ s, infeed system 1 began to recover. However, due to the rapid rise of the DC current, the active power of the system is restored quickly, and the reactive power demand soars, and the commutation failure of the feed system occurs again at $t = 1.75$ s for the original control strategy. While the proposed control method includes the local AC voltage into the input signal, which obviously limits the rise rate of the DC current. It can be seen from **Figure 3A** that the proposed VDCOL has a faster recovery speed and smaller overshoot than the original control system. There were continuous commutation failures ($\gamma_{min} = 7°$) at the extinction angle in the scenario with no recovery control, while the addition of recovery control avoided subsequent commutation failures in **Figure 3A**.

Due to the interaction between subsystem 2 and subsystem 1, when a three-phase short-circuit fault occurs to the converter bus of subsystem 1, the bus voltage will drop to a certain value, the DC voltage will drop, and the DC current will rise at a high speed. At this time, the current deviation control will be started, and the switch from extinction angle control to current control. It can be seen from **Figure 2B** that the DC current recovery speed of the proposed VDCOL is almost the same as that of the control system without recovery. When $t = 1.75$ s, the subsequent commutation failure occurred in infeed system 2 under the original control method. On the one hand, the VDCOL makes the DC current recover quickly and increases the demand for reactive power. On the other hand, the SCR of infeed system 2 is small, and the poor AC voltage support capacity of the commutator bus. Under

A Simulation result of infeed subsystem 1 (LCC-HVDC)



B Simulation result of infeed subsystem 2 (LCC-HVDC)



C Simulation result of infeed subsystem 3 (MMC-HVDC)

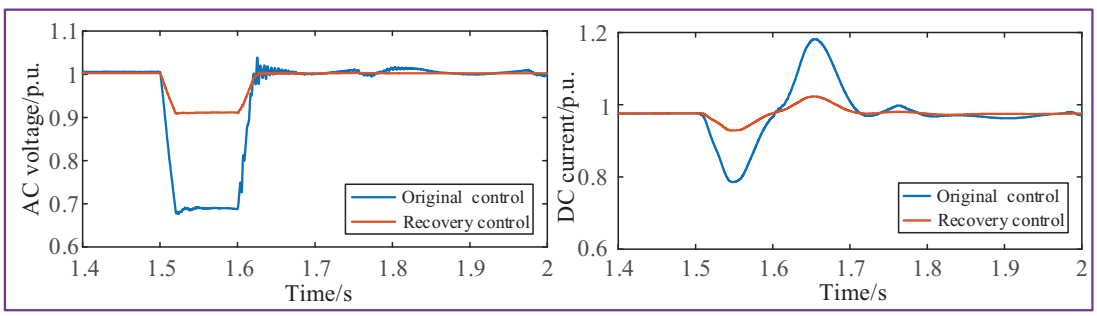


FIGURE 3 | Simulation results of HMIDC after failure. **(A)** Simulation result of infeed subsystem 1 (LCC-HVDC). **(B)** Simulation result of infeed subsystem 2 (LCC-HVDC). **(C)** Simulation result of infeed subsystem 3 (MMC-HVDC).

the proposed recovery control method, the voltage recovery of the infeed system 1 with large SCR and VSC suppressed the DC current mutation of the infeed system 2, thus avoiding the subsequent commutation failure.

The simulation results of DC voltage, DC current, and inverter side converter bus voltage of subsystem 3 can be seen in **Figure 3C**. When the subsystem 1 fault occurs ($t = 1.5$ s), the voltage on the inverter side of MMC-HVDC decreases, and the transmission power decreases as well, while the rectifier is still in active power control, leading to the rise of DC voltage and the decrease of DC current. The recovery control improves DC current fluctuation degree of MMC-HVDC through the interaction between multi-infeed systems and has a positive effect on the stability of AC and DC systems.

In conclusion, The HMIDC rapid coordinated recovery method based on the improved VDCOL links the multiple DC transmissions together, makes full use of the voltage support capacity of the strong power grid and VSC, and makes the incoming power grid orderly recover the DC current and commutation bus voltage, and reduces the subsequent commutation failure.

CONCLUSION

In this article, a rapid recovery control strategy based on improved VDCOL is proposed for HMIDC. The simulation results of the three-terminal hybrid multi-infeed DC system show that the improved VDCOL rapid recovery control method can avoid the subsequent commuting failure to a great extent, and at the same time improve the recovery speed and quality of the multi-infeed system compared with the traditional VDCOL. Especially for the short circuit relatively small power grid such as the China Guangdong power grid area, the improved VDCOL rapid recovery coordination control will greatly enhance the disturbance resistance ability of regional weakly infeed power grid and improve the recovery characteristics after failure.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

CM: conceptualization. XL: methodology. QL: software. ZX: validation. YX: formal analysis. TW: data curation. All authors contributed to the article and approved the submitted version.

FUNDING

This research was financially supported by the Science and Technology Project of China Southern Power Grid (CGYKJXM20180388).

REFERENCES

Bakas, P., Okazaki, Y., Shukla, A., Patro, S. K., Ilves, K., Dijkhuizen, F., et al. (2020). Review of hybrid multilevel converter topologies utilizing thyristors for HVDC applications. *IEEE Transactions on Power Electronics* 36, 174–190. doi: 10.1109/TPEL.2020.2997961

Denis, L. H. A., and Göran, A. (2013). Analysis of voltage and power interactions in multi-infeed HVDC systems. *IEEE Transactions on Power Delivery* 28, 816–824. doi: 10.1109/TPWRD.2012.2227510

Hong, L., Zhou, X., Liu, Y., Haitao, X., Hanhang, Y., Yandong, C., et al. (2020a). Analysis and improvement of the multiple controller interaction in LCC-HVDC for mitigating repetitive commutation Failure. *IEEE Transactions on Power Delivery* 2, 1.

Hong, L., Zhou, X., Liu, Y., Xia, H., Yin, H., Chen, Y., et al. (2020b). Analysis and improvement of the multiple controller interaction in LCC-HVDC for mitigating repetitive commutation Failure. *IEEE Transactions on Power Delivery* 2, 1. doi: 10.1109/TPWRD.2020.3017802

Huang, M., and Wang, J. (2018). Review on optimization control strategy for voltage dependent current order limiter. *Guangdong Electric Power* 31, 18–27.

Jiang, N., and Chiang, H. D. (2013). Energy function for power system with detailed DC model: construction and analysis. *IEEE Transactions on Power Systems* 28, 3756–3764. doi: 10.1109/TPWRS.2013.2265402

Lee, G., Kwon, D., and Moon, S. (2020). Method for determining the droop coefficients of hybrid multi-terminal HVDC systems to suppress AC voltage fluctuations. *IEEE Transactions on Power Systems* 35, 4944–4947. doi: 10.1109/TPWRS.2020.3014805

Li, G., An, T., Liang, J., Jun, L. W., Joseph, T., Jingjing, L., et al. (2020). Power reversal strategies for hybrid LCC/MMC HVDC systems. *CSEE Journal of Power and Energy Systems* 6, 203–212.

Li, Y., Li, Y., Li, G., Zhao, D., and Chen, C. (2018). Two-stage multi-objective OPF for AC/DC grids with VSC-HVDC: Incorporating decisions analysis into optimization process. *Energy* 147, 286–296. doi: 10.1016/j.energy.2018.01.036

Lin, S., Liu, J., Liu, L., Thunderstorm, and Fu, C. (2020). *A review of commutation failure suppression methods for HVDC systems based on control protection measures. Proceedings of the CSEE*. 1–16. [2020-09-12]. http://kns.cnki.net/kcms/detail/11.2107.TM.20200410.1109.026.html

Liu, X., Wang, Z., Zheng, B., et al. (2020a). Mechanism analysis and mitigation measures for continuous commutation failure during the restoration of LCC-HVDC. *Proceedings of the CSEE* 40, 31631–33171.

Liu, X., Zengping, W., Bowen, Z., et al. (2020b). Mechanism analysis and mitigation measures for continuous commutation failure during the restoration of LCC-HVDC. *Proceedings of the CSEE* 40, 31631–33171.

Liu, Y., and Chen, Z. (2013). A flexible power control method of VSC-HVDC link for the enhancement of effective short-circuit ratio in a hybrid multi-infeed HVDC System. *IEEE Transactions on Power Systems* doi: 10.1109/TPWRS.2012.2222057

Nguyen, M. H., Saha, T. K., and Eghbal, M. (2017). Master self-tuning VDCOL function for hybrid multi-terminal HVDC connecting renewable resources to a large power system. *IET Generation Transmission & Distribution* 11, 3341–3349. doi: 10.1049/iet-gtd.2017.0094

Song, J., Li, Y., Zeng, L., and Zhang, Y. (2020). *A review of commutation failure for LCC-HVDC. Automation of Electric Power Systems*. 1–12. [2020-09-12]. http://kns.cnki.net/kcms/detail/32.1180.TP.20200718.1726.006.html

Wei, Z., Fang, W., and Jun, L. (2020). Variable extinction angle control strategy based on virtual resistance to mitigate commutation failures in HVDC system. *IEEE Access* 8, 1–10. doi: 10.1109/ACCESS.2020.2994245

Xiao, H., and Li, Y. (2020). Multi-infeed voltage interaction factor: a unified measure of inter-inverter interactions in hybrid multi-infeed HVDC systems.

IEEE Transactions on Power Delivery 35, 2040–2048. doi: 10.1109/TPWRD.2019.2960393

Xiao, H., Li, Y., He, X., et al. (2017). A rapid assessment method of commutation failure immunity level for hybrid multi-infeed HVDC transmission systems. *Proceedings of the CSEE* 37, 4986–4998. (in Chinese). doi: 10.1049/joe.2017.0424

Xiao, H., Zhang, Y., Duan, X., and Li, Y. (2020). Evaluating Strength of Hybrid Multi-Infeed HVDC Systems for Planning Studies Using Hybrid Multi-Infeed Interactive Effective Short-Circuit Ratio. *IEEE Transactions on Power Delivery* 2, 1. doi: 10.1109/TPWRD.2020.3020957

Zeng, Q., Li, X., Feng, M., Li, B., and Zeng, X. (2017). Coordinated control method of VDCOLs in MIDC system based on the wide-area measurement system. *High Voltage Engineering* 043, 1168–1174.

Zhang, W., Xiong, Y., Li, C., Yao, W., Wen, J., and Gao, D. (2020). Continuous commutation failure suppression and coordinated recovery of multi-infeed DC system based on improved VDCOL. *Power System Protection and Control* 63–72.

Conflict of Interest: CM, QL, and ZX were employed by Maintenance Test Center of EHV Transmission Company.

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

frontiers in Energy Research

ORIGINAL RESEARCH
published: 14 April 2021
doi: 10.3389/fenrg.2021.659154

Triple-Mode Grid-Balancing Plants via Biomass Gasification and Reversible Solid-Oxide Cell Stack: Economic Feasibility Evaluation via Plant Capital-Cost Target

Yumeng Zhang[1,2], *Ningling Wang*[2], *Chengzhou Li*[1,2], *Mar Pérez-Fortes*[3], *Liqiang Duan*[2], *Jan Van herle*[3], *François Maréchal*[4], *Tzu-En Lin*[5], *Ligang Wang*[1,3,4]* and *Yongping Yang*[6]

[1] *Innovation Research Institute of Energy and Power, North China Electric Power University, Beijing, China,* [2] *National Research Center for Thermal Power Engineering and Technology, North China Electric Power University, Beijing, China,* [3] *Group of Energy Materials, Swiss Federal Institute of Technology in Lausanne, Lausanne, Switzerland,* [4] *Industrial Process and Energy Systems Engineering, Swiss Federal Institute of Technology in Lausanne, Lausanne, Switzerland,* [5] *Institute of Biomedical Engineering, National Chiao Tung University, Hsinchu, China,* [6] *Key Laboratory of Power Station Energy Transfer Conversion and System (North China Electric Power University), Ministry of Education, Beijing, China*

OPEN ACCESS

Edited by:
Chao Long,
Cranfield University, United Kingdom

Reviewed by:
Haoshui Yu,
Massachusetts Institute of Technology, United States
Rui Jing,
Cardiff University, United Kingdom

***Correspondence:**
Ligang Wang
ligang.wang@ncepu.edu.cn

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

Received: *27 January 2021*
Accepted: *12 March 2021*
Published: *14 April 2021*

Citation:
Zhang Y, Wang N, Li C, Pérez-Fortes M, Duan L, Van herle J, Maréchal F, Lin T-E, Wang L and Yang Y (2021) Triple-Mode Grid-Balancing Plants via Biomass Gasification and Reversible Solid-Oxide Cell Stack: Economic Feasibility Evaluation via Plant Capital-Cost Target. Front. Energy Res. 9:659154. doi: 10.3389/fenrg.2021.659154

Electricity production and consumption must be balanced for the electrical grid. However, the rapidly growing intermittent power sources are now challenging the supply-demand balance, leading to large flexibility needs for grid management. The plant integrating biomass gasification and reversible solid-oxide cell stacks can be potential means of flexibility, which could flexibly switch among power generation, power storage, and power neutral modes. This paper investigates the economic feasibility of such grid-balancing plants, i.e., plant capital expenditure (CAPEX) target, *via* a systematic overall decomposition-based methodology for real geographical zones and flexibility-need scenarios. The plant CAPEX target (€/ref-stack) is defined as the maximum affordable investment cost for each reference stack (active cell area 5,120 cm^2). The results show that, for a 5-year payback time, 5-year stack lifetime, and 40 €/MWh grid balancing price, the plant concept with 10–100 MWth gasifier has high economic potential with target reaching 17,000 €/ref-stack; however, the plant concept with 100–1,000 MWth gasifier has a limited commercialization potential with the target reaching below 1,000 €/ref-stack due to high biomass supply costs. Considering the sale of chemical product, plant CAPEX target can reach up to 22,000 and 3,000–12,000 €/ref-stack for the plants with 10–100 and 100–1,000 MWth, respectively. The plant CAPEX target is decreased by increasing the total capacities of all plants deployed since more and more capacities will be put into power neutral mode (isolated from the electrical grid) *via* the coordination of multiple plants. The plant CAPEX target can be further increased by higher grid up/down-regulating price and longer payback years.

Keywords: biomass-to-chemical, grid balancing, biomass-to-electricity, reversible solid-oxide cell, economic feasibility, plant capital expenditure target

1. INTRODUCTION

Sustainable renewable energy sources are urgently required to satisfy the rapidly growing energy demand and limit greenhouse gas emissions. The penetration of renewables in the global electricity supply has reached a record of 27% in 2019 and is expected to be 49% in 2030 (IEA, 2020). This rapid growth is largely contributed by wind and solar power (IRE Agency, 2014). The high penetration of the intermittent renewables challenges the electricity market in terms of supply-demand balance, transient and frequency stability, thus the Transmission System Operators (TSOs) will require large flexibility needs for grid management. In Denmark, with a 10 GW offshore wind plant built, the grid flexibility needs will vary between a power shortage of 8 GW to a power surplus of 20 GW (Energinet, 2020).

There is a portfolio of supply- and demand-side options for advanced TSOs evolving toward high flexibility by means of, e.g., flexible power generators as a capacity reserve, demand-side management, cross-region interconnections, and crucial alternative energy storage (Koltsaklis et al., 2017). Particularly, when the penetration of renewable power becomes high enough that energy is no longer a limiting factor, while thermal power plants and nuclear power plants will still give a firm supply and not be dispatched frequently, there will be a considerable power surplus (Energinet, 2020). The excess electricity could be addressed by energy storage technologies, which store excess energy and release it when needed. A renewable energy penetration of up to 50% can be addressed in Texas by employing a storage capacity of average daily energy demand, which is 15% higher than only employing flexible power generators as reserve (Denholm and Hand, 2011).

Various energy storage technologies are available. Currently, physical storage options, including pumped-hydro storage and thermal storage, are dominating, i.e., 50% of the global flexibility needs (European Commission, 2016); however, they are generally limited by geographical and environmental restrictions, or low energy conversion efficiency (Frate et al., 2021). Electrochemical storage options, particularly lithium batteries and flow batteries, have gained large development and reached commercialization. Batteries can participate in primary frequency response and also energy management (Lee et al., 2019), expected to account for 25% of the 2030 flexibility needs in Belgium (Elia, 2019). However, large-scale, long-term energy management may be better handled by electrolysis-based power-to-hydrogen technologies, which convert excess electricity to hydrogen and its derivatives. The stored chemical energy can be converted back to electricity *via* fuel-to-power. This energy storage-release cycle is named as power-to-x-to-power (PXP). It has been expected that over 20% of Danish electricity will be converted to hydrogen or hydrogen-based fuels in 2030 for the down-regulation of the electrical grid (Energinet, 2020), while the fuels generated can be converted to electricity for the up-regulation of the grid, or injected to chemical and transportation sectors.

Economic flexibility of power-to-hydrogen system for down-regulation of the grid has been investigated in the literature, which mainly focuses on the systems enabled by proton exchange membrane or alkaline electrolyzers (Jorgensen and Ropenus, 2008; Paulus and Borggrefe, 2011; Kroniger and Madlener, 2014; Guinot et al., 2015; Grueger et al., 2017). The power-to-hydrogen systems can hardly be economically viable due to (1) low hydrogen production efficiency of 65% (lower heat value, LHV) (Grueger et al., 2017), (2) high investment cost 3,400 €/kW (Jorgensen and Ropenus, 2008; Guinot et al., 2015), and (3) expensive hydrogen storage (Kroniger and Madlener, 2014). Those power-to-hydrogen systems for down-regulation of the grid can further attach a hydrogen-to-power system for up-regulation of the grid *via* fuel cells. However, using proton exchange membrane fuel cell for Germany secondary control reserve market is also not economically potential (Grueger et al., 2017), because (1) the threshold of hydrogen price is calculated to be 1.1 €/kg, which is far from the current market price (around 5 €/kg; Apostolou, 2020); and (2) the annual equivalent full-load hours are as low as 150 h, due to the expensive bidding balance service using hydrogen.

The investment costs of PXP are mainly contributed by the fuel cell and electrolyzer stacks, which could be largely reduced by employing solid-oxide technology. The same solid-oxide stack can operate "reversibly" by switching between fuel cell mode and electrolysis mode. Thus, the solid-oxide cell can also be called reversible solid-oxide cell (RSOC) and enables a single RSOC-based plant to work for power generation (PowGen) or power storage (PowSto) at different time periods. This reversible operation of the same plant can potentially (1) enhance the utilization hours, (2) reduce the total capital expenditure (CAPEX), (3) achieve high round-trip efficiency of 55–70% (LHV) (Butera et al., 2020; Mogensen, 2020; Motylinski et al., 2021), and (4) utilize hydrocarbons as fuels directly, such as methane (Mogensen, 2020) and methanol (Butera et al., 2020) to avoid expensive hydrogen storage. Additional CO_2 sources needed for the case of hydrogen derivatives can be solved by combining the biomass with RSOC systems.

Using renewable, carbon-neutral, and widely available biomass, RSOC-based plants are potential to provide grid balancing services, which is called as Waste2GridS (Waste2GridS, 2020) (W2G) plants. In the literature, only a few case studies of such a plant concept can be found, e.g., RSOC-based plants for power generation and power storage fueled by wood chips (Sigurjonsson and Clausen, 2018) or municipal solid wastes (Perna et al., 2018) with an improvement in Butera et al. (2020), where four operating modes are proposed to cope with different grid flexibility needs. These studies focus mainly on the thermodynamic perspective. The authors have recently proposed a triple-mode plant concept for grid balancing in Wang L. et al. (2020) by connecting the sectors of electrical grid, gas grid, and chemical market (**Figure 1**). The plant concept integrates biomass/waste gasification and RSOC technology, and can interact with or be isolated from the electrical grid as capacity reserves, thus is capable of non-stop operation all year round. The optimal conceptual plant design has been investigated by focusing on the thermodynamic performances, leading to an application-free pool of optimal plant designs.

This paper is a follow-up of our previous study (Wang L. et al., 2020). The objective is to evaluate in a systematically manner from both the technological and application viewpoints,

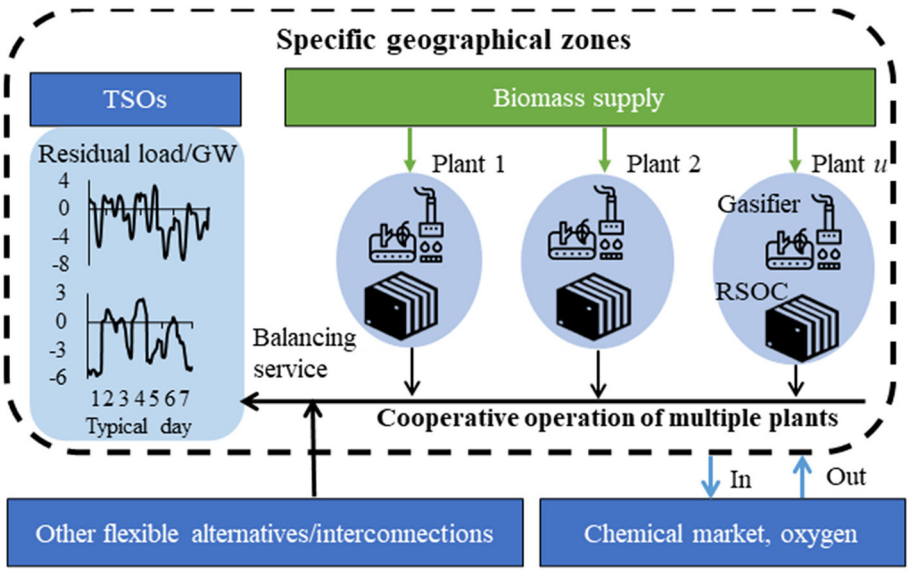


FIGURE 1 | Grid balancing services provided by the W2G concept.

the economic feasibility of the W2G plants in 2030, with a decomposition-based optimization method. The maximum affordable plant CAPEX, named as plant CAPEX target, is evaluated by considering (1) different conversion technology combinations, (2) different geographical zones, (3) different magnitudes of grid flexibility needs, (4) a number of plant design alternatives, (5) optimal plant sizing and cooperative operation of multiple plants, and (6) minimum cost for biomass supply chain. Sensitivity analysis of key influencing factors are also performed, including (1) grid regulating prices, (2) payback time of capital expenditure, and (3) revenue from chemical product sale.

The remaining paper is organized as follows: In section 2, the W2G plant is introduced with the description of plant concept. Then, method for evaluating the economic feasibility of W2G plants is described in section 3 with the overall method introduced first followed by detailed mathematical formulations. In section 4, the applications of W2G plants for addressing grid flexibility needs are specified. Economic feasibility of W2G plants in applications is discussed comprehensively in section 5. Finally, conclusions are drawn in section 6.

2. CONCEPT OF THE TRIPLE-MODE GRID-BALANCING PLANT

The generic plant concept proposed in our previous work (Wang L. et al., 2020) is illustrated in **Figure 2**. The plant is comprised of the processes of onsite biomass pretreatment, gasification, syngas cleaning, RSOC stacks, methanation, heat exchanger network, and steam turbine network for heat recovery. The biomass is pretreated and gasified to produce syngas, which is further cleaned in the cleaning section. The clean syngas is fed to the RSOC-methanation subsystems for electricity or methane production.

The biomass pretreatment, gasification, and syngas cleaning sections are considered to operate at full load all the time with no load shifting, so that the same amount of clean syngas can be produced continuously for the RSOC-methanation subsystems. The mode switch of the plant is enabled by the change of operating modes of the RSOC-methanation subsystems. Specifically, the RSOC subsystems involves two RSOC blocks, whose coordination as shown in **Figure 2** enables the switch among the PowGen, PowSto, and PowNeu modes:

- PowGen mode: biomass-to-electricity with both RSOC blocks under the fuel cell mode, offering up-regulation service.
- PowSto mode: biomass-to-chemicals with both RSOC blocks under electrolyzer mode powered by electricity, offering down-regulation service.
- PowNeu mode: biomass-to-chemical with one RSOC block under fuel cell mode to power the other block under electrolyzer mode for chemical production. The plant is not interacted with the electrical grid.

The thermodynamic performance is affected by multiple degrees of freedom of plant designs, including (1) the combination of gasification technology (entrained flow gasifier or circulating fluidized bed gasifier), syngas cleaning technology (hot/cold), electrolysis mode (steam electrolysis or co-electrolysis), (2) the design operating points of the key components, particularly the stacks (under both fuel cell and electrolyzer modes), (3) optimal heat cascade utilization, and (4) the optimal placement of steam turbine network. Two gasifier options were considered: entrained-flow gasifier with direct heating (EFG) for large-scale applications (100–1,000 MWth), and fast internally circulating fluidized-bed (FICFB) for medium-scale applications (10–100 MWth).

These degrees of freedom of process design have been considered in the optimal conceptual plant design performed

in Wang L. et al. (2020), which employed a well-established multi-period, multi-objective process optimization method. The optimization considers three objective functions: the PowGen efficiency, the PowSto efficiency, and the specific cell area needed to process 1 kW-LHV dry basis biomass. A pool of optimal Pareto designs is obtained, revealing the trade-off between the three objectives. Thus, for different plant designs, the plants with the same stack size can provide up- and down-regulation services at different capacities.

3. METHOD FOR IDENTIFYING POTENTIAL CASE STUDIES

3.1. Overall Method

For the deployment of W2G plants, a case study must based on a specific geographical zone to consider realistic (or reasonably predicted) grid flexibility needs and biomass availability. However, it is difficult to simultaneously consider the non-linear programming for optimal conceptual plant design and the mixed-integer programming for optimal plant scheduling to cope with a specific imbalance profile, not even to mention the computation-expensive supply chain optimization. Thus, an overall decomposition-based methodology has been initially proposed by Wang L. et al. (2020) and is slightly adapted as shown in **Figure 3**. The method involves the following seven steps with detailed described in Wang L. et al. (2020):

- Step 1: Assessment of future grid flexibility needs of the specific geographical zones, referring to Olsen et al. (2020),
- Step 2: Assessment of future biomass availability (i.e., a geodatabase with biomass weight, characteristics and location coordinates) of the specific geographical zones, referring to Laínez et al. (2011), Pérez-Fortes et al. (2014), and Puigjaner (2011),
- Step 3: Generation of plant design pool and preselection of plant design, referring to Wang L. et al. (2020),
- Step 4: Optimal design selection, and plant sizing and scheduling to maximize the profits from grid-balancing services when addressing the grid flexibility needs identified in step 1,
- Step 5: Minimization of biomass supply chain cost given the plant sizes resulted from step 4,
- Step 6: Evaluation of the economic indicator, i.e., Plant CAPEX Target, by considering the profits (step 4) and costs (step 5),
- Step 7: Business case evaluation with actual plant CAPEX evaluation to determine the economic prerequisites for potential business cases.

This work focuses on the step 4–6, i.e., red-colored task in **Figure 3** to calculate the plant CAPEX target, so that potential case studies can be listed. The step 4 performs (1) optimal design selection from design pool, (2) optimal sizing and scheduling (i.e., mode switch) of the plants deployed, and (3) optimal capacity and scheduling of the auxiliaries, particularly the oxygen tank. Particularly, the deployment of multiple plants are considered to increase the profits from the grid balancing services. Thus, it concludes the grid-balancing profits, and the corresponding design, size, and operation of each plant. Then, for each plant with its size obtained, the costs of biomass supply chain are minimized in step 6 to enhance the economic feasibility and determine the plant location. Then, the maximum affordable plant investment cost, i.e., the Plant CAPEX Target (€/ref-stack) defined as the maximum affordable total plant investment costs divided by the equivalent number of reference stacks (ref-stack, each with 5,120 cm^2 active cell area) of all plants deployed in a case study, can be calculated for a payback time of l years:

$$\begin{aligned} &\textit{Plant CAPEX target}\,(l) \\ &= \frac{\textit{Profit}\,(l) - \textit{biomass supply chain cost}\,(l)}{\textit{Total number of reference stacks of all plants installed}}. \end{aligned} \quad (1)$$

The plant CAPEX target is an economic index of the W2G plant, including the investment costs of components employed in the RSOC blocks and biomass gasification. Those case studies with high plant CAPEX target are considered to be economically potential, and each case study provides exact information on geographical zones (grid flexibility needs and biomass availability), the plants deployed (design, location, size, and operation), and biomass supply chain.

3.2. Mathematical Formulations

The mathematical formulations for step 4, optimal plant design selection and optimal plant sizing and scheduling, are presented in this section. To reduce the computation efforts, the hourly imbalance profiles with 8,760 points are processed to typical days.

3.2.1. Objective Function

The profit of the W2G plants deployed for l payback years (Equation 2) is associated with (1) the revenue from grid-balancing services $R^{\text{bal}}_{d,i}$, €, (2) additional revenue (positive) or cost (negative) of oxygen trade with the chemical market $R^{\text{oxy}}_{d,i}$, €, and (3) the cost of oxygen storage tank R^{tank}, €:

$$\textit{Profit}\,(l) = \sum_{l} \frac{\sum_{d,i} \alpha_d \left(R^{\text{bal}}_{d,i} + R^{\text{oxy}}_{d,i}\right)}{\left(1+k\right)^{l}} - R^{\text{tank}} \quad (2)$$

where α_d is the repetition time of the typical day d in an entire year, i is the hour in a day, and r is the interest rate (0.05).

The revenue from grid-balancing services $R^{\text{bal}}_{d,i}$ can be calculated by the balancing energy $E_{u,d,i,n}$ (MWh) and the regulating price θ^{bal}_{n} (€/MWh) for both PowGen and PowSto modes:

$$R^{\text{bal}}_{d,i} = \sum_{u,n} \theta^{\text{bal}}_{n} E_{u,d,i,n} \quad (3)$$

where n represents PowGen, PowSto, or PowNeu mode, u is the index of plant design in the design pool.

The additional oxygen revenue is calculated based on the trade with market, outflow $F^{\text{out}}_{td,i}$, inflow $F^{\text{in}}_{td,i}$, kg/h, and the oxygen price θ^{oxy}, €/kg:

$$R^{\text{oxy}}_{d,i} = \theta^{\text{oxy}} \left(F^{\text{out}}_{d,i} - F^{\text{in}}_{d,i}\right) \quad (4)$$

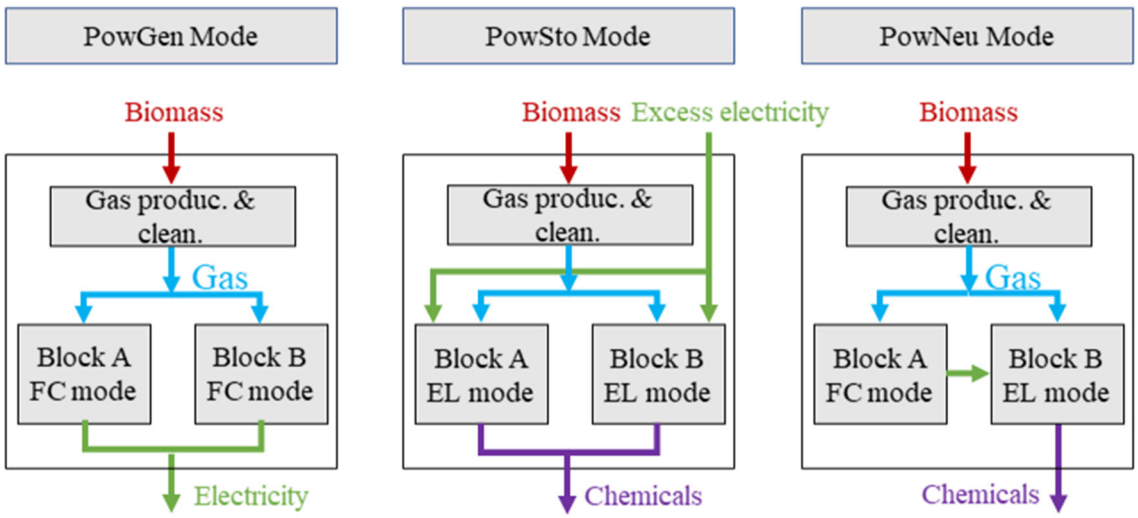


FIGURE 2 | The generic concept of W2G plant (Wang L. et al., 2020). The chemical produced in the PowSto and PowNeu modes is synthesis natural gas, due to the existing large infrastructure.

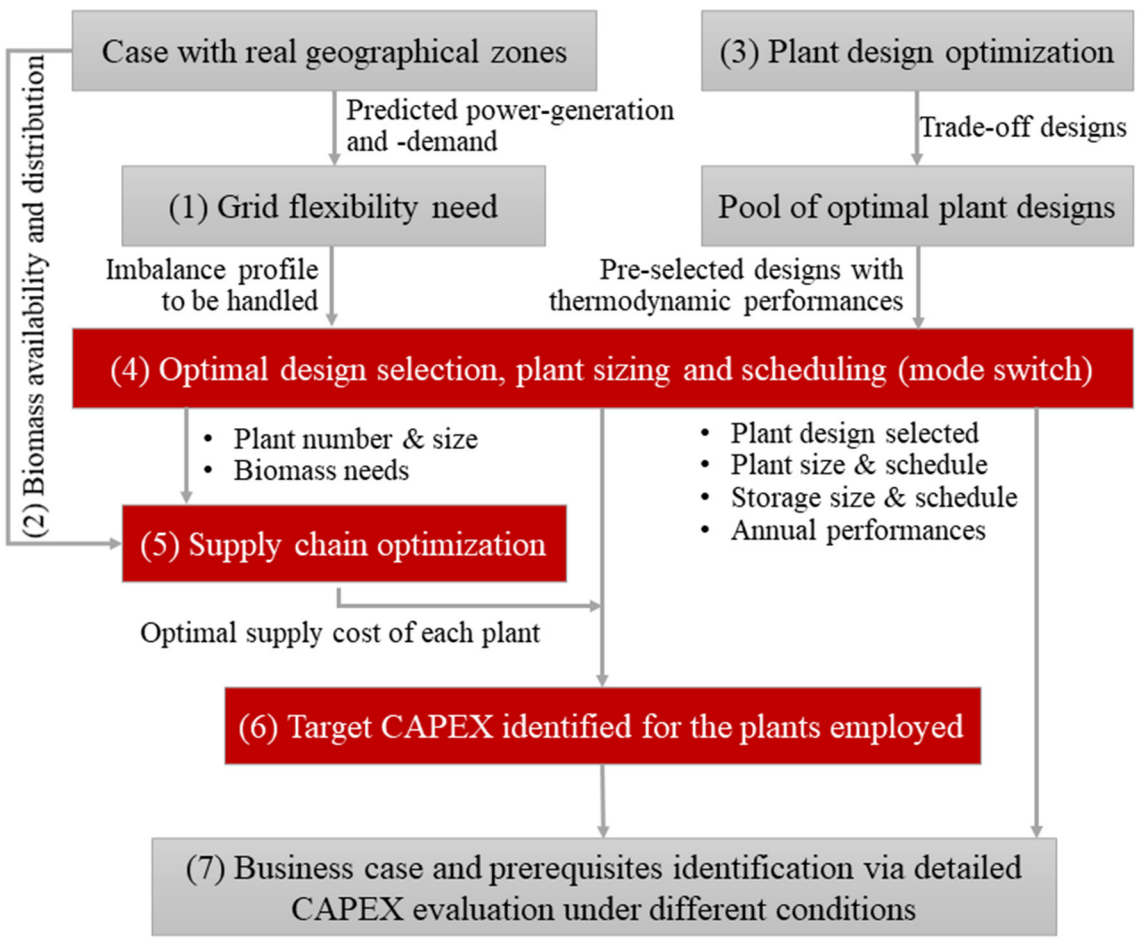


FIGURE 3 | The decomposition-based method to identify potential business cases for the triple-mode grid-balancing plants, adapted from Wang L. et al. (2020).

The cost of oxygen storage tank $R^{\text{tank}}$is evaluated by the tank capacity m (kg) and the tank price θ^{tank} (€/kg):

$$R^{\text{tank}} = \theta^{\text{tank}} m \tag{5}$$

3.2.2. Constraints

The grid balancing power provided by the plants deployed at each time step (d, i) is calculated based on the mode and capacity of each plant, and only mode switch but no load shifting is considered:

$$P_{u,d,i,n} = Y_u r_u \dot{P}_{u,n} Z_{u,d,i,n} \tag{6}$$

where $\dot{P}_{u,n}$ (MW) is the capacity of the plant using design u in mode n (PowGen or PowSto), the binary variable Y_u represents whether the design u is selected or not, r_u is the sizing factor scaling from the reference size of the design to the size of the actual plant deployed, and the binary variable $Z_{u,d,i,n}$ represents the status of the PowGen and PowSto modes. The coordination among multiple plants is considered by employing multiple plants U:

$$\sum_u Y_u = U \tag{7}$$

The grid balancing power provided by the plants in the PowGen and PowSto modes should be lower than the up-regulation and down-regulation needs $\hat{P}_{d,i,n}$ (MW):

$$\sum_u P_{u,d,i,n} \leq \hat{P}_{d,i,n} \tag{8}$$

The status of each plant is unique in a time step (d, i):

$$\sum_n Y_u Z_{u,d,i,n} = 1 \tag{9}$$

The oxygen storage level $\bar{m}_{d,i}$ (kg) in the gas tank is limited by the tank capacity m (kg):

$$\bar{m}_{d,i} \leq m \tag{10}$$

The storage level of the tank in time (d, i) is related to the the level in previous hour $(d, i-1)$ and the oxygen generated/consumed by plants ($F^{\text{gen}}_{d,i}$, $F^{\text{con}}_{d,i}$, kg/h), and the trade with the chemical market ($F^{\text{in}}_{d,i}$, $F^{\text{out}}_{d,i}$, kg/h):

$$\bar{m}_{d,i} = \bar{m}_{d,i-1} + F^{\text{gen}}_{d,i} - F^{\text{con}}_{d,i} + F^{\text{in}}_{d,i} - F^{\text{out}}_{d,i} \tag{11}$$

The oxygen production and generation in time step (d, i) are calculated based on the design value $\dot{m}$, kg/h, and the sizing factor:

$$F^{\text{gen}}_{d,i} = \sum_{u,n} \dot{m}^{\text{gen}}_{u,n} Y_u Z_{u,d,i,n} r_u \tag{12}$$

$$F^{\text{con}}_{d,i} = \sum_{u,n} \dot{m}^{\text{con}}_{u,n} Y_u Z_{u,d,i,n} r_u \tag{13}$$

The storage levels in the first and the last hour in each typical day are set to be equal for continuous operation of the storage tanks:

$$\bar{m}_{d,i=1} = \bar{m}_{d,i=24} \tag{14}$$

4. APPLICATION

Existing wind and solar extreme zones (with a high penetration of variable renewable energy sources) in Europe (e.g., Denmark and Italy) are identified to provide typical grid flexibility needs. The zones in-and-around the electricity market zones DK-DK1 (northern Denmark) and IT-SUD (southern Italy), as shown in **Figure 4A**, are selected for deploying the overall methodology (section 3) to determine the economic feasibility of the W2G plants. The W2G plants are expected to provide up-regulation under the PowGen mode (biomass-to-electricity), down-regulation under the PowSto mode (biomass-to-chemical), or reserve capacity under the PowNeu mode (biomass-to-chemical). All evaluations are for the timeline 2030.

4.1. Grid Flexibility Needs

The theoretical flexibility needs in 2030 of the two zones have been predicted in Olsen et al. (2020) based on historical data and available forecast data by using the multi-timescale data-driven method. The hourly profiles (365 days, 8,760 values) are clustered into a set of typical days using the *k*-means method developed in Fazlollahi et al. (2014) to reduce the number of data points and hence the computational load. The theoretical flexibility needs of DK-DK1, IT-SUD in 2030 represented by typical days are shown in **Figure 4B**, which vary within -6–2 GW and -7–3 GW.

Real grid flexibility needs to be coped by the W2G plants will be smaller or even much smaller than the theoretical flexibility needs, due to multiple flexibility means. In 2030, cross-country interconnections are expected to play an important role in coping with the flexibility needs, while conventional technologies will not be completely replaced by new technologies, as agreed by several TSOs (European Commission, 2016; Elia, 2019; Energinet, 2020). Unfortunately, Terna (the Italian TSO) and Energinet (the Danish TSO) have not provided their official prediction for the real flexibility needs for 2030.

The grid flexibility needs addressed by the W2G plants are thus scaled from the theoretical flexibility needs, with the sizing factors taken from the Adequacy and Flexibility Study of Belgian TSO (Elia, 2019), which predicts the 2030 flexibility needs of Belgium. It is reported that in 2030, 34% of the up-regulation and 32% of the down-regulation can be handled by cross-country interconnections. Considering also batteries and classical thermal power plants, only 14% of the up-regulation and 30% of the down-regulation may be handled by the W2G plants and other competing technologies. Thus, two scenarios are defined to represent the flexibility needs for the W2G technology:

- S1: excluding interconnections: 66% of theoretical up-regulation needs and 68% of theoretical down-regulation needs,
- S2: excluding interconnections, batteries, and classic plants: 14% of theoretical up-regulation needs and 30% of theoretical down-regulation needs.

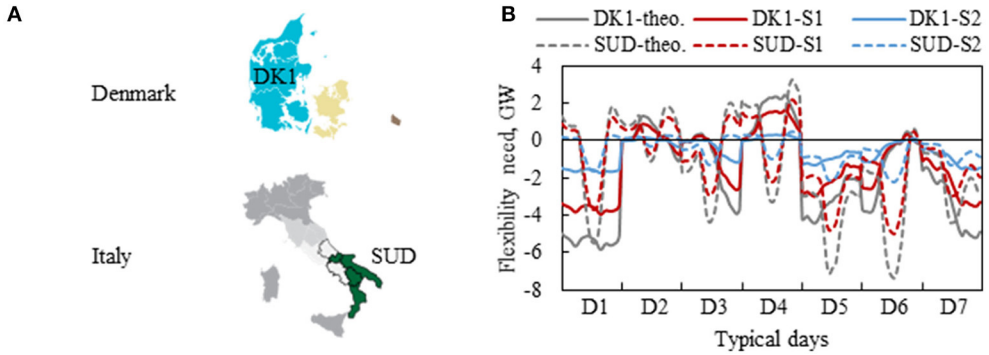


FIGURE 4 | The grid flexibility needs: **(A)** Geographic locations and **(B)** theoretical and scaled flexibility needs.

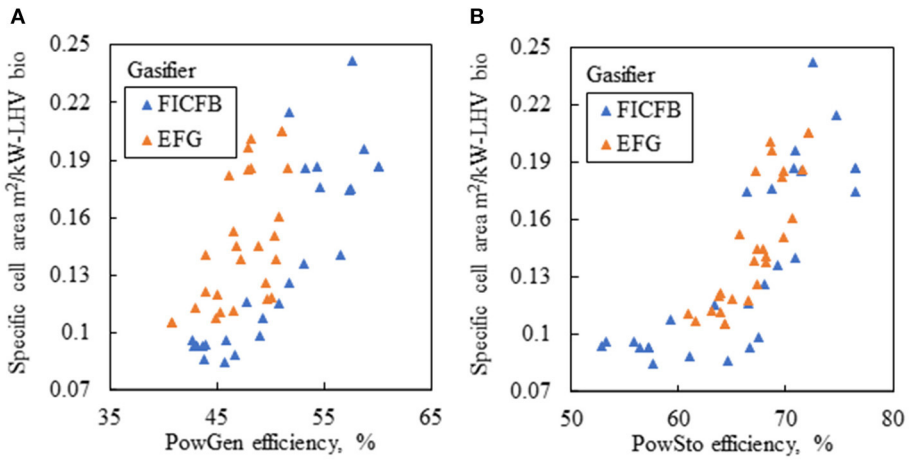


FIGURE 5 | The pool of plant design with FICFB and EFG option: **(A)** PowGen efficiency vs. specific cell area **(B)** PowSto efficiency vs. specific cell area.

The theoretical flexibility needs and the two scenarios defined have been given in **Figure 4B**. The scenarios S1 and S2 will result in the deployment of W2G plants at different scales. Considering the economy of scale, a capacity factor x is introduced to set a basis for the comparison of the deployment under the two scenarios. It represents the contribution of the W2G plants installed to address the target flexibility needs:

$$x = \frac{\sum_{n,u} Y_u r_u \dot{P}_{u,n}}{\sum_n max(\hat{P}_{d,i,n})}, \tag{15}$$

with the symbols introduced for Equation (6).

4.2. Biomass Availability

Available biomass in-and-around the specific zones are assessed considering the resources of the zones. In IT-SUD, the main available biomass comes from agriculture residues, forestry, and municipal solid waste (organic, paper, and wood). The amount of the available biomass in 2030 is predicted by a geodatabase built based on the future Europe directives in The European Parliament and The Council Of The European Union (2018), with energy (LHV) evaluated as 19 TWhth in 2030. In DK-DK1, agricultural residues, forestry, and municipal solid waste organic fractions are available, which are able to supply 22 TWhth in 2030 referring to work in S2Biom project (S2Biom, 2016). The detailed distributions of each type of biomass are fed to the optimization of biomass supply chain.

4.3. Plant Design

The 27 plant designs considered in step 4–6 are preselected from the Pareto-front obtained in our previous study (Wang L. et al., 2020). These designs consider the trade-off between specific cell area m^2/kW-LHV biomass, PowGen efficiency and PowSto efficiency, as shown in **Figure 5**.

The characteristics of each design used for optimal deployment (step 4 in **Figure 3**) include:

- Specific PowGen capacity: Net electricity generation, kWe/ref-stack,
- Specific PowSto capacity: Total electricity consumption, kWe/ref-stack,
- Oxygen consumed/generated in each mode, (kg/h)/ref-stack,
- Synthesis natural gas (SNG) produced in each mode, (kg/h)/ref-stack,

- Biomass energy need, in MWth/ref-stack.

4.4. Key Assumptions

The key economic parameters specified in the optimal plant deployment are given in **Table 1**.

TABLE 1 | Parameters specifications.

Parameters	Description	Fixed value	References
$\theta^{bal}_{PowGen}/\theta^{bal}_{PowSto}$	Regulating price, €/MWh	40	Skytte, 1999
$\theta^{oxy,in}/\theta^{oxy,out}$	Oxygen price, €/kg	0.1/0.06	Bellotti et al., 2017
θ^{tank}	Oxygen tank price, €/kg	8	Wang C. et al., 2020

5. RESULTS AND DISCUSSION

The economic feasibility of the W2G plants is evaluated with the plant CAPEX target (€/ref-stack), with an overview given first and a subsequent investigation of the effects of the influential factors.

5.1. Overview of the Plant CAPEX Target

The plant CAPEX targets (Equation 1) with respect to the capacity factor x (Equation 15) are shown in **Figure 6** for a payback time of 5 years, a stack lifetime of 5 years, different flexibility-need scenarios and plant technologies. The longest acceptable duration of the W2G plants is set as 5 years, which approaches the stack lifetime of cutting-edge technologies (SOLIDpower, 2017). Overall, the FICFB-based W2G plants with typical gasifier size of 10–100 MWth biomass feed with the installed capacity factor x below 8% show high economic-potential, with the plant CAPEX target reaching over 16,000 €/ref-stack. While the EFG based W2G plant with typical gasifier size of 100–1,000 MWth biomass feed can hardly be economically feasible due to the low and even negative plant CAPEX target. The FICFB-based W2G plant achieves a plant CAPEX target of 2,000–17,000 €/ref-stack, which is significantly higher than that of the EFG based plant (i.e., -5,000 to 1,000 €/ref-stack). This is mainly due to the profit from grid-balancing services and the costs of biomass supply, which will be further analyzed below.

The plant CAPEX target decreases with the increase in the capacity factor x, i.e., an increase in the plant capacities for a given imbalance to be handled. Particularly, when the capacity factor of the FICFB-based W2G plant is around 3%, the plant CAPEX target can be up to 17,000 €/ref-stack for the case study SUD-S1. Increasing the capacity factor to 10%, the plant CAPEX target is sharply decreased to 5,000 €/ref-stack.

For the plants with the same gasification technology, there is no big difference in plant CAPEX target among different magnitudes of grid flexibility needs (S1 and S2) and different zones (DK-DK1 or IT-SUD). For the FICFB-based W2G plant, the variation of plant CAPEX targets among different cases with similar sizes of plants deployed is less than 1,000 €/ref-stack.

Referring to the definitions of plant CAPEX target (Equations 1 and 2), the major influential factors leading to such observations are (1) the total size of all plants deployed, represented by the total number of reference stacks employed, (2) the revenue from grid balancing services, (3) the costs of onsite oxygen storage, (4) the cost/revenue from oxygen trade with the chemical market, and (5) the costs of biomass supply. The effects of these factors will be further discussed below.

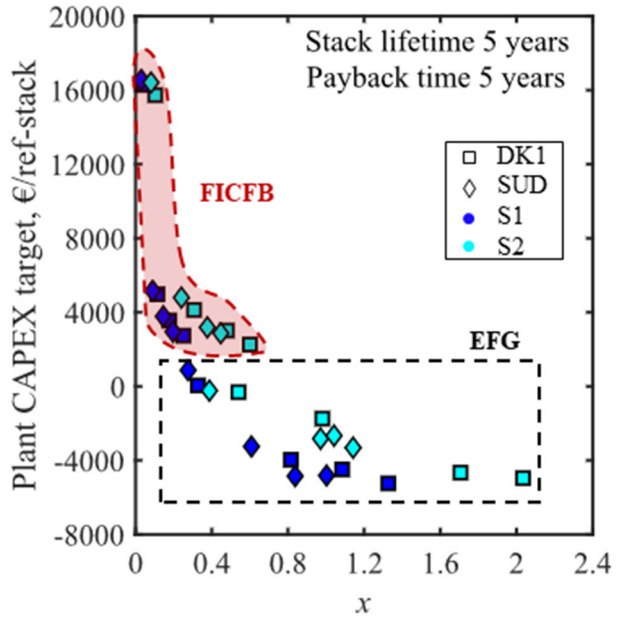


FIGURE 6 | Plant CAPEX target of the case studies with different plant technologies, geographical zones, and flexibility-need scenarios. Note the profit from methane sale is not considered in this figure. The reference stack is defined as a stack with a total active cell area of 5,120 cm^2.

5.2. Plant Operation

The W2G plants switch between PowGen and PowSto modes to interact with the grid by providing or consuming electricity. Since the W2G plant is set to operate full load under all modes, when the grid flexibility needs are lower than the PowGen or PowSto capacities of a plant, the plant will be put under PowNeu mode. The optimal operating strategy of the W2G plants for SUD-S2 is taken as an example as shown in **Figure 7**. The maximum flexibility needs are 450 MWe for up-regulation and 2,200 MWe for down regulation.

For FICFB-based W2G plant (**Figure 7A**), the power capacity of the plant is limited to 10–100 MWth biomass feed. When employing a single plant, the optimal plant capacity is with a gasifier capacity of 100 MWth providing 60 MWe for the PowGen mode and 160 MWe for the PowSto mode. Such power capacities are far below the grid flexibility needs, with the capacity factor of only 3%. Thus, the plant is rarely put into PowNeu mode due to the constraint of grid flexibility needs (Equation 8).

For EFG-based W2G plant (**Figure 7B**) with a capacity of 100–1,000 MWth biomass feed, seven plants can provide 760 MWe for the PowGen mode and 2,270 MWe for the PowSto mode,

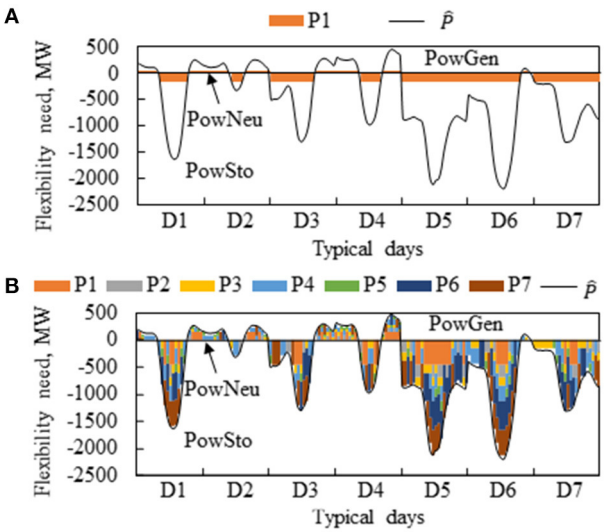


FIGURE 7 | Dispatch of the plant deployed to cope with the flexibility needs in SUD-S2: **(A)** single FICFB-based plant, **(B)** seven EFG-based plants.

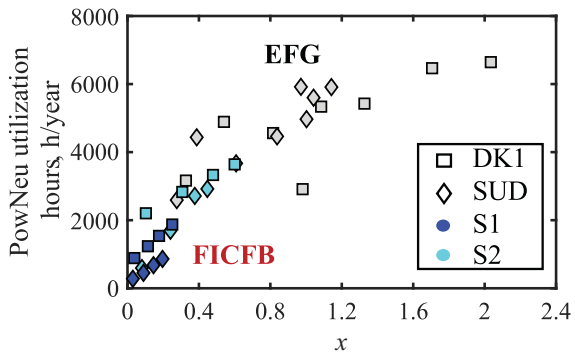


FIGURE 8 | Average annual utilization hours of the PowNeu mode with the case of FICFB-based plants highlighted.

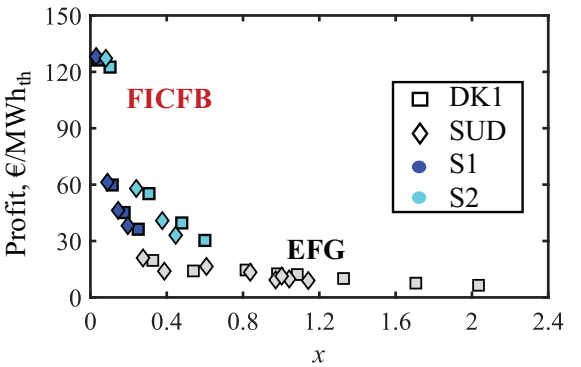


FIGURE 9 | Maximum profit of W2G plants under different scenarios.

reaching a capacity factor of up to 110%. Over 90% grid flexibility needs are covered by the cooperative operation of plants, as shown in **Figure 7B**; however, this coordination also increases the operating hours of PowNeu mode.

The utilization hours of the PowNeu mode increase with the increasing capacity factor x, i.e., the increased capacity of all plants deployed. This remains the same for different geographical zones (DK1 and SUD) and magnitudes of flexibility needs (S1, S2) as shown in **Figure 8**. For the FICFB-based plant, the annual utilization hour of the PowNeu mode of scenario S1 is lower than that of scenario S2. For the scenario S1, the annual utilization hours of the PowNeu mode are within 300–1,900 h with capacity factors in 3–25%; Under the scenario S2 with less flexibility needs, the capacity factors are increased to 8–60%, resulting in an annual PowNeu operation of 600–3,600 h. For the EFG-based plant, the annual PowNeu utilization hours reach over 2,500 h and even up to 6,500 h, because larger EFG-based plants allow for coping with larger flexibility needs and thus higher profit, with the capacity factor achieving 27–200%.

5.3. Profit

The profits of the W2G plants defined in Equation (2) are affected by (1) the revenue from providing grid flexibility, (2) the cost/revenue from oxygen trade, and (3) the cost of oxygen storage tank. The maximum profits under different geographical zones and magnitudes of flexibility needs are illustrated in **Figure 9**. Increasing the total capacities of all plants deployed with an increase in the number of plant, the maximum profit obtained by 1 MWhth biomass will be decreased. This is mainly due to the high utilization of PowNeu mode under high capacity factor. High annual PowNeu utilization hours indicate low utilization of the PowSto and PowGen modes. This results in a reduced profit, due to that (1) there is no extra grid balancing gain under the PowNeu mode and (2) the PowNeu mode reaches a lower efficiency than the PowSto and PowGen modes.

The profit of the FICFB-based plants is from 40 €/MWhth (capacity factor 25% for DK1 and 20% for SUD) to 130 €/MWhth (capacity factor 4% for DK1, 3% for SUD) of scenario S1, and from 30 €/MWhth (capacity factor of 60% for DK1, and 45% for SUD) to 130 €/MWhth (capacity factor 10% for DK1 and 8% for SUD) from scenario S2. The revenue from addressing the flexibility needs contributes over 75% of the profits, with the remaining part contributed by oxygen sale. Increasing the capacity factor, the contribution of oxygen sale is increased from 6 to 25% due to the utilization of PowNeu mode. For the FICFB-based plant, the plants have oxygen generation on the PowNeu mode. For the EFG-based plant, the profits are lower, 6–21 €/MWhth due to high utilization hours of PowNeu mode as shown in **Figure 8**, and higher oxygen consumption of the PowNeu mode.

5.4. Biomass Supply Chain Cost

The costs of biomass supply chain, considering biomass collection from geographical zones, pre-treatment, and storage, should be minimized. The superstructure-based method

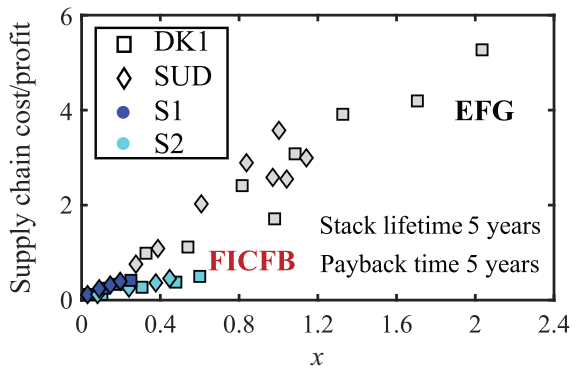


FIGURE 10 | The proportion of biomass supply chain cost on profit with the FICFB-based plant highlighted.

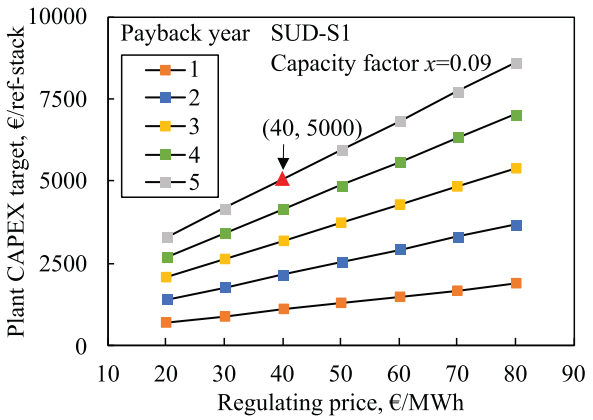


FIGURE 11 | Plant CAPEX target with different payback years and regulating prices for the cases of SUD-S1.

presented in Puigjaner (2011), Laínez et al. (2011), and Pérez-Fortes et al. (2014) is adopted. Biomass is required all the time as long as the W2G plants are under operation, based on the plant concept in **Figure 2**. For the FICFB-based plant, the costs of biomass supply chain account for 10–50% of the profits as shown in **Figure 10**, supplying 1–6 TWhth biomass a year. For the larger EFG-based plant, the annual biomass needs are sharply increased to 5–37 TWh/year, costing at least 75% of the profit and even five times more than the profit. Such big biomass supply costs make the EFG-based plants hardly economically-feasible. When the biomass collection is over 10 TWh/year, biomass production costs, i.e., dry and storage, will be significantly increased, since biomass with a lower LHV and high moisture content has to be chosen. Moreover, when biomass need cannot be satisfied by the available waste in the zones considered, i.e., 19 TWhth/year in DK-DK1 and 22 TWh/year in IT-SUD, additional biomass has to be collected from the neighboring zones with potentially high transportation costs.

5.5. Sensitivity Analysis

This section presents a sensitivity analysis with respect to grid regulating price, payback time, and the revenue from SNG sale. Those parameters are critical for determining the economic feasibility.

5.5.1. Regulating Price and Payback Time

The plant CAPEX target is evaluated with a up/down-regulating price of 20–80 €/MWh (Skytte, 1999) and a payback time of 1–4 years as shown in **Figure 11** for the case SUD-S1 with a capacity factor of 9%.

The plant CAPEX target decreases if reducing the up/down-regulating prices or the payback time. It may reach 8,600 €/ref-stack with 5 payback years if the regulating price is up to 80 €/MWh. While, if the regulating price is down to 20 €/MWh, the plant CAPEX target should be lower than 700 €/ref-stack to have the chance of being profitable for a 1-year payback time. The plant CAPEX target evaluated with a payback time of 1–4 years is affected by the oxygen tank costs incurred in the first year and the sum of the revenue from providing grid balancing service occurring before reaching the set payback year. With a grid balancing price of 20 €/MWh, the plant CAPEX targets are reduced from 3,300 €/ref-stack for a payback time of 5 years to 1,400 €/ref-stack for a payback time of 2 years. However, even under the most potential situation (80 €/MWh, five payback years), the cases of employing three plants is less economically-competitive than those with one single plant.

5.5.2. Revenue From SNG Sale

The revenue from the sale of SNG generated during the PowSto and PowNeu modes will further enhance the economic feasibility of the case studies, as shown in **Figure 12**. The FICFB-based plant achieves a higher plant CAPEX target than the EFG option under the same scenario. The plant CAPEX target is in 8,000–22,000 €/ref-stack for the FICFB-based plant but is limited to 12,000 €/ref-stack for the EFG-based plant, which can even be down to 3,000 €/ref-stack. The plant CAPEX target also decreases as the increase of plant capacities deployed. There is still no big difference in the plant CAPEX target among different zones and different flexibility-need scenarios when using the same gasifier type and plant number.

6. CONCLUSIONS

The economic feasibility of a novel grid-balancing plants integrating biomass gasification and reversible solid-oxide cell stacks is evaluated by the plant CAPEX target, defined as the maximum affordable plant investment cost (€/ref-stack, a reference stack with 5,120 cm^2 active cell area). A decomposition-based methodology is implemented to calculate the economic indicator considering different grid flexibility needs, biomass availability, and plant designs. The profit from providing grid-balancing services is maximized by cooperating multiple plants, which vary in plant design, size, and operation. Further, with the minimization of biomass supply chain, the plant CAPEX target is carried out for a set of cases in different geographical

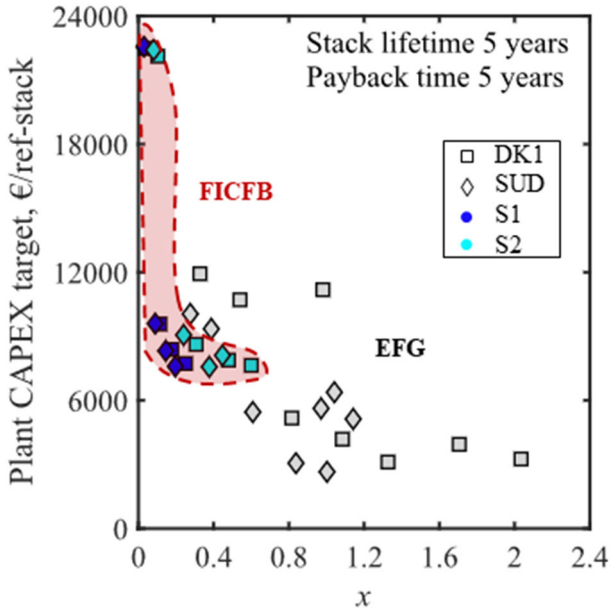


FIGURE 12 | Plant CAPEX target recalculated by considering the revenue from SNG sale to market, considering the price of SNG being 0.6 €/kg (IEA, 2019).

zones (DK1, SUD) and flexibility-need scenarios. The key conclusions are:

- The plants with the size of 10–100 MWth are potential to realize high economic feasibility with the plant CAPEX target reaching up to 17,000 €/ref-stack (5-year payback time, 5-year stack lifetime, and 40 €/MWh regulating price). Counting the chemical-sale profit, plant CAPEX target can even reach 22,000 €/ref-stack.
- The plants with the size of 100–1,000 MWth are much less economically feasible due to the high biomass supply chain costs, which take at least 70% of and even more than the profit gained from energy balancing of the grid. The plant CAPEX target of the plants of such sizes is 3,000–12,000 €/ref-stack with the gain from the sale of synthesis natural gas.
- For the plant enabled by the same type of gasification technology, there is no big difference in plant CAPEX target in different zones and flexibility-need scenarios. The plant CAPEX target can be further increased by higher grid up/down regulating price and longer payback years.
- The economic feasibility is affected by their contribution to address the targeted grid flexibility needs. The higher the contribution, the more plant capacities deployed will be coordinated to operate under the PowNeu mode. The increased annual utilization hours of the PowNeu mode, which does not gain from energy-balancing services, cause a decrease in the plant CAPEX target.

The plant CAPEX should be calculated in future work based on the process flow diagram and the component sizes of each plant deployed. Comparison of the plant CAPEX and plant CAPEX target can further reveal the economic feasibility of W2G plants.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

YZ: data curation, formal analysis, and writing - original draft. NW and LD: supervision, validation, and writing - review, editing. CL: data curation and writing - review, editing. JV and YY: funding acquisition, resources, and writing - review, editing. FM: resources, software, and writing - review, editing. TL: Writing - review, editing. LW: conceptualization, methodology, supervision, and writing - review, editing. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

LW and JV have received funding from the European Union's Horizon 2020 under grant agreement No. 826161 (Waste2GridS), 826234 (Waste2Watts), 815284 (BLAZE), and 735692 (CH2P), and the support from the Fuel Cells and Hydrogen Joint Undertaking, Hydrogen Europe and Hydrogen Europe research. LD has received funding from the National Nature Science Foundation of China (Grant No. 52076078) and the National Science and Technology Major Project (Grant No. 2017-I-0002-0002). YZ has received funding from the Fundamental Research Funds for the Central Universities (2019QN029). YY has received funding from the National Nature Science Foundation of China (Grant No. 51821004).

REFERENCES

Apostolou, D. (2020). Optimisation of a hydrogen production – storage – re-powering system participating in electricity and transportation markets. a case study for denmark. *Appl. Energy* 265:114800. doi: 10.1016/j.apenergy.2020.114800

Bellotti, D., Rivarolo, M., Magistri, L., and Massardo, A. (2017). Feasibility study of methanol production plant from hydrogen and captured carbon dioxide. *J. CO2 Util.* 21, 132–138. doi: 10.1016/j.jcou.2017.07.001

Butera, G., Jensen, S. H., Gadsbøll, R. Ø., Ahrenfeldt, J., and Clausen, L. R. (2020). Flexible biomass conversion to methanol integrating solid oxide cells and TwoStage gasifier. *Fuel* 271:117654. doi: 10.1016/j.fuel.2020.117654

Denholm, P., and Hand, M. (2011). Grid flexibility and storage required to achieve very high penetration of variable renewable electricity. *Energy Policy* 39, 1817–1830. doi: 10.1016/j.enpol.2011.01.019

Elia (2019). *Adequacy and Flexibility Study for Belgium 2020-2030*. Technical report, Elia.

Energinet (2020). *System Perspectives for the 70% Target and Large-Scale Offshore Wind*. Technical report, Energinet.

European Commission (2016). *The Role and Need of Flexibility in 2030: Focus on Energy Storage*. Technical report, European Commission.

Fazlollahi, S., Bungener, S. L., Mandel, P., Becker, G., and Maréchal, F. (2014). Multi-objectives, multi-period optimization of district energy systems: I. selection of typical operating periods. *Comput. Chem. Eng.* 65, 54–66. doi: 10.1016/j.compchemeng.2014.03.005

Frate, G. F., Ferrari, L., and Desideri, U. (2021). Energy storage for grid-scale applications: technology review and economic feasibility analysis. *Renew. Energy* 163, 1754–1772. doi: 10.1016/j.renene.2020.10.070

Grueger, F., Mhrke, F., Robinius, M., and Stolten, D. (2017). Early power to gas applications: reducing wind farm forecast errors and providing secondary control reserve. *Appl. Energy* 192, 551–562. doi: 10.1016/j.apenergy.2016.06.131

Guinot, B., Montignac, F., Champel, B., and Vannucci, D. (2015). Profitability of an electrolysis based hydrogen production plant providing grid balancing services. *Int. J. Hydrogen Energy* 40, 8778–8787. doi: 10.1016/j.ijhydene.2015.05.033

IEA (2019). *Gas2019*. Technical report, IEA.

IEA (2020). *World Energy Outlook 2020*. Technical report, IEA, Paris, France.

IRE Agency (2014). *Remap 2030 - A Renewable Energy Roadmap*. Agency I R E. REmap 2030 - A renewable energy roadmap[J]. International Renewable Energy Agency.

Jorgensen, C. and Ropenus, S. (2008). Production price of hydrogen from grid connected electrolysis in a power market with high wind penetration. *Int. J. Hydrogen Energy* 33, 5335–5344. doi: 10.1016/j.ijhydene.2008.06.037

Koltsaklis, N. E., Dagoumas, A. S., and Panapakidis, I. P. (2017). Impact of the penetration of renewables on flexibility needs. *Energy Policy* 109, 360–369. doi: 10.1016/j.enpol.2017.07.026

Kroniger, D., and Madlener, R. (2014). Hydrogen storage for wind parks: a real options evaluation for an optimal investment in more flexibility. *Appl. Energy* 136, 931–946. doi: 10.1016/j.apenergy.2014.04041

Laínez, J. M., Pérez-Fortes, M., Bojarski, A. D., and Puigjaner, L. (2011). "Raw materials supply," in *Syngas From Waste*, editor L. Puigjaner (London: Springer), 23–54.

Lee, R., Homan, S., Dowell, N. M., and Brown, S. (2019). A closed-loop analysis of grid scale battery systems providing frequency response and reserve services in a variable inertia grid. *Appl. Energy* 236, 961–972. doi: 10.1016/j.apenergy.2018.12.044

Mogensen, M. B. (2020). Materials for reversible solid oxide cells. *Curr. Opin. Electrochem.* 21, 265–273. doi: 10.1016/j.coelec.2020.03.014

Motylinski, K., Kupecki, J., Numan, B., Hajimolana, Y. S., and Venkataraman, V. (2021). Dynamic modelling of reversible solid oxide cells for grid stabilization applications. *Energy Convers. Manage.* 228:113674. doi: 10.1016/j.enconman.2020.113674

Olsen, K. P., Zong, Y., You, S., Bindner, H., Koivisto, M., and Gea-Bermúdez, J. (2020). Multi-timescale data-driven method identifying flexibility requirements for scenarios with high penetration of renewables. *Appl. Energy* 264:114702. doi: 10.1016/j.apenergy.2020.114702

Paulus, M., and Borggrefe, F. (2011). The potential of demand-side management in energy-intensive industries for electricity markets in germany. *Appl. Energy* 88, 432–441. doi: 10.1016/j.apenergy.2010.03.017

Pérez-Fortes, M., Laínez-Aguirre, J. M., Bojarski, A. D., and Puigjaner, L. (2014). Optimization of pre-treatment selection for the use of woody waste in co-combustion plants. *Chem. Eng. Res. Design* 92, 1539–1562. doi: 10.1016/j.cherd.2014.01.004

Perna, A., Minutillo, M., Lavadera, A. L., and Jannelli, E. (2018). Combining plasma gasification and solid oxide cell technologies in advanced power plants for waste to energy and electric energy storage applications. *Waste Manage.* 73, 424–438. doi: 10.1016/j.wasman.2017.09.022

Puigjaner, L. (2011). *Syngas From Waste: Emerging Technologies*. London: Springer.

S2Biom (2016). *Delivery of Sustainable Supply of Non-food Biomass to Support a Resource-Efficient Bioeconomy in Europe*. Technical report, European commission.

Sigurjonsson, H. Æ., and Clausen, L. R. (2018). Solution for the future smart energy system: a polygeneration plant based on reversible solid oxide cells and biomass gasification producing either electrofuel or power. *Appl. Energy* 216, 323–337. doi: 10.1016/j.apenergy.2018.02.124

Skytte, K. (1999). The regulating power market on the nordic power exchange nord pool: an econometric analysis. *Energy Econ.* 21, 295–308. doi: 10.1016/s0140-9883(99)00016-x

SOLIDpower (2017). SOLIDpower expands production capacity, SOFC runtime record. *Fuel Cells Bull.* 2017:9. doi: 10.1016/S1464-2859(17)30421-2

The European Parliament And The Council Of The European Union (2018). *Directive (eu) 2018/851 of the European Parliament and of the Council*.

Wang, C., Akkurt, N., Zhang, X., Luo, Y., and She, X. (2020). Techno-economic analyses of multi-functional liquid air energy storage for power generation, oxygen production and heating. *Appl. Energy* 275:115392. doi: 10.1016/j.apenergy.2020.115392

Wang, L., Zhang, Y., Li, C., Pérez-Fortes, M., Lin, T.-E., Maréchal, F., et al. (2020). Triple-mode grid-balancing plants via biomass gasification and reversible solid-oxide cell stack: concept and thermodynamic performance. *Appl. Energy* 280:115987. doi: 10.1016/j.apenergy.2020.115987

Waste2GridS (2020). Technical report.

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

NOMENCLATURE

Abbreviations

CAPEX Capital expenditure

DK1 Northern Denmark

EFG Entrained-flow gasifier

FICFB Fast internally circulating fluidized-bed

LHV Lower heat value

PowGen Power generation

PowNeu Power neutral

PowSto Power storage

PXP Power-to-x-to-power

RSOC Reversible solid-oxide cell

SNG Synthesis natural gas

SUD Southern Italy

TSO Transmission System Operator

W2G Waste2GridS

Mathematical Symbols

α Repetition time

$\dot{m}$ Plant design oxygen production/consumption

$\dot{p}$ Plant design capacity

$\hat{P}$ Grid balancing need

$\overline{m}$ Oxygen tank storage level

θ Price

E Energy flow

F Mass flow

k Interest rate

l Payback year

m Tank capacity

P Power

R Revenue/cost

r Sizing factor of plant design

U Plant number

Y Plant selection

Z Plant operating status

Subscripts

d Typical day

i Time step

n Plant mode

u Plant design

bal Balance service

con Oxygen consumption

gen Oxygen generation

oxy Oxygen

frontiers in Energy Research

MINI REVIEW
published: 21 April 2021
doi: 10.3389/fenrg.2021.671133

Application of Blockchain Technology in Energy Trading: A Review

Hongbiao Li[1], *Fan Xiao*[2], *Lixin Yin*[3]* *and Fengtong Wu*[4]

[1] School of Computer Science, Northeast Electric Power University, Jilin, China, [2] School of Electrical and Electronic Engineering, North China Electric Power University, Beijing, China, [3] State Grid Xinyuan Fengman Training Center, Jilin, China, [4] School of Mechatronic Engineering and Automation, Shanghai University, Shanghai, China

As blockchain and energy trading have become hot topics in industry and academia, this paper presents a brief literature regarding the blockchain-based energy trading in the fields of energy trading with blockchain. At first, the background and development process is presented, and then the applications of blockchain in energy trading are surveyed and analyzed. Finally, conclusions are summarized and important directions are highlighted in this field.

Keywords: energy trading, blockchain, review, transaction mechanism, platform construction

OPEN ACCESS

Edited by:
Liang Chen,
Nanjing University of Information Science and Technology, China

Reviewed by:
Wei He,
Nanjing University of Information Science and Technology, China
Jia Cui,
Shenyang University of Technology, China

***Correspondence:**
Lixin Yin
ylx_SGXFTC@163.com

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

Received: *23 February 2021*
Accepted: *15 March 2021*
Published: *21 April 2021*

Citation:
Li H, Xiao F, Yin L and Wu F (2021) Application of Blockchain Technology in Energy Trading: A Review. Front. Energy Res. 9:671133. doi: 10.3389/fenrg.2021.671133

INTRODUCTION

At present, as the environmental pollution issue is becoming more and more serious, the disadvantages of the centralized energy system, such as the heavy loss of energy in the long-distance transmission and distribution process, and the low fault-tolerant ability of the system, are becoming prominent. The development of decentralized renewable energy technology has gained attention; however, its storage and redistribution process can not be satisfied through the current centralized system. Meanwhile, some researchers put forward the idea of blockchain-based energy trading (Zhang et al., 2017; Pee et al., 2019), and some projects have been successfully implemented (Laszka et al., 2018b; Mengelkamp et al., 2018). The construction of an energy trading platform and trading mechanism based on blockchain has become a hot topic. This paper presents a systematic literature review of studies and projects based on blockchain-based energy trading, as proposed in papers published in recent years. Considering that energy trading based on blockchain is still at an early stage, this paper divides the current research focus into four aspects: (1) construction of trading platform; (2) economy, privacy, and security of transaction mechanism; (3) redundancy and scalability of trading platform; (4) implementation of the specific technology of trading platform.

The rest of this paper is organized as follows. In section The Background and Development Process, the background of the research problem and the development process are introduced. Section Application of Blockchain in Energy Trading describes four key issues and analyzes the corresponding research status. Conclusions and future development trends are explained in section Conclusion.

THE BACKGROUND AND DEVELOPMENT PROCESS

Microgrid Energy Markets

Renewable energy plays an important role in reshaping the future of energy industry, which can be integrated into power systems, in various forms, such as active distribution networks

(Li et al., 2018a), integrated energy systems (Li et al., 2020), and microgrids (Li et al., 2018d). In this context, how to maximize the utilization of renewable energy by managing micro-grid markets is becoming a hot topic from academia to industry. Papaefthymiou and Dragoon (2016) elaborated on how to transform the traditional power systems into systems with 100% renewable energy. Similarly, Hanna et al. (2017) stressed the importance of policy support; to find the optimal operation mode to model the microgrid, Hanna et al. (2017) found that only with the support of policies, the microgrid will be able to achieve the lowest cost operation and environment friendly factors. Li et al. (2018c) presented a scheduling scheme of microgrids with an electric vehicle battery swapping station, taking into consideration, the real-time pricing mechanism. Li et al. (2019) put forward a multi-objective microgrid dispatch strategy considering the user experience. To ensure the benefit of microgrid participants, Kuznetsova et al. (2014) put forward the individual goal of stakeholders to optimize microgrid energy management framework. Furthermore, in terms of economic efficiency optimization, Montuori et al. (2014) proposed the optimization model, hybrid optimization of multiple energy resources (HOMER), to evaluate the economic efficiency of the microgrid with a biomass gasification power plant. Demand response management (DSM) has experienced a renaissance since microgrids require a flexible demand-side to simplify system operations (Palensky and Dietrich, 2011), but the use of DSM in microgrids does not take the advantage of the development of renewable energy sources in the long run, and reflects the needs of socio-economic development. An improved scheme based on the blockchain was put forward by Noor et al. (2018).

Li and Li (2019) proposed a microgrid dispatch strategy taking into account, the demand response of electric vehicles. By combing the advantages of DSM and blockchain technology, Li and Li (2019) presented a game-theoretic model for DSM within microgrid networks that were enhanced by blockchain, and which realized payment mechanisms and intelligent decentralized control.

Blockchain Technology

Blockchain technology provides a powerful tool for implementing energy trading. In 2019, Nakamoto (2019) presented a peer-to-peer (P2P) network, which employed proof-of-work to record a public history of the transaction, and this consensus mechanism can enforce any needed rules and incentives. That was the beginning of the blockchain-based research. Dong et al. (2018) described the blockchain as a distributed, redundant, chain-connected, ledger-sharing database, in which each node in the network is fault-tolerant and can achieve point-to-point communication.

There are four key characters of blockchain: (1) decentralized distributed nodes and storage; (2) consensus, smart contract, and asymmetric encryption, which enable it to have huge potential in many domains, such as finance, computer software, and computer applications; (3) the information economy and postal economy, such as investment and securities; (4) the shared health-care data framework, generation, and distribution in the citizen-level microgrid which may benefit from the widespread dissemination of blockchain transactions (Giungato et al., 2017).

Andoni et al. (2019) presented an overall review of blockchain technology, which involved 140 blockchain research projects and initiatives. However, most of the related studies are still in their infancy. Some social factors, such as laws and policies will also have an impact on the later development of blockchain technology. More specifically, Ali et al. (2018) presented an extensive survey of the application of blockchain in the internet of things (IoT), demonstrating the potential advantages of blockchain in some aspects, such as privacy, secured communications, identity, and data management as well as monetization of IoT data and resources. Fan et al. (2018) modeled the pricing and transaction of energy-internet electricity, based on the blockchain and big data, which provided a reference for the parties involved, including producers, consumers, and managers.

APPLICATION OF BLOCKCHAIN IN ENERGY TRADING

Since most studies related to energy trading based on blockchain are still in the initial stage, according to different research focuses, this paper divides relevant studies into the following four aspects: (1) construction of trading platform; (2) study on the economy, privacy, and security of transaction mechanism; (3) the latency and scalability of trading platform; (4) implementation of the specific technology of trading platform.

Design and Construction of a Trading Platform

Integrated demand response (IDR) has been proven to be effective in improving the operating flexibility of the system and energy utilization efficiency through optimization of the operations of flexible loads, energy conversion, and storage equipment on the demand side (Li et al., 2021). To make the full use of the integrated scattered IDR resources, Zhao et al. (2018a,b) presented an energy transaction mechanism based on the blockchain technology. Moreover, Mannaro et al. (2017) have launched the Crypto-Trading Project, where, they highlighted the key role of blockchain technology and smart contracts in the management and control of innovation typology of the energy market.

Privacy, Security, and Economy of Transaction Mechanism

To solve the privacy problem of energy transactions, Zhou et al. (2018), Laszka et al. (2018a), and Tan et al. (2019) proposed different solutions, respectively. Aiming at the security and privacy issues of large scale vehicle to grid (V2G) energy trading, Zhou et al. (2018) developed a consortium of blockchain-based energy trading mechanism and an edge computing-based task offloading for V2G and local energy aggregators (LEAGs), respectively, and the validation of the proposed framework was proved from the perspectives of task offloading and security. Laszka et al. (2018a), considered not only privacy but also

resilience problem, and they provided a novel distribution application platform. Tan et al. (2019) designed a model for privacy-preserving energy scheduling based on the energy blockchain network in which they solved the problem by using Lagrangian relaxation and smart contracts.

In terms of the economy of the trading platform, Park et al. (2018) suggested a blockchain-based P2P energy transaction platform and provided simulation results that calculated and compared the economic benefit of the platform. Considering the huge operational overhead resulted from a high-frequency transaction, Hou et al. (2019) designed a scheme that enabled nodes to satisfy their power loads through locally stored energy (self-sufficiency), before participating as sellers, if they still had considerable surplus electricity. Alcarria et al. (2018) presented a blockchain-based authorization system for trustworthy resource monitoring and trading.

Latency and Scalability of Trading Platform

To solve the issues of latency resulted from processing the energy trading decisions at remote control centers and the security concerns, while trading the energy, Jindal et al. (2019) proposed a blockchain-based edge-as-a-service framework, which used a software-defined network (SDN) architecture to reduce the latency and secured the underlying trading transactions by blockchain. Blom and Farahmand (2018) modeled a local energy market using Ethereum platform and concluded that the given market with 600 participants and a trading frequency of one transaction every 5 min can be processed by the Ethereum protocol. Besides, Liu et al. (2019) provided the off-chain energy trading method and asynchronous transaction recoding mechanism, which are supported by a local energy trading cyber-physical system.

Implementation of Specific Technology of Trading Platform

In addition to the research and analysis of the characteristics of blockchain, the implementation of the algorithm is also the focus of the current research problem. Kang et al. (2018) exploded a renewable energy trading platform using smart contract of Ethereum, proved its scalability and adaptability utilizing coding transaction process and contents of smart contracts, and presented a simple scenario for two nodes. Pipattanasomporn et al. (2018) presented the laboratory scale implementation of the blockchain network for the exchange of solar energy. Tai et al. (2016) provided an improved algorithm of distributed security checking and proved the feasibility by a case consisting of six nodes.

However, all the schemes mentioned above are in the initial exploration stage, and the verified schemes and scale are in an ideal environment or laboratory environment. Therefore, there is still a way to go before the blockchain-based energy trading platform can be widely used in practical applications.

CONCLUSION

As an emerging and powerful technology, energy trading based on blockchain has attracted a growing attention of many scholars. After studying the existing literature, this paper summarizes the key issues into the following four points: (1) construction of trading platform; (2) economy, privacy, and security of transaction mechanism; (3) redundancy and scalability of trading platform; (4) implementation of the specific technology of trading platform.

Since most studies are in the primary stage, the construction of an energy trading platform and efficient algorithm implementation will be important research directions in the future. Another interesting topic is to apply machine learning to blockchain-based applications (Shi et al., 2008; Li et al., 2018b; Tanwar et al., 2019).

AUTHOR CONTRIBUTIONS

HL, FX, and LY contributed to conception and design of the study. HL and FX performed the statistical analysis. HL wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

REFERENCES

Alcarria, R., Bordel, B., Robles, T., Martín, D., and Manso-Callejo, M. Á. (2018). A blockchain-based authorization system for trustworthy resource monitoring and trading in smart communities. *Sensors* 18:3561. doi: 10.3390/s18103561

Ali, M. S., Vecchio, M., Pincheira, M., Dolui, K., Antonelli, F., and Rehmani, M. H. (2018). Applications of blockchains in the internet of things: a comprehensive survey. *IEEE Commun. Surv. Tutor.* 21, 1676–1717. doi: 10.1109/COMST.2018.2886932

Andoni, M., Robu, V., Flynn, D., Abram, S., Geach, D., Jenkins, D., et al. (2019). Blockchain technology in the energy sector: a systematic review of challenges and opportunities. *Renew. Sustain. Energy Rev.* 100, 143–174. doi: 10.1016/j.rser.2018.10.014

Blom, F., and Farahmand, H. (2018). "On the scalability of blockchain-supported local energy markets," in *2018 International Conference on Smart Energy Systems and Technologies (SEST)* (Seville: IEEE), 1–6. doi: 10.1109/SEST.2018.8495882

Dong, Z., Luo, F., and Liang, G. (2018). Blockchain: a secure, decentralized, trusted cyber infrastructure solution for future energy systems. *J. Mod. Power Syst. Clean Energy* 6, 958–967. doi: 10.1007/s40565-018-0418-0

Fan, T., He, Q., Nie, E., and Chen, S. (2018). "A study of pricing and trading model of Blockchain & Big data-based Energy-Internet electricity," in *IOP Conference Series: Earth and Environmental Science* (Chongqing: IOP Publishing), 052083. doi: 10.1088/1755-1315/108/5/052083

Giungato, P., Rana, R., Tarabella, A., and Tricase, C. (2017). Current trends in sustainability of bitcoins and related blockchain technology. *Sustainability* 9:2214. doi: 10.3390/su9122214

Hanna, R., Ghonima, M., Kleissl, J., Tynan, G., and Victor, D. G. (2017). Evaluating business models for microgrids: interactions of technology and policy. *Energy Policy* 103, 47–61. doi: 10.1016/j.enpol.2017.01.010

Hou, W., Guo, L., and Ning, Z. (2019). Local electricity storage for blockchain-based energy trading in industrial internet of things. *IEEE Trans. Ind. Inform.* 15, 3610–3619. doi: 10.1109/TII.2019.2900401

Jindal, A., Aujla, G. S., and Kumar, N. (2019). SURVIVOR: a blockchain based edge-as-a-service framework for secure energy trading in SDN-enabled vehicle-to-grid environment. *Comput. Netw.* 153, 36–48. doi: 10.1016/j.comnet.2019.02.002

Kang, E. S., Pee, S. J., Song, J. G., and Jang, J. W. (2018). "A blockchain-based energy trading platform for smart homes in a microgrid," in *2018 3rd*

International Conference on Computer and Communication Systems (ICCCS) (IEEE), 472–476. doi: 10.1109/CCOMS.2018.8463317

Kuznetsova, E., Li, Y. F., Ruiz, C., and Zio, E. (2014). An integrated framework of agent-based modelling and robust optimization for microgrid energy management. *Appl. Energy* 129, 70–88. doi: 10.1016/j.apenergy.2014.04.024

Laszka, A., Eisele, S., Dubey, A., Karsai, G., and Kvaternik, K. (2018b). "TRANSAX: a blockchain-based decentralized forward-trading energy exchanged for transactive microgrids," in *2018 IEEE 24th International Conference on Parallel and Distributed Systems (ICPADS)* (Singapore: IEEE), 918–927.

Laszka, A., Eisele, S., Dubey, A., Karsai, G., and Kvaternik, K. (2018a). "TRANSAX: a blockchain-based decentralized forward-trading energy exchanged for transactive microgrids," in *2018 IEEE 24th International Conference on Parallel and Distributed Systems (ICPADS)* (IEEE), 918–927. doi: 10.1109/PADSW.2018.8645001

Li, Y., Feng, B., Li, G., Qi, J., Zhao, D., and Mu, Y. (2018a). Optimal distributed generation planning in active distribution networks considering integration of energy storage. *Appl. Energy* 210, 1073–1081. doi: 10.1016/j.apenergy.2017.08.008

Li, Y., and Li, K. (2019). Incorporating demand response of electric vehicles in scheduling of isolated microgrids with renewables using a bi-level programming approach. *IEEE Access* 7, 116256–116266. doi: 10.1109/ACCESS.2019.2936487

Li, Y., Wang, C., Li, G., and Chen, C. (2021). Optimal scheduling of integrated demand response-enabled integrated energy systems with uncertain renewable generations: a Stackelberg game approach. *Energy Convers. Manag.* 235:113996. doi: 10.1016/j.enconman.2021.113996

Li, Y., Wang, C., Li, G., Wang, J., Zhao, D., and Chen, C. (2020). Improving operational flexibility of integrated energy system with uncertain renewable generations considering thermal inertia of buildings. *Energy Convers. Manag.* 207:112526. doi: 10.1016/j.enconman.2020.112526

Li, Y., Wang, J., Zhao, D., Li, G., and Chen, C. (2018b). A two-stage approach for combined heat and power economic emission dispatch: Combining multi-objective optimization with integrated decision making. *Energy* 162, 237–254. doi: 10.1016/j.energy.2018.07.200

Li, Y., Yang, Z., Li, G., Mu, Y., Zhao, D., Chen, C., et al. (2018c). Optimal scheduling of isolated microgrid with an electric vehicle battery swapping station in multi-stakeholder scenarios: a bi-level programming approach via real-time pricing. *Appl. Energy* 232, 54–68. doi: 10.1016/j.apenergy.2018.09.211

Li, Y., Yang, Z., Li, G., Zhao, D., and Tian, W. (2018d). Optimal scheduling of an isolated microgrid with battery storage considering load and renewable generation uncertainties. *IEEE Trans. Ind. Electron.* 66, 1565–1575. doi: 10.1109/TIE.2018.2840498

Li, Y., Yang, Z., Zhao, D., Lei, H., Cui, B., and Li, S. (2019). Incorporating energy storage and user experience in isolated microgrid dispatch using a multi-objective model. *IET Renew. Power Gener.* 13, 973–981. doi: 10.1049/iet-rpg.2018.5862

Liu, S., Chen, F., Shen, L., Hu, Y., and Ding, Y. (2019). A high-performance local energy trading cyber-physical system based on blockchain technology. *Earth Environ. Sci.* 227:032009. doi: 10.1088/1755-1315/227/3/032009

Mannaro, K., Pinna, A., and Marchesi, M. (2017). "Crypto-trading: blockchain-oriented energy market," in *2017 AEIT International Annual Conference* (Cagliari: IEEE), 1–5. doi: 10.23919/AEIT.2017.8240547

Mengelkamp, E., Gärttner, J., Rock, K., Kessler, S., Orsini, L., and Weinhardt, C. (2018). Designing microgrid energy markets: a case study: the brooklyn microgrid. *Appl. Energy* 210, 870–880. doi: 10.1016/j.apenergy.2017.06.054

Montuori, L., Alcázar-Ortega, M., Álvarez-Bel, C., and Domijan, A. (2014). Integration of renewable energy in microgrids coordinated with demand response resources: economic evaluation of a biomass gasification plant by Homer Simulator. *Appl. Energy* 132, 15–22. doi: 10.1016/j.apenergy.2014.06.075

Nakamoto, S. (2019). *Bitcoin: A Peer-to-Peer Electronic Cash System*. Manubot.

Noor, S., Yang, W., Guo, M., van Dam, K. H., and Wang, X. (2018). Energy demand side management within micro-grid networks enhanced by blockchain. *Appl. Energy* 228, 1385–1398. doi: 10.1016/j.apenergy.2018.07.012

Palensky, P., and Dietrich, D. (2011). Demand side management: demand response, intelligent energy systems, and smart loads. *IEEE Trans. Ind. Inform.* 7, 381–388. doi: 10.1109/TII.2011.2158841

Papaefthymiou, G., and Dragoon, K. (2016). Towards 100% renewable energy systems: uncapping power system flexibility. *Energy Policy* 92, 69–82. doi: 10.1016/j.enpol.2016.01.025

Park, L. W., Lee, S., and Chang, H. (2018). A sustainable home energy prosumer-chain methodology with energy tags over the blockchain. *Sustainability* 10:658. doi: 10.3390/su10030658

Pee, S. J., Kang, E. S., Song, J. G., and Jang, J. W. (2019). "Blockchain based smart energy trading platform using smart contract," in *2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)* (Okinawa: IEEE), 322–325. doi: 10.1109/ICAIIC.2019.8668978

Pipattanasomporn, M., Kuzlu, M., and Rahman, S. (2018). "A blockchain-based platform for exchange of solar energy: laboratory-scale implementation," in *2018 International Conference and Utility Exhibition on Green Energy for Sustainable Development (ICUE)* (IEEE), 1–9. doi: 10.23919/ICUE-GESD.2018.8635679

Shi, Z. B., Yu, T., Zhao, Q., Li, Y., and Lan, Y. B. (2008). Comparison of algorithms for an electronic nose in identifying liquors. *J. Bionic Eng.* 5, 253–257. doi: 10.1016/S1672-6529(08)60032-3

Tai, X., Sun, H., and Guo, Q. (2016). Electricity transactions and congestion management based on blockchain in energy internet. *Power Syst. Technol.* 40, 3630–3638. doi: 10.13335/j.1000-3673.pst.2016.12.002

Tan, S., Wang, X., and Jiang, C. (2019). Privacy-preserving energy scheduling for ESCOs based on energy blockchain network. *Energies* 12:1530. doi: 10.3390/en12081530

Tanwar, S., Bhatia, Q., Patel, P., Kumari, A., Singh, P. K., and Hong, W. C. (2019). Machine learning adoption in blockchain-based smart applications: the challenges, and a way forward. *IEEE Access* 8, 474–488. doi: 10.1109/ACCESS.2019.2961372

Zhang, J., Gao, W. Z., Zhang, Y. C., Zheng, X. H., Yang, L. Q., Hao, J., et al. (2017). Blockchain based intelligent distributed electrical energy systems: needs, concepts, approaches and vision. *Acta Automat. Sin.* 43, 1544–1554. doi: 10.16383/j.aas.2017.c160744

Zhao, S., Li, Y., Wang, B., and Su, H. (2018a). "Research on the blockchain-based integrated demand response resources transaction scheme," in *2018 International Power Electronics Conference (IPEC-Niigata 2018-ECCE Asia)* (IEEE), 795–802. doi: 10.23919/IPEC.2018.8507866

Zhao, S., Wang, B., Li, Y., and Li, Y. (2018b). Integrated energy transaction mechanisms based on blockchain technology. *Energies* 11:2412. doi: 10.3390/en11092412

Zhou, Z., Tan, L., and Xu, G. (2018). "Blockchain and edge computing based vehicle-to-grid energy trading in energy internet," in *2018 2nd IEEE Conference on Energy Internet and Energy System Integration (EI2)* (IEEE), 1–5. doi: 10.1109/EI2.2018.8582652

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

frontiers in Energy Research

BRIEF RESEARCH REPORT
published: 21 April 2021
doi: 10.3389/fenrg.2021.666130

Coordinated Cyber-Attack Detection Model of Cyber-Physical Power System Based on the Operating State Data Link

Lei Wang[1,2], Pengcheng Xu[3], Zhaoyang Qu[1,2], Xiaoyong Bo[1,2], Yunchang Dong[1,2], Zhenming Zhang[1,2] and Yang Li[1]*

[1] School of Electrical Engineering, Northeast Electric Power University, Jilin, China, [2] Jilin Engineering Technology Research Center of Intelligent Electric Power Big Data Processing, Jilin, China, [3] Siping Power Supply Company of State Grid Jilin Electric Power Company Limited, Siping, China

OPEN ACCESS

Edited by:
Liang Chen,
Nanjing University of Information Science and Technology, China

Reviewed by:
Donglai Wang,
Shenyang Institute of Engineering, China
Ruyi Dong,
Jilin Institute of Chemical Technology, China
Shuaibing Lu,
Beijing University of Technology, China

***Correspondence:**
Lei Wang
752953593@qq.com

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

Received: *09 February 2021*
Accepted: *24 February 2021*
Published: *21 April 2021*

Citation:
Wang L, Xu P, Qu Z, Bo X, Dong Y, Zhang Z and Li Y (2021) Coordinated Cyber-Attack Detection Model of Cyber-Physical Power System Based on the Operating State Data Link. Front. Energy Res. 9:666130. doi: 10.3389/fenrg.2021.666130

Existing coordinated cyber-attack detection methods have low detection accuracy and efficiency and poor generalization ability due to difficulties dealing with unbalanced attack data samples, high data dimensionality, and noisy data sets. This paper proposes a model for cyber and physical data fusion using a data link for detecting attacks on a Cyber–Physical Power System (CPPS). The two-step principal component analysis (PCA) is used for classifying the system's operating status. An adaptive synthetic sampling algorithm is used to reduce the imbalance in the categories' samples. The loss function is improved according to the feature intensity difference of the attack event, and an integrated classifier is established using a classification algorithm based on the cost-sensitive gradient boosting decision tree (CS-GBDT). The simulation results show that the proposed method provides higher accuracy, recall, and F-Score than comparable algorithms.

Keywords: cyber-physical power system, coordinated cyber-attack, cluster analysis, oversampling, gradient boosting decision tree

INTRODUCTION

In recent years, a new type of coordinated cyber-physical attack has caused blackouts of the power grid and disrupted power systems. The main reason is that the coordinated attack on the power grid by hackers was not detected in time, and effective measures to prevent major accidents could not be implemented at the optimum time (Haes Alhelou et al., 2019; Lai et al., 2019). In the 2015 attack on the Ukrainian power grid, the attack point was not the power infrastructure, and the 0-day vulnerability was not used. Its attack cost is significantly lower than that of Stuxnet, Equation, and other attacks, but it is also more effective (Zhang et al., 2016; Koopman et al., 2019). Therefore, traditional security protection methods for power systems have their limitations, and it is urgent to research detection and defense methods for coordinated attacks on the Cyber–Physical Power System (CPPS) to identify attack types and intentions. It is crucial to establish a comprehensive active defense system to ensure the security of power systems (Chen et al., 2011; Dai et al., 2019; Wang X. et al., 2019).

Many scholars have investigated the detection and identification of coordinated attacks on the CPPS. The coupling relationship between the cyber side and the physical side has been considered

in several studies (Drayer and Routtenberg, 2019; Shen et al., 2019), which focused on the fusion of the attack path on the information side and the attack object on the physical side. Xu and Abur (2017) combined *a priori* and *a posteriori* bad data detection and proposed a new decomposition method to solve the state estimation data corruption in cyber-attacks. Kurt et al. (2018) and Basin et al. (2016) used a dynamic equation of the measured variables with a joint transformation to detect false data injection (FDI) attacks in real time to improve the detection accuracy.

In summary, existing detection methods for cyber-attacks on the CPPS have the following limitations: (1) the cyberspace and the physical space are closely coupled and interact with each other. An attack detection from the cyber side or the physical side alone is not sufficient (Lin et al., 2016; Nath et al., 2019). (2) Attack detection methods based on physical power grid data ignore the impact of cyber network attacks on the performance of smart grids. The effects of power grid failures and cyber-attacks on the physical side are similar, and it is difficult to distinguish them based on data characteristics (Liu et al., 2016; Huang and Zhu, 2020). (3) A cyber-attack is characterized by unbalanced attack samples, high data dimensionality, and noise, and data with a long tail are common. Low detection accuracy of attacks and low real-time detection efficiency are typical (Osanaiye et al., 2018; Tian et al., 2019).

In this paper, the cyber-side alarm data and the physical-side measurement data are merged to establish a cyber-physical coupling state chain. A clustering method is designed to classify and distinguish different operating states of the CPPS. An oversampling algorithm is used to reduce the imbalance in the operating states' samples. Subsequently, a coordinated cyber-attack detection algorithm based on the improved gradient boosting decision tree (GBDT) is proposed. The algorithm optimizes the cost-sensitive (CS) loss function, minimizing the error associated with the small sample size of attack data and providing high accuracy of attack detection and a high recall rate and F1-score.

DETECTION MODEL FOR COORDINATED CYBER-ATTACKS ON THE CYBER–PHYSICAL POWER SYSTEM

The framework of the coordinated cyber-attack detection model is shown in **Figure 1**. The model exploits the data characteristics in different states, such as normal operation, fault operation, and coordinated attack of the CPPS. First, the data link of the cyber–physical operation state is established according to the coupling relationship. A clustering algorithm is used to classify the state data link, and a feature set is obtained under different operating conditions. Then, the adaptive synthetic sampling algorithm (ADASYN) is used to balance the majority of the samples and the minority of the samples in different state data sets. Finally, new CS conditions are added using the GBDT's CS loss function to detect different coordinated cyber-attacks.

ESTABLISHMENT OF THE DATA LINK OF THE CYBER–PHYSICAL OPERATING STATE

Data Link of the Operating State of the Physical Power Grid

The physical grid measurement data reflect the real-time operating status of the grid under different working conditions. The measurement data of each section of the grid reflect the operating status at that moment. We do not consider the reasons for changes in the grid state (caused by cyber-attacks or general equipment failures); it can be described as a specific interval $_{\Delta t}(t_1 \sim t_n)$. According to the acquisition sequence, all state data fragments $S_p(t_i)$ consisting of the physical grid operating data link $_{Q_p(\Delta t)}$ are defined as follows:

$$\begin{cases} X_p(t_i) = \{x_1, x_2, \ldots..x_h\} \\ S_p(t_i) = \{X_{p_1}, X_{p_2}, \ldots X_{p_m}\} \\ Q_p(\Delta t) = \{S_p(t_1), S_p(t_2)..., S_p(t_n)\} \end{cases} \quad (1)$$

where $X_p(t_i)$ represents the h measured attributes obtained from the physical-side device X_p at time t_i, including the voltage, current, phase angle, active power, and reactive power; $S_p(t_i)$ represents all the measurement data collected by m devices on the physical side at time t_i.

Data Link of the Operating State of the Cyber Network

The transmission delay and data packet loss rate typically reflect the performance status of a cyber-network. When the control signal or status information is lost during the transmission of the data packet because it exceeds the allowable proportion, the control of the device has been lost due to a network attack (Davarikia and Barati, 2018; Wang Q. et al., 2019). Three indicators (delay rate, packet loss rate, and threat degree) are established to characterize the operating data link of the cyber network.

1. The delay ratio (DR) is defined as follows:

$$R_{\mathrm{dr}}(n) = \frac{\sum_{k=1}^{n} \left| \frac{P'_k}{P_k} - P_T \right|}{n} \times 100\% \quad (2)$$

where n is the number of communication links transmitting data, P' is the number of data packet losses for link k, P_k is the number of data packets sent by link k, and P_T is the threshold of the packet loss rate of link k.

1. The packet loss ratio (PR) indicator is defined as follows:

$$R_{pr}(m) = \frac{\sum_{l=1}^{m} (T_l^{\mathrm{send}} - T_l^{\mathrm{receive}})}{m} \times 100\% \quad (3)$$

where m is the number of devices that send information, T_l^{send} is the sending time of data packet l, and $T_l^{receive}$ is the receiving time of data packet l.

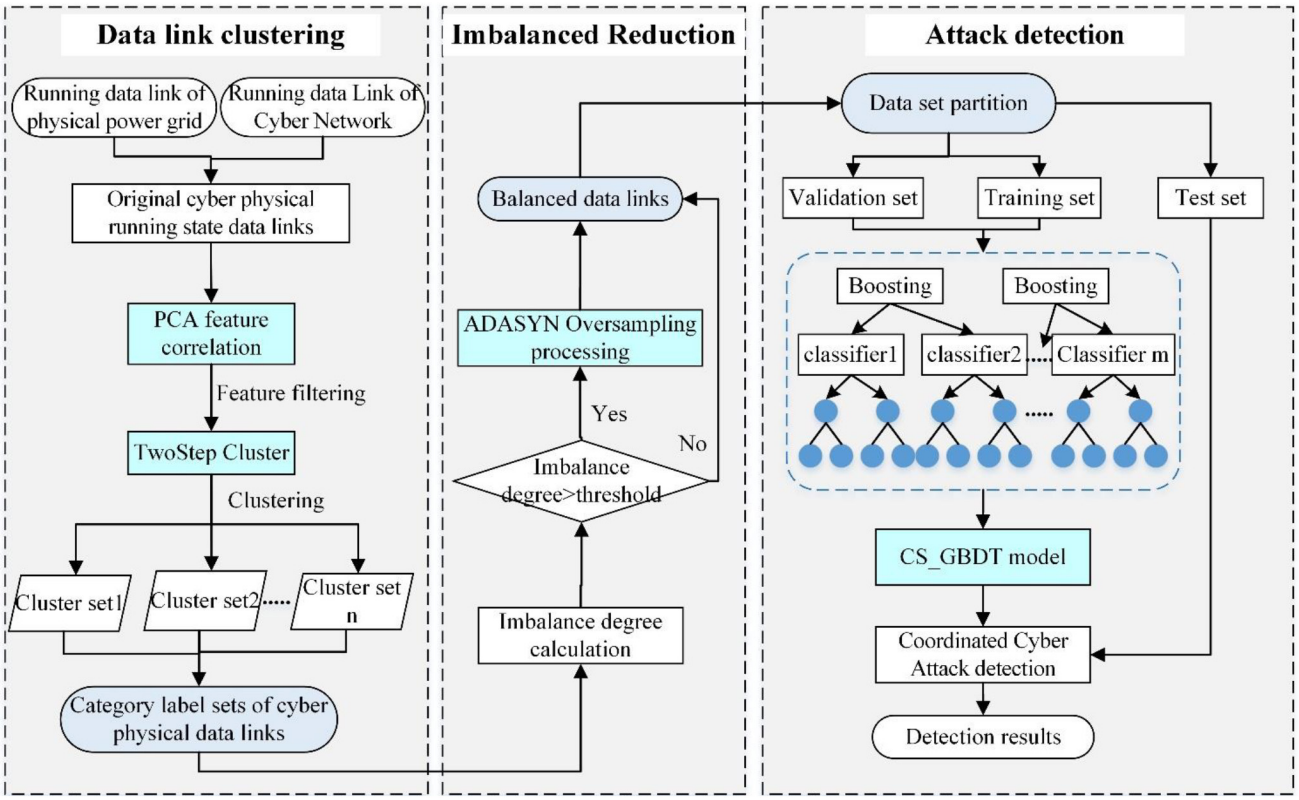


FIGURE 1 | Framework of the coordinated cyber-attack detection model for the Cyber–Physical Power System (CPPS). CS-GBDT, CS gradient boosting decision tree; PCA, principal component analysis.

1. Threat degree W_{th} $(a_{i,j})$. Assuming that n alarm events are generated within the sampling time window $_{\Delta t,}$ the address set of the information equipment is $\{IP_1, IP_2, \ldots IP_m\}$, and $a_{i,j}$ indicates that IP_i contains j alarm events. The intrusion detection system (IDS) deployed in the power cyber network indicates that the original threat degree is W. The threat degree is redefined as follows to determine the impact of alarm events on the attack risk of the entire system:

$$W_{th}=\frac{\sum_{i=1}^{n_i}\left|w_{ij}-\overline{w}_i\right|}{n_i}\times 100\% \quad (4)$$

where w_{ij} is the threat degree of alarm events $a_{i,j}$, w_i is the average value of the threat degrees of all alarm events in IP_i, and n_i is the number of all alarm events in IP_i.

The three performance indicators of the operating status of the cyber network are used to establish the cyber system operating data link Q_c (Δt) in the interval $\Delta t(t_1 \sim t_n)$:

$$\begin{cases} Y_c(t_i) = \{R_{dr},\ R_{pr},\ W_{th}\} \\ S_c(t_i) = \{Y_{c1},\ Y_{c2},\ \ldots Y_{ck}\} \\ Q_c(\Delta t) = \{S_c(t_1), S_c(t_2), \ldots, S_c(t_n)\} \end{cases} \quad (5)$$

where $Y_c(t_i)$ represents the R_{dr}, R_{pr}, and W_{th} obtained from the cyber-side device Y_c at time t_i; $S_c(t_i)$ represents the status data obtained from k devices on the cyber side at time t_i.

Coupled Mapping of the Operating State of the Cyber–Physical System

We use topological mapping to couple and map the data links of the two heterogeneous networks to form a data link of the cyber–physical operating state. The grid can be divided into m areas according to the physical grid connectivity, and each area has n transmission lines. It is assumed that a line consists of k electrical components $\{X_1, X_2, \ldots, X_k\}$, each line is connected to n communication devices $\{Y_1, Y_2, \ldots, Y_n\}$, and each communication device has a unique IP address $\{IP_1, IP_2, \ldots, IP_n\}$ in the cyber network. We sequentially connect each electrical component number, line number, and connected area in the data chain to create an index table linking the <connected area number Area, line number Line, electrical component ID number, and information component IP address>. The cyber network operating data link Q_c and the physical power grid operating data link Q_p in the interval are compared using the index table, and the data are stored in the corresponding index.

The cyber network clock with a collection period of T is used, and we set the sampling time window to $\varepsilon = [T - \alpha T', T]$, where $_{\alpha}$ is the window size parameter. The larger the value, the longer the collection period is. In the sampling time window $_{\varepsilon,}$ many identical state events may occur in the cyber–physical coupling state chain. Therefore, these repetitive events are filtered and compressed to form the cyber–physical operating state data link,

which is expressed as follows:

$$Q(\varepsilon) = \{x_1[Q_p(t_1), Q_c(t_1)], x_2[Q_p(t_2), Q_c(t_2)], \ldots, x_n[Q_p(t_n), Q_c(t_n)]\} \quad (6)$$

COORDINATED CYBER-ATTACK DETECTION MODEL OF THE CYBER–PHYSICAL POWER SYSTEM

Operating State Clustering Based on the Two-Step Principal Component Analysis

There are no labels for the different state categories in the original cyber–physical operating state data link $Q(\varepsilon)$. It is necessary to distinguish the different state categories using cluster analysis. In this paper, the two-step principal component analysis (PCA) clustering algorithm is proposed. The PCA algorithm is used to cluster, transform, and filter the correlated attributes to extract linear uncorrelated attributes (Jian et al., 2004). The two-step algorithm is used to cluster the attribute set; it reduces the computational complexity and provides high clustering accuracy (Dom et al., 2003; Northrup et al., 2004; Phelps et al., 2009). The algorithm steps are as follows:

Input: cyber–physical operating state data link $Q(\varepsilon) = \{x_1, x_2, \ldots, x_n\}$.

Output: D = $\{x_i, C_i\}$, where C_i is the operating state of the clusters $C = \{C_1, C_2, \ldots, C_k\}$.

Step 1: Feature selection for clustering. The PCA algorithm is used to map n attributes in the data link to m dimension ($m <$ n). The correlated attributes are filtered using an orthogonal transformation to obtain m-dimensional new features, $A = \{A_1, A_2, \ldots A_m\}$. The centralizing mean $x_{\mathrm{i}} = x_{\mathrm{i}} - \frac{1}{n}\sum_{i=1}^{n} x_{\mathrm{i}}$ is used to derive the covariance matrix $_{XX^T}$, whose eigenvalues and eigenvectors are obtained. The data link set $Q'(\varepsilon)$ is obtained after dimensionality reduction.

Step 2: Calculate the number of category clusters in the operating state. After the traversal process, the clustering feature (CF tree) growth in the balanced iterative reducing and clustering using hierarchies (BIRCH) algorithm is applied to the data link set $Q'(\varepsilon)$. The data points in the data set are evaluated one by one to collect all data points in the dense area while generating the CF tree. The log-likelihood distance $d(C_s, C_t) = \zeta_s + \zeta_t - \zeta_{<s,t>}$ between the two clusters is used to create many small subclusters. The Bayes information criterion (BIC) is used to calculate the number of possible division schemes for the state category.

Step 3: Determine the number of categories C_J in the $Q'(\varepsilon)$. The agglomerative hierarchical clustering (AHC) method is used to merge the subclusters one by one, and the desired number of clusters is reached according to the $R(k)$ between the two clusters.

$$R(k) = \frac{d_{\min}(C_k)}{d_{\min}(C_{k+1})} \quad (7)$$

where C_k and C_{k+1} is a partition scheme with k or k+1 cluster numbers; $d_{min}(C_{k+1})$ and $d_{min}(C_k)$ is the distance between the two smallest clusters in the scheme.

Step 4: Label the sample data in each operating state cluster. The data points in each cluster are determined; the data points x_i in the state data link set $Q'(\varepsilon)$ are regarded as single-point clusters according to the clustering results C_J. The logarithm similarity between x_i and each cluster in C_J is determined. Given the distance $d\{\{x_i\}, C_J\}$, x_i is placed into the nearest cluster, and labels are generated for each operating state category $C = \{C_1, C_2, \ldots, C_k\}$.

Algorithm to Reduce the Imbalance of the Operating State Classes

A coordinated cyber-attack event of the CPPS has a small probability and high risk. In the data link $_{Q(\varepsilon)}$, normal operation data account for the largest proportion, whereas the proportion of attack data is relatively small, resulting in unbalanced data. Therefore, the ADASYN algorithm is used to deal with the imbalance of the operating state classes (Qu et al., 2018; Wang et al., 2020). Balanced data distribution is obtained by adaptive synthetic oversampling. Different minority samples are given different weights to generate different numbers of samples. The algorithm process is as follows:

Input:$D = \{x_i, C_i\}$, where x_i is the cyber–physical operating state data link $Q(\varepsilon)$, C_i is the class label. α is the imbalance threshold, C_k is a minority class, and C_l is the majority class.

Output: Balanced data set D'.

Step 1: Calculate the class imbalance, where $\text{Imbalance} = \frac{\text{Lagre num}(C_l)}{\text{Small num}(C_k)}$. Calculate the number of samples to be synthesized based on the degree of imbalance $G = \left(\text{Lagre num}(C_l) - \text{Small num}(C_k)\right) \times \beta$, $\beta \in [0, 1]$.

Step 2: Calculate the proportion of the majority class in the K-nearest neighbors (KNNs). $r_i = \Delta_i / K$, where Δ_i is the number of samples of the majority class in the KNN.

Step 3: Calculate the majority class surrounding each minority sample.

$$\hat{r_i} = \frac{r_i}{\sum_{\mathrm{i}=1}^{\text{samll num}(C_k)} r_i} \quad (8)$$

Step 4: Calculate the number of samples that need to be generated for each minority sample C_k.

$$g_i = \hat{r_i} \times G \quad (9)$$

Step 5: Select a minority sample among k neighbors around each minority sample and synthesize using Eq. (10). Repeat the synthesis g_i times until the desired number of synthesized samples is obtained.

$$s_{\mathrm{i}} = X_i + (X_{zi} - X_i) \times \eta \quad (10)$$

where s_i is the composite sample, $_{X_i}$ is the i-th sample in the minority sample, $X_i \in [0, 1]$, X_{zi} is a randomly selected minority sample among the KNNs of X_i.

Repeat the synthesis until the desired number of synthesized samples in Eq. (5) has been obtained.

Classification Algorithm of Coordinated Cyber-Attacks Based on Cost-Sensitive Gradient Boosting Decision Tree

The purpose of attack detection is to minimize the harm to the power grid caused by the attack. The harm caused by misinterpreting an attack as a normal event is far greater than that caused by misinterpreting a normal event as an attack (Huang et al., 2018). We propose using the CS function to improve the GBDT (Sakhnovich, 2011; Liao et al., 2016). The CS loss function replaces the standard cost loss function to prevent attack event misclassification. The improved CS loss function is defined as follows:

$$\text{Loss}\left(C, \text{f(x)}\right) = -\sum_{k=1}^{K} w_k C_k \log\left(p_k(x)\right) \tag{11}$$

where K is the class of all attacks, $_{C_k}$ is the sample of the k-th attack, and $_{p_k(x)}$ is the probability of the k-th attack, $_{w_k}$ is the CS function, it can be divided into two costs, i.e., the missed detection cost $w_{(-,+)}(1-p(x)), p(x) \geq \frac{w_-}{w_+ + w_-}$ and the misdetection cost $w_{(+,-)}p(x), p(x) < \frac{w_-}{w_+ + w_-}$.

Coordinated cyber-attack detection is a multi-classification task. A total of K types of attacks are assumed. The sample x in the cyber–physical operating state set is obtained, and the CS-GBDT algorithm is used to determine which class the x sample belongs to. The specific steps of the algorithm are as follows:

Input: Balanced data set $D' = \{(x_1, C_1), (x_2, C_2), ..., (x_N, C_N)\}$, loss function Loss $(C_k, f_k(x))$, and the number of classifiers M.

Output: A strong learner for attack classification $F(x)$.

Step 1: Initialize $f_{k0}(x) = 0$, the number of categories classified $k = 1,2,...K$.

Step 2: Starting from $t = 1$ to $t = M$, there are M iterations in total, repeating steps 3 through 6, at last building M classifiers.

Step 3: The one-hot code for each class y_i is generated. We calculate the probability of sending the k-th attack sample $p_k(x)$.

$$p_k(x) = \frac{e^{f_k(x)}}{\sum_{l=1}^{K} e^{f_l(x)}} \tag{12}$$

Step 4: Start from $k = 1$ to $k = K$, repeating steps 5 through 6, we generate K different CART classification trees $f_1(x)$, $f_2(x),...f_K(x)$.

Step5: Calculate the negative gradient of each class in the m class and obtain the negative gradient error of the i-th sample corresponding to category k in the t-th iteration:

$$r_{ki} = C_{ki} - p_k(x) = -\left[\frac{\partial Loss\left(C_k, f_k(x)\right)}{\partial f_k(x)}\right], i = 1, 2, ..., N \tag{13}$$

where N is the number of sample data.

We use the estimated residual $\{(x_1, r_{k1}),...(x_N, r_{kN}))\}$ as an input to calculate the leaf node area of the m-th decision tree:

$$C_{mkj} = \frac{K-1}{K} \frac{\sum_{x_i \in R_{mkj}} r_{ki}}{\sum_{x_i \in R_{mkj}} |C_{ki}| \left(1 - |C_{ki}|\right)} \tag{14}$$

where R_{mkj} is the leaf node region R_{mj} of the m-th tree. K is the number of categories.

Step 6: Update the classifier $f_{mk}(x)$.

$$f_{mk}(x) = f_{k,m-1}(x) + \sum_{j=1}^{J} C_{mkj} I,\ x \in R_{mkj} \tag{15}$$

where J is the number of leaf nodes per tree.

Step 7: Build final classification tree with high accuracy used for attack detection.

$$F_{mk}(x) = \sum_{m=1}^{M} \sum_{j=1}^{J} C_{mkj} I, x \in R_{mkj} \tag{16}$$

EXPERIMENTAL ANALYSIS

Experimental Environment and Data

We simulate the different fault states of the physical power grid caused by cyber-attacks on the IEEE39-bus system in the RT-LAB and OPNET co-simulation environment. We collect the DR, PR, and threat information at different times on the cyber side. The voltage, current, impedance, and other data are collected on the physical side. The 10 data sets are obtained. Each set contains 56 attributes, and the cumulative number of records is about 50,000, including five types of operating states in the CPPS, as follows:

(1) Normal operating state (S1): there is no network attack on the cyber side, and the power grid on the physical side is operating normally. (2) Distributed denial of service (DDOS) attack state (S2): the data in the communication system are blocked by a DDOS attack, affecting the normal operation of the power system, measurement acquisition, and control commands. (3) Data injection attack state (S3): malicious data injection into physical power grid disguised as a normal fault, resulting in the operator mistakenly assuming a short-circuit fault. (4) Protection device parameter tampering attack state (S4): the attacker tampers with the distance parameter of the protection device, causing a failure of the protection device to disconnect the fault area. (5) Fault operation state (S5): the physical power grid has a single-phase, two-phase, or short-circuit fault.

Results of the Operating State Classification of the Cyber–Physical Power System

The data set 1 with 4,966 records is selected in the experiment. After implementing the two-step PCA clustering algorithm, the number of outliers is 89, and there are five operating states, as shown in **Figure 2A**. Cluster-3 (S2) has the largest number of records (1,325). The clustering superiority is 0.93, and the clustering importance is 0.85, accounting for 29.3% of the records. The smallest cluster is Cluster-4 (S5), with 59 records, accounting for 1.3% of all records. The clustering superiority is 0.97, and the clustering importance is 0.93. Clustering superiority is a measure of cluster separation (−1∼0.2 poor| 0.2∼0.5 medium| 0.5∼1 good), and clustering importance is a measure

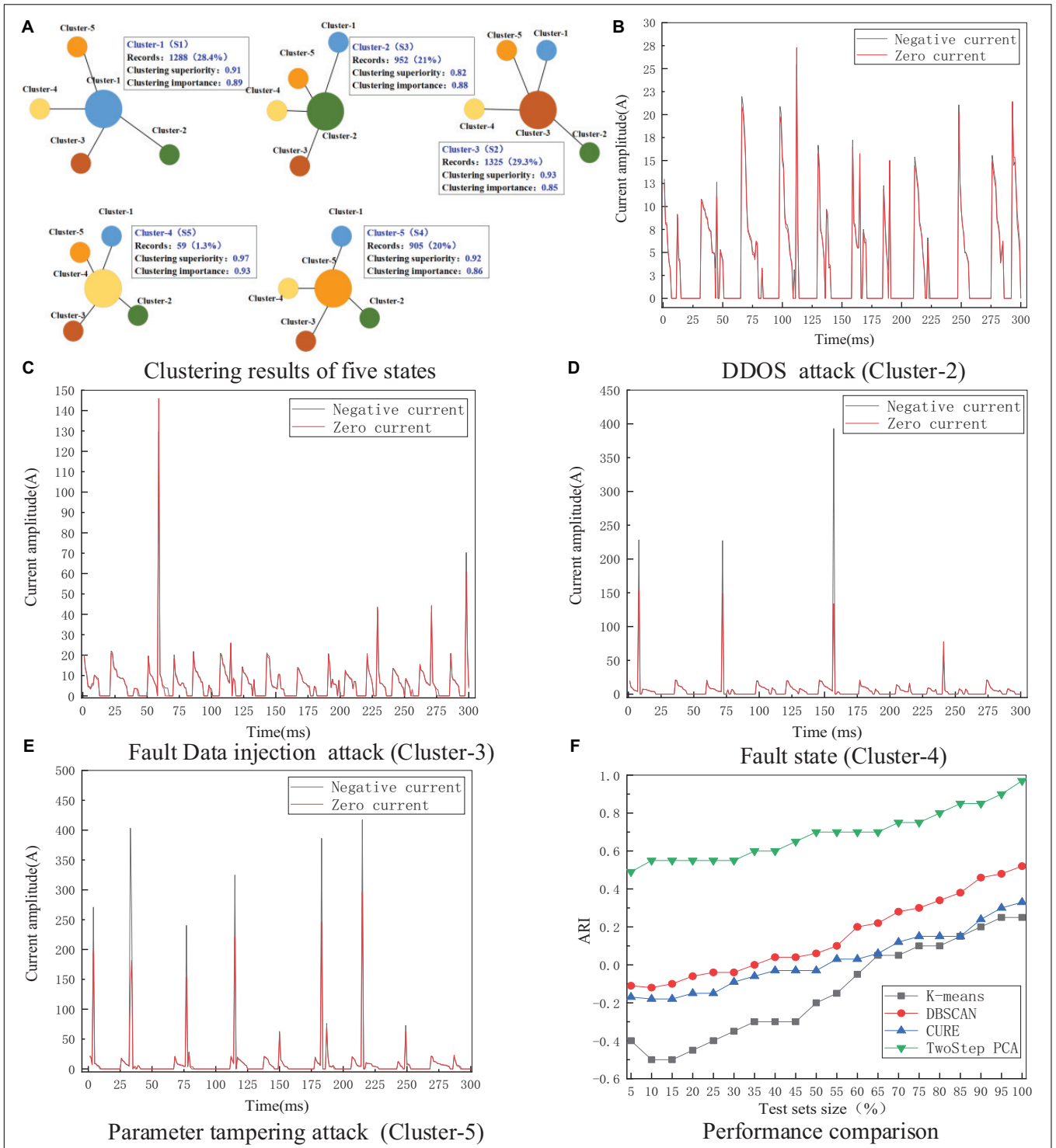


FIGURE 2 | Clustering results of five states **(A)**, Current amplitude under DDOS attack **(B)**, Current amplitude under fault data injection attack **(C)**, Current amplitude under fault state **(D)**, Current amplitude under parameter tampering attack **(E)**, Performance comparison **(F)**.

of cluster cohesion (0∼0.2 poor| 0.2∼0.6 medium| 0.6∼1 good) (Nair and Narendran, 1997).

According to the negative sequence current and zero sequence current amplitude of each cluster in the experiment, the curves of the three attack states and the fault state are obtained, as shown in **Figures 2B–E**. Cluster-2 is significantly different from the other four states, while Cluster-4 and Cluster-5 have high similarities. The reason is that Cluster-2 is an

attack that causes network blocking and delay, which is significantly different from the other types of data tampering attacks. Cluster-3 and Cluster-5 are physical power grid failures caused by information tampering attacks. These states are similar to the changes occurring in the Cluster-4 power grid normal fault.

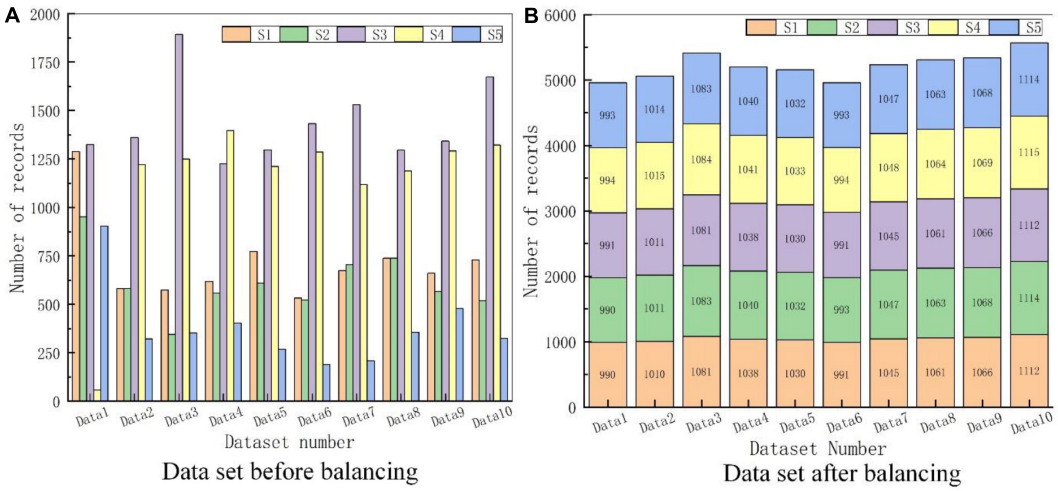


FIGURE 3 | Data set before balancing **(A)**, Data set after balancing **(B)**.

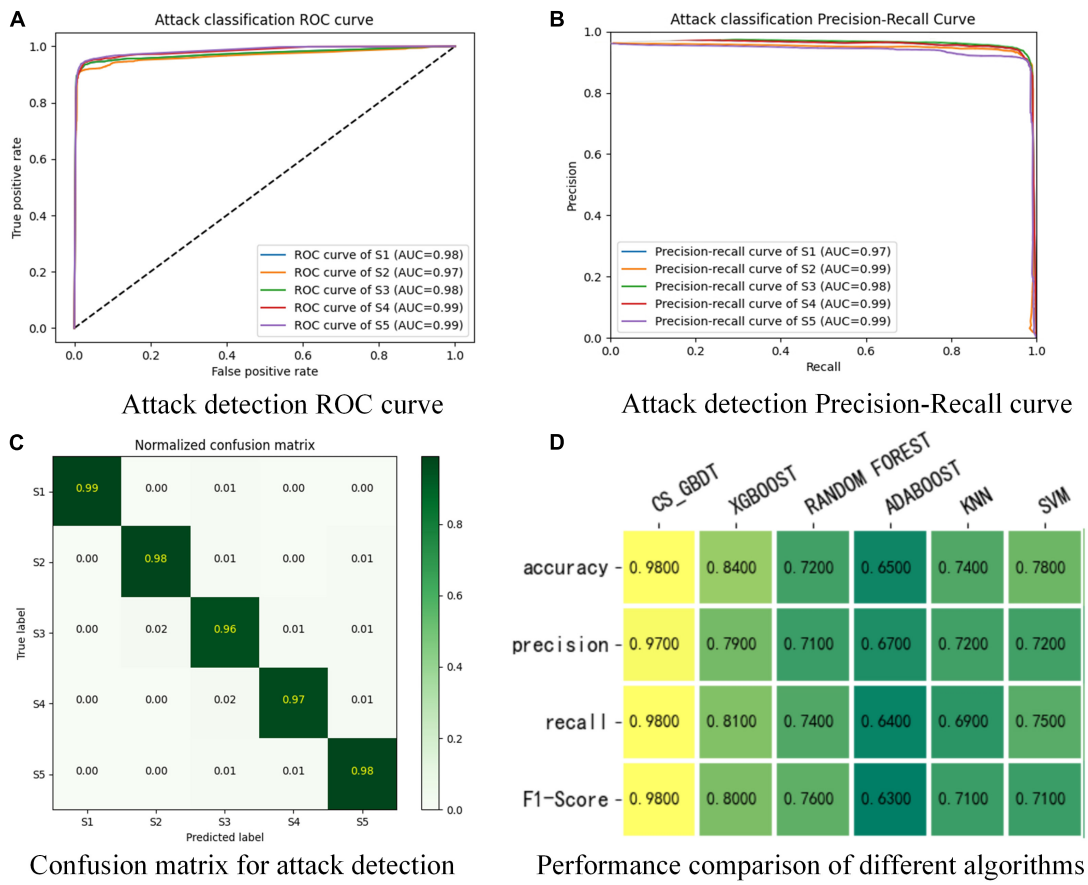


FIGURE 4 | Attack detection receiver operating characteristic (ROC) curve **(A)**, Precision-Recall curve **(B)**, Confusion matrix **(C)**, and Performance comparison **(D)**. AUC, area under the ROC curve; CS-GBDT, CS gradient boosting decision tree; KNN, *K*-nearest neighbors; SVM, support vector machine.

The adjusted Rand index (ARI) is used to measure the accuracy of the clustering results; ARI $\in [-1, 1]$, the closer the value is to 1, the better the clustering performance is. The index is calculated as follows:

$$\mathrm{ARI} = \frac{RI - E(RI)}{\max(RI) - E(RI)} \tag{17}$$

where RI is the Rand coefficient, and $E(RI)$ is the expected value of each class.

Four typical clustering algorithms are selected for performance comparison, i.e., K-means, density-based spatial clustering of applications with noise (DBSCAN), clustering using representatives (CURE), and BIRCH. In the experiment, the sample size of the test data set is randomly selected and ranges from 5 to 100% of the data set. The ARI values of the different algorithms are shown in **Figure 2F**. As the proportion of the test data set increases, the ARI increases significantly. The accuracy of the proposed two-step PCA method is 97% for a sample size of 100%, demonstrating the excellent performance of this method. The K-means algorithm has the lowest ARI values.

Result of Balancing the Operation State Classes

The number of samples in the operating state classes in 10 data sets before implementing the algorithm is shown in **Figure 3A**. The number of samples is imbalanced in the different operating states. The largest number of records (143,766) occurs in the S3 state, and the fewest number (3,080) is observed in the S5 state. The maximum class imbalance is 3.77. There are multiple minority and majority categories in the joint data set, showing multi-category imbalance.

The ADASYN algorithm is used to oversample the categories whose number is less than the threshold. We set the maximum imbalance threshold to 1.2. The results in the 10 data sets are shown in **Figure 3B**. The proportion of records in each dataset is close to 20%. The ADASYN algorithm uses local screening and sampling to reduce the influence of data imbalance on the false alarm rate of coordinated cyber-attack detection.

Performance Verification of the Coordinated Cyber-Attack Detection in the Cyber–Physical Power System

The balanced data set is divided into a training set (70% samples) and a test set (30% samples). The model loss parameters are set according to the improved CS loss function. There are 130 integrated base classifiers, and the depth of each independent tree (max_depth) is seven.

The receiver operating characteristic (ROC) curve obtained by classifying the test data set is shown in **Figure 4A**. The curves of the five categories are close to the (0,1) position, and the average area under the ROC curve (AUC) is 0.982. This result shows that the attack detection model has a low false alarm rate and high accuracy.

The precision-recall curve obtained by classifying the test data set is shown in **Figure 4B**. The precision-recall curve are all close to the (1,1) position, indicating that the attack detection model has high recall and accuracy, even when the ratio of positive and negative samples is large. Therefore, the proposed attack detection model has a high classification accuracy for unbalanced data.

The confusion matrix of the attack detection results is shown in **Figure 4C**. The detection accuracy for the DDOS blocking attack (S2) is 98%, that of the data injection attack (S3) is 96%, that of the protection device parameter tampering attack (S4) is 97%, that of the normal operation (S1) is 99%, and that of the fault operation (S5) is 98%. These results demonstrate that the proposed coordinated cyber-attack detection model accurately detects coordinated attack events on the network and distinguishes attack states from the fault operation state, with a maximum false-positive rate of only 4%.

Finally, the proposed model is compared with typical classification algorithms, including the KNN, Xgboost, Random Forest, Adaboost, and support vector machine (SVM). The overall accuracy, average recall, average precision, and average F1-score of the algorithms are shown in **Figure 4D**. The recall and precision of the CS-GBDT algorithm are higher than 97%. The algorithm performance is stable, and it provides better performance for detecting various attack events than comparable algorithms.

CONCLUSION

In this paper, a cyber–physical operating state data link was established using data fusion mapping. The two-step PCA clustering algorithm is proposed for accurate labeling of the different operating states of the network. A coordinated cyber-attack classifier based on the CS-GBDT was established that considers the imbalance of the attack status categories and the cost sensitivity of the attack event. The algorithm can detect attacks on the CPPS and distinguish different attack types. The proposed model has a low false alarm rate and high accuracy for attack detection. It is suitable for the detection of coordinated cyber-attack events with unbalanced attack sample data and high data dimensionality.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

LW designed the model framework of the manuscript and experimental verification. PX contributed to the construction method of the cyber–physical operation state data link. ZQ contributed to design the two-step PCA algorithm for operating state clustering. XB completed the simulation experiment of attack classification detection. YD performed the data collection and researched balance processing algorithm of operating

state classes. ZZ studied the classification algorithm of cyber cooperative attack based on CS_GBDT. YL built a simulation environment and improved the grammar and sentence structure of the full manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This manuscript was supported in part by the science and technology innovation development plan project of Jilin (20200401097GX).

REFERENCES

Basin, D., Cremers, C., Kim, T. H. J., Perrig, A., Sasse, R., and Szalachowski, P. (2016). Design, analysis, and implementation of ARPKI: an attack-resilient public-key infrastructure. *IEEE Trans. Depend. Sec. Comput.* 15, 393–408. doi: 10.1109/tdsc.2016.2601610

Chen, T. M., Sanchez-Aarnoutse, J. C., and Buford, J. (2011). Petri net modeling of cyber-physical attacks on smart grid. *IEEE Trans. Smart Grid* 2, 741–749. doi: 10.1109/tsg.2011.2160000

Dai, Q., Shi, L., and Ni, Y. (2019). Risk assessment for cyberattack in active distribution systems considering the role of feeder automation. *IEEE Trans. Power Syst.* 34, 3230–3240. doi: 10.1109/tpwrs.2019.2899983

Davarikia, H., and Barati, M. (2018). A tri-level programming model for attack-resilient control of power grids. *J. Modern Power Syst. Clean Energy* 6, 918–929. doi: 10.1007/s40565-018-0436-y

Dom, G., Shaw-Jackson, C., Matis, C., Bouffioux, O., Picard, J. J., Prochiantz, A., et al. (2003). Cellular uptake of Antennapedia Penetratin peptides is a two-step process in which phase transfer precedes a tryptophan-dependent translocation. *Nucleic Acids Res.* 31, 556–561. doi: 10.1093/nar/gkg160

Drayer, E., and Routtenberg, T. (2019). Detection of false data injection attacks in smart grids based on graph signal processing. *IEEE Syst. J.* 14, 1886–1896. doi: 10.1109/jsyst.2019.2927469

Haes Alhelou, H., Hamedani-Golshan, M. E., Njenda, T. C., and Siano, P. (2019). A survey on power system blackout and cascading events: Research motivations and challenges. *Energies* 12:682. doi: 10.3390/en12040682

Huang, L., and Zhu, Q. (2020). A dynamic games approach to proactive defense strategies against advanced persistent threats in cyber-physical systems. *Comput. Sec.* 89:101660. doi: 10.1016/j.cose.2019.101660

Huang, T., Satchidanandan, B., Kumar, P. R., and Xie, L. (2018). An online detection framework for cyber-attacks on automatic generation control. *IEEE Trans. Power Syst.* 33, 6816–6827. doi: 10.1109/tpwrs.2018.2829743

Jian, Y., David, Z., and Frangi, A. F. (2004). Two-dimensional PCA: a new approach to appearance-based face representation and recognition. *IEEE Trans. Pattern Analysis Machine Intelligence* 26, 131–137. doi: 10.1109/tpami.2004.1261097

Koopman, G., Mortier, D., Michels, S., Hofman, S., Fagrouch, Z., Remarque, E. J., et al. (2019). Influenza virus infection as well as immunization with DNA encoding haemagglutinin protein induces potent antibody-dependent phagocytosis (ADP) and monocyte infection-enhancing responses in macaques. *J. Gen. Virol.* 100, 738–751. doi: 10.1099/jgv.0.001251

Kurt, M. N., Yılmaz, Y., and Wang, X. (2018). Distributed quickest detection of cyber-attacks in smart grid. *IEEE Trans. Inform. Forensics Sec.* 13, 2015–2030. doi: 10.1109/tifs.2018.2800908

Lai, K., Illindala, M., and Subramaniam, K. (2019). A tri-level optimization model to mitigate coordinated attacks on electric power systems in a cyber-physical environment. *Appl. Energy* 235, 204–218. doi: 10.1016/j.apenergy.2018.10.077

Liao, Z., Huang, Y., Yue, X., Lu, H., Xuan, P., and Ju, Y. (2016). In silico prediction of gamma-aminobutyric acid type-A receptors using novel machine-learning-based SVM and GBDT approaches. *BioMed. Res. Int.* 2016:2375268.

Lin, H., Slagell, A., Kalbarczyk, Z. T., Sauer, P. W., and Iyer, R. K. (2016). Runtime semantic security analysis to detect and mitigate control-related attacks in power grids. *IEEE Trans. Smart Grid* 9, 163–178. doi: 10.1109/tsg.2016.2547742

Liu, X., Li, Z., and Li, Z. (2016). Optimal protection strategy against false data injection attacks in power systems. *IEEE Trans. Smart Grid* 8, 1802–1810. doi: 10.1109/tsg.2015.2508449

Nair, G. J., and Narendran, T. T. (1997). Cluster goodness: a new measure of performance for cluster formation in the design of cellular manufacturing systems. *Int. J. Prod. Econ.* 48, 49–61. doi: 10.1016/s0925-5273(96)00067-9

Nath, S., Akingeneye, I., Wu, J., and Han, Z. (2019). Quickest detection of false data injection attacks in smart grid with dynamic models. *IEEE J. Emerg. Selected Top. Power Electron.* 99, 1–10.

Northrup, A. B., Mangion, I. K., Hettche, F., and MacMillan, D. W. (2004). Enantioselective organocatalytic direct aldol reactions of α−oxyaldehydes: step one in a two-step synthesis of carbohydrates. *Angewandte Chem. Int. Edition* 43, 2152–2154. doi: 10.1002/anie.200453716

Osanaiye, O. A., Alfa, A. S., and Hancke, G. P. (2018). Denial of service defence for resource availability in wireless sensor networks. *IEEE Access* 6, 6975–7004. doi: 10.1109/access.2018.2793841

Phelps, R. A., Chidester, S., Dehghanizadeh, S., Phelps, J., Sandoval, I. T., Rai, K., et al. (2009). A two-step model for colon adenoma initiation and progression caused by APC loss. *Cell* 137, 623–634. doi: 10.1016/s9999-9994(09)00528-5

Qu, Z., Zhang, Y., Qu, N., Wang, L., Li, Y., and Dong, Y. (2018). Method for quantitative estimation of the risk propagation threshold in electric power CPS based on seepage probability. *IEEE Access* 6, 68813–68823. doi: 10.1109/access.2018.2879488

Sakhnovich, A. L. (2011). The time-dependent Schrödinger equation of dimension k+ 1: explicit and rational solutions via GBDT and multinodes. *J. Phys. A: Math. Theoret.* 44:475201. doi: 10.1088/1751-8113/44/47/475201

Shen, Y., Zhang, W. A., Ni, H., Zhang, D., and Yu, L. (2019). Guaranteed cost control of networked control systems with DoS attack and time-varying delay. *Int. J. Control Automat. Syst.* 17, 811–821. doi: 10.1007/s12555-018-0324-2

Tian, W., Ji, X., Liu, W., Liu, G., Lin, R., Zhai, J., et al. (2019). Defense strategies against network attacks in cyber-physical systems with analysis cost constraint based on honeypot game model. *Comput. Mater. Continua* 60, 193–211. doi: 10.32604/cmc.2019.05290

Wang, L., Qu, Z., Li, Y., Hu, K., Sun, J., Xue, K., et al. (2020). Method for extracting patterns of coordinated network attacks on electric power CPS based on temporal–topological correlation. *IEEE Access* 8, 57260–57272. doi: 10.1109/access.2020.2982057

Wang, Q., Tai, W., Tang, Y., Ni, M., and You, S. (2019). A two-layer game theoretical attack-defense model for a false data injection attack against power systems. *Int. J. Elect. Power Energy Syst.* 104, 169–177. doi: 10.1016/j.ijepes.2018.07.007

Wang, X., Tian, M., Cao, M., Li, X., Zhao, Y., Zhao, X., et al. (2019). Countermeasures to false data injection attacks on power system state estimation based on protecting measurements. *J. Nanoelectron. Optoelectron.* 14, 626–634. doi: 10.1166/jno.2019.2590

Xu, C., and Abur, A. (2017). A massively parallel framework for very large scale linear state estimation. *IEEE Trans. Power Syst.* 33, 4407–4413. doi: 10.1109/tpwrs.2017.2788360

Zhang, Y., Xiang, Y., and Wang, L. (2016). Power system reliability assessment incorporating cyber-attacks against wind farm energy management systems. *IEEE Trans. Smart Grid* 8, 2343–2357. doi: 10.1109/tsg.2016.2523515

Conflict of Interest: PX was employed by Siping Power Supply Company of State Grid Jilin Electric Power Company Limited, China.

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

frontiers in Energy Research

METHODS
published: 22 April 2021
doi: 10.3389/fenrg.2021.641518

Optimized Energy Storage System Configuration for Voltage Regulation of Distribution Network With PV Access

Qiang Li[1]*, *Feijie Zhou*[2], *Fuyin Guo*[2], *Fulin Fan*[3] and *Zhengyong Huang*[1]

[1]School of Electrical Engineering, Chongqing University, Chongqing, China, [2]Key Laboratory of Modern Power System Simulation and Control and Renewable Energy Technology, Ministry of Education (Northeast Electric Power University), Jilin, China, [3]Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, United Kingdom

With the large-scale integration of renewable energy such as wind power and PV, it is necessary to maintain the voltage stability of power systems while increasing the use of intermittent renewable energy sources. The rapid development of energy storage technologies permits the deployment of energy storage systems (ESS) for voltage regulation support. This paper develops an ESS optimization method to estimate the optimal capacity and locations of distributed ESS supporting the voltage regulation of a distribution network. The electrical elements of the network integrated with PV and ESS are first modelled to simulate the voltage profile of the network. Then an improved multi-objective particle swarm optimization (PSO) algorithm is employed to minimise a weighted sum of the overall nodal voltage deviation from the nominal level across the network and across the time horizon and the energy capacity of ESS reflecting the associated investment. The improved PSO algorithm adaptively adjusts the inertia weight associated with each particle based on its distance from the best known particle of the population and introduces the cross-mutation operation for a small distance to avoid falling into local optimal solutions. Then the dynamic dense distance arrangement is taken to update the non-inferior solution set and indicate potential global optimal solutions so as to keep the scale and uniformity of the optimal Pareto solution set. To mitigate the impact of decision makers' preference, the information entropy based technique for order of preference by similarity to ideal solution is used to select the optimal combination of the ESS access scheme and capacity from the Pareto solution set. The proposed ESS optimization method is tested based on the IEEE 24-bus system with additional imports from high-voltage power supply. The voltage profile of the network simulated without the ESS or with the random or optimized ESS placement is compared to illustrate the effectiveness of the optimized ESS in performing voltage regulation under normal operation and supporting emergency power supply during high-voltage transmission failures.

Keywords: distribution network, energy storage system, particle swarm optimization, photovoltaic energy, voltage regulation

OPEN ACCESS

Edited by:
Yang Li,
Northeast Electric Power University, China

Reviewed by:
Liang Chen,
Nanjing University of Information Science and Technology, China
Shaoyan Li,
North China Electric Power University, China

***Correspondence:**
Qiang Li
13844230965@163.com

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

Received: *14 December 2020*
Accepted: *08 February 2021*
Published: *22 April 2021*

Citation:
Li Q, Zhou F, Guo F, Fan F and Huang Z (2021) Optimized Energy Storage System Configuration for Voltage Regulation of Distribution Network With PV Access. Front. Energy Res. 9:641518. doi: 10.3389/fenrg.2021.641518

INTRODUCTION

Due to the continuous consumption of fossil fuels and the resulting aggravation of environmental pollution, the utilization of renewable energy sources (RES) has developed rapidly in recent years. A high proportion of RES integrated within distribution networks separately features the superposition of the output randomness and volatility which may cause violent fluctuations of the node voltage (Jung et al., 2014; Watson et al., 2016; Hasheminamin et al., 2018) and thus seriously affect the power quality of networks. Therefore, it is of great significance to study the problem of voltage limit exceedance given the large-scale integration of RES.

With the rapid development of electrochemical energy storage technologies, the technical and economic maturity of battery energy storage technology have been greatly improved. The energy storage systems (ESS) installed within electrical grids can effectively improve the grid's ability to absorb renewable energy and deal with integration problems such as the voltage limit violation caused by the high penetration of renewable energy. Meanwhile, it can provide emergency power supply when a short-circuit fault occurs, mitigating the system impacts caused by serious failures.

The voltage regulation of active distribution networks have been investigated in considerable research which can be divided into three main aspects: 1) changing electric power distribution; 2) adjusting transformer taps; and 3) installing reactive power compensation devices. References (Zhou and Bialek, 2007; Tonkoski and Lopes, 2011; Yap et al., 2014) have proposed strategies to reduce the PV power output to improve the voltage level at the cost of a decreased efficiency of solar energy conversion. References (Jashfar and Esmaeili, 2014; Collins and Ward, 2015; Zad et al., 2015) put forward voltage regulation approaches by using PV system inverters for reactive power compensation. In addition, references (Esslinger and Witzmann, 2012; El Moursi et al., 2014) improved voltage levels based on the on-load regulating transformer. From the perspective of model predictions, Zhao et al. (Zhao et al., 2016) developed a method for voltage regulation through installing reactive power compensation devices and adjusting wind turbine outputs. In (Yoshida et al., 2000), voltage regulation was implemented by adjusting generator outputs, transformer taps, and the capacity of reactive power compensation devices. A coordinated control method for on-load voltage regulation and reactive power compensation was proposed in (Salih and Chen, 2016; Kulmala et al., 2014) for distribution systems with RES. Similarly, reference (Brenna et al., 2013) developed a distributed voltage control method to automatically regulate the voltage and reactive power of substations. Reference (Senjyu et al., 2008) optimized the control of distribution voltage with the coordination of distributed installations including the load ratio control transformer, step voltage regulator, shunt capacitor, shunt reactor, and static var compensator (SVC). In (Bakir and Kulaksiz, 2019), four proportional-integral controllers in STATCOM were optimized to reduce the voltage fluctuation and improve the voltage stability. Dong et al. (Dong et al., 2005) made full use of the reactive power margin of the generator for voltage regulation. In (Azzouz et al., 2015), the vehicle-to-grid technology, distributed power supply, and transformer tap were coordinated to regulate the voltage in real time. Daratha et al. (Daratha et al., 2014) proposed a two-stage control strategy comprising transformer tap adjustment and SVC compensation. In (Alam et al., 2012), a local charge and discharge controller was designed for distributed ESS to deal with the over-voltage problem at the point of common coupling caused by PV generation. Based on the consistency algorithm, a distributed coordinated control method was developed in (Mokhtari et al., 2012) which, however, did not consider the differences of state of charge (SOC) and constraints on charge and discharge rates between distributed ESS. A mathematical model of the ESS within power system was described in (Bachurin et al., 2018). In addition, Sugihara et al. (Sugihara et al., 2013) assessed the feasibility of employing the user-side ESS to suppress voltage fluctuations from the economic point of view. In (Mehmood et al., 2017), the ESS location within the system was optimised to improve the service life and voltage regulation of ESS. Liu et al. (Liu et al., 2012) proposed a coordinated control of distributed ESS with traditional voltage regulators including on-load tap changer transformers and step voltage regulators to alleviate the voltage rise problem caused by the high PV penetration in the low-voltage distribution network. In (Nara et al., 2005), voltage control effects of distributed generators (DGs) were discussed based on a simplified radial distribution network. A voltage control method performing the partial generation rejection of PV systems was proposed in (Lin et al., 2012), based on which the installed capacity of PV systems was optimized to maximize the net present value of the systems. Muttaqi et al. (Muttaqi et al., 2013) proposed a control coordination technique which exploited the DG ability as a voltage regulator and minimized the interaction with other active devices. Li et al. (Yang et al., 2018) proposed a two-stage optimization method for the optimal distributed generation planning with the ESS integration. Mehmood et al. (Mehmood et al., 2017) optimised the location and size of ESS considering the battery life and the voltage quality of distribution network. Jannesar et al. (Jannesar et al., 2018) optimised the location and capacity of a battery ESS as well as day-ahead schedules based on a cost function including energy arbitrage, environmental emission, energy losses, transmission access fee, as well as capital and operating costs of the battery ESS.

To alleviate the voltage limit violation caused by the increased use of RES, many literatures regulate the node voltage from the perspective of conventional generator outputs and reactive power compensation devices but rarely consider the optimization of ESS locations and sizes for voltage regulation. Therefore, this paper will propose an ESS optimization method to estimate the best capacity and locations of distributed ESS that support the voltage regulation and ensure the voltage stability of a distribution network. An improved multi-objective particle swarm optimization algorithm (PSO) is first used to minimize a weighted sum of the ESS energy capacity and the overall node voltage deviation from the nominal level simulated across the network and across the time horizon, producing the Pareto solution set which comprises potential global optimal

solutions. Then the information entropy based technique for order of preference by similarity to ideal solution (TOPSIS) is used to determine the best combination of ESS capacity and locations from the Pareto solution set in order to mitigate the impact of decision-makers' preference. The effectiveness of the proposed ESS optimization method is examined based on the IEEE 24-bus system through a comparison between the voltage profiles of the network without the ESS deployment or with the random or optimized ESS placement. In addition to the voltage regulation under normal operating conditions, the ESS performance in the emergency power supply is assessed when the network fails to import from additional high-voltage power supply.

This paper is structured as follows: describes the *Modeling of Distribution Network Integrated With ESS and PV*; establishes a multi-objective *Optimization Model of ESS Capacity and Locations for Voltage Regulation*; develops an *Improved Multi-Objective Particle Swarm Optimization Algorithm* for the ESS access scheme optimization based on the TOPSIS; *Case Study* implements simulation experiments based on the IEEE 24-bus system to validate the performance of the optimized ESS in the voltage regulation and emergency power supply; and presents *Conclusion and Recommendations for Future Work*.

MODELING OF DISTRIBUTION NETWORK INTEGRATED WITH ESS AND PV

This section presents the modelling of distribution network, ESS and PV system. The distribution network model includes the network structure, load component and voltage distribution across the network without or with the connection of the ESS which is represented by an equivalent circuit of battery in combination with associated parameters.

Distribution Network Model

Most of the distribution networks in China have radial chain structures. The active and reactive power of the load within a distribution network are changing over time. A constant power static model is employed here for each instant to facilitate the simulation of voltage regulation. **Supplementary Figure A1** shows a specific distribution network structure where a particular feeder is adopted for the ESS access.

Without the connection of the ESS, the voltage U_m at the mth node ($m = 1, \ldots, N$) can be calculated as the difference between the voltage U_0 at the node zero (i.e., a local substation) and the aggregate of the voltage drop ΔU_p along each distribution line ($p = 1, \ldots, m$) that transfers electricity from the node zero to the mth node:

$$U_m = U_0 - \sum_{p=1}^{m} \Delta U_p = U_0 - \sum_{p=1}^{m} \frac{R_p \sum_{n=p}^{N} P_n + X_p \sum_{n=p}^{N} Q_n}{U_p} \quad (1)$$

where ΔU_p along the pth distirbution line that supports active P_n and reactive power Q_n of the load at the downstream nth node ($n = p, \ldots, N$) is determined based on its resistance R_p and inductance X_p combined with the voltage U_p at the pth node.

When the ESS accesses to the ith node, the voltage U_m of the mth node located upstream of the ith node is formulated by:

$$U_m = U_0 - \sum_{p=1}^{m} \frac{R_p\left(\sum_{n=p}^{N} P_n - P_{ess}\right) + X_p\left(\sum_{n=p}^{N} Q_n - Q_{ess}\right)}{U_P} \quad (2)$$

where P_{ess} and Q_{ess} denote active and reactive power outputs of the ESS. When the mth node is located downstream of the ith node, the voltage of the mth node can be determined as the difference between the voltage U_i at the ith node and the voltage drop $\Delta U_{i,m}$ along distribution lines connecting the ith node to the mth node:

$$\begin{aligned} U_m &= U_i - \Delta U_{i,m} \\ &= U_0 - \sum_{p=1}^{i} \frac{R_p\left(\sum_{n=p}^{N} P_n - P_{ess}\right) + X_p\left(\sum_{n=p}^{N} Q_n - Q_{ess}\right)}{U_P} \\ &\quad - \sum_{p=i}^{m} \frac{R_p \sum_{n=p}^{N} P_n + X_p \sum_{n=p}^{N} Q_n}{U_p} \end{aligned} \quad (3)$$

Eqs 1–3 show that the load distribution across the network, active and reactive power outputs of DGs and ESS as well as their locations within the network all affect the voltage profile of the network.

ESS Model

The widely employed lithium battery ESS is modelled in this study. The lithium battery is an electrochemical energy storage device which realizes the conversion between chemical and electric energy through its charging and discharging processes. In addition, it has high storage energy density and does not produce harmful heavy metal elements and substances, showing the characteristics of lightweight and green environmental protection. The terminal voltage of the battery generally changes with its current, i.e., the higher the battery current, the smaller the terminal voltage. Given a particular current level, the terminal voltage remains the same when the SOC is between 20 and 80% where a battery ESS mostly operates in practice. The SOC is an important basis for charging and discharging control and requires to be updated in real time (Fan et al., 2021). Given the battery ESS with energy capacity of Ah_{nom} starting to operate from a full SOC, its SOC at a particular time t is calculated from the integral of its discharge and charge currents I over time:

$$SOC(t) = \frac{Ah_{nom} - \int_0^t I(\tau)d\tau}{Ah_{nom}} \quad (4)$$

Figure 1 shows a commonly used battery model consisting of an ideal voltage source E_0 and an equivalent internal resistance r (Rosewater et al., 2019). According to the full circuit Ohm's law, the terminal voltage V is formulated by:

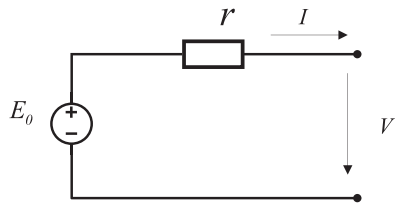


FIGURE 1 | Simplified equivalent circuit diagram of the energy storage battery.

$$V = E_0 - Ir \tag{5}$$

A more detailed equivalent circuit of the battery (Rosewater et al., 2019) is shown in **Figure 2**. The open-circuit voltage V_{OCV} is a function of the SOC and the internal resistance R_0 of the battery in combination with additional impedance parameters (R_1, C_1, R_2, and C_2) that are used to describe dynamic characteristics of the battery. The terminal voltage V of the battery can be formulated by **Eq. 6** where only V_{OCV} is related to I and SOC. **Supplementary Table S1** lists typical values of model parameters in **Eq. 6** based on A123-M1 LiFePO4 battery (Cheng et al., 2016).

$$\begin{aligned} V &= V_{OCV}(SoC) - \left(R_0 + \frac{R_1}{1+sR_1C_1} + \frac{R_2}{1+sR_2C_2}\right)I \\ &= V_{OCV}(SoC) - R_0 I - \left(\frac{K(1+sK_1)}{(1+sK_2)(1+sK_3)}\right)I \end{aligned} \tag{6}$$

PV System Model

A PV system is composed of several PV panels that are combined according to certain rules. The aggregate power output of a PV system fluctuates with the change of light intensity which also has uncertainties. A particular model of the grid-connected PV system co-located with a battery ESS is shown in **Supplementary Figure A2**.

According to (Atwa et al., 2010; Teng et al., 2013), power outputs of a PV system could be modelled by the beta probability density function based on PV module parameters. Part of the electricity produced by the PV system during peak sunshine hours can be absorbed by the ESS through "peak shaving and valley filling"; when the PV system output is small or at night, then the ESS will put the electric energy onto the network for power supply, helping match the PV-ESS system output curve with the load profile in daily operation.

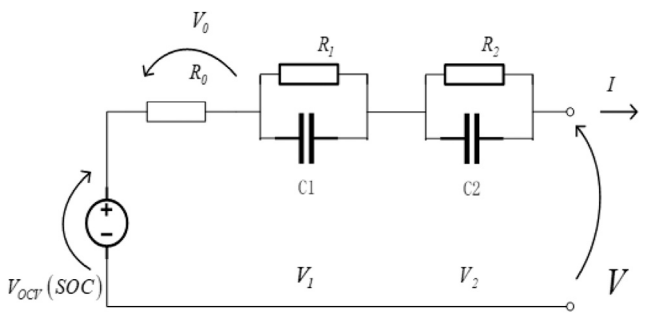


FIGURE 2 | Detailed equivalent circuit diagram of the energy storage battery.

OPTIMIZATION MODEL OF ESS CAPACITY AND LOCATIONS FOR VOLTAGE REGULATION

When a distribution network connects large-scale RES via ultra-high-voltage (UHV) transmission channels, its node voltage fluctuations will increase due to the randomness and volatility of renewable generation. In general, the ESS can be introduced to regulate the voltage profile of the main network and suppress the voltage fluctuations caused by groups of Distributed Generation Integration. Therefore, the selection of ESS locations is important for an optimal ESS configuration. It determines the efficiency of voltage regulation support from ESS to the greatest extent and helps the main network to maintain stable and reliable operation. This section will establish a multi-objective optimization model including the ESS capacity and locations for an efficient voltage regulation. The effective control of charging and discharging of the ESS placed at more suitable access points can better reduce the node power fluctuation and improve the voltage stability of the network.

Objective Function

The access locations of ESS within the distribution network are optimized here taking into account a trade-off between the voltage regulation performance and the total energy capacity of ESS which are described by the normalised functions f_1^* and f_2^*, respectively. The weight coefficients α and β are introduced to transform the functions f_1^* and f_2^* into a single objective function f in **Eq. 7**. The voltage regulation performance f_1 of ESS is inversely quantified based on the overall deviation between the node voltage U_i ($i = 1, \ldots, N$) and the nominal level U_0 across the network and across the time horizon ($t = 1, \ldots, T$) in **Eq. 8**. The total ESS energy capacity E_{TC} representing f_2 in **Eq. 9** and f_1 are normalised by **Eq. 10** based on their respective minimum $f_{i\,\min}$ and maximum $f_{i\,\max}$ values such that f_1^* and f_2^* range from 0 to 1.

$$f = \min\left(\alpha f_1^* + \beta f_2^*\right) \tag{7}$$

$$f_1 = \left(\sum_{i=1}^{N}\sum_{t=1}^{T}\left(U_i(t) - U_0\right)^2\right) \tag{8}$$

$$f_2 = E_{\mathrm{TC}} \tag{9}$$

$$f_i^* = \frac{f_i - f_{i\min}}{f_{i\max} - f_{i\min}} \tag{10}$$

Constraints

Operation Constraints of Distribution Network

(1) Power balance constraints of distribution network:

$$\begin{cases} P_{Gi}(t) + P_{RESi}(t) - P_{Li}(t) + \Delta P_i(t) + P_{Ei}(t) = 0 \\ Q_{Gi}(t) - Q_{Li}(t) + \Delta Q_i(t) + Q_{Ei}(t) = 0 \end{cases} \tag{11}$$

$$\begin{cases} \Delta P_i(t) = U_i(t)\sum_{j=1}^{N} U_j(t)\left(G_{ij}\cos\theta_{ij} + B_{ij}\sin\theta_{ij}\right) \\ \Delta Q_i(t) = U_i(t)\sum_{j=1}^{N} U_j(t)\left(G_{ij}\sin\theta_{ij} + B_{ij}\cos\theta_{ij}\right) \end{cases} \tag{12}$$

where subscripts *Gi*, *RESi*, *Li*, and *Ei* in **Eq. 11** represent the conventional generating unit, RES, load and ESS connected to the *i*th node; terms $\Delta P_i(t)$ and $\Delta Q_i(t)$ are the active and reactive power flowing into the *i*th node; terms G_{ij} and B_{ij} in **Eq. 12** denote the real and imaginary parts of the element in the bus admittance matrix corresponding to the *i*th and *j*th nodes and θ_{ij} is the difference in voltage angle between the *i*th and *j*th nodes.

(2) Node voltage constraints where $\underline{U_i}$ and $\overline{U_i}$ denote lower and upper limits of node voltage:

$$\underline{U_i} \le U_i(t) \le \overline{U_i} \tag{13}$$

(3) Constraints on distribution network branch power where $\overline{P_{i,j}}$ is the power transfer limit on the line connecting the *i*th node to the *j*th node:

$$\begin{cases} 0 \le \left|P_{i,j}(t)\right| \le \overline{P_{i,j}} \\ \left|P_{i,j}(t)\right| = \left|U_i(t)U_j(t)\left(G_{ij}\cos\theta_{ij} + B_{ij}\sin\theta_{ij}\right) - U_i(t)^2 G_{ij}\right| \end{cases} \tag{14}$$

(4) Ramp rate constraints of conventional generating units where RD_i or RU_i is the limit on ramping down or up:

$$-RD_i \le P_{Gi}(t) - P_{Gi}(t-1) \le RU_i \tag{15}$$

(5) Export limits of conventional generating units where subscript *Gimin* or *Gimax* denotes the minimum or maximum limit:

$$\begin{cases} P_{Gi\min} \le P_{Gi}(t) \le P_{Gi\max} \\ Q_{Gi\min} \le Q_{Gi}(t) \le Q_{Gi\max} \end{cases} \tag{16}$$

ESS Configuration Constraints

(1) The constraints of ESS access points:

$$E_{\mathrm{BC}} N_{\max} \le E_{\mathrm{TC}} \tag{17}$$

$$0 \le \sum_{i=1}^{N} N_{\mathrm{ESS}}(i) \le N_{\max} \tag{18}$$

where E_{BC} denotes the energy capacity of a single ESS, and given the limit of the total energy capacity E_{TC}, $N_{\max}$ is the maximum number of ESS that could be installed within the network which in turn limits the number $N_{\mathrm{ESS}}(i)$ of ESS connected to each node.

(2) ESS capacity and power constraints:

$$E_{\mathrm{sto}}(i) = E_{\mathrm{BC}} N_{\mathrm{ESS}}(i) \tag{19}$$

$$\begin{cases} P_{Ei}(t) = P_{Ei}^{D}(t) - \mathrm{P}_{Ei}^{C}(t) \\ P_{Ei}^{D}(t) \times P_{Ei}^{C}(t) = 0 \end{cases} \tag{20}$$

$$\begin{cases} 0 \le P_{Ei}^{C}(t) \le \overline{P_{Ei}^{C}} \\ 0 \le P_{Ei}^{D}(t) \le \overline{P_{Ei}^{D}} \end{cases} \tag{21}$$

$$\begin{cases} \sqrt{\left(P_{Ei}(t)\right)^2 + \left(\mathrm{Q}_{Ei}(t)\right)^2} \le S_{\max i} \\ -Q_{Ei\max} \le Q_{Ei}(t) \le Q_{Ei\max} \end{cases} \tag{22}$$

where E_{sto} is the total ESS energy capacity at a particular node and superscripts *D* and *C* denote discharging and charging phases of ESS, respectively, which must not violate the maximum allowable discharge rate $\overline{P_{Ei}^{D}}$ and charge rate $\overline{P_{Ei}^{C}}$. The total reactive power Q_{Ei} of ESS at the *i*th node should meet the reactive power limit $Q_{Ei\max}$ as well as the apparent power limit $S_{\max i}$ given their aggregate active power output P_{Ei}.

(3) The constraints of ESS operation:

$$\begin{cases} \underline{E_i} \le E_i(t) \le \overline{E_i} \\ \underline{E_i} = 0.2E_{\mathrm{sto}}(i), \ \ \overline{E_i} = E_{\mathrm{sto}}(i) \end{cases} \tag{23}$$

$$E_i(\mathrm{t}) = E_{i0} + \sum_{\tau=1}^{t}\left(P_{Ei}^{C}(\tau)\eta_c - P_{Ei}^{D}(\tau)/\eta_d\right) \tag{24}$$

where $E_i(t)$ is the total energy stored in ESS at the *i*th node at time step *t* which is determined based on their charge and discharge rates over previous time steps in combination with charging and discharging efficiencies denoted by η_c and η_d, respectively, and is maintained between 20 and 100% of $E_{\mathrm{sto}}(i)$.

IMPROVED MULTI-OBJECTIVE PARTICLE SWARM OPTIMIZATION ALGORITHM

Particle Swarm Optimization

The particle swarm optimization (PSO) is a stochastic optimization algorithm that minimizes an objective function starting from a group of randomly produced particles which travel in the problem space until convergence is reached (Kennedy and Eberhart, 1995). The particles jointly adjust speeds and directions of their motion based on their own and group information and search for the optimal solution through iteration. In the *k*th iteration, each particle updates its speed $v_{id}^{(k+1)}$ and position $x_{id}^{(k+1)}$ by tracking the best known positions of the particle itself and the group denoted by $p_{id}^{(k)}$ and $g_{id}^{(k)}$, respectively:

$$v_{id}^{(k+1)} = \omega v_{id}^{(k)} + c_1 r_1\left(p_{id}^{(k)} - x_{id}^{(k)}\right) + c_2 r_2\left(g_d^{(k)} - x_{id}^{(k)}\right) \tag{25}$$

$$x_{id}^{(k+1)} = x_{id}^{(k)} + v_{id}^{(k+1)} \tag{26}$$

where ω is an inertia weight associated with the present particle speed $v_{id}^{(k)}$ and parameters c_1, c_2, r_1, and r_2 are, respectively, two positive constants and two random parameters between 0 and 1

associated with the deviations from the present particle position $x_{id}^{(k)}$ to $p_{id}^{(k)}$ and $g_{id}^{(k)}$.

The conventional PSO-based method dealing with a multi-objective problem combines the Pareto sorting mechanism with the basic PSO algorithm to determine the best known solution of the particle and update the non-inferior solution set based on the dominating relationship between particles. The conventional multi-objective PSO algorithm has the following problems in the solution process: 1) lack of guidance on the value of ω; 2) poor diversity and distribution of the Pareto solution set due to the improper update and maintenance strategy of the non-inferior solution set; 3) lack of guidance on the selection of the global optimal solution; and that 4) the loss of population diversity is fast and the PSO is likely to fall into local optima. Therefore, the paper makes the following improvements on the basis of the Pareto sorting based multi-objective PSO algorithm.

Improved Multi-Objective Particle Swarm Optimization

Adaptive Inertia Weight

The inertia weight ω has a significant impact on the convergence performance of the PSO algorithm. Most of the commonly used methods linearly or non-linearly reduce the ω value with the increasing number of iterations without considering the particle characteristics in the iterative process, and thus lack guidance on the adjustment of ω. The difference between the particle position vector and the best known solution of the population can reflect the distance from a particular particle to the best known particle of the population. When there is a larger gap between the current particle and the best known particle of the population, a greater ω should be used to permit the particle having better global searchability. When the distance from the best known particle of the population is reduced, a smaller ω will be adopted for better local searchability. In this study, the gap $X_i^{(k)}$ between a particular particle and the best known particle of the population estimated by **Eq. 27** is used to determine the value of ω as formulated by **Eq. 28**. **Supplementary Figure A3** shows the non-linear adjustment of $\omega_i^{(k)}$ between $\omega_{start} = 0.9$ and $\omega_{end} = 0.4$ based on $X_i^{(k)}$.

$$X_i^{(k)} = \frac{1}{x_{\max} - x_{\min}} \frac{1}{D} \sum_{d=1}^{D} \left| g_d^{(k)} - x_{id}^{(k)} \right| \tag{27}$$

$$\omega_i^{(k)} = \omega_{start} - (\omega_{start} - \omega_{end}) \left(X_i^{(k)} - 1 \right)^2 \tag{28}$$

Cross-Mutations

The PSO algorithm suffering from a premature convergence problem in iterative optimization may produce a local optimal solution. In order to increase the diversity of the population, the paper incorporates the cross-mutation operation of the genetic algorithm (Senjyu et al., 2006) into the PSO to cross-mutate the position vector of particles based on the difference X between the position vector of the particle and the best known position of the population. The specific steps are as follows:

(1) specify the threshold $X_{\min} = 0.1$ for X, mutation rate $p_m = 0.05$, and crossover rate $p_c = 0.1$;

(2) determine X_i for the particle i, and for $X_i < X_{\min}$, carry out crossover mutation on the particle; otherwise, go to step 5;

(3) select a random number r_{id} within [0,1] for the position component of the particle i along each dimension, and for $r_{id} < p_m$, initialize the d-dimensional position component by **Eq. 29**;

(4) after the mutation, for $r_{id} < p_c$, cross the d-dimensional position component with the best known position of the population (i.e., the crossover object g_d), as shown in **Supplementary Figure A4**;

(5) the cross-mutation is completed.

$$x_{id} = x_{\min} + (x_{\max} - x_{\min}) \cdot r \tag{29}$$

Non-Inferior Solution Set Update Based on Dynamic Dense Distance

When solving a multi-objective problem, the non-inferior solution set is updated in each iteration. To keep the scale of the Pareto solution set and the uniformity of solution distribution, it is necessary to select the best Pareto solution. The dense distance is used here to describe the density between particles and their surrounding particles as well as the uniformity of solutions. For a multi-objective optimization problem consisting of three objective functions f_1, f_2, and f_3 as shown in **Supplementary Figure A5**, the dense distance $I(x_i)$ of the particle i is calculated by:

$$I(x_i) = \frac{\left| \left[f_1(x_j) - f_1(x_k) \right] \right|}{f_{1\max}} + \frac{\left| \left[f_2(x_j) - f_2(x_k) \right] \right|}{f_{2\max}} + \frac{\left| \left[f_3(x_j) - f_3(x_k) \right] \right|}{f_{3\max}} \tag{30}$$

If the number of objective functions is n, then the dense distance of the particle i can be written as:

$$I(x_i) = \frac{\sum_{m=1}^{n} \left| \left[f_m(x_j) - f_{\mathrm{m}}(x_k) \right] \right|}{f_{m\max}} \tag{31}$$

Once the dense distance of each Pareto solution is solved, it is sorted in a descending order and then screened. The common method is to select N solutions with large dense distances in sequence. Although this method is fast and only needs to calculate the dense distance of Pareto solution once in each iteration, it may result in the poor diversity and uniformity of Pareto solutions. The "one by one removal" method is used here to update the non-inferior solution, that is, the solution with the smallest dense distance is removed after sorting by cluster density, followed by calculating the dense distance of the remaining Pareto solutions until there are N Pareto solutions left.

Selection of Global Optimal Solution of Population

The PSO algorithm needs to track the best known positions of the particles and the population when updating the population. In a single-objective PSO process, the best known positions can be obtained directly by comparing the

size of the particle fitness function. When the PSO deals with a multi-objective optimization problem, the result of each iteration is a group of independent Pareto solutions. This requires determining the global optimal solution from the Pareto solution set. To ensure the diversity and uniformity of Pareto solutions, the potential global optimal solution of the population could be randomly selected from the first 20% of Pareto solutions with larger dense distances based on their dense distance ranking.

Multi-Attribute Decision Making Based on TOPSIS

The final optimization result of the improved multi-object PSO algorithm is a set of Pareto solutions from which decision-makers need to select the optimal solution according to the preference information, which is essentially a multi-attribute decision-making problem. Instead of randomly selecting the potential global optimal solution from the Pareto solution set, the paper uses the technique for order of preference by similarity to ideal solution (TOPSIS) to choose the optimal solution based on information entropy. The TOPSIS assists in decision-making by calculating distances between alternative schemes and positive or negative ideal schemes and then determining the alternative that minimizes the distances from positive ideal schemes and maximizes the distances from negative ideal schemes. Given N schemes x_1, x_2, ... , x_N composed of non-inferior solutions in the Pareto solution set, each scheme x_i ($i = 1$, ... , N) comprises n attributes (i.e., the number of objective functions) with the mth attribute value denoted by $f_m(x_i)$ ($m = 1$, ... , n). Considering the dimensional errors among the attributes, the attributes under different types require to be transformed into dimensionless attributes by normalizing their $f_m(x_i)$ based on the attribute values of the same type in all schemes. This is formulated by **Eq. 32** where the normalized attribute value is denoted by $f_m^{'}(x_i)$.

$$f_m^{'}(x_i) = \frac{f_m(x_i)}{\sqrt{\sum_{i=1}^{N} f_m^2(x_i)}} \quad m = 1, 2, \cdots, n \tag{32}$$

Then the relative distance $d(x_i)$ of the scheme x_i is calculated based on the differences in normalized attribute values between x_i and positive and negative ideal schemes by **Eqs 33–35**:

$$d(x_i) = \frac{d_+(x_i)}{d_+(x_i) + d_-(x_i)} \tag{33}$$

$$d_+(x_i) = \sqrt{\sum_{m=1}^{n} \left[\lambda_m f_m^{'}(x_i) - \lambda_m f_{m+}^{'}\right]^2} \tag{34}$$

$$d_-(x_i) = \sqrt{\sum_{m=1}^{n} \left[\lambda_m f_m^{'}(x_i) - \lambda_m f_{m-}^{'}\right]^2} \tag{35}$$

The TOPSIS method will assign weight to each objective in the calculation process where the selection of the weight has higher requirements on the experience and knowledge of decision-makers. To mitigate the influence of the decision-maker on the final decisions, the paper uses the information entropy method to determine the weight assigned to each objective by judging the value deviation of each objective within a Pareto solution set. Given a particular objective having a smaller value deviation within the solution set, the influence of the objective value on the final decision-making or the associated weight should be small.

Optimization of ESS Locations and Operation

When the number and capacity of distributed ESS are determined, their locations and import/export require to be optimized. Given $N^{'}$ ESS being placed on the network, the vector of optimization variables x comprises the ESS access nodes denoted by x_1,..., $x_{N^{'}}$ and their active power at each time step ($t = 1$, ... , T) denoted by $y_{(t-1)\times N^{'}+1}$,..., $y_{t\cdot N^{'}}$:

$$x = \left[x_1, x_2, \cdots, x_{N^{'}}, y_1, \cdots, y_{N^{'}}, \cdots, y_{t\cdot N^{'}+i}, \cdots, y_{T\cdot N^{'}}\right] \tag{36}$$

To restore the SOC of each ESS back to its initial SOC level at the end of the time horizon, the active power of a particular ESS at the final time step T is corrected by:

$$y_{(T-1)\cdot N^{'}+i} = -\sum_{t-1}^{T-1} y_{(t-1)\cdot N^{'}+i} \tag{37}$$

A flowchart describing the process of the improved multi-objective PSO algorithm for the combined estimation of location and capacity of distributed ESS is shown in **Supplementary Figure A6**.

CASE STUDY

Simulation Case

The proposed improved multi-objective PSO based approach to optimizing ESS capacity and locations within a 12.66 kV distribution network is tested here based on the IEEE 24-bus system (see **Supplementary Figure A7**) with associated network parameters tabulated in **Table 1**. For detailed network topology parameters (e.g., line parameters and node load), the reader is referred to the standard calculation example in (Chang, 2012).

TABLE 1 | IEEE 24-bus based network parameters.

Parameter	Bus/Branch	Maximum total network load	Maximum capacity of generating units	Renewable energy penetration rate (%)
Value	24/38	2.85 GW, 0.58 GVar	3.075 GVA	40

Supplementary Figure A8 shows a typical daily curve of 15 min average total network demand with a maximum of 2.85 GW and 0.58 GVar. In addition to the local generating units supplying a maximum of 3.075 GW, the distribution network can import from the 200 MW UHV AC supply and the 400 MW UHV DC supply through nodes 17 and 6, respectively, as shown in **Supplementary Figure A7**. The daily profile of the normalized 15 min average power transfer across UHV transmission lines is shown in **Figure 3** where UHV transmission line failures are presumed to occur around the middle of the day. The total energy capacity of ESS connected to each node is limited to 100 MWh, permitting an installation of up to five 20 MWh ESS. The proposed optimization method requires around 9.26 s to compute the optimal combination of ESS capacity and locations within the network.

Impacts of PV Access on Voltage Stability

To investigate the impacts of distributed generation on the voltage profile of the distribution network, four particular nodes, e.g., nodes 2, 9, 17, and 21 in **Supplementary Figure A7**, are selected here as the potential access points for distributed PV plants with a total installed capacity of 1 GW. Presuming that distributed PV plants export at their rated power, **Figures 4**, **5** compare the node voltage levels simulated without or with the PV access at two particular network load levels (i.e., the demand trough at time step 22 and the peak demand at time step 72), respectively, showing that the integration of PV plants aggravates the voltage deviation between nodes and reduces the overall voltage stability of the distribution network.

The impact of the PV access on voltage stability is additionally investigated based on the modelled PV power outputs on a typical sunny day. The resulting daily voltage profiles of the four PV access nodes are shown in **Figure 6**. Without the use of any voltage control method, the voltage levels at the PV access nodes 17 and 21 would experience larger variations within a day and exceed the upper and lower limits at some particular time steps. The significant exceedance over the upper voltage limit at the relevant PV access points is mainly caused by the reverse power flow of the PV system which results in a voltage rise. Therefore, when the voltage regulation is not performed, the DG access point is prone to the voltage limit violation which threatens the safe and reliable system operation.

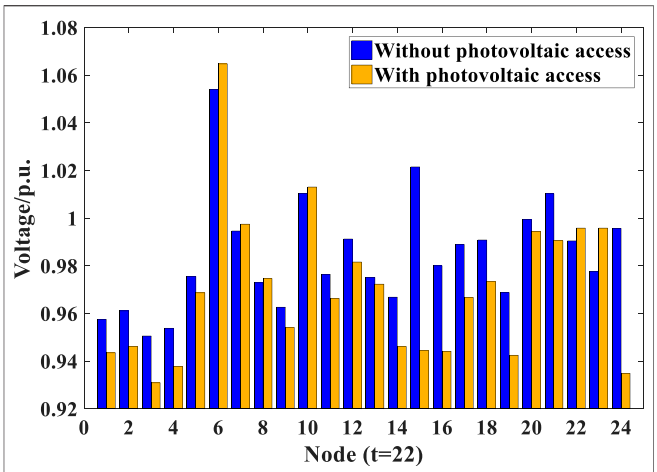


FIGURE 4 | Node voltage levels (p.u.) simulated without or with the PV access at off-peak time (t = 22).

Simulation Scenarios Design

To verify the effectiveness of the proposed ESS optimization method, the following four scenarios are simulated to compare the voltage profile of the network without the ESS support or under the voltage regulation performed by distributed ESS

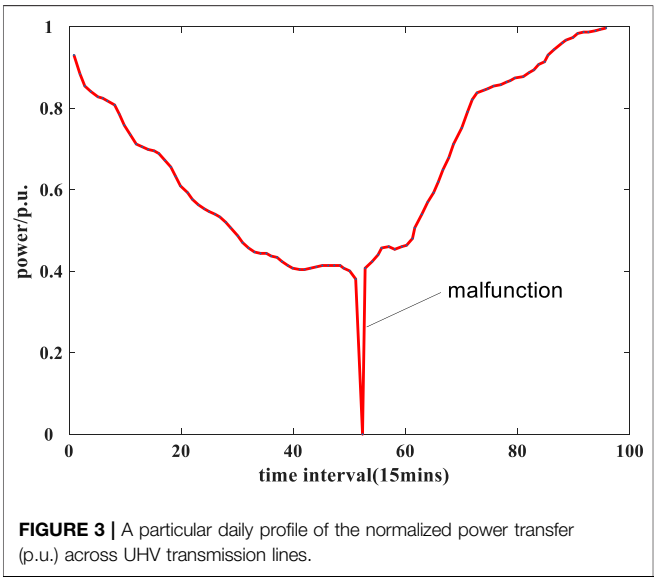


FIGURE 3 | A particular daily profile of the normalized power transfer (p.u.) across UHV transmission lines.

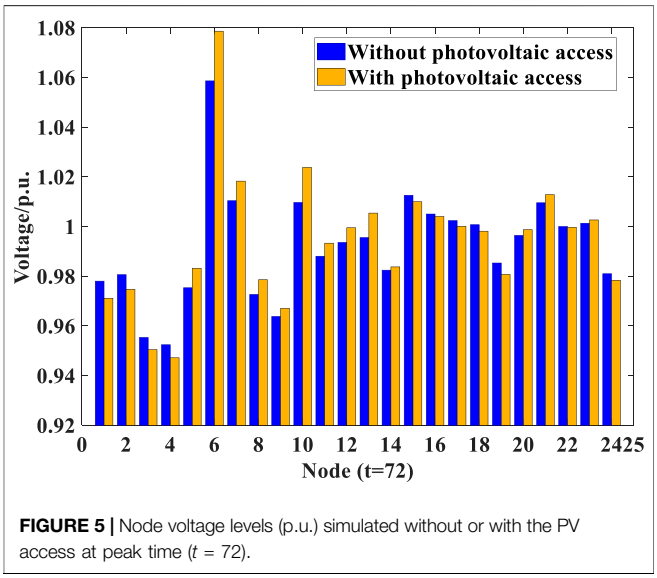


FIGURE 5 | Node voltage levels (p.u.) simulated without or with the PV access at peak time (t = 72).

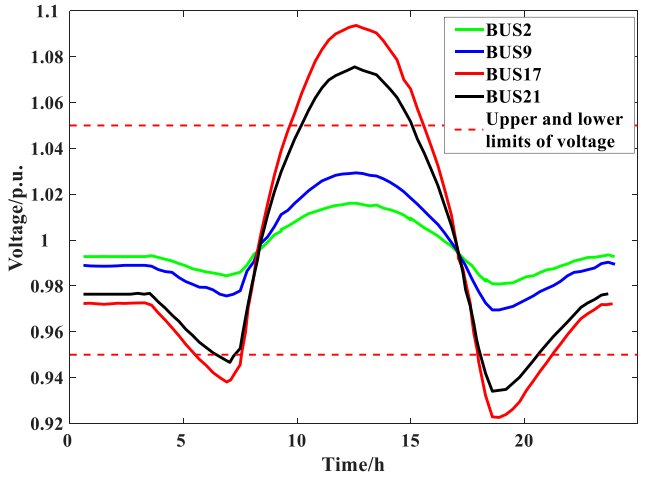


FIGURE 6 | Daily voltage profiles of the four PV access nodes on a typical sunny day.

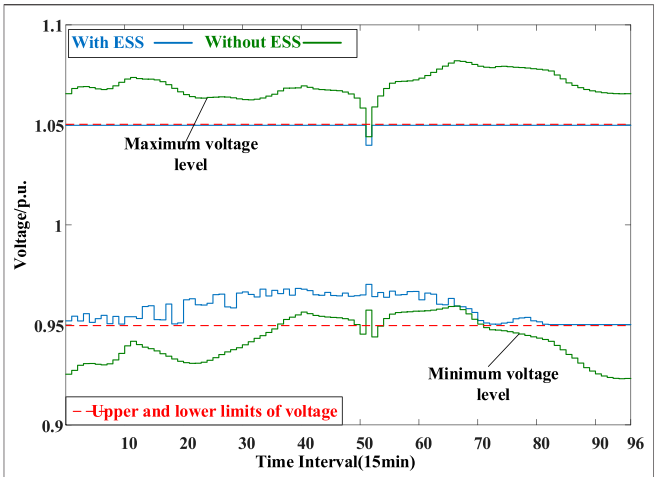


FIGURE 7 | The maximum and minimum voltage levels (p.u.) across the network without or with the ESS deployment in a day.

which are randomly configurated or optimised by the proposed method:

Scenario 1: The distribution network with the PV integration but without the ESS deployment;
Scenario 2: The ESS capacity is optimized but with randomly selected locations within the network;
Scenario 3: The ESS capacity and locations within the network are optimized by the proposed method;
Scenario 4: The operation of the optimized ESS given short-circuit faults on UHV transmission lines.

Simulation Results

Impacts of ESS Deployment on Voltage Stability

The node voltage levels simulated in Scenarios 1 and 3 are compared here to investigate the impacts of the ESS deployment on voltage stability. **Table 2** lists the optimal ESS locations and sizes determined by the improved multi-objective PSO algorithm. The distributed ESS with a total capacity of around 95 MW, 220 MWh are located at three access nodes 4, 6, and 10. The ESS connected to node 4 has a capacity of 8.62 MW, 20 MWh while the ESS connected to nodes 6 and 19 both have a total capacity of 43.08 MW, 100 MWh. As shown in **Supplementary Figure A7**, these ESS access nodes are concentrated around the distribution area of the UHV DC supply point.

TABLE 2 | Optimization results of ESS capacity and locations in Scenario 3.

Serial number	Access node	Power capacity/MW	Energy capacity/MWh
ESS1	4	8.62	20
ESS2	6	43.08	100
ESS3	10	43.08	100

The maximum and minimum node voltage levels across the network without or with the ESS placement at each time step are plotted in **Figure 7**. Without the ESS support in Scenario 1, the upper and lower node voltage limits are violated for the majority of the time. When the optimized ESS is placed on the network in Scenario 3, the difference between the maximum and minimum node voltage levels at each time step as well as the voltage fluctuations within the day are reduced. In addition, the node voltage levels across the network are well maintained within the upper and lower limits. **Figure 8** compares histograms of node voltage levels across the network and across 96 15 min time steps between Scenarios 1 and 3. The node voltage levels in Scenario 3 are shown to be more concentrated around the nominal value than those in Scenario 1. The overall characteristics of node voltage levels in the two Scenarios are summarized in **Table 3**. The deployment of the optimised ESS is shown to not only reduce

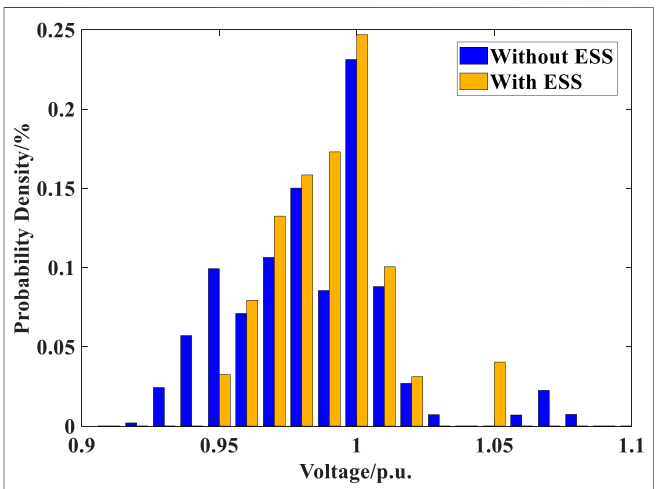


FIGURE 8 | Histograms of node voltage (p.u.) across the network without or with the ESS deployment in a day.

TABLE 3 | Overall characteristics of node voltage in Scenarios 1 and 3.

Items	Voltage deviation $\sum_{i=1}^{N}\sum_{t=1}^{T}\lvert U_{i,t}-U_0\rvert$	Maximum voltage/p.u.	Minimum voltage/p.u.	Limit violation /Total samples
Scenario 1	58.5084	1.0821	0.9231	401/2304
Scenario 3	40.7696	1.0500	0.9500	0/2304
Change	−30.32%	−0.0321	0.0269	−17.40%

the overall voltage deviation from the nominal level by about 30.3% but also help to keep the node voltage within the limits, avoiding the voltage limit violation in around 17.4% of the voltage samples. The improvement of voltage stability in Scenario 3 illustrates the effective voltage regulation performed by the optimized ESS.

The voltage profiles of the network simulated in Scenarios 1 and 3 during off-peak and peak periods (e.g., time steps 22 and 72) are compared in **Figures 9**, **10**, respectively. Though conventional generators can support voltage regulation, their regulation capability is constrained by the ramp rate and export limits. In addition, the conventional generators distributed at some particular nodes cannot effectively deal with the fluctuations of the network load and the UHV power supply. When the network load is at a low level at time step 22, nodes 1–4, 6, 14–16, 19, and 24 suffer from large voltage deviations (see **Figure 9**) which cannot meet the network voltage requirements. Though the lower voltage limit violation is alleviated at peak time, the rise above the upper voltage limit at node 6 increases (see **Figure 10**). The operation of the optimized ESS can provide active support to the network and alleviate the power fluctuations at UHV transmission access nodes, improving the voltage stability of the network as shown in **Figures 9**, **10**.

Impacts of ESS Optimization on Voltage Stability

The node voltage levels simulated in Scenarios 2 and 3 are compared here to analyze the advantages of the proposed ESS optimization method in voltage regulation over the unoptimized ESS placement. Though considering the same number of ESS access nodes and the same ESS capacity as those obtained by the ESS optimization method in **Table 2**, the unoptimized method randomly determines the ESS access nodes within the network, e.g., nodes 3, 20, and 21 in **Supplementary Figure A7**, rather than the optimized access nodes 4, 6, and 10. **Table 4** summarizes the overall characteristics of node voltage levels in Scenarios 2 and 3. Though the ESS placed at randomly selected access nodes in Scenario 2 slightly reduce the overall voltage deviation compared to Scenario 1, the maximum/minimum node voltage level still exceeds the upper/lower limit. (In total, 357 out of 2,304 voltage samples violate the voltage limits in Scenario 2). When the ESS are connected to the optimal access nodes in Scenario 3, the overall voltage deviation is further reduced without any voltage limit violation, showing more effective voltage regulation than the random ESS placement.

The daily variations of the maximum and minimum node voltage levels across the network with the optimized or random placement of ESS are plotted in **Figure 11** where the random placement results in the voltage limit violation for most of the time while the optimized placement maintains the node voltage within the upper and lower limits. **Figure 12** compares histograms of the node voltage simulated based on the unoptimized or optimized ESS placement. The node voltage based on the unoptimized ESS placement shows a more spread distribution than

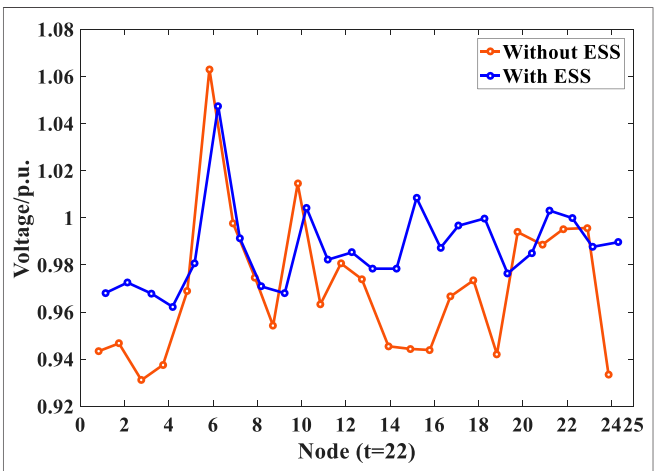


FIGURE 9 | The voltage profile (p.u.) of the network without or with the ESS deployment at off-peak time (time step 22).

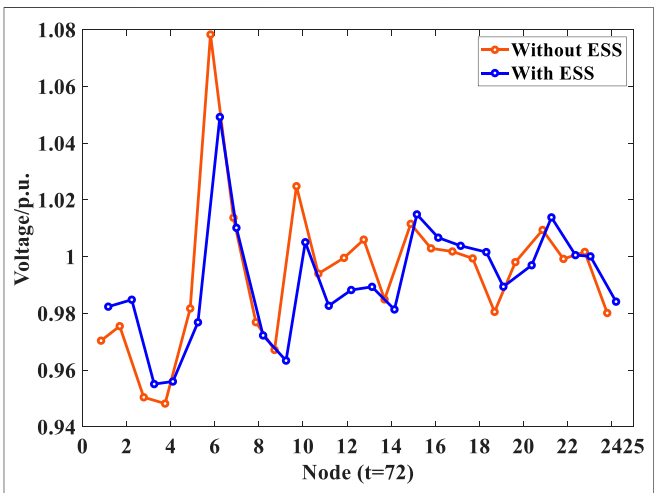


FIGURE 10 | The voltage profile (p.u.) of the network without or with the ESS deployment at peak time (time step 72).

TABLE 4 | Overall characteristics of node voltage in Scenarios 2 and 3.

Item	Voltage deviation $\sum_{i=1}^{N}\sum_{t=1}^{T}\lvert U_{i,t}-U_0\rvert$	Maximum voltage/p.u	Minimum voltage/p.u.	Limit violation/Total samples
Scenario 2	55.7215	1.0818	0.9293	357/2,304
Scenario 3	40.7696	1.0500	0.9500	0/2,304

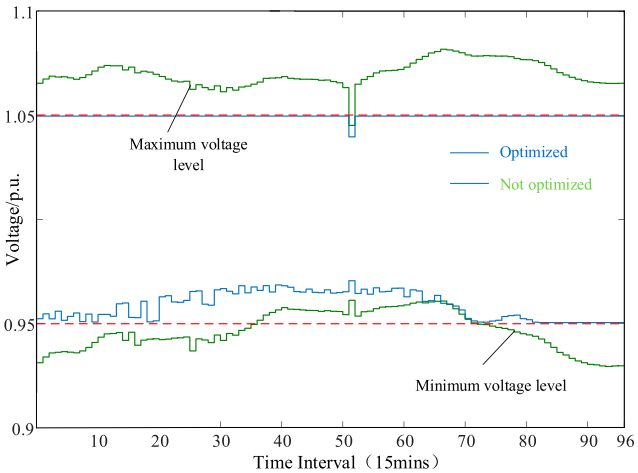


FIGURE 11 | The maximum and minimum voltage levels (p.u.) across the network with the random or optimized ESS placement in a particular day.

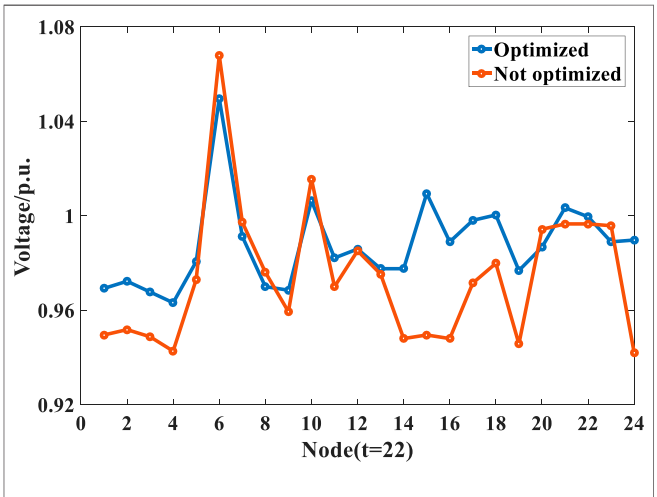


FIGURE 13 | The voltage profile (p.u.) of the network with the random or optimized ESS placement at off-peak time (time step 22).

those based on the optimized placement. Therefore, the voltage stability of the network cannot be effectively improved when the ESS are randomly connected to the network without the location optimization.

The voltage profiles of the network simulated in Scenarios 2 and 3 at two particular time steps 22 and 72 with different load levels are compared in **Figures 13**, **14**, respectively. The ESS connected to different nodes have various regulation effects on the power flow within the network. When the site selection is not optimized, though imports/exports of access nodes can be adjusted by the ESS, the regulation effect of the ESS is not consistent with the direction of improving the voltage

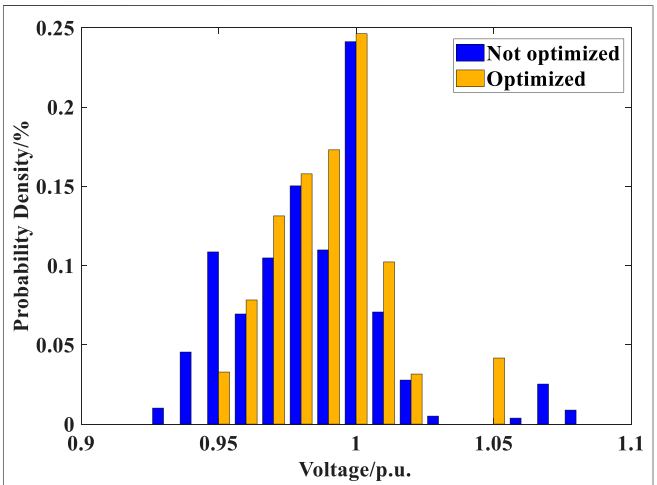


FIGURE 12 | Histograms of node voltage (p.u.) across the network with the random or optimized ESS placement in a particular day.

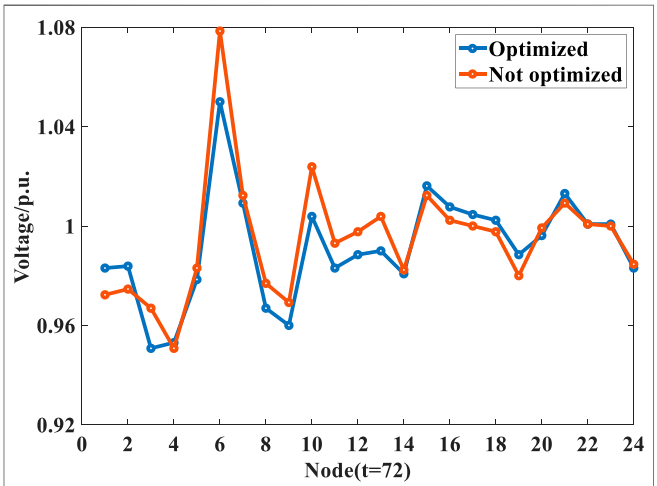


FIGURE 14 | The voltage profile (p.u.) of the network with the random or optimized ESS placement at peak time (time step 72).

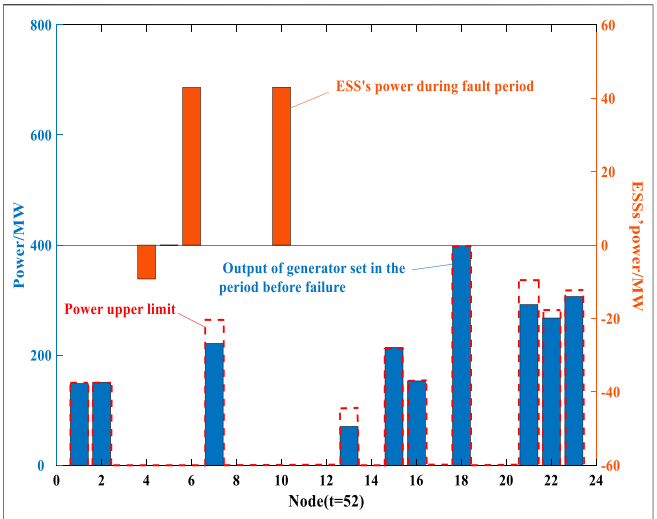


FIGURE 15 | Power outputs (MW) of generating units before UHV transmission failures and discharge (+ve) and charge (−ve) rates (MW) of ESS during fault occurrences.

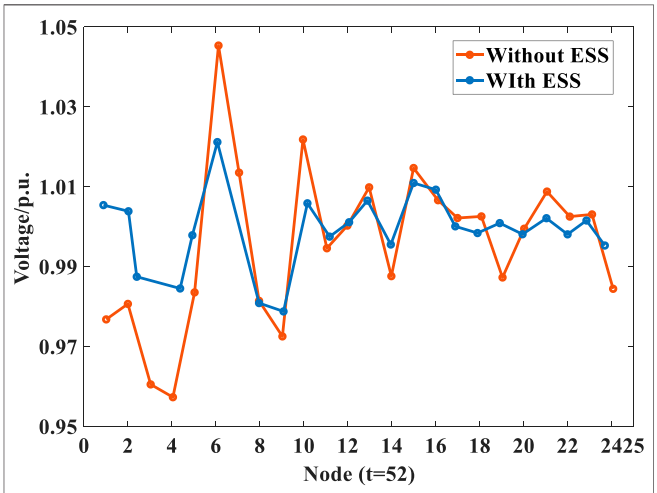


FIGURE 16 | The voltage profile (p.u.) of the network without or with the ESS operation during UHV transmission failures (time step 52).

regulation of the network due to the influence of their locations within the network. Therefore, the proposed ESS optimization method with its advantages of access location optimization plays a significant role in voltage regulation.

Emergency Power Supply by ESS Under Fault Conditions

To examine the ESS role in voltage regulation and power support under fault conditions, the ESS operation and the resulting voltage profile of the network are simulated when short-circuit faults are assumed to occur on the UHV transmission at time step 52 in Scenario 4. The fault occurrence forces the total UHV transmission power which would otherwise supply 240.6 MW to zero (see **Figure 3**). The power outputs of generating units on the network before UHV transmission failures and the ESS operation during fault occurrences are shown in **Figure 15**. To maintain the normal system operation before fault occurrences, the generating unit export is kept at such a high level that the generating unit sets at nodes 1, 2, 15, 16, and 18 reach their respective maximum allowable export limits and have no spare regulation capacity to deal with a power shortage. When the UHV transmission failures happen, the significant drops in the power injections of nodes 6 and 17 have certain impacts on power transfers across the lines connected to them. Since the generating unit sets with certain headroom in the regulation capacity are located at specific nodes, they cannot completely support the voltage regulation while the voltage profile of the network without the ESS operation would suffer from large variations as shown in **Figure 16**. To assist in the emergency power supply and the stable network operation during the fault period, **Figure 15** shows that the ESS at nodes 6 and 10 export at their rated power to mitigate the impacts of UHV transmission failures on the power flow distribution within the network while the ESS at node 4 imports at its rated power to reduce the power fluctuation of the associated line, both helping to improve the voltage stability of the network (see **Figure 16**).

CONCLUSION AND FUTURE WORK

This paper has proposed an improved multi-objective particle swarm optimization (PSO) based method to estimate the best combination of sizes and locations of distributed energy storage systems (ESS) that effectively support the voltage regulation of a distribution network with PV access. The improved multi-objective PSO algorithm produces an optimal Pareto solution set by minimising a weighted sum of the overall deviation between the voltage profile of the network and the nominal level across the time horizon and the energy capacity of ESS reflecting their investment. The adaptive adjustment of inertia weight and the cross-mutation operation have been introduced into the improved PSO algorithm to avoid falling into local optimal solutions while the dynamic dense distance of the particle has been employed to update the Pareto solution set in each iteration. Compared with the traditional PSO algorithm, the improvements have shown superiority in computation speed and permitted considering the voltage offset. Then the information entropy based technique for order of preference by similarity to ideal solution is used to determine the optimal access scheme of distributed ESS from the Pareto solution set, reducing the impact of decision-makers' preference.

The effectiveness of the proposed ESS optimization method has been assessed by comparing voltage profiles of the modified IEEE 24-bus network without the ESS deployment or with the random or optimized ESS placement. Though the randomly placed ESS can slightly reduce the overall node voltage deviation from the nominal level, the upper and lower limits on the node voltage are still violated. When the ESS are connected

to the optimized nodes around the access point of the ultra-high-voltage DC power supply, the overall node voltage deviation is largely reduced while the voltage limit violation is completely avoided in this work. The optimized deployment of distributed ESS has effectively coped with nodal power fluctuations caused by variations of local demand and PV generation and also supported the voltage regulation of the network to the maximum extent and the emergency power supply during ultra-high-voltage transmission failures.

Building on the present work, the proposed ESS optimization method will be applied to an expanded power system with more nodes and lines, finding the most suitable capacity and installation locations of distributed ESS for the system. Furthermore, the proposed method will be further examined based on the practical network operation of cities or provinces with high renewable energy penetration rates (e.g., "Three North regions" in China). Moreover, the follow-on research will address the location and capacity determination for grid-side ESS with high proportion of wind power connected to the distribution network.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

QL performed the experiment and simulation; FZ and FG performed analysis and prepared the manuscript; FF revised the manuscript; and ZH contributed to the conception and supervision of the work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fenrg.2021.641518/full#supplementary-material.

REFERENCES

Alam, M. J. E., Muttaqi, K. M., and Sutanto, D. (2012). "Distributed energy storage for mitigation of voltage-rise impact caused by rooftop solar PV," in Power and energy society general meeting, San Diego, CA, July 22–26, 2012 (San Diego, CA: IEEE), 1–8.

Atwa, Y. M., El-Saadany, E. F., Salama, M. M. A., and Seethapathy, R. (2010). Optimal renewable resources mix for distribution system energy loss minimization. *IEEE Trans. Power Syst.* 25 (1), 360–370. doi:10.1109/tpwrs.2009.2030276

Azzouz, M. A., Shaaban, M. F., and El-Saadany, E. F. (2015). Real-time optimal voltage regulation for distribution networks incorporating high penetration of PEVs. *IEEE Trans. Power Syst.* 30 (6), 3234–3245. doi:10.1109/tpwrs.2014.2385834

Bachurin, P. A., Zyryanov, V. M., Kiryanova, N. G., Kuchak, S., Metalnikov, D. G., Nesterenko, G. B., et al. (2018). "Mathematical model of the ESS in the power system," in 2018 XIV international scientific-technical conference on actual problems of electronics instrument engineering, Novosibirsk, Russia, October 14, 2018 (Novosibirsk, Russia: APEIE), 41–47.

Bakir, H., and Kulaksiz, A. A. (2019). "Modelling and voltage control of the solar-wind hybrid micro-grid with optimized STATCOM," in 2019 23rd international conference electronics, Palanga, Lithuania, June 17–19, 2019 (Palanga, Lithuania: IEEE), 1–6.

Brenna, M., Berardinis, E. D., Carpini, L. D., Foiadelli, F., Pietro, P., Paulon, P., et al. (2013). Automatic distributed voltage control algorithm in smart grids applications. *IEEE Trans. Smart Grid* 4 (2), 877–885. doi:10.1109/TSG.2012.2206412

Chang, Y. C. (2012). Multi-objective optimal SVC installation for power system loading margin improvement. *IEEE Trans. Power Syst.* 27 (2), 984–992. doi:10.1109/tpwrs.2011.2176517

Cheng, P., Zhou, Y., Song, Z., and Qu, Y. (2016). "Modeling and SOC estimation of LiFePO4 battery," in International conference on robotics and biomimetics (ROBIO), Qingdao, China, December 3–7, 2016 (Qingdao, China: IEEE), 2140–2144.

Collins, L., and Ward, J. K. (2015). Real and reactive power control of distributed PV inverters for overvoltage prevention and increased renewable generation hosting capacity. *Renew. Energ.* 81, 464–471. doi:10.1016/j.renene.2015.03.012

Daratha, N., Das, B., and Sharma, J. (2014). Coordination between OLTC and SVC for voltage regulation in unbalanced distribution system distributed generation. *IEEE Trans. Power Syst.* 29 (1), 289–299. doi:10.1109/tpwrs.2013.2280022

Dong, F., Chowdhury, B. H., Crow, M. L., and Acar, L. (2005). Improving voltage stability by reactive power reserve management. *IEEE Trans. Power Syst.* 20 (1), 338–345. doi:10.1109/tpwrs.2004.841241

El Moursi, M. S., Zeineldin, H. H., Kirtley, J. L., and Alobeidli, K. (2014). A dynamic master/slave reactive power-management scheme for smart grids with distributed generation. *IEEE Trans. Power Deliv.* 29 (3), 1157–1167. doi:10.1109/tpwrd.2013.2294793

Esslinger, P., and Witzmann, R. (2012). "Regulated distribution transformers in low-voltage networks with a high degree of distributed generation," in PES innovative smart grid technologies Europe (ISGT Europe), Berlin, Germany, October 14–17, 2012 (Berlin, Germany: IEEE), 1–7.

Fan, F., Zorzi, G., Campos-Gaona, D., Burt, G., Anaya-Lara, O., Nwobu, J., et al. (2021). Sizing and coordination strategies of battery energy storage system co-located with wind farm: the UK perspective. *Energies* 14 (5), 1439. doi:10.3390/en14051439

Hasheminamin, M., Agelidis, V. G., Ahmadi, A., Siano, P., and Teodorescu, R. (2018). Single-point reactive power control method on voltage rise mitigation in residential networks with high PV penetration. *Renew. Energy* 119, 504–512. doi:10.1016/j.renene.2017.12.029

Jannesar, M. R., Sedighi, A., Savaghebi, M., and Guerrero, J. M. (2018). Optimal placement, sizing, and daily charge/discharge of battery energy storage in low voltage distribution network with high photovoltaic penetration. *Appl. Energy* 226, 957–966. doi:10.1016/j.apenergy.2018.06.036

Jashfar, S., and Esmaeili, S. (2014). Volt/var/THD control in distribution networks considering reactive power capability of solar energy conversion. *Int. J. Electr. Power Energy Syst.* 60, 221–233. doi:10.1016/j.ijepes.2014.02.038

Jung, J., Onen, A., Arghandeh, R., and Broadwater, R. P. (2014). Coordinated control of automated devices and photovoltaic generators for voltage rise mitigation in power distribution circuits. *Renew. Energy* 66, 532–540. doi:10.1016/j.renene.2013.12.039

Kennedy, J., and Eberhart, R. (1995). "Particle swarm optimization," in Proceeding international conference on neural networks, Perth, Australia, Novovember–December 1–27, 1995 (Perth, Australia: IEEE), 1942–1948.

Kulmala, A., Repo, S., and Jarventausta, P. (2014). Coordinated voltage control in distribution networks including several distributed energy resources. *IEEE Trans. Smart Grid* 5 (4), 2010–2020. doi:10.1109/tsg.2014.2297971

Lin, C.-H., Hsieh, W.-L., Chen, C.-S., Hsu, C.-T., and Ku, T.-T. (2012). Optimization of photovoltaic penetration in distribution systems considering annual duration curve of solar irradiation. *IEEE Trans. Power Syst.* 27 (2), 1090–1097. doi:10.1109/tpwrs.2011.2178619

Liu, X., Aichhorn, A., Liu, L., and Li, H. (2012). Coordinated control of distributed energy storage system with tap changer transformers for voltage rise mitigation

under high photovoltaic penetration. *IEEE Trans. Smart Grid* 3 (2), 897–906. doi:10.1109/tsg.2011.2177501

Mehmood, K. K., Khan, S. U., Lee, S., Haider, Z. M., Rafique, M. K., and Kim, C. (2017). Optimal sizing and allocation of battery ESSs with wind and solar power DGs in a distribution network for voltage regulation considering the lifespan of batteries. *IET Renew. Power Gener.* 11 (10), 1305–1315. doi:10.1049/iet-rpg.2016.0938

Mokhtari, G., Nourbakhsh, G., and Ghosh, A. (2012). Smart coordination of energy storage units(ESUs) for voltage and loading management in distribution networks. *IEEE Trans. Power Syst.* 28 (4), 4812–4820. doi:10.1109/TPWRS.2013.2272092

Muttaqi, K. M., Le, A. D., Negnevitsky, M., and Ledwich, G. (2013). A coordinated voltage control approach for coordination of OLTC voltage regulator and DG to regulate voltage in a distribution feeder. *IEEE Trans. Ind. Appl.* 51 (2), 1073–1081. doi:10.1109/TIA.2014.2354738

Nara, K., Ishizu, S., and Mishima, Y. (2005). "Voltage control availability of distributed generators in power distribution system," in Russia power tech, St. Petersburg, Russia, June 27–30 2005 (Petersburg, Russia: IEEE), 1–6.

Rosewater, D. M., Copp, D. A., Nguyen, T. A., Byrne, R. H., and Santoso, S. (2019). Battery energy storage models for optimal control. *IEEE Access* 7, 178357–178391. doi:10.1109/access.2019.2957698

Salih, S. N., and Chen, P. (2016). On coordinated control of OLTC and reactive power compensation for voltage regulation in distribution systems with wind power. *IEEE Trans. Power Syst.* 31 (5), 4026–4035. doi:10.1109/tpwrs.2015.2501433

Senjyu, T., Hayashi, D., Urasaki, N., and Funabashi, T. (2006). Optimum configuration for renewable generating systems in residence using genetic algorithm. *IEEE Trans. Energy Convers.* 21 (2), 459–466. doi:10.1109/tec.2006.874250

Senjyu, T., Miyazato, Y., Yona, A., Urasaki, N., and Funabashi, T. (2008). Optimal distribution voltage control and coordination with distributed generation. *IEEE Trans. Power Deliv.* 23 (2), 1236–1242. doi:10.1109/tpwrd.2007.908816

Sugihara, H., Yokoyama, K., Saeki, O., Tsuji, K., and Funaki, T. (2013). Economic and efficient voltage management using customer-owned energy storage systems in a distribution network with high penetration of photovoltaic systems. *IEEE Trans. Power Syst.* 28 (1), 102–111. doi:10.1109/tpwrs.2012.2196529

Teng, J.-H., Luan, S.-W., Lee, D.-J., and Huang, Y.-Q. (2013). Optimal charging/discharging scheduling of battery storage systems for distribution systems interconnected with sizeable PV generation systems. *IEEE Trans. Power Syst.* 28 (2), 1425–1433. doi:10.1109/tpwrs.2012.2230276

Tonkoski, R., and Lopes, L. A. C. (2011). Impact of active power curtailment on overvoltage prevention and energy production of PV inverters connected to low voltage residential feeders. *Renew. Energy* 36 (12), 3566–3574. doi:10.1016/j.renene.2011.05.031

Watson, J. D., Watson, N. R., Santos-Martin, D., Wood, A. R., Lemon, S., and Miller, A. J. V. (2016). Impact of solar photovoltaics on the low-voltage distribution network in New Zealand. *IET Gener. Transm. Distrib.* 10 (1), 1–9. doi:10.1049/iet-gtd.2014.1076

Yang, L., Feng, B., Li, G., Qi, J., Zhao, D., and Mu, Y. (2018). Optimal distributed generation planning in active distribution networks considering integration of energy storage. *Appl. Energy* 210, 1073–1081. doi:10.1016/j.apenergy.2017.08.008

Yap, W. K., Havas, L., Overend, E., and Karri, V. (2014). Neural network-based active power curtailment for overvoltage prevention in low voltage feeders. *Expert Syst. Appl.* 41 (4), 1063–1070. doi:10.1016/j.eswa.2013.07.103

Yoshida, H., Kawata, K., Fukuyama, Y., Takayama, S., and Nakanishi, Y. (2000). A particle swarm optimization for reactive power and voltage control considering voltage security assessment. *IEEE Trans. Power Syst.* 15 (4), 1232–1239. doi:10.1109/59.898095

Zad, B. B., Lobry, J., Vallee, F., and Hasanvand, H. (2015). Optimal reactive power control of DGs for voltage regulation of MV distribution systems using sensitivity analysis method and PSO algorithm. *Int. J. Electr. Power Energy syst.* 68, 52–60. doi:10.1016/j.ijepes.2014.12.046

Zhao, H., Wu, Q., Guo, Q., Sun, H., Huang, S., and Xue, Y. (2016). Coordinated voltage control of a wind farm based on model predictive control. *IEEE Trans. Sustain. Energy* 7 (4), 1440–1451. doi:10.1109/tste.2016.2555398

Zhou, Q., and Bialek, J. W. (2007). Generation curtailment to manage voltage constraints in distribution networks. *IET Gener. Transm. Distrib.* 1 (3), 492–498. doi:10.1049/iet-gtd:20060246

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

frontiers in Energy Research

BRIEF RESEARCH REPORT
published: 26 April 2021
doi: 10.3389/fenrg.2021.635010

Determining Region Boundaries of Critical Commutation Failures in Multi-Infeed HVDC Systems Under Unbalanced Short Circuit Faults

Guoqing Li, Song Zhang, Shuguang Li, Xianchao Liu and Xintong Liu*

Department of Electrical Engineering, Northeast Electric Power University, Jilin, China

This paper presents a region boundaries determination method for critical commutation failures (CF) in multi-infeed high-voltage direct-current (HVDC) systems under unbalanced short circuit faults. By using the nodal impedance matrix, a calculation approach for the converter extinction angles is deduced. First, the extinction angles under unbalanced short circuit faults are calculated. If the extinction angle of a bus is less than the minimum extinction angle, the bus is a failed bus set. By this means, region boundaries of critical commutation failures are demarcated by examining each bus in the AC system. Finally, the effectiveness of the proposed approach was verified by using a modified IEEE 39-bus system.

Keywords: commutation failure, muti-infeed HVDC, extinction angle, three-phase short circuit fault, unbalanced short circuit fault

OPEN ACCESS

Edited by:
Chao Long,
Cranfield University,
United Kingdom

Reviewed by:
Yang Li,
Northeast Electric Power University, China
Chen Liang,
Nanjing University of Information Science and Technology, China
Shaoyan Li,
North China Electric Power University, China

****Correspondence:***
Song Zhang
zhangsongneepu@aliyun.com

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

Received: *29 November 2020*
Accepted: *09 February 2021*
Published: *26 April 2021*

Citation:
Li G, Zhang S, Li S, Liu X and Liu X (2021) Determining Region Boundaries of Critical Commutation Failures in Multi-Infeed HVDC Systems Under Unbalanced Short Circuit Faults. Front. Energy Res. 9:635010. doi: 10.3389/fenrg.2021.635010

INTRODUCTION

Multi-infeed high-voltage direct-current (HVDC) is widely used in power grids for its advantages in long-distance power transmission (Yao et al., 2020; Mirsaeidi et al., 2018). Commutation failure (CF) at the inverter side is a very frequent dynamic event in multi-infeed HVDC systems (Wang Q. et al., 2019; Zheng et al., 2019). An AC system failure may cause commutation failure (CF) at the convertor station near a fault location (Wang F. et al., 2017; Xue et al., 2018). If the commutation function is not recovered quickly, it may lead to a bulk drop of HVDC transmission power and the continuous CFs will occur at the local and adjacent inverter station, which will lead to a wide-range power outage. However, not all AC system faults can trigger the CF of a DC system. In view of the rapid construction and deployment of DC transmission systems, it is important to establish an accurate and fast solution to detect commutation faults and reduce or avoid simultaneous commutation failures in multiple DC transmission systems.

A great amount of research in the field of CF analysis has proposed that a decrease in the extinction angle of the inverter and the drop of commutation voltage are the primary causes of the CF (Yao et al., 2020; Xiao et al., 2016; Xue et al., 2016). In reference (Wang J. et al., 2019), based on the expected value of reactive power consumption of inverter stations, a method for calculating DC command values by real-time measurement of AC bus voltage is proposed. In reference (Guo et al., 2015), a method of setting different current command values for different fault types of AC system is proposed. Similarly, in order to reduce reactive power consumption when the system recovers from CF (Wei et al., 2014), HVDC recovery is promoted by reducing the current command value. Reference (Li et al., 2017) proposes a method of demarcating critical failure impedance boundary of

multi-infeed HVDC systems based on the minimum extinction angle, but the types of short circuit failures it considered were not comprehensive enough.

In this paper, the method of demarcating critical failure impedance boundary in reference (Li et al., 2017) is further supplemented. The Extinction Angle of a converter station is calculated when any bus has a two-phase fault and line fault in an AC system. When an unbalanced short circuit fault occurs in an AC system, the critical impedance boundary of a multi-feed DC system can be identified from the network topology by taking the critical extinction Angle as the criterion. An unbalanced short-circuit fault of a node lying within the critical fault impedance boundary is more likely to cause CF in a DC system, whereas an unbalanced fault of a node outside the boundary will not cause CF. The method can quickly identify the critical fault impedance boundary and graphically identify where the AC system failure may cause CF.

UNBALANCED SHORT-CIRCUIT FAULT COMMUTATION FAILURE CRITERION

When unbalanced faults occur in the AC system, the voltage drops and phase shift of CB are related to the electrical distance of the fault location. To verify the CF criterion of an unbalanced fault, the vector of each phase voltage should be calculated, respectively. Then the line-to-line voltage which corresponds to different commutation processes is calculated. Unbalanced short-circuit faults include single line-to-ground faults, double line-to-ground faults, line-to-line faults, etc. In this section, the CF criterion under different kinds of faults conditions will be introduced in detail.

When the system is symmetric, the inverter extinction angle can be expressed as (Wang Z. et al., 2017):

$$\gamma_i = \arccos\left(\frac{\sqrt{2}k_i I_d X_{Li}}{U_L} + \cos\beta\right) \tag{1}$$

where I_d is the DC current, β is the advance trigger angle, k is the transformer tap ratio, and γ_i is the extinction angle of the converter i.

When an unbalanced short-circuit fault happens at bus j, the voltages of CB i can be formulated as:

$$\begin{cases} \dot{U}_{i(1)} = \dot{U}_{i0} - \dot{I}_{j(1)} Z_{ij(1)} \\ \dot{U}_{i(2)} = -\dot{I}_{j(2)} Z_{ij(2)} \\ \dot{U}_{i(0)} = -\dot{I}_{j(0)} Z_{ij(0)} \end{cases} \tag{2}$$

where U_{i0} *is* the initial voltages of CB i and $U_{i(1)}$, $U_{i(2)}$, and $U_{i(0)}$ are the positive-sequence, negative-sequence, and zero-sequence voltage components respectively, $I_{j(1)}$, $I_{j(2)}$, and $I_{j(0)}$ are the three sequences short-circuit currents and $Z_{ij(1)}$, $Z_{ij(2)}$, and $Z_{ij(0)}$ are the three sequence mutual impedances of converter CB i and bus j, respectively.

According to the symmetrical component method, the three-phase voltage of CB i can be derived as:

$$\begin{cases} \dot{U}_{ia} = \dot{U}_{i(1)} + \dot{U}_{i(2)} + \dot{U}_{i(0)} \\ \dot{U}_{ib} = \alpha^2 \dot{U}_{i(1)} + \alpha \dot{U}_{i(2)} + \dot{U}_{i(0)} \\ \dot{U}_{ic} = \alpha \dot{U}_{i(1)} + \alpha^2 \dot{U}_{i(2)} + \dot{U}_{i(0)} \end{cases} \tag{3}$$

Single Line-To-Ground Fault

When a single line-to-ground fault occurs on phase A of bus j, the fault conditions can be written as:

$$\dot{I}_{j(1)} = \dot{I}_{j(2)} = \dot{I}_{j(0)} = \frac{U_{j0}}{Z_{jj(1)} + Z_{jj(2)} + Z_{jj(0)}} \tag{4}$$

where $Z_{jj(1)}$, $Z_{jj(2)}$, and $Z_{jj(0)}$ are the impedances of the three sequences of bus j respectively.

Substitute **(4)** into **(2)** and **(3)**, the voltage variations of phase A at CB i is:

$$\Delta U_{ia} = \left| \frac{Z_{ij(1)} + Z_{ij(2)} + Z_{ij(0)}}{Z_{jj(1)} + Z_{jj(2)} + Z_{jj(0)}} \cdot \dot{U}_{j0} \right| \tag{5}$$

As it is shown in **Figure 1A**, after a single line-to-ground fault happens, the phase voltage of phase B and phase C remains unchanged. According to the triangle relation, the line-to-line voltage between phase A and B, and phase A and C is:

$$U'_{ab} = U'_{ca} = \sqrt{3 - 3\Delta U_a + \Delta U_a^2} \tag{6}$$

Therefore, the zero-crossing phase shift of the commutating voltage can be formulated as:

$$\phi = \arctan\frac{\Delta U_a}{\sqrt{3}(2 - \Delta U_a)} \tag{7}$$

At the moment of the fault, the converter transformer ratio remains unchanged, thus the percentage of the commutation voltage drop and that of the voltage drop at the CB are the same. Therefore, the extinction angle of each converter valve at the moment of phase A to ground short-circuit is:

$$\begin{cases} \gamma_{iV1,V4} = \arccos\left(\dfrac{\sqrt{2}kI_d X_{Li}}{\sqrt{3 - 3\Delta U_a + \Delta U_a^2} U_L \Big/ \sqrt{3}} + \cos\beta\right) - \arctan\dfrac{\Delta U_a}{\sqrt{3}(2 - \Delta U_a)} \\ \gamma_{iV2,V5} = \arccos\left(\dfrac{\sqrt{2}kI_d X_{Li}}{\sqrt{3 - 3\Delta U_a + \Delta U_a^2} U_L \Big/ \sqrt{3}} + \cos\beta\right) + \arctan\dfrac{\Delta U_a}{\sqrt{3}(2 - \Delta U_a)} \\ \gamma_{iV3,V6} = \arccos\left(\dfrac{\sqrt{2}kI_d X_{Li}}{U_L} + \cos\beta\right) \end{cases} \tag{8}$$

According to **(8)**, when a single-phase to ground short-circuit fault happens, the corresponding converter extinction angle γ_i reaches the minimum value. When the corresponding converter extinction angle $\gamma_i \le \gamma_{\min}$, CF will happen at the converter station of DC i.

Double Line-To-Ground Fault and Line-To-Line Fault

The fault conditions for a double line-to-ground fault occurring on phase A and B of Bus j can be written as:

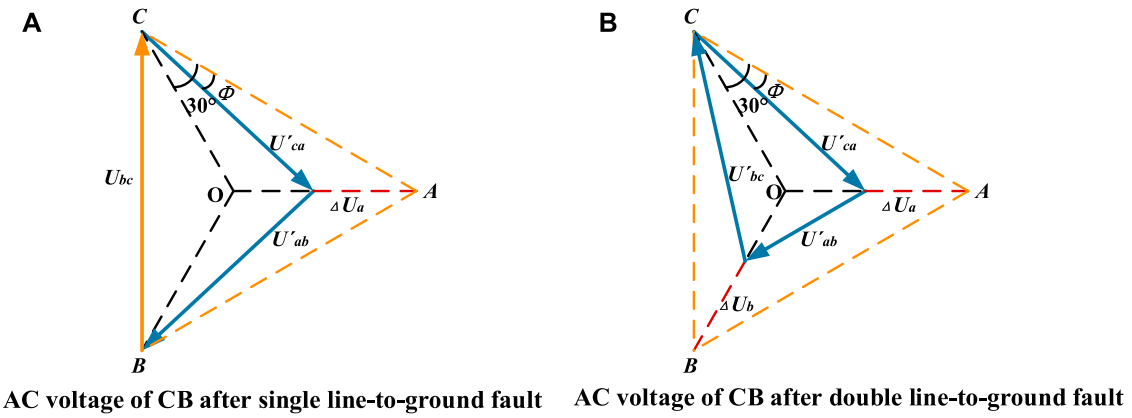


FIGURE 1 | AC voltage of converter bus after an unbalanced short circuit. **(A)** AC voltage of CB after single line-to-ground fault. **(B)** AC voltage of CB after double line-to-ground fault.

$$\begin{cases} \dot{I}_{j(1)} = \dfrac{\dot{U}_{j|0|}}{Z_{jj(1)} + \dfrac{Z_{jj(2)}Z_{jj(0)}}{Z_{jj(2)} + Z_{jj(0)}}} \\ \dot{I}_{j(2)} = -\dot{I}_{j(1)}\dfrac{Z_{jj(0)}}{Z_{jj(2)} + Z_{jj(0)}} \\ \dot{I}_{j(0)} = -\dot{I}_{j(1)}\dfrac{Z_{jj(2)}}{Z_{jj(2)} + Z_{jj(0)}} \end{cases} \tag{9}$$

The **Equation 9** can be reduced as:

$$\begin{cases} \dot{I}_{j(1)} = \dfrac{Z_{jj(2)} + Z_{jj(0)}}{k}\dot{U}_{j|0|} \\ \dot{I}_{j(2)} = \dfrac{Z_{jj(0)}}{k}\dot{U}_{j|0|} \\ \dot{I}_{j(0)} = \dfrac{Z_{jj(2)}}{k}\dot{U}_{j|0|} \end{cases} \tag{10}$$

where $k = Z_{jj(1)}\,Z_{jj(2)} + Z_{jj(1)}\,Z_{jj(0)} + Z_{jj(2)}\,Z_{jj(0)}$.

As is shown in **Figure 1B**, the phase voltage of phase *C* remains unchanged after a double line-to-ground fault happens,; the drop voltages and angle changing of phase *A* are equal to phase *B*. Substitute **(10)** into **(10)** and **(11)**, the voltage variations of phase *A* at CB *i* is:

$$\Delta U_{ia} = \left|\left[Z_{jj(0)}\left(Z_{ij(2)} - Z_{ij(1)}\right) + Z_{jj(2)}\left(Z_{ij(0)} - Z_{ij(1)}\right)\right] \cdot \dot{U}_{j0}/k\right| \tag{11}$$

According to the peculiarity of the parallel triangle, after a double line-to-ground fault, the line-to-line voltage between phase A and phase B is:

$$U'_{ab} = \sqrt{3}\,(1 - \Delta U_a) \tag{12}$$

Furthermore, it can be concluded from **Figure 1B** that the angle of $\dot{U}'_{ab}$ is not changed and the relationship of size between each line-to-line voltage is:

$$U'_{ca} = U'_{bc} > U'_{ab} \tag{13}$$

Accordingly, the extinction angle $\gamma_{V1,V4}$ which is related to U'_{ab} is smaller than the others. To reduce workload and improve the efficiency of calculation, only the line-to-line voltage between phase *A* and phase *B* will be calculated when the double line-to-ground fault happens. Substitute **Eqs 11**, **12** into **(1)**, the extinction angle $\gamma_{V1,V4}$ after double line-to-ground fault can be formulated as:

$$\gamma_{iV1,V4} = \arccos\left(\frac{\sqrt{2}k_i I_d X_{Li}}{\sqrt{3}\,(1 - \Delta U_a)} + \cos\beta\right) \tag{14}$$

The fault conditions for a line-to-line fault occurring on phase A and B can be represented as:

$$\dot{I}_{j(1)} = -\dot{I}_{j(2)} = \frac{\dot{U}_{j|0|}}{Z_{jj(1)} + Z_{jj(2)}} \tag{15}$$

Like single line-to-ground faults and double line-to-ground faults, substitutes Eq. 9 into 2 and **3**, the voltage variations of phase *A* and phase *B* of CB *i* by line-to-line fault can be formulated as:

$$\Delta U_{ia} = \Delta U_{ib} = \left|\frac{Z_{ij(1)} + Z_{ij(2)}}{Z_{jj(1)} + Z_{jj(2)}} \cdot \dot{U}_{j0}\right| \tag{16}$$

When the line-to-line fault occurs in the AC system, the relationship of the voltage phasor of CB is similar to the phasor which has occurred after a double line-to-ground fault, as is shown in **Figure 1B**. Using the same process as in *Single Line-To-Ground Fault*, the line-to-line voltage between phase *A* and phase *B* can be formulated as **Eq. 12**, and the extinction angle can also be formulated by an equation.

When the double line-to-line fault or line-to-line fault occurs at bus *j*, the extinction angle $\gamma_{V1,V4}$ of the corresponding converter at CB *i* can be calculated by **Eqs. 11–14**. CF will happen at the converter which is connected with CB *i* while $\gamma_j \le \gamma_{\min}$. This serves as a criterion of CF in detecting the line-to-line fault.

THE DEMARCATION OF THE CRITICAL FAILURE IMPEDANCE BOUNDARIES

In a multi-infeed HVDC system which contains n buses and k HVDC systems, given static models of generators, transmission lines, transformers, loads, and HVDC systems, the steps of demarcating the critical failure impedance boundaries are as follows:

1) Solve AC/DC network load flow and calculate the voltage of each bus.
2) Create the impedance matrix of the system.
3) Calculate extinction angles γ (γ_1, γ_2 ... γ_k) of every converter of the multi-infeed HVDC system in different fault conditions. To improve the computational efficiency, when the unbalanced fault happened on bus B_i, only $\gamma_{V1,V4}$ ($\gamma_{1V1,V4}$, $\gamma_{2V1,V4}$ ... $\gamma_{mV1,V4}$) which corresponds to the fault phase should be calculated. When a single line-to-ground fault happens, $\gamma_{mV1,V4}$ could be calculated by **Eq. 8**. When a double line-to-ground fault happens, $\gamma_{mV1,V4}$ could be calculated by **Eq. 14**. When a line-to-line fault happens, $\gamma_{mV1,V4}$ could be calculated by **Eq. 16**,
4) Form the bus sets based on the criterion $\gamma_{ij} \le \gamma_{\min}$ which detects the CF. The bus sets are defined in **Eq. 17** as follows:

$$\begin{aligned} B_{Tfail-i} &= \{B_1, B_2, \cdots B_k\}\left(\gamma_i \le \gamma_{\min}\right) \\ B_{Sfail-i} &= \{B_1, B_2, \cdots B_k\}\left(\gamma_{iV1,V4} \le \gamma_{\min}\right) \\ B_{Dfail-i} &= \{B_1, B_2, \cdots B_k\}\left(\gamma_{iV1,V4} \le \gamma_{\min}\right) \\ B_{Lfail-i} &= \{B_1, B_2, \cdots B_k\}\left(\gamma_{iV1,V4} \le \gamma_{\min}\right) \end{aligned} \tag{17}$$

$B_{Tfail-i}$ is the three-phase short circuit fault bus set; when a three-phase short circuit fault occurs on any bus which is included in this bus set, it may cause the CF fault of a DC converter station i. $B_{Sfail-i}$ is the single line-to-ground fault bus set when a single line-to-ground fault occurs on any bus which is included in this bus set; it may cause the CF of a DC converter station i. $B_{Dfail-i}$ is the double line-to-ground fault bus set when a double line-to-ground fault occurs on any bus which is included in this bus set; it may cause the CF of a DC converter station i. $B_{Lfail-i}$ is the line-to-line fault bus set when a line-to-line fault occurs on any bus which is included in this bus set, and it may cause the CF of a DC converter station i.

5) The corresponding critical failure impedance boundaries are drawn on the system topology diagram, and the communication failure area is defined according to the bus set in step 4). The area within the critical failure impedance boundary is defined as the corresponding critical failure impedance region. When the bus in the critical failure impedance area has a three-phase metal short circuit fault or a single line to ground relative to ground fault, it will lead to CF of the corresponding DC system. On the other hand, if a fault occurs on a bus outside the area, it will not result in CF for the corresponding DC system.

Compared with the common method which detects CF utilizing electromagnetic simulation software, the proposed method uses the minimum extinction angle as the criterion and identifies the critical impedance area quickly and accurately through a simple calculation. The identified critical impedance area provides a clear and direct perspective to the system operators of the area where faults will be critical to the DC CF of the DC system. Moreover, it will provide valuable information for system planning and protection design.

CASE STUDIES

In this section, the proposed method of detecting CFs critical impedance boundaries is validated on the IEEE 39-bus. Based on the CIGRE HVDC standard model, the quasi-steady-state model is adopted in the Multi-infeed DC system. Three-infeed HVDC systems are established in the IEEE 39-bus test system. Parameters of the three HVDC systems are listed in **Table 1**.

Following the steps in *Case Studies*, the extinction angles of the three-infeed HVDC systems in different fault conditions are calculated by substituting the above parameters of the three-infeed HVDC systems into **Eqs 8**, **14**. While the calculated extinction angle is less than $\gamma_{\min}$, these buses can be put into fault bus sets of different fault type. According to general engineering experience, $\gamma_{\min}$ is taken as 10 in this paper. Take HVDC 1 as example; its three-phase short circuit fault bus set $B_{Tfail-1}$ contains 33 buses, the single line-to-ground fault bus set $B_{Sfail-1}$ contains 12 buses, the double line-to-ground fault bus set $B_{Dfail-1}$ contains 28 buses, and the line-to-line fault bus set $B_{Lfail-1}$ contains 14 buses. The result is consistent with the severity of the bus fault. When a three-phase short circuit fault happens, the voltage drop of CB is largest in four fault conditions. The second largest voltage drop occurs when a double line-to-ground fault happens. However, for a single line-to-ground fault and line-to-line fault, the number of fault bus sets cannot be directly compared and needs to be calculated based on the method proposed in this paper. For example, the number of $B_{Sfail-1}$ is less than the number of $B_{Lfail-1}$, but the number of $B_{Sfail-3}$ is greater than the number of $B_{Lfail-3}$.

In this paper, single line-to-ground faults and line-to-line faults are selected as the focus to analyze the critical failure impedance boundaries. Then, the critical failure impedance boundaries of the three-infeed HVDC system in different fault conditions based on these fault bus sets are demarcated. The detailed results are shown in **Figure 2**.

As it is shown in **Figure 2**, some buses are included in the overlap region of different critical failure impedance boundaries. When the corresponding fault occurs on these buses, it may cause CF in more than one HVDC system. Some buses are included in non-overlap regions. When the corresponding fault occurs on these buses, it may cause CF in only one HVDC system. Therefore, the effectiveness of critical failure impedance

TABLE 1 | Parameter of the three-infeed HVDC system.

Subsystem	Rectifier bus	Inverter bus	*Control mode*	P_{DC}/ MW	V_{DC}/ kV	R_{DC}/ Ω	X_L % (%)	β	K
HVDC 1	Bus 25	Bus 18	Constant-power	200	500	7.854	15	38.4°	1.0
HVDC 2	Bus 29	Bus 17	Constant-power	200	500	7.854	15	38.4°	1.0
HVDC 3	Bus 36	Bus 13	Constant-power	200	500	7.854	15	38.4°	1.0

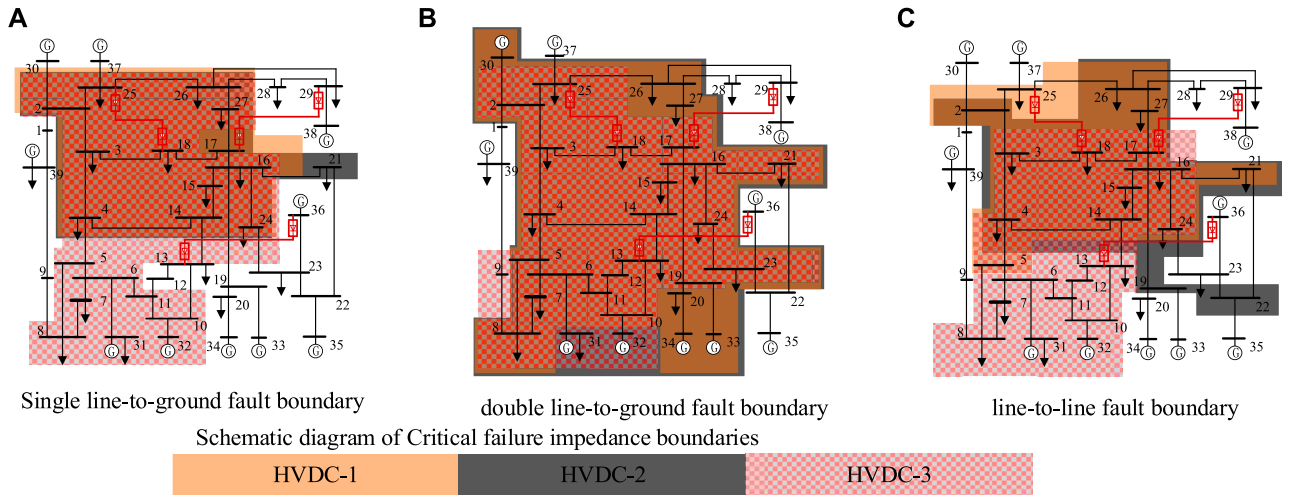


FIGURE 2 | Critical failure impedance boundaries of IEEE 39-bus test system. **(A)** Single line-to-ground fault boundary. **(B)** Double line-to-ground fault boundary. **(C)** Line-to-line fault boundary.

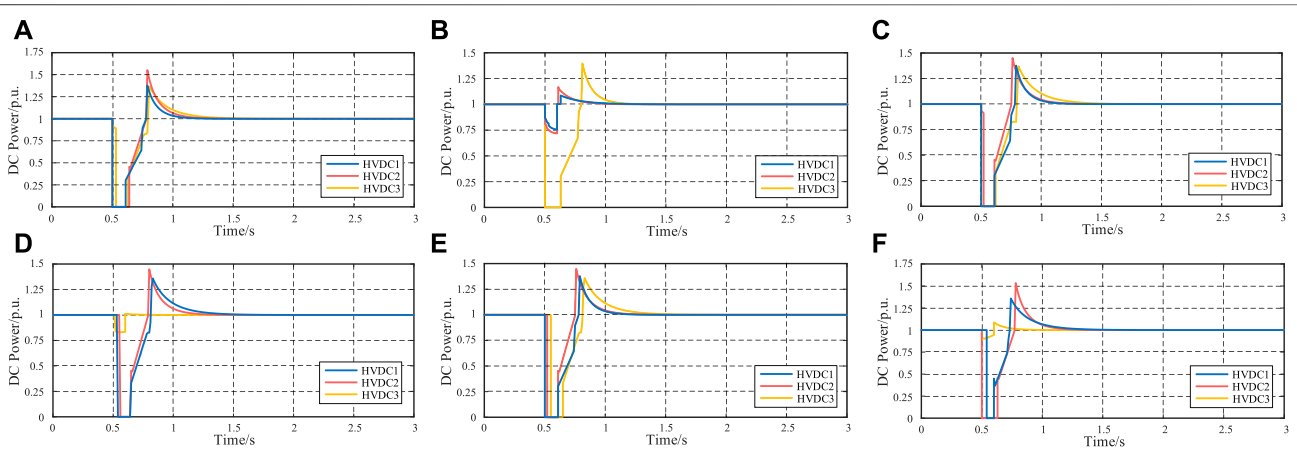


FIGURE 3 | Simulation waveform of different fault. **(A)**: DC powers while single line-to-ground fault occurs on Bus 15. **(B)**: DC powers while single line-to-ground fault occurs on Bus 5. **(C)**: DC powers while double line-to-ground fault occurs on Bus 14. **(D)**: DC powers while double line-to-ground fault occurs on Bus 20. **(E)**: DC powers while line-to-line fault occurs on Bus 15. **(F)**: DC powers while line-to-line fault occurs on Bus 27.

boundaries can be verified by conducting different types of fault simulations of the buses in the overlap region and some other buses in the non-overlap region. The single line-to-ground fault simulation is conducted on bus 5 and bus 15, the line-to-line fault simulation is conducted on bus 3 and bus 27, and the double line-to-ground fault simulation is conducted on bus 14 and bus 20.

The calculation results are obtained by using the method presented in this paper, while the single line-to-ground fault simulation is conducted on bus 15, $\gamma_{1,15} = 9.40^{\circ}$, $\gamma_{2,15} = 8.61^{\circ}$, $\gamma_{3,15} = 9.72^{\circ}$, they are all less than $\gamma_{\min}$. According to the calculation results, when a single line-to-ground fault is applied at Bus 15, CF fault will occur on three HVDC systems at the same time, the three inverters will be blocked, and their DC powers drop to 0 MW. The calculation results are the same as the simulation waveforms which are shown in **Figure 3A**. While the single line-to-ground fault simulation is conducted on bus 5, $\gamma_{1,5} = 12.96^{\circ}$, $\gamma_{2,5} = 11.45^{\circ}$, $\gamma_{3,5} = 8.70^{\circ}$, only $\gamma_{3,5}$ is less than $\gamma_{\min}$. According to the calculation results, when a single line-to-ground fault is applied at Bus 5, the DC powers of HVDC 1 and HVDC 2 drop slightly and CF only occurs on HVDC 3. The calculation results are the same as the simulation waveforms, which are shown in **Figure 3B**. Therefore, the results of the dynamic simulation are consistent with the results of critical impedance boundaries for a single line-to-ground fault proposed in this paper.

The calculation results are obtained by using the method presented in this paper, while the double line-to-ground fault simulation is conducted on bus 14, $\gamma_{1,14} = 8.94^{\circ}$, $\gamma_{2,14} = 8.96^{\circ}$, $\gamma_{3,14} = 8.28^{\circ}$, they are all less than $\gamma_{\min}$. According to the calculation results, when a double line-to-ground fault is applied at Bus 14, CF fault will occur on three HVDC systems at the same time, the three inverters will be blocked, and their DC powers drop to 0 MW. The calculation results are the same as the simulation waveforms which are shown in **Figure 3C**. While the single line-to-ground fault simulation is conducted on bus 20, $\gamma_{1,20} = 9.24^{\circ}$, $\gamma_{2,20} = 9.30^{\circ}$, $\gamma_{3,20} = 12.14^{\circ}$, only $\gamma_{1,20}$ and $\gamma_{2,20}$ are less than $\gamma_{\min}$. According to the calculation results, when a double line-to-ground fault is applied at Bus 20, the DC powers of HVDC 3 drops slightly and CF only occurs on HVDC 1 and HVDC 2. The calculation results are the same as the simulation waveforms which are shown in **Figure 3**. Therefore, the results of the dynamic simulation are consistent with the results of critical impedance boundaries for a double line-to-ground fault proposed in this paper.

The calculation results are obtained by using the method presented in this paper, while the line-to-line fault simulation is conducted on bus 3, $\gamma_{1,3} = 8.04^{\circ}$, $\gamma_{2,3} = 7.90^{\circ}$, $\gamma_{3,3} = 9.72^{\circ}$, they are all less than $\gamma_{\min}$. According to the calculation results, when a line-to-line fault is applied at Bus 15, CF fault will occur on three HVDC systems at the same time, the three inverters will be blocked, and their DC powers drop to 0 MW. The calculation results are the same as the simulation waveforms which are shown in **Figure 3E**. While the line-to-line fault simulation is conducted on bus 27, $\gamma_{1,27} = 7.68^{\circ}$, $\gamma_{2,27} = 7.75^{\circ}$, $\gamma_{3,27} = 12.14^{\circ}$, only $\gamma_{1,27}$ and $\gamma_{2,27} = 7.75^{\circ}$ are less than $\gamma_{\min}$. According to the calculation results, when a line-to-line fault is applied at Bus 27, the DC powers of HVDC 3 drop slightly and CF occurs on HVDC 1.and HVDC2. The calculation results are the same as the simulation waveforms which are shown in **Figure 3F**. Therefore, the results of the dynamic simulation are consistent with the results of critical impedance boundaries for a line-to-line fault proposed in this paper.

CONCLUSION

In this paper, a region boundaries' determination method for critical commutation failures in multi-infeed HVDC systems is proposed with consideration of unbalanced short circuit faults. A criterion of critical extinction angle is put forward. Numerical simulation results on the modified IEEE 39 bus system demonstrate that the presented method manages to identify the boundaries of the critical commutation failures. Compared with the actual power grid operation, the method presented in this paper is a programmatic calculation method and the results are more rigorous. The calculated boundaries of commutation failures will be larger than the actual boundaries.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

GL instructed and proposed the methodology used in this paper. SZ conceptualized and implemented this study, and wrote the original draft. SL dedicated his time to validate the method. XCL and XTL assisted in the conceptualization, investigation, and data.

REFERENCES

Guo, C., Liu, Y., Zhao, C., Wei, X., and Xu, W. (2015). Power component fault detection method and improved current order limiter control for commutation failure mitigation in HVDC. *IEEE Trans. Power Deliv.* 30 (3), 486–495. doi:10.1109/tpwrd.2015.2411997

Li, G., Zhang, S., and Jiang, T., Chen, H., and Li, X. (2017). "A method of detecting commutation failure in multi-infeed HVDC systems based on critical failure impedance boundary. "in IEEE Power & Energy Society General Meeting, Chicago, IL, July 16–20, 2017, 1–5.

Mirsaeidi, S., Dong, X., Tzelepis, D., Said, D. M., Dysko, A., and Booth, C. (2018). A predictive control strategy for mitigation of commutation failure in LCC-Based HVDC systems. *IEEE Trans. Power Elect.* 34 (1), 160–172. doi:10.1109/TPEL.2018.2820152

Wang, F., Liu, T. Q., and Li, X. Y. (2017). Decreasing the frequency of HVDC commutation failures caused by harmonics. *IET Power Elect.* 10 (2), 215–221. doi:10.1049/iet-pel.2016.0230

Wang, Z., Liu, X., Li, L., and Yang, Y. (2017). Boundary conditions of commutation failure in multi-infeed HVDC systems. *Trans. China Electrotech. Soc.* 32 (10), 12–19.

Wang, J., Huang, M., Fu, C., Li, H., Xu, S., and Li, X. (2019). A new recovery strategy of HVDC system during AC faults. *IEEE Trans. Power Deliv.* 34 (2), 486–495. doi:10.1109/tpwrd.2019.2892410

Wang, Q., Zhang, C., Wu, X., and Tang, Y. (2019). Commutation failure prediction method considering commutation voltage distortion and DC current variation. *IEEE Access* 7, 96531–96539. doi:10.1109/access.2019.2929301

Wei, Z., Yuan, Y., Lei, X., Wang, H., Sun, G., and Sun, Y. (2014). Direct-current predictive control strategy for inhibiting commutation failure in HVDC converter. *IEEE Trans. Power Syst.* 29 (5), 2409–2417. doi:10.1109/tpwrs.2014.2302010

Xiao, H., Li, Y., Zhu, J., and Duan, X. (2016). Efficient approach to quantify commutation failure immunity levels in multi-infeed HVDC systems. *IET Gene., Transm. Distri.* 10 (4), 1032–1038. doi:10.1049/iet-gtd.2015.0800

Xue, Y., Zhang, X., and Yang, C. (2016). Elimination of commutation failures of LCC HVDC system with controllable capacitors. *IEEE Trans. Power Syst.* 31 (4), 3289–3299. doi:10.1109/tpwrs.2015.2491784

Xue, Y., Zhang, X. P., and Yang, C. (2018). Commutation failure elimination of LCC HVDC systems using thyristor-based controllable capacitors. *IEEE Trans. Power Deliv.* 33 (3), 1448–1458. doi:10.1109/tpwrd.2017.2776867

Yao, W., Liu, C., Fang, J., Ai, X., Wen, J., and Cheng, S. (2020). Probabilistic analysis of commutation failure in LCC-HVDC system considering the CFPREV and the initial fault voltage angle, *IEEE Trans. Power Deliv.* 35 (2), 715–724. doi:10.1109/tpwrd.2019.2925399

Zheng, A., Guo, C., Cui, P., Jiang, W., and Zhao, C. (2019). Comparative study on small-signal stability of LCC-HVDC system with different control strategies at the inverter station. *IEEE Access* 7, 34946–34953. doi:10.1109/access.2019.2904395

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

The reviewer YL declared a shared affiliation with one of the authors, SZ, to the handling editor at time of review.

frontiers in Energy Research

BRIEF RESEARCH REPORT
published: 27 April 2021
doi: 10.3389/fenrg.2021.674865

A Power Customer Data Relational Algorithm Based on Magnanimity Fuzzy Address Matching

Peng Jin, Jing Yang, Zongwei Wang, Xiaoyang Bu and Peng Wu*

State Grid Customer Service Center, Tianjin, China

According to the short text and unstructured characteristics of customer address, a data association fusion method for address has been proposed. In this method, the address was mapped to a digital fingerprint by improved Simhash technology, which effectively reduced the dimension of massive addresses and simplified the similarity-matching process of multi-source heterogeneous addresses. Furthermore, the weight setting of the eigenvector of the simhash algorithm was improved by introducing special weight gain. A two-level index mechanism was established by the characteristics of address division and data structure of digital fingerprints; the time-consuming digital fingerprint comparison was greatly reduced. The experimental results showed that calculation efficiency was greatly optimized; accuracy and coverage of the comparison were ensured. Through address matching of different databases, information fusion can be completed and the goal which power customers' demands is connected to power grid equipment is achieved.

Keywords: improved simhash, multi-source heterogeneous data, address matching, data associations, digital fingerprint, data of electric client

OPEN ACCESS

Edited by:
Liang Chen,
Nanjing University of Information Science and Technology, China

Reviewed by:
Jie Sheng,
Soochow University, China
Jianzhong Xu,
North China Electric Power University, China

****Correspondence:***
Peng Jin
470292065@qq.com

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

Received: *02 March 2021*
Accepted: *24 March 2021*
Published: *27 April 2021*

Citation:
Jin P, Yang J, Wang Z, Bu X and Wu P (2021) A Power Customer Data Relational Algorithm Based on Magnanimity Fuzzy Address Matching. Front. Energy Res. 9:674865. doi: 10.3389/fenrg.2021.674865

INTRODUCTION

With the deepening of electric power reform, grid enterprises have gradually begun to establish a modern customer-centered service mode in recent years. They were eager to explore the potential demands of customers to support the development of new formats and optimize the allocation of service resources. However, the characteristics of scattered and massive customer information bring great challenges to grid enterprises to carry out customer behavior mining and accurate service.

Big data technology, as a means to efficiently process massive data with complex sources, has brought revolutionary changes and has made the association analysis possible. Many cases have proved that the value will be brought into full play by integrating data in different fields, specialties, and channels (Wang et al., 2018; Shen et al., 2019; Zhou et al., 2020). Large grid enterprises, which set up big data organizations and big data platforms, have taken the initiative to carry out digital transformation (Sun, 2019) and promote the integration of grid business and customer electricity behavior information rapidly (Teeraratkul et al., 2018; Wang et al., 2019; Li et al., 2021).

Unlike banks or other financial service industries, customers rarely provide information, such as name and power electricity number, in the process of power-related services; power grid companies must carry out emergency repair services according to the address fed back by customers. Because of the fragmentation and unstructured characteristics of the information provided by the customers (Song, 2013), there are great differences between the information provided by the customers and the electric power standardized data in the same entity expression (Wang, 2012). In the process of

fusion with power grid data, it is necessary to make pair-wise comparison of all entities between the customer information database and the power marketing database. Because of the diversity and heterogeneity of linked data (Xie et al., 2015; Wu et al., 2016), there are problems of low coverage and precision in association matching. In addition, the amount of customer information is massive, and the calculation complexity, which achieves data alignment of similar address entities among different databases, is high (Shen and Feng, 2018; Kang et al., 2019).

In order to realize the information fusion between two databases, it is generally necessary to rely on the reliable similarity function (Liu et al., 2017). The comparison texts were decomposed into a set of tokens, and then, the token set was transformed into an n-dimensional vector; the similarity of comparison texts was evaluated by calculating the cosine similarity between vectors. The order of the token was not considered in this method, but the quality of tokens seriously affected the accuracy of comparison (Ye, 2011). At present, the mainstream algorithms mainly use word segmentation technology to form tokens, but segmentation ambiguity has a great impact on the results. In order to solve the problems caused by word segmentation, the q-gram similarity function was proposed (Sun and Wang, 2014). Since the length of the substring was fixed, there was overlap between the tokens, which could effectively avoid the problem caused by segmentation divergence. However, the q-gram method will greatly increase the amount of calculation, where q is generally <4. In the minimum edit distance method (Belazzougui and Venturini, 2016), it is not necessary to segment the word vector; this method transformed the target text into the comparison text by inserting, deleting, and exchanging operations. The operation cost was used as the similarity function, and the difference between matched texts could be better quantified. In addition, this method was not sensitive to local missing characters. Unstructured text entities were more likely to be converted into structured data for comparison, so the local sensitive hashing algorithm was proposed (Can et al., 2017), which could map the text data to a fixed-length fingerprint set through a special hash function. Compared with matching the text entity, the comparison of a fingerprint could greatly reduce the complexity of data storage and calculation and provide an effective choice for mass data comparison. In addition, the computational complexity of matching text entity was proportional to the quadratic power of the entity size in the database (Zhuang et al., 2016). For the entity alignment of massive data, the computing resources and computational efficiency were unbearable, and further dimensionality reduction must be carried out. Partitioned indexes (Qu et al., 2018) were used to reduce computational complexity, and the entities with the same key value were placed in the same block. The similarity comparison was only carried out in the same block, and it was no comparison between blocks, which should greatly reduce the amount of calculation.

At present, there are few cases of data fusion through the address in the field of the electric power industry. Therefore, this study proposes a data fusion method based on the improved simhash algorithm, which considers the short text characteristics of the address comprehensively. Through the fuzzy address-matching technology, customer behavior data can be associated with data of power grid equipment. Furthermore, through the two-level partition index mechanism, the dimensionality reduction of comparison is realized, which can support the massive data matching of more than 100 million levels.

ADDRESS FUZZY MATCHING ASSOCIATION MODEL BASED ON IMPROVED SIMHASH ALGORITHM

Association Model Based on Address Data

State Grid Corporation of China (SGCC) has constructed complete file information of customers, metering devices, and power grid equipment, which can be linked through a unique power consumption number. Customers only provide the detailed address of the fault when calling the customer service hotline but cannot provide valuable information such as the power consumption number. Therefore, the information provided by power customers is difficult to directly relate to the power supply point, which has a negative impact on fault research and data analysis.

Figure 1 shows the data model structure. The work order of the customer service can be associated with power grid equipment only through customer file information and metering point information. Besides the customer electricity number, the site address of the work order can also be associated with the customer file information reservation. However, the above data are distributed in different data fields. In order to associate the data through the address, the address data must be standardized to improve the coverage of data matching.

Standardization of Customer Address

At present, the site address of SGCC repair work order is divided into two parts: one is structured data, including the information of provincial-, municipal- and county-level power supply units; the other part is unstructured data, covering the string information of village, town, street, road, community, and door combination.

Affected by the accuracy of the expression and local dialects of customers, some addresses may have some deviation. The first step is to remove invalid information, such as separators, spaces, and English letters in the address, and convert the Arabic numerals information into Chinese characters; the second step is to filter and remove the house number and other information in the text to achieve the comparison of address dimensions, such as community, village, and street; and the third step is to convert the cleaned address into Chinese pinyin.

Principle of Simhash Algorithm

The simhash algorithm was proposed by Charikar in 2002 (Naumann, 2010). As a digital fingerprint algorithm, it aimed to solve the problem of removing the duplicate of massive web pages. Due to digit limitation of a fingerprint, the matching objects could be detected quickly in a large-scale database. So the data dimension was compressed, and the number of similarity comparison was greatly reduced. In addition, local sensitivity

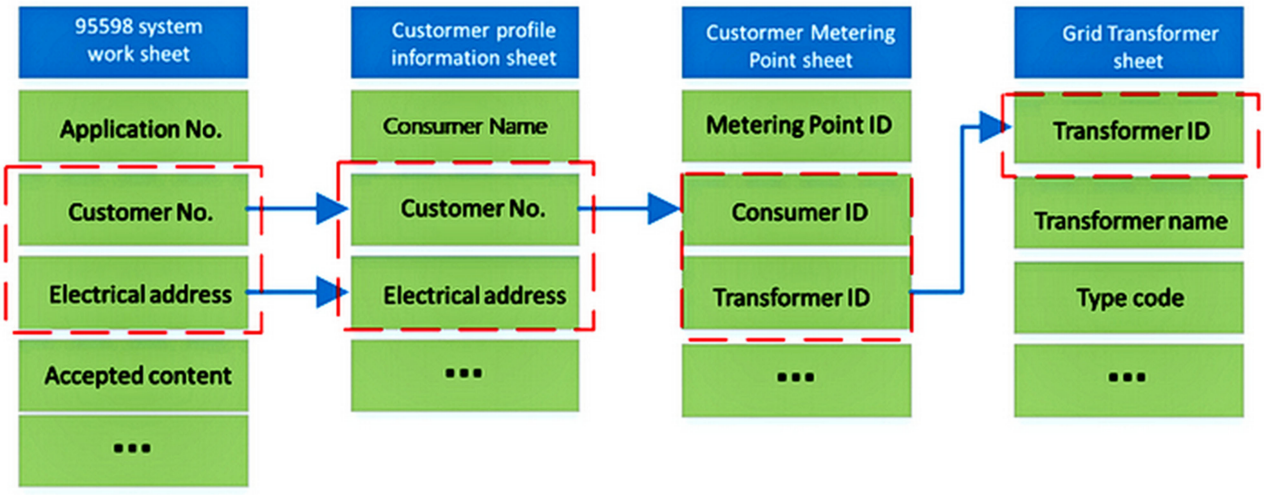


FIGURE 1 | Model of customer demands associated with power grid equipment.

was demonstrated in this method, and the similarity of any two entities was proportional to the similarity of the generated fingerprint, which helped to quantify the differences between the contrast entities.

The core idea of the simhash algorithm is to map the unstructured text set to the unique signature value generated from the original data. The generated digital fingerprint is a series of fixed-length binary codes, as shown in **Figure 2**. The whole process can be divided into four steps: text feature extraction, fingerprint generation, fingerprint index, and matching calculation.

By using the traditional MD5 hash technology, the word vector is mapped to the digital fingerprint and weighted according to the given word vector.

$$f(hash_{pv}) = \begin{cases} weight_p, & hash_{pv} = 1 \\ -weight_p, & hash_{pv} = 0 \end{cases} \quad (1)$$

where $weight_p$ is the weight of the p-th word vector; $Hash_{pv}$ is the value of the v-th bit of the hash fingerprint that the p-th word vector is mapped.

Then, the weighted values of all word vectors in the text are accumulated and merged to form a new sequence T_j. The dimension is reduced to form the final simhash digital fingerprint according to Equation (2).

$$Simhash_j = \begin{cases} 1, & T_j > 0 \\ 0, & T_j < 0 \end{cases} \quad (2)$$

After the batch formation of digital fingerprints, the similarity between fingerprints needs to be further judged. Generally, the similarity between digital fingerprints is measured by hamming distance. There are two binary strings, **x** and **y,** of length n. The hamming distance between them can be calculated as follows:

$$Hamming(\mathbf{x}, \mathbf{y}) = \sum_{1}^{n} (x_i \oplus y_i) \quad (3)$$

where $\oplus$ is XOR operation. Different digits of digital fingerprints can be calculated by hamming distance. Generally, the higher the coincidence degree of two text sets, the higher the similarity between digital fingerprints. If two 64-bit binary strings have less than three different characters, they can be regarded as similar text sets.

Improved Simhash Algorithm Considering Address Characteristics

There are no predicate, attribute, adverbial, and complement in Chinese addresses. The average length of the addresses of 1 million customers is 19.7 characters. Generally, the length of an address is <45 characters, which basically follows the normal distribution. The number of address characters is relatively small, and the traditional word vector is not used as the simhash feature vector. Instead, the character is used as the feature vector to avoid the misjudgment and the time cost caused by word segmentation.

The structure of an address is basically extended from large administrative divisions to fine addresses with strong regularity. The accuracy of fine address matching often determines the matching of the whole address. In order to improve the resolution accuracy of a terminal fine address and the weight component of a fine address, the weight of the simhash feature vector is adjusted according to the following Equation (4):

$$K_{\mathrm{i}} = \lfloor i^k \rfloor (i \in [1, addr_{len}]) \quad (4)$$

where K_i is the weight gain of the improved eigenvector, $addr_{\mathrm{len}}$ is the word length of the address text, i is the character bit of the eigenvector, k is the power exponent parameter, and $\lfloor\rfloor$ represents the down rounding function.

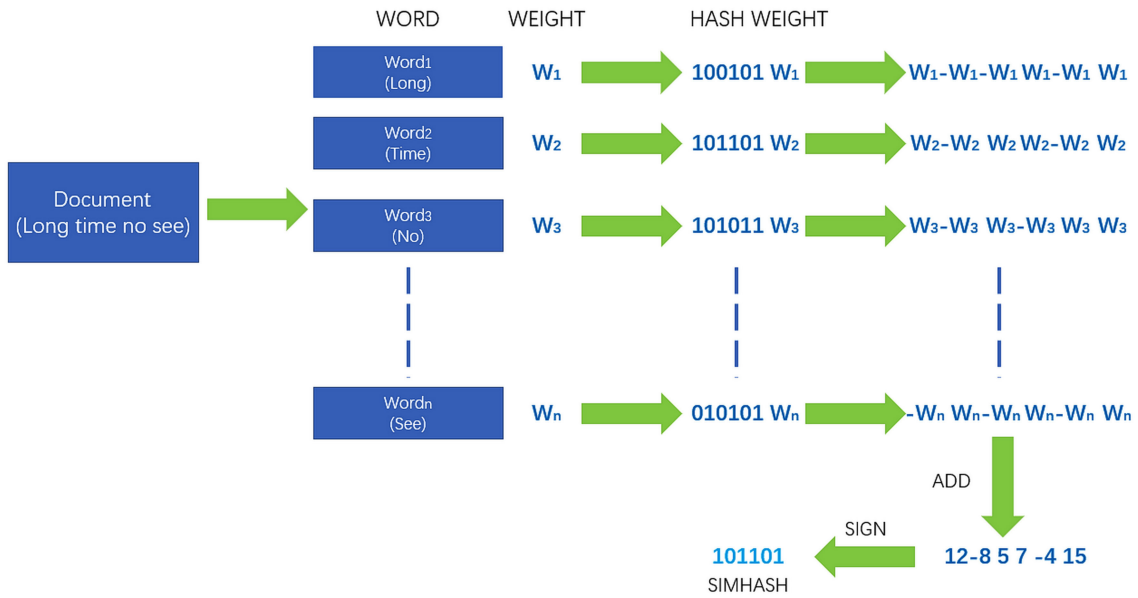


FIGURE 2 | Simhash algorithm flow.

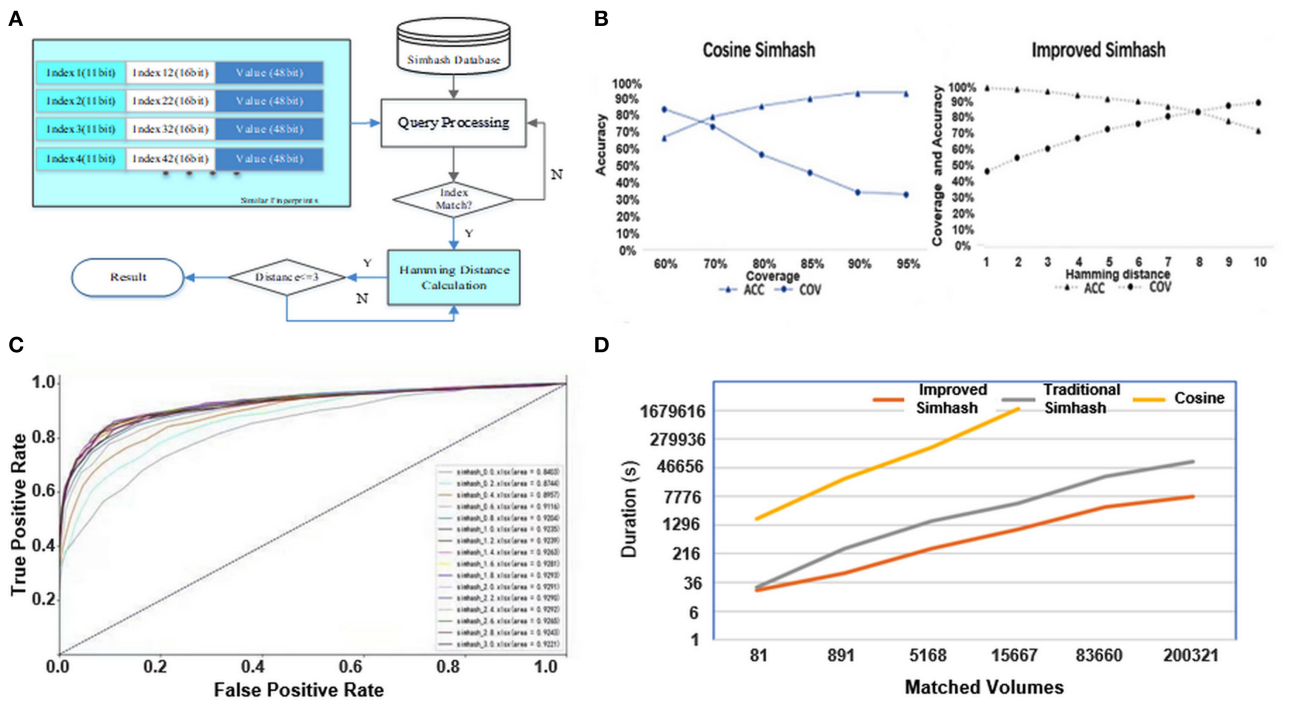


FIGURE 3 | (A) Hierarchical index architecture. **(B)** Algorithm performance evaluation**. (C)** Influence of parameter change on ROC curve. **(D)** Efficiency comparison of different methods.

Segmented Index Method Based on Improved Simhash Algorithm

Text information dimensionality reduction can be achieved by Simhash technology, but the workload of text comparison between the two types of data is very huge. It is necessary to compress the comparison amount through the segmented index method to support the ability of massive data processing. In the address structure, the detailed address text is unstructured, while the addresses of provinces, cities, districts, and counties are usually structured.

TABLE 1 | Comparison of data classifying quality.

Indicators / Hamming distance		1 (%)	2 (%)	3 (%)	4 (%)	5 (%)	6 (%)
The proposed method	Accuracy	99.07	98.05	96.86	94.29	92.34	90.17
	Coverage	47.35	55.53	61.39	67.59	73.34	75.01
The traditional simhash method	Accuracy	98.62	95.67	89.95	85.41	77.89	70.76
	Coverage	15.82	24.45	37.61	52.43	66.26	72.43

Therefore, the index mechanism is established based on the dimension of a district and a county, and the simhash comparison is limited to the same district and the county, which greatly reduces the amount of unnecessary comparison. At present, there are more than 1,400 county-level power supply units within the scope of SGCC. The resolution of the 11-bit binary code is 2,048, so all district- and county-level units can be indexed by the 11-bit binary code.

If the hamming distance is <3, two strings, which are converted to the 64-bit digital fingerprint, are considered to be similar. According to the drawer principle, if two similar texts are serialized into 64-bit digital fingerprints, there must be 16 bits that are exactly the same. Therefore, the same 16-bit binary code can be used as the index; the actual comparison bit will be reduced to 48 bits.

By combining the above two-tier index construction methods, the matching workload will be greatly reduced. In addition, the repeated calculation can be avoided effectively when the new address is added to the database, as shown in **Figure 3**.

EXAMPLE ANALYSIS

Sample Selection and Experimental Environment

In order to prove the address matching verification, the standard addresses of power customer file information and the address of customer demand work order are processed centrally as the basic data for comparison. Address processing includes eliminating a house number and other detailed information.

The algorithm test is based on the big data platform, with 10 physical machines, which are dual-channel 12 core Xeon e5-2650 V4 processors (2.2 GHz). The Hadoop cluster of this platform has 24 nodes, 1.6T memory, and 235T physical storage.

Experimental Results

In the simhash algorithm, the more the same number of the digital fingerprint, the more similar. But the higher the similarity, the smaller the hamming distance, which leads to lower coverage in similar matching. The performance of the proposed method is further analyzed by comparing the traditional simhash algorithm. The test results are as shown in **Table 1**. The test results are as follows:

With the same accuracy, the coverage of the proposed method is better than the traditional simhash method. In the case of short text, the improvement is more obvious, and the accuracy and the coverage are more than 80%.

In addition, compared with the cosine similarity comparison method proposed in (Ye, 2011), the improved simhash method has obvious advantages in accuracy and coverage, as shown in **Figure 3B**.

In the cosine similarity comparison method, the feature vector is determined by the word segmentation technology. Once the unknown words appear in the address, it is easy to produce segmentation divergence, which greatly affects the accuracy and success rate. In the improved simhash method, the feature vectors are determined by characters to avoid the problems that may be caused by word segmentation, so the matching accuracy and the coverage are improved.

The concept of weight gain is proposed in (Sun, 2019). In order to get a more suitable weight gain, we traversed the k value between [0.2, 3] and further analyzed the influence of k value change on accuracy and coverage. As can be seen from the **receiver operating characteristic** (ROC) curve in **Figure 3C**, with the increase of k value, the weight of the terminal address, such as village, town, and cell name, also increased. In addition, the area under the ROC curve continued to increase, and the generalization ability gradually improved. When $k = 1.8$, the area under the ROC curve reached the maximum (AUC $= 0.9294$). Then, the generalization ability of the model decreased with the increase of k.

Cosine similarity, traditional simhash, and improved simhash are tested by matching the address of the village and the county level, and the specific time consumption is shown in **Figure 3D**. The cosine similarity method contains word segmentation, cosine calculation, and other steps, and dimensionality reduction cannot be realized. When the amount of calculation increases greatly, the results may not be calculated.

When the number of matching jobs is low, the efficiency of traditional simhash and improved simhash is almost the same. With the increase of the number of matching, the efficiency advantage of the improved simhash method appears. The main reason is that the traditional simhash method needs to segment the address, and the time consumption of text-dimension reduction is higher than the improved simhash method.

MAIN CONCLUSIONS

An address-matching method based on the improved simhash algorithm is proposed to realize the association between

unstructured addresses. There are two innovations in this method: first, the traditional simhash algorithm is optimized by improving the weight gain of simhash and adjusting the text vector composition so as to ensure the coverage and accuracy of address matching; second, according to the characteristics of the hash fingerprint and address text, a two-level index mechanism is constructed to reduce the complexity of the address-matching algorithm and improve the efficiency of data fusion. The mechanism solves massive address-matching problems and helps to link up the non-standard addresses reflected by power customers to power grid equipment so as to enhance the application value of power customer data.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

PJ led the analysis and wrote the manuscript. JY performed the experiment. ZW carried out the study and collected important background information. XB contributed to the research of algorithm concept. PW helped perform the analysis with constructive discussions. All authors contributed to the article and approved the submitted version.

REFERENCES

Belazzougui, D., and Venturini, R. (2016). Compressed string dictionary search with edit distance one. *Algorithmica* 74, 1099–1122. doi: 10.1007/s00453-015-9990-0

Can, L., Qian, J., and Dong, Y. (2017). M2LSH:an LSH based technique for approximate nearset neighbor searching on high dimensional data. *Acta Electron. Sin.* 45, 1431–1442. doi: 10.3969/j.issn.0372-2112.2017.06.022

Kang, S., Ji, L., Liu, S., and Ding, Y. (2019). Cross-lingual entity alignment model based on the similarities of entity descriptions and knowledge embeddings. *Acta Electron. Sin*. 47, 1841–1847.

Li, Y., Wang, C., Li, G., and Chen, C. (2021). Optimal scheduling of integrated demand response-enabled integrated energy systems with uncertain renewable generations: a Stackelberg game approach. *Energy Convers. Manage.* 235:113996. doi: 10.1016/j.enconman.2021.113996

Liu, Z., Chen, J., Zheng, J., Hua, J., and Xiao, L. (2017). Research on aggregation model for chinese short texts. *J. Softw*. 28, 2674–2692. doi: 10.13328/j.cnki.jos.005147

Naumann, F. (2010). *An Introduction to Duplicate Detection*. San Rafael, CA: Morgan and Claypool.

Qu, Z., Fan, M., Zhou, R., Wang, H., and Zhu, D. (2018). Inverted index query technique of non-primary key for mass dispatch and monitoring information of distribution network. *Power Syst. Protection Control.* 46, 162–168. doi: 10.7667/PSPC171742

Shen, B., and Feng, J. (2018). Crowdsourcing knowledge base index alignment. *Chin. J. Comput*. 41, 1814–1826. doi: 10.11897/SP.J.1016.2018.01814

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, 43–55+319.

Song, Z. (2013). Address matching algorithm based on chinese natural language understanding. *J. Remote Sens*. 17, 788–801. doi: 10.11834/jrs.20132164

Sun, D., and Wang, X. (2014). Q-gram index for approximate string matching with multi-seeds. *Comput. Sci.* 41, 279–284. doi: 10.11896/j.issn.1002-137X.2014.09.053

Sun, Y. (2019). Accelerate digital transformation to create a new pattern of coordinated development. *Energy Res. Utilization* 1, 4–5+7. doi: 10.16404/j.cnki.issn1001-5523.2019.03.001

Teeraratkul, T., O'Neill, D., and Lall, S. (2018). Shape-based approach to household electric load curve clustering and prediction. *IEEE Trans. Smart Grid* 9, 5196–5206. doi: 10.1109/TSG.2017.2683461

Wang, D. (2012). Power data center infrastructure based on cloud computing and its key technologies. *Automation Electric Power Syst*. 36, 67–71+107.

Wang, Q., Li, F., Tang, Y., and Xue, Y. (2018). On-line prediction method of transient frequency characteristics for power grid based on physical-statistical model. *Automation Electric Power Systems*. 42, 1–11. doi: 10.7500/AEPS20171001001

Wang, Y., Zhang, N., Kang, C., Xi, W., and Huo, M. (2019). Electrical consumer behavior model:basic concept and research framework. *Trans. China Electrotech. Soc*. 34, 2056–2068. doi: 10.19595/j.cnki.1000-6753.tces.190073

Wu, Q., Gao, J., Hou, G., Han, B., Wang, K., and Li, G. (2016). Short-term load forecasting support vector machine algorithm based on multi-source heterogeneous fusion of load factors. *Automation Electric Power Syst*. 40, 67–72+92. doi: 10.7500/AEPS20160229012

Xie, G., Hu, Y., Chen, J., Yu, N., and Zhou, H. (2015). A fusion method for multi-source and heterogeneous parameters of power grid and its engineering application. *Automation Electric Power Syst.* 39, 121–127. doi: 10.7500/AEPS20140424004

Ye, J. (2011). Cosine similarity measures for intuitionistic fuzzy sets and their applications. *Math. Comput. Model.* 53:91–97. doi: 10.1016/j.mcm.2010.07.022

Zhou, F., Zhou, H., and Diao, Y. (2020). Development of intelligent perception key technology in the ubiquitous internet of things in electricity. *Proc. CSEE* 40, 70–82+375. doi: 10.13334/j.0258-8013.pcsee.191198

Zhuang, Y., Li, G., and Feng, J. (2016). A survey on entity alignment of knowledge base. *Comput. Res. Dev*. 53, 165–192. doi: 10.7544/issn1000-1239.2016.20150661

Conflict of Interest: PJ, JY, PW, ZW, and XB are employed by the same company State Grid Customer Service Center.

frontiers in Energy Research

BRIEF RESEARCH REPORT
published: 20 May 2021
doi: 10.3389/fenrg.2021.686138

Improved Short-Circuit Current Calculation of Doubly Fed Wind Turbines With Uninterrupted Excitation

*Jun Yin**

School of Electric Power, North China University of Water Resources and Electric Power, Zhengzhou, China

With the enlarging scale of a doubly fed induction generator (DFIG) connected to a power system, the influence of short-circuit current on the system relay protection could not be ignored. Setting and configuring relay protection would be affected by an imprecise short-circuit current calculation. However, some existing studies only consider the condition that the input is the Crowbar and the rotor excitation is blocked. China's new network standard requires the output reactive support current of a DFIG and will change the characteristics of short-circuit current. To solve this problem, on the basis of analyzing the characters of the transient equivalent potential of a DFIG, the transient model of a DFIG with uninterrupted excitation is provided. Based on the characteristics of a non-abrupt change of flux linkage and the requirement of a new grid standard reactive support current, the short-circuit current calculation method of a DFIG with uninterrupted excitation is put forward. Based on the real-time digital simulator (RTDS), a digital-analog experimental platform containing the actual control unit of the DFIG converter is founded, the proposed short-circuit current root mean square (RMS) value calculating method is validated.

Keywords: DFIG, short-circuit current, fault analysis, relay protection, rotor excitation

OPEN ACCESS

Edited by:
Liang Chen,
Nanjing University of Information Science and Technology, China

Reviewed by:
Yunbing Wei,
Shanghai University of Engineering Sciences, China
Xiangqian Zhu,
Zhengzhou University of Light Industry, China

***Correspondence:**
Jun Yin
yinjun0371@126.com

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

Received: *26 March 2021*
Accepted: *13 April 2021*
Published: *20 May 2021*

Citation:
Yin J (2021) Improved Short-Circuit Current Calculation of Doubly Fed Wind Turbines With Uninterrupted Excitation. Front. Energy Res. 9:686138. doi: 10.3389/fenrg.2021.686138

INTRODUCTION

In recent years, as represented by wind power, new energy sources have been developing rapidly in China. In 2019, there are 26.79 million kilowatts wind power newly installed in China. And the grid connected wind power capacity has reached 270 million kilowatts. However, with rapid development in the generation of wind power, its impact on the power grid has become more and more obvious. In particular, the impact on the power system relay protection has become a major concern in the current power system field (Haj-Ahmed et al., 2018; Telukunta et al., 2018; Duan et al., 2020).

In China, doubly fed induction generators (DFIGs) were widely used in MW-class wind turbines whose converter capacity is only about one-third of the rated capacity. They have the advantages of low cost and the decoupled control of active and reactive power. However, with a continuous increase in the capacity of a DFIG, its short-circuit current can no longer be neglected in the

protection setting and configuration. Therefore, it is necessary to deeply study the short-circuit current calculation method of a DFIG after being connected to the grid.

At present, some scholars have studied the short-circuit current calculation method of a DFIG. A few literature studies (Ouyang et al., 2017; Garibay et al., 2018; Pan et al., 2019; Rahimi and Azizi, 2019) analyze the condition that the input is the rotor Crowbar after the fault. And a DFIG is equivalent to a squirrel-cage asynchronous generator. Then, the analytical expression of the short-circuit current is given. However, the above study only considers the case in which the input is the Crowbar and the rotor-side converter is locked. China has decreed a new wind power grid-connected standard GB/T19963-2011 "Technical Regulations for Wind Farm Access to Power System." In the standard, it is required that a DFIG needs to provide reactive current to support system voltage during the fault. In this situation, the rotor-side converter is no longer locked and will provide continuous excitation during the fault.

At present, there are also few literature studies, which have studied the short-circuit current calculation method of a DFIG with continuous excitation after the fault. The simulation of the literature (Firouzi and Gharehpetian, 2017; Ying et al., 2020) verifies that there is a big difference between the short-circuit current of a DFIG under continuous excitation and the input being the Crowbar. A few literature studies (Tamaarat and Benakcha, 2014; Wang et al., 2015; Liu et al., 2020) assume that the rotor excitation current does not change before and after the fault, and give the expression of the short-circuit current of a DFIG, in the case of remote fault. However, the above studies consider that the excitation current of the rotor converter is constant before and after the fault. The new grid-connected standard requires that the rotor converter needs to adjust the excitation current according to the degree of voltage dip. However, the rotor excitation current will change before and after the fault, which will affect the output characteristics of short-circuit current. Therefore, it is necessary to analyze the requirements for reactive support power in the new grid-connected standard and propose a method for calculating the short-circuit current of a DFIG.

In order to solve this problem, based on analyzing the equivalent potential characteristics of a DFIG during the fault, this paper firstly proposes a DFIG transient equivalent model DFIG under continuous excitation. By using the non-abrupt change theorem of the flux linkage, the short-circuit current calculation method of a DFIG at the initial fault time is proposed. And the requirement of the reactive support power in the grid-connected standard is analyzed, then the short-circuit current calculation method of a DFIG at the steady-state fault time is proposed. Finally, the short-circuit current calculation model of a DFIG under continuous excitation is established. Based on the real-time digital simulator (RTDS), a digital-analog experimental platform containing the actual control unit of the DFIG converter is established, and the accuracy of the short-circuit current calculation method is verified.

DFIG TRANSIENT MODEL UNDER CONTINUOUS EXCITATION

The main circuit topology of a DFIG is shown in **Figure 1**. A stator is directly connected to a grid and a DFIG is excited by a rotor-side converter.

In the previous studies, it was considered that the input was the Crowbar of a DFIG after the occurrence of the fault and the converter was locked. But (National Electricity Regulatory Standardization Technical Committee, 2011) China's new wind power grid-connected standard GB/T19963-2011 "Technical Regulations for Wind Farm Access to Power Systems," requires that a DFIG needs to output reactive current to provide voltage support for the system during the fault. At this point, the rotor converter is no longer locked and provides continuous excitation for a DFIG during the fault.

Ignoring the magnetic saturation and assuming that the rotation speed does not change during a transient process, the space vector model of a DFIG in the synchronous rotating coordinate system is shown in Equations (1, 2) (Muljadi et al., 2013):

$$\begin{bmatrix} u_s \\ u_r \end{bmatrix} = \begin{bmatrix} R_s & 0 \\ 0 & R_r \end{bmatrix} \begin{bmatrix} i_s \\ i_r \end{bmatrix} + \begin{bmatrix} \frac{d\psi_s}{dt} \\ \frac{d\psi_r}{dt} \end{bmatrix} + \begin{bmatrix} j\omega_s & 0 \\ 0 & j\omega_{s-r} \end{bmatrix} \begin{bmatrix} \psi_s \\ \psi_r \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} \psi_s \\ \psi_r \end{bmatrix} = \begin{bmatrix} L_s & L_m \\ L_m & L_r \end{bmatrix} \begin{bmatrix} i_s \\ i_r \end{bmatrix} \quad (2)$$

where, u_s,u_r,i_s,i_r,ψ_s, and ψ_r are the stator voltage, rotor voltage, stator current, rotor current, stator flux, and rotor flux, which are converted to the stator side, respectively. L_s, L_r, and L_m are the stator inductance, rotor inductance, and mutual inductance, respectively. $L_{s\sigma}$ and $L_{r\sigma}$ are the stator leakage and rotor leakage, R_s and R_r are the stator and rotor resistance, ω_s is the synchronous speed, and ω_{s-r} is the slip speed.

When a three-phase metallic short-circuit fault occurs in the power grid, assuming that the line reactance of a DFIG to the short-circuit point is X_e. According to Equations (1, 2), the fault DFIG equivalent circuit under continuous excitation can be obtained as shown in **Figure 2**.

Eliminating the rotor current in Equation (2), the stator flux linkage ψ_s can be obtained in Equation (3):

$$\psi_s = L_s i_s + L_m i_r = \frac{L_m}{L_r}\psi_r + (\frac{L_s L_r - L_m^2}{L_r}) i_s \quad (3)$$

where the terminal voltage is $u_s = i_s X_e$, then the stator flux linkage ψ_s in Equation (3) is put into Equation (1):

$$i_s X_e = R_s i_s + j\omega_s \frac{L_m}{L_r}\psi_r + j\omega_s \left(\frac{L_s L_r - L_m^2}{L_r}\right) i_s + \frac{d\psi_s}{dt} \quad (4)$$

In the transient process of a DFIG, a component of the short-circuit current generated by $d\psi_s/dt$ (the transient of the stator flux linkage) is a DC attenuation component. Therefore, when analyzing the fundamental frequency component of the short-circuit current, which is caused by continuous excitation, the influence of stator flux linkage transients can be ignored (Sulla

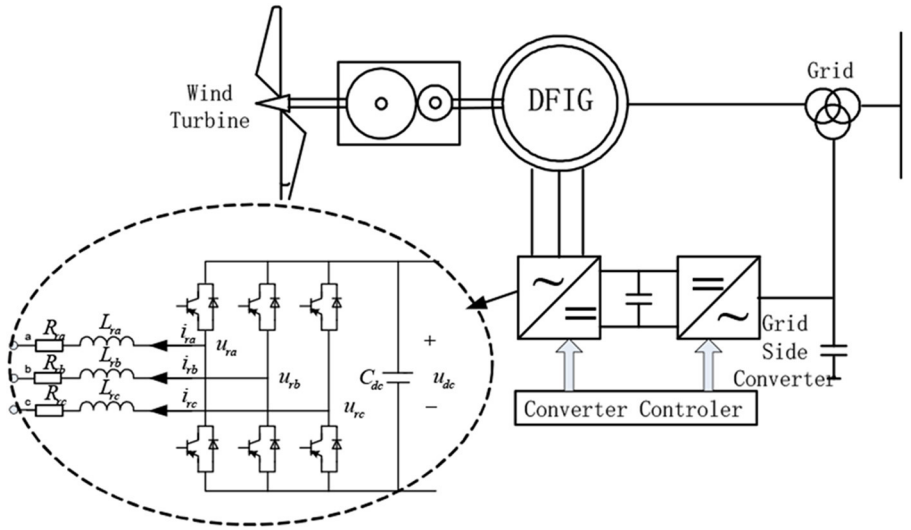


FIGURE 1 | Circuit structure of a doubly fed induction generator (DFIG) wind generator.

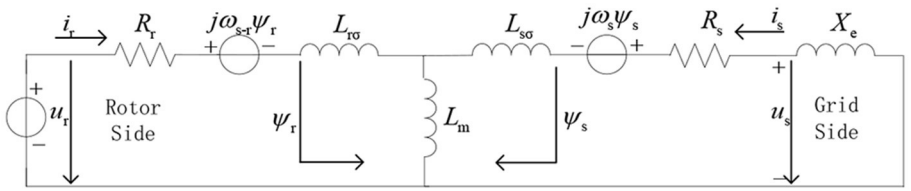


FIGURE 2 | Equivalent circuit of a DFIG after the fault.

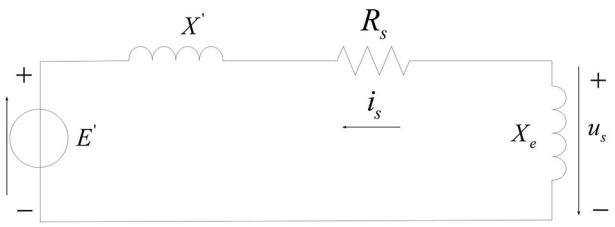


FIGURE 3 | Simplified equivalent circuit of DFIG.

et al., 2011; Swain and Ray, 2017). Assuming that the equivalent potential of a DFIG is $E' = j\omega_s L_m L_r^{-1}\psi_r - u_0 e^{-\frac{t}{\tau_s}}$, and transient reactance is $L_s' = L_s - L_m L_r^{-1} L_m$, $X' = j\omega_s L_s'$. A simplified diagram of the transient equivalent circuit of a DFIG after the fault can be obtained, as shown in **Figure 3**.

Simplifying Equation (4), the short-circuit current of a DFIG can be obtained during the fault:

$$i_s = \frac{E'}{-(R_s + X' + X_e)} \tag{5}$$

From Equation (5), it can be seen that the short-circuit current of a DFIG can be obtained from E', R_s, X', and X_e, where R_s, X', and X_e are the known quantities. Since the equivalent potential of a DFIG is $E' = j\omega_s L_m L_r^{-1}\psi_r$, the short-circuit current is the same as the change law of the rotor flux linkage ψ_r after the fault. Then, the short-circuit current attenuates according to the rotor time constant τ_r after the fault. Among them, τ_r is the rotor time constant $\tau_r = (L_r L_s - L_m^2)/R_r L_s$. From the above analysis, it can be seen that in order to calculate the short-circuit current during the entire fault period, the initial value and the steady-state value of the short-circuit current should be obtained firstly.

CALCULATION METHOD OF DFIG SHORT-CIRCUIT CURRENT AT THE INITIAL FAULT TIME

The flux linkage is non-abrupt before and after the fault, so it can be seen that the stator flux $d\psi_s/dt = 0$ at the initial time. Simplifying Equation (4), and the short-circuit current at the initial moment can be obtained:

$$i_{s0}' = \frac{j\omega_s \frac{L_m}{L_r}\psi_{r0}}{-(R_s + X' + X_e)} \tag{6}$$

As it can be seen from Equation (6), the initial short-circuit current of a DFIG can be obtained by R_s, ψ_{r0}, X_e, and X', among them, R_s, X', and X_e are the known quantities.

Because the rotor flux is also non-abrupt at the moment of failure, the initial value of the rotor flux ψ_{r0} can be obtained from the pre-fault working conditions. Before the fault occurs, the output active and the reactive power of a DFIG are:

$$\begin{cases} P_0 = \frac{3}{2}\left(u_{sq}i_{sq} + u_{sd}i_{sd}\right) = \frac{3}{2}u_{sq}i_{sq} = \frac{3}{2}u_s i_{sq} \\ Q_0 = \frac{3}{2}\left(u_{sq}i_{sd} - u_{sd}i_{sq}\right) = \frac{3}{2}u_{sq}i_{sd} = \frac{3}{2}u_s i_{sd} \end{cases} \tag{7}$$

In Equation (7), i_{sq} and i_{sd} are the active and reactive components of the stator current before the fault; P_0 and Q_0 are the output active and reactive power of a DFIG before the fault; u_{sq} is the q-axis component of the stator voltage. As the rotor-side converter of a DFIG adopts the stator flux orientation, namely: $\psi_{sd} = \psi_s = \frac{u_s}{j\omega_s}$, $\psi_{sd} = 0$, and Equation (8) can be obtained:

$$\begin{cases} \psi_{sd} = L_s i_{sd} + L_m i_{rd} = \psi_s = \frac{u_s}{j\omega_s} \\ \psi_{sq} = L_s i_{sq} + L_m i_{rq} = 0 \\ \psi_{rd} = L_m i_{sd} + L_r i_{rd} \\ \psi_{rq} = L_m i_{sq} + L_r i_{rq} \end{cases} \tag{8}$$

Combining Equations (7, 8), the rotor current can be eliminated. The rotor flux at the initial moment can be obtained by the output active and reactive power of a DFIG and the voltage before the fault.

$$\begin{cases} \psi_{rd0} = \frac{2Q_0}{3u_s}\left(\frac{L_m^2 - L_s L_r}{L_m}\right) + \frac{L_r}{L_m}\frac{u_{s0}}{j\omega_s} \\ \psi_{rq0} = \frac{2P_0}{3u_s}\left(\frac{L_m^2 - L_s L_r}{L_m}\right) \\ \psi_{r0} = \psi_{rd0} + j\psi_{rq0} \end{cases} \tag{9}$$

In Equation (9), ψ_{rq0} and ψ_{rd0} are the *q-axis* and *d*-axis components of the rotor flux before the fault. According to Equation (9), the rotor flux can be obtained by P_0, Q_0, u_{s0} at the initial moment of fault. Among them, the active and reactive power are determined by the working conditions before the failure. And the stator voltage is close to the rated value before the fault. Therefore, the short-circuit current at the initial fault time can be obtained in Equation (10):

$$i_{s0'} = \frac{j\omega_s \frac{L_m}{L_r}\left[\frac{2Q_0}{3u_s}\left(\frac{L_m^2 - L_s L_r}{L_m}\right) + \frac{L_r}{L_m}\frac{u_{s0}}{j\omega_s} + j\frac{2P_0}{3u_s}\left(\frac{L_m^2 - L_s L_r}{L_m}\right)\right]}{-(R_s + X' + X_e)} \tag{10}$$

CALCULATION METHOD OF DFIG SHORT-CIRCUIT CURRENT AT THE STEADY STATE

Based on China's grid-connected standard, the rotor converter needs to adjust the rotor current according to voltage drop degree. Therefore, the rotor excitation current will change before and after the fault, which will affect the output characteristics of short-circuit current.

In previous research, it was considered that the input was the Crowbar of a DFIG and the converter was locked. And the short-circuit current of a DFIG would reduce to zero. Under the new grid-connected standard, the rotor-side converter will provide continuous excitation current during the fault. In this situation, the short-circuit current of a DFIG will not reduce to zero, but the continuous short-circuit current will be output.

If there is a fault at the power grid, the voltage drop of the machine will be detected by a DFIG. And the rotor converter will adjust the excitation current, according to the requirement of the new grid standard. Then, DFIG will output the reactive current, which can provide support for system voltage. Through a dynamic process, the rotor excitation current will reach the reference value under the new grid-connected standard.

When the dynamic process is over, the fault steady state will be reached, in which $d\psi_s/dt = 0$. Simplifying Equation (1), the following equation can be obtained:

$$u_s = R_s i_s + j\omega_s \psi_s \tag{11}$$

When the fault steady state is reached, the excitation current of the rotor converter i_r attains $i_{r\infty}$ (the reference value of the excitation current in the new grid-connected standard), through the dynamic process. And at the moment, the stator current i_s also has reached $i_{s\infty}$ (the steady-state short-circuit current). Combining Equations (1, 11), the following equation can be obtained:

$$u_s = R_s i_{s\infty} + j\omega_s(L_s i_{s\infty} + L_m i_{rd\_ref} + jL_m i_{rq\_ref}) \tag{12}$$

In Equation (12), $i_{r\infty}$ is the rotor excitation current at the steady state, $i_{s\infty} = i_{rd,ref} + ji_{rq,ref}$, ($i_{rq\_refand}$ i_{rd_ref} are the reference value of active and reactive current of the rotor, respectively).

China has issued the new wind power grid-connected standard GB/T19963-2011 "Technical Regulations for Wind Farm Access to Power System." In the standard, the output

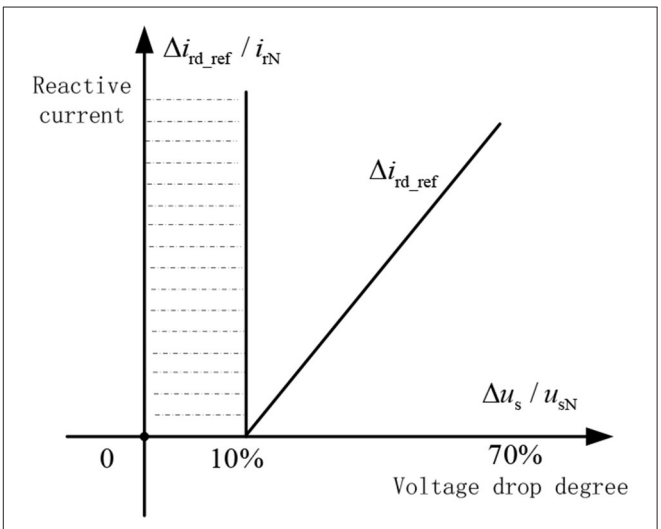


FIGURE 4 | Relationship between the deviation of u_s and rotor reactive current.

reactive power during the fault is demanded as shown in Equation (13):

$$\begin{cases} i_{rd_r ef} = f(u_s) = i_{rd0} + K_d(0.9 - u_s)\, i_{rN},\ (K_d \geq 1.5) \\ i_{rq_r ef} = -\frac{2}{3}\frac{L_s P_0}{L_m u_{s0}},\ (0 \leq i_{rq_r ef} \leq \sqrt{i_{\max}^2 - i_{rd_r ef}^2}) \end{cases} \tag{13}$$

In Equation (13), K_{d} is the gain coefficient of reactive current, i_{rmax} and i_{rN} are the maximum current and rated current of a rotor, respectively.

It can be seen from Equation (13) that the rotor converter gives priority to adjust the reactive current reference value $i_{\mathrm{rd\_ref}}$, during the fault. Its proportion is determined by the degree of voltage drop and the gain coefficient K_{d}. And the active current is determined by the output power and voltage before the rotor current surpasses the maximum current. **Figure 4** shows the relationship between the reference value of reactive current and the degree of voltage drop during the fault.

Before the fault, the power factor of a DFIG is 1, $i_{\mathrm{rd0}} = 0$. Using Equations (12, 13), the steady-state short-circuit current of a DFIG under continuous excitation can be obtained:

$$i_{s\infty} = \frac{j\frac{2X_m L_s P_0}{3L_m u_{s0}} - X_m K_d 0.9 i_{rN}}{R_s + X + X_e + X_m K_d X_e i_{rN}} \tag{14}$$

In Equation (14), X is the steady-state stator reactance.

EXPERIMENTAL RESULTS AND ANALYSIS

Based on RTDS, a digital-to-analog experimental platform for a DFIG with actual converter control units was established. Among them, the control units of the rotor converter are designed by using the field-programmable gate array (FPGA) chips, which are the controller core, the DFIG, and an RTDS-built grid model. The data transmission is realized through a parallel communication interface, which realizes the real-time control of the rotor converter.

In **Figure 5**, an actual DFIG wind farm connected to a grid was considered as an example to verify the experiment. The DFIG is connected to a 35-kV collector line through the transformer T2 at a terminal and is then connected to a 220-kV power system through the wind farm main transformer T1. There are 15 sets of 2.0-MW DFIG wind turbine generators in the DFIG wind farm. The stator and rotor resistances are 0.015 and 0.0099 p.u., respectively, and the stator and rotor leakage inductances are 0.168 and 0.152 p.u., respectively. The excitation mutual inductance is 3.49 p.u. The equivalent impedance of lines AB, BC, and DE is $(1.95 + j5.53)\ \Omega$, $(1.46 + j4.16)\ \Omega$, $(0.13 + j0.11)\ \Omega$, and the system equivalent impedance is $j0.5\ \Omega$. Because the wind farm adopts the same type of the DFIG wind turbine in this paper, and the distance between the DFIGs is short, the transient characteristics of DFIGs are almost the same when they are faulty. The DFIGs on each collection line can be used instead by a single DFIG with equal capacity.

Assuming that a DFIG is working under rated operating conditions before the fault, a three-phase metal short-circuit fault occurs at the terminal B of the AB line at $t = 0.5$ s, and the fault lasting for 0.2 s. Firstly, the rotor excitation current of a DFIG under the new grid-connected standard has been analyzed. **Figure 6A** shows a comparison between the measured rotor current and reference rotor current when a three-phase metal short-circuit fault occurs at the terminal B.

It can be seen from **Figure 6A**, after the fault occurs, the rotor excitation current reaches the reference value under the new grid-connected standard through the dynamic process, which is consistent with the result of an analysis in section Calculation Method of DFIG Short-Circuit Current at the Initial Fault Time. **Figure 6B** shows the experimental result of the short-circuit current of a DFIG, when a fault occurs at the terminal B. Due to the influence of continuous excitation current during the fault, a DFIG will output the steady-state short-circuit current, which is a big difference compared with the traditional DFIG. Using the full-cycle Fourier algorithm to extract the root mean square (RMS) of the short-circuit current, and the extracted RMS is shown in **Figure 6C**. In addition, the RMS of the short-circuit current is calculated by using the method proposed in this paper, which is also shown in **Figure 6C**.

It can be seen from **Figure 6C**, at the initial time of fault ($t = 0.5$ s), the RMS of the short-circuit current of a DFIG suddenly increases to 3.01 p.u., the calculation result obtained by using the proposed method is 3.07 p.u., and the error from the experimental result is 2.01%. After the fault reaches steady state, the experimental result is 1.94 p.u. The calculation result obtained by using the proposed method is 1.91 p.u., and the error from the experimental results is 1.6%. The similarity between the calculation results and the experimental results is extremely high during the attenuation period, and the extreme results fluctuate above and below the calculation result. From the above analysis, the proposed method not only accurately calculates the initial value and the steady-state value of the short-circuit current but also accurately describes the variation law of the short-circuit current in the attenuation process.

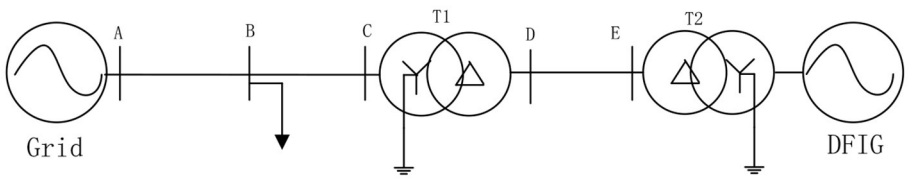


FIGURE 5 | The grid structure with DFIG wind farms.

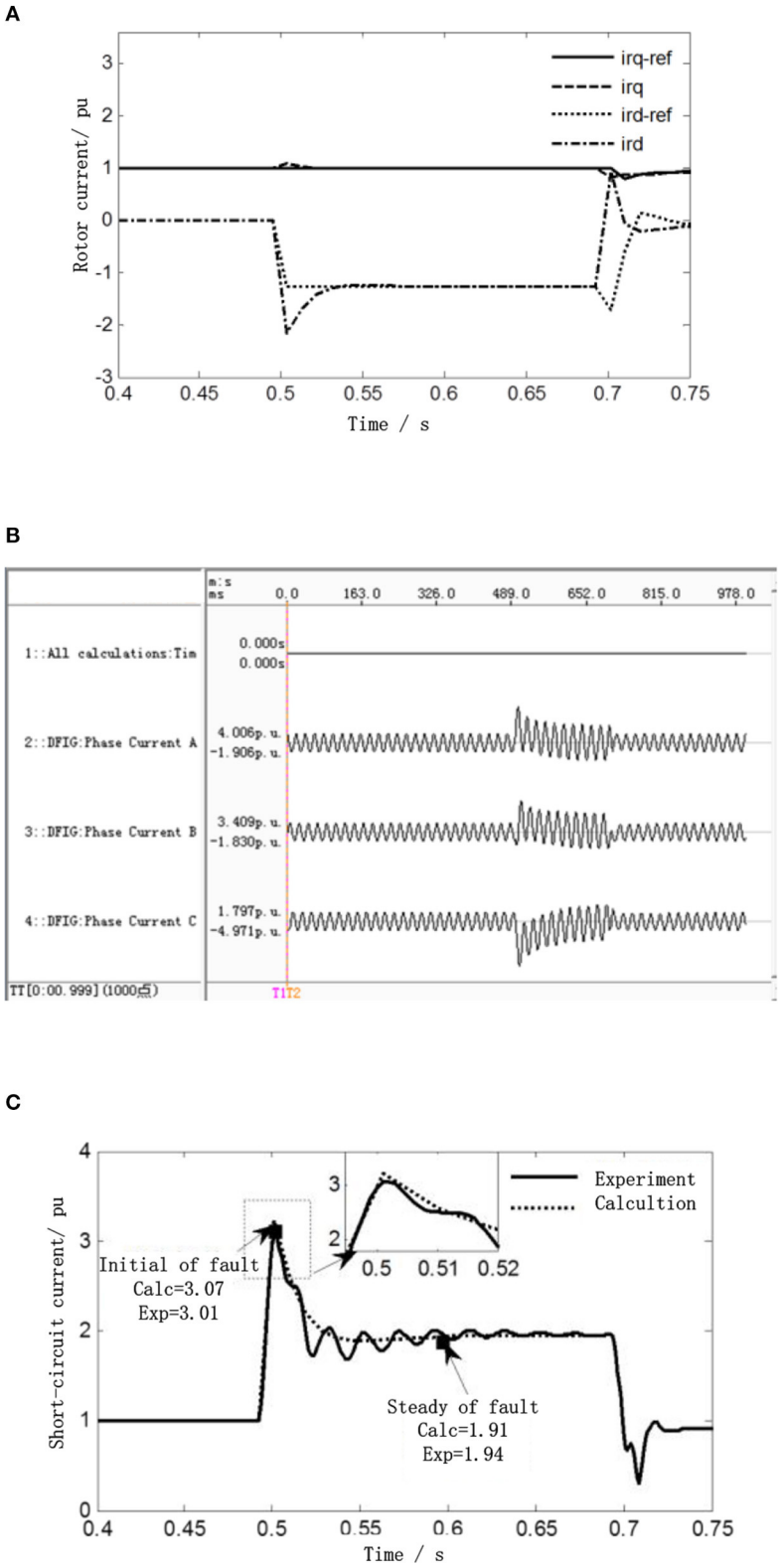


FIGURE 6 | Comparison figure between the calculation result and experimental result of DFIG three-phase short circuit at bus B. **(A)** Comparison of the rotor current control reference value and actual rotor value. **(B)** Experimental result of short-circuit current of three-phase short-circuit DFIG at bus B under rated conditions. **(C)** Comparison of the calculation result and experimental result of root mean square (RMS) three-phase short circuit at bus B.

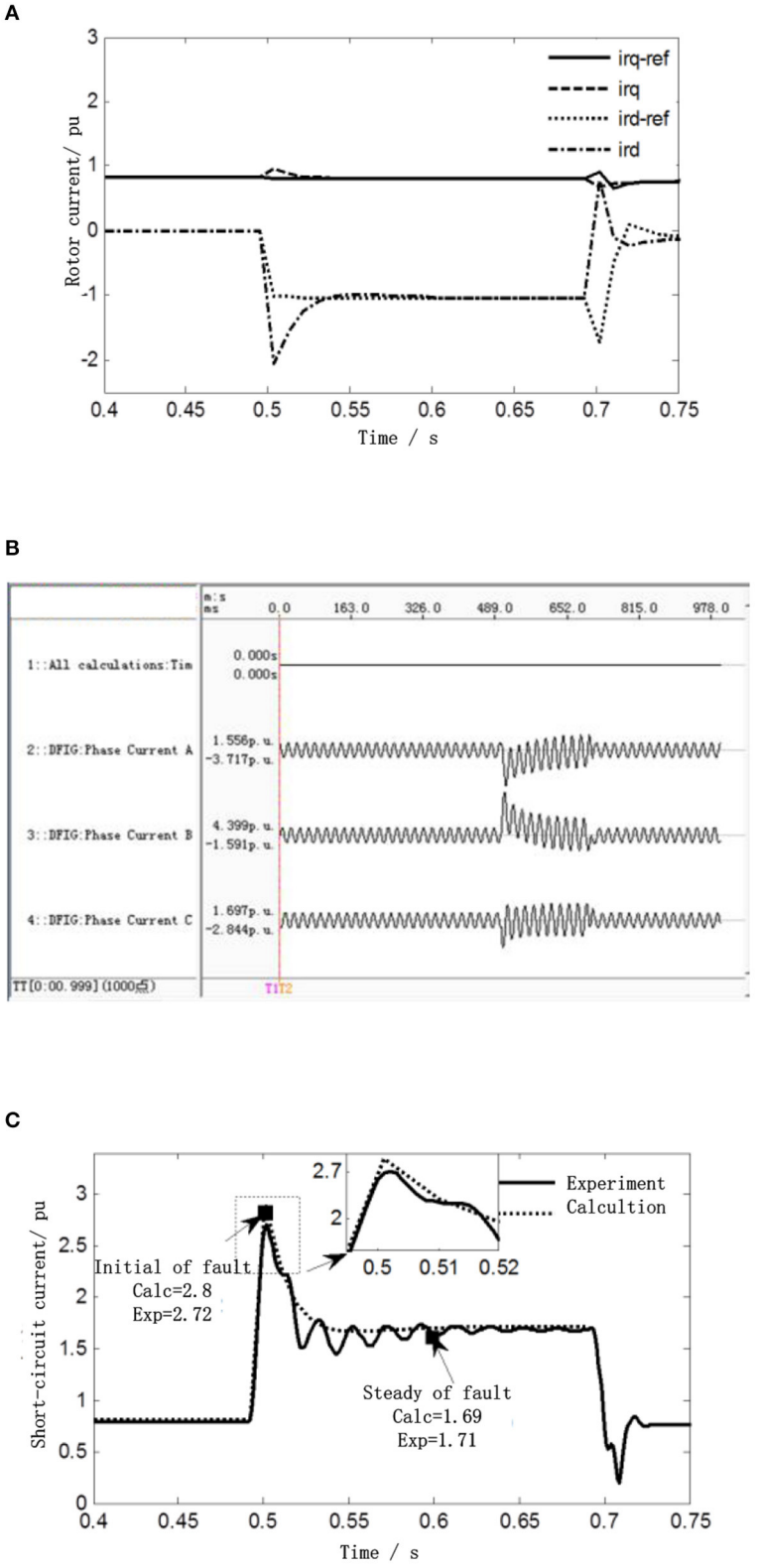


FIGURE 7 | Comparison figure between the calculation result and experimental result of DFIG three-phase short circuit at bus A. **(A)** Comparison of the rotor current control reference value and actual rotor value. **(B)** Experimental result of short-circuit current of three-phase short-circuit DFIG at bus A under 0.8 rated conditions. **(C)** Comparison of the calculation result and experimental result of RMS three-phase short circuit at bus A.

Assuming that a DFIG is working under 0.8 times of rated operating conditions before the fault, a three-phase metal short-circuit fault occurs at the terminal A of the AB line at $t = 0.5$ s, and the fault lasts for 0.2 s. Firstly, the rotor excitation current of a DFIG under the new grid-connected standard has been analyzed. **Figure 7A** shows a comparison between the measured rotor currents and reference rotor currents, when a three-phase metal short-circuit fault occurs at the terminal A.

Figure 7B shows the experimental result of the short-circuit current of a DFIG, when a fault occurs at the terminal A. Using the full-cycle Fourier algorithm to extract the RMS of the short-circuit current, and the extracted RMS is shown in **Figure 7C**. In addition, the RMS of the short-circuit current is calculated by using the method proposed in this paper, which is also shown in **Figure 7C**.

According to **Figure 7C**, at the initial time of fault ($t = 0.5$ s), the RMS of the short-circuit current of a DFIG suddenly increases to 2.72 p.u. The calculation result obtained by using the proposed method is 2.8 p.u., and the error from the experimental result is 2.9%. After the fault reaches steady state, the experimental result is 1.71 p.u. The calculation result obtained by using the proposed method is 1.69 p.u., and the error from the experimental results is 1.2%. From the above analysis, it can be seen that the proposed method has high accuracy in the calculation of the short-circuit currents, under different fault locations and under different working conditions.

About 24 sets of experimental tests were carried out, respectively, for different working conditions (the output power before the fault is 1, 0.9, 0.8, and 0.7 p.u.) and different fault locations (at 20, 30, 40, 50, 60, and 70% of line AB). The experimental results and the calculation results obtained by using the proposed method are shown in **Figure 8**. At the initial time ($t = 0.5$ s), steady-state time ($t = 0.6$ s), $t = 0.52$ s, and $t = 0.55$ s in the dynamic process, the errors between the calculation results and experimental results were compared. From **Figure 8**, it can be seen that the errors between the calculation results and experimental results at the initial time are $< 3\%$, under different fault conditions. And the errors at steady-state time are $< 2\%$. In the whole attenuation process, the errors between the calculation results and experimental results do not exceed 6%, which can correctly describe the change law of the short-circuit current of a DFIG.

CONCLUSION

In previous research, the problem of continuous excitation was not considered. To solve the problem, a DFIG transient equivalent model under the condition of continuous excitation is established. According to a non-abrupt change of flux linkage and the China's new grid standard, the short-circuit current calculation method of a DFIG under continuous excitation condition is put forward. The

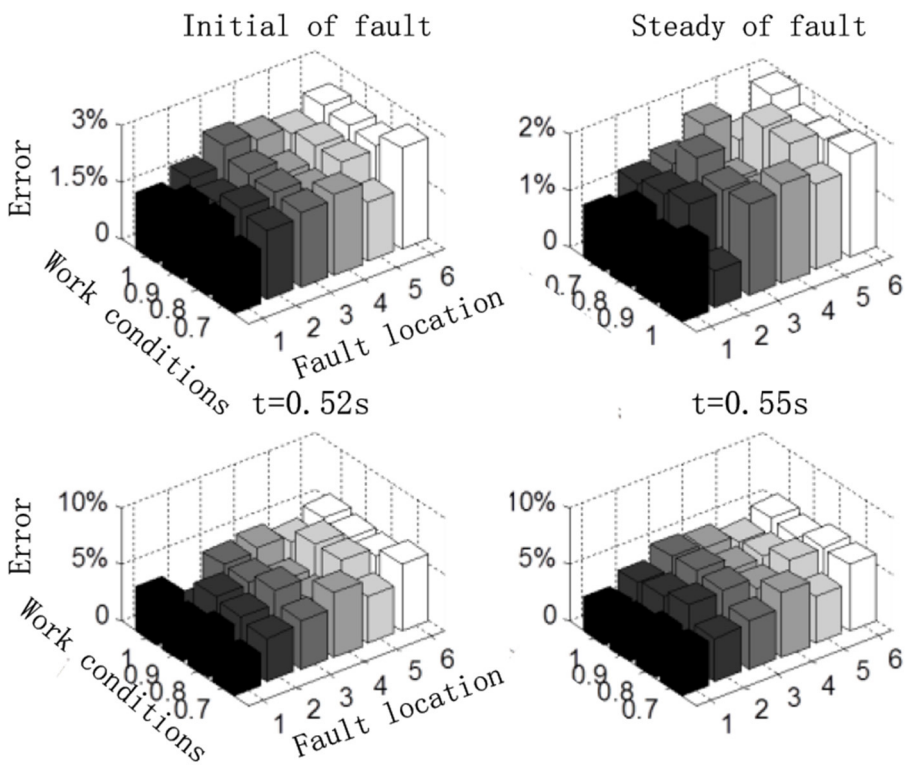


FIGURE 8 | Comparison between the calculation result and experimental result of the short-circuit current.

experimental results of an RTDS experimental platform show that:

(1) In this paper, the characters of the equivalent potential of a DFIG are analyzed during the fault period. And the DFIG transient equivalent model under the constant excitation is established, which describes the transient process of a DFIG accurately.
(2) The requirements of reactive power in the China's new grid-connected standard are analyzed, and the short-circuit current calculation method of a DFIG under the condition of continuous excitation is put forward. The proposed method not only accurately calculates the initial value and steady-state value of short-circuit current, but also accurately describes the change law of short-circuit current, which is verified by the experimental result.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

JY: data curation, writing, and software.

FUNDING

This work was supported by the key scientific research projects of colleges and universities in Henan Province (19A470003).

REFERENCES

Duan, J., Chen, T., Shang, D., Cui, S., and He, Y. (2020). Current attenuation factor based line protection scheme for distribution network of DFIG wind power integration system. *Proc. CSEE* 40, 1915–1924. doi: 10.13334/j.0258-8013.pcsee.X190386

Firouzi, M., and Gharehpetian, G. (2017). LVRT performance enhancement of DFIG-based wind farms by capacitive bridge-type fault current limiter. *IEEE Trans. Sustain. Energy* 9, 1118–1125. doi: 10.1109/TSTE.2017.2771321

Garibay, A., Rodriguez, J., and Correa, J. (2018). Analysis of rotor current impact on DFIG-WECS under fault condition. *IEEE Latin Am. Trans*. 16, 329–334. doi: 10.1109/TLA.2018.8327383

Haj-Ahmed, M. A., Feilat, E. A., Khasawneh, H. J., Abdelhadi, A. F., and Awwad, A. (2018). Comprehensive protection schemes for different types of wind generators. *IEEE Trans. Ind. Appl*. 54, 2051–2058. doi: 10.1109/TIA.2018.2789865

Liu, T., Zhang, Y., Wang, M., Zhang, Y., Qin, X., and Wang, Y. (2020). Short-circuit current calculation method of doubly fed induction generator considering control strategy switching. *Proc. CSEE* 40, 17–25. doi: 10.13334/j.0258-8013.pcsee.200890

Muljadi, E., Samaan, N., Gevorgian, V., Li, J., and Pasupulati, S. (2013). Different factors affecting short circuit behavior of a wind power plant. *IEEE Trans. Ind. Appl*. 49, 284–292. doi: 10.1109/TIA.2012.2228831

National Electricity Regulatory Standardization Technical Committee. (2011). *GBT_19963-2011 Technical Rule for Connecting Wind Farm to Power System*. Beijing: Standard Press of China.

Ouyang, J., Tang, T., Zheng, D., Ren, W., Xiong, X., and Zhong, J. (2017). Characteristics and calculation method of short-circuit current of doubly fed wind generator under lower voltage ride through. *Trans. China Electrotech. Soc.* 32, 216–224. doi: 10.19595/j.cnki.1000-6753.tces.160823

Pan, W., Liu, M., Zhao, K., Zhang, Y., and Liu, T. (2019). A practical short-circuit current calculation method for DFIG-based wind farm considering voltage distribution. *IEEE Access* 7, 31774–31781. doi: 10.1109/ACCESS.2019.2902848

Rahimi, M., and Azizi, A. (2019). Transient behavior representation, contribution to fault current assessment, and transient response improvement in DFIG based wind turbines assisted with crowbar hardware. *Int. Trans. Electr. Energy Syst.* 29:e2698. doi: 10.1002/etep.2698

Sulla, F., Svensson, J., and Samuelsson, O. (2011). Symmetrical and unsymmetrical short-circuit current of squirrel cage and doubly fed induction generators. *Electr. Power Syst. Res*. 8, 1610–1618. doi: 10.1016/j.epsr.2011.03.016

Swain, S., and Ray, P. K. (2017). Short circuit fault analysis in a grid connected DFIG based wind energy system with active crowbar protection circuit for ridethrough capability and power quality improvement. *Int. J. Electr. Power Energy Syst*. 84, 64–75. doi: 10.1016/j.ijepes.2016.05.006

Tamaarat, A., and Benakcha, A. (2014). Performance of PI controller for control of active and reactive power in DFIG operating in a grid-connected variable speed wind energy conversion system. *Front. Energy* 8, 371–378. doi: 10.1007/s11708-014-0318-6

Telukunta, V., Pradhan, J., Agrawal, A., Singh, M., and Srivani, S. G. (2018). Protection challenges under bulk penetration of renewable energy resources in power systems: a review. *CSEE J. Power Energy Syst*. 3, 365–379. doi: 10.17775/CSEEJPES.2017.00030

Wang, Z., Liu, Y., Lei, M., Bian, S., and Shi, Y. (2015). Doubly-fed induction generator wind farm aggregated model based on crowbar and integration simulation analysis. *Trans. China Electrotech. Soc.* 30, 45–52. doi: 10.19595/j.cnki.1000-6753.tces.2015.04.006

Ying, Y., Sun, Y., Yang, J., Hong, M. (2020). Load characteristic analysis of grid fault ride-through process for DFIG based large wind turbine. *Automat. Electr. Power Syst*. 44, 127–138. doi: 10.7500/AEPS20190911006

Conflict of Interest: The 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.

frontiers in Energy Research

BRIEF RESEARCH REPORT
published: 28 May 2021
doi: 10.3389/fenrg.2021.686146

Research on Short-Circuit Current Calculation Method of Doubly-Fed Wind Turbines Considering Rotor Dynamic Process

*Jun Yin**

School of Electric Power, North China University of Water Resources and Electric Power, Zhengzhou, China

With the enlarging scale of doubly-fed induction generators (DFIGs) connected to power systems, it is important to analyze the influence of a short-circuit current to system relay protection. Due to the correct evaluation of the protection operation characteristics, the DFIG short-circuit current needs to be calculated accurately. But the current research on the short-circuit current of DFIG is based on the following assumption: the rotor excitation current is zero after the rotor crowbar is put, and the influence of its dynamic process is ignored. This will bring errors to the calculation results. This paper takes into account the influence of rotor current dynamics by studying the mechanism of the potential transient change of DFIG. The stator rotor flux linkage of DFIG in the event of a three-phase short-circuit is accurately calculated, and an improved RMS calculation method of doubly-fed wind turbine short circuit current is proposed. A physical experiment platform with an actual controller of a doubly-fed fan is established, based on RTDS. It can be seen from the experiment that the short-circuit current calculation method proposed in this paper is more accurate than those methods that ignore the rotor dynamic process. This study lays a foundation for further study of the influence of DFIG on the protection operation characteristics.

Keywords: DFIG, short circuit current, fault analysis, relay protection, rotor dynamic process

OPEN ACCESS

Edited by:
Liang Chen,
Nanjing University of Information Science and Technology, China

Reviewed by:
You Zhou,
Changsha University of Science and Technology, China
Jianhui Meng,
North China Electric Power University, China

***Correspondence:**
Jun Yin
yinjun0371@126.com

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

Received: *26 March 2021*
Accepted: *21 April 2021*
Published: *28 May 2021*

Citation:
Yin J (2021) Research on Short-Circuit Current Calculation Method of Doubly-Fed Wind Turbines Considering Rotor Dynamic Process. Front. Energy Res. 9:686146. doi: 10.3389/fenrg.2021.686146

INTRODUCTION

In recent years, the wind power in the world has developed rapidly. Since 2010, China has become the country with the largest installed capacity of wind power. Especially in Inner Mongolia, Gansu, and Liaoning Province, wind power has developed rapidly. However, large-scale wind power access to the power grid has made a great influence, and the impact of grid-connected wind power on relay protection has become a major concern in the current power system field.

DFIG are widely used in wind farms as their main turbines, because of the advantages of the wide range of operating wind speeds and the decoupled control of active and reactive power (Tamaarat and Benakcha, 2014; Firouzi and Gharehpetian, 2017; Sun and Wang, 2018; Okedu, and Barghash, 2020). However, with the increase of DFIG, the influence of the transient process of DFIG on the short-circuit current calculation cannot be ignored. Inaccurate calculation of the short-circuit current will cause deviation in the protection setting, affect the accuracy of the protection, and even cause the phenomenon of protection rejection or mis-operation in severe cases. Therefore, in order to improve the accuracy of the protection action, it is necessary to study the methods of short-circuit current calculation of DFIG thoroughly.

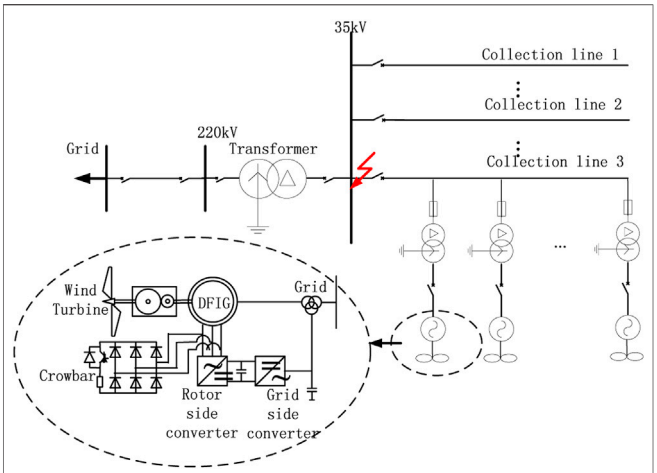


FIGURE 1 | Circuit structure of DFIG wind farm.

At present, there are some works of literature studying the short-circuit current calculation after the DFIG connected to the grid on different perspectives. In the literature (Howard, et al., 2012; Muljadi et al., 2013; Liu et al., 2018; Telukunta et al., 2018), in the no-load case, when the three-phase metallic short-circuit occurred, the crowbar was input. DFIG was equivalent to an asynchronous generator, and the rotor excitation current was assumed to be zero. A simplified short-circuit current calculation formula was presented in this paper. The method in this paper simply assumes that the rotor current after the Crowbar is accessed is zero, ignoring the dynamic process of the rotor current, and this method does not match the actual rotor current after a short-circuit fault.

In the literature (Sulla, et al., 2011; Zhai, et al., 2013; Wang et al., 2015; Fu et al., 2017; Rahimi and Azizi, 2019), the flux linkage of the stator was solved simply after the faults at different positions, and in this paper the calculation formula of the DFIG short-circuit current was given under different drop degrees of the terminal voltage, further. However, this method does not take into account rotor current dynamic process effects. Actually, after the fault of the power grid, the crowbar will be input to protect the rotor converter. However, the rotor flux cannot be mutated at the moment of the fault, as the rotor winding will induce a big rotor current. The rotor current may reach three to five times that of the rated value, and decay to zero gradually after 30–50 milliseconds. Ignoring the dynamic process of the rotor current will cause a certain error in the calculation result of the short circuit current, and it will affect the accuracy of the current protection of the collector line. Therefore, this paper presents an accurate calculation method of a DFIG short-circuit current, which accounts for the dynamic process of the rotor current.

In this paper, from the perspective of the mechanism of the DFIG's transient potential, the stator flux linkage of the DFIG in a three-phase short circuit is calculated. The influence of rotor current dynamics is taken into account, and an improved method of the short-circuit current of the DFIG is proposed. In order to verify the correctness of the calculation method, a physical experiment platform with an actual controller of the DFIG current transformer is established based on the RTDS. The experimental results show that the short-circuit current calculation method proposed in this paper is more accurate, compared with the methods which ignore the dynamic process of the rotor. This study lays a foundation for further research on the influence of the DFIG's short-circuit current on the protection operation characteristics.

DFIG ELECTROMAGNETIC TRANSIENT PROCESS

Figure 1 shows the main circuit topology of a grid-connected DFIG with a rotor Crowbar circuit. The DFIG is connected to the power grid by the collector line. The main protection on the collector line is a three-section current protection. For accurate evaluation of the current protection action, the key is accurately calculating the short-circuit current, which is provided to the collector wire by the DFIG, after the fault occurs.

Because the fault of the power grid will cause the electromagnetic transience of the DFIG, the electromagnetic transient process will cause the DFIG to output a big short-circuit current. Therefore, in order to calculate the short-circuit current accurately, the electromagnetic transient process of DFIG after the fault should be analyzed first. That means the relationship among the short-circuit current, the transient reactance, and the equivalent potential of DFIG should be obtained.

Ignoring the phenomenon of magnetic saturation, assuming the rotor speed does not change in transient process.

The mathematical model of the DFIG in the form of space vectors in the synchronous rotating coordinate system is (Muljadi et al., 2013):

$$\begin{bmatrix} u_s \\ u_r \end{bmatrix} = \begin{bmatrix} R_s & 0 \\ 0 & R_r \end{bmatrix} \begin{bmatrix} i_s \\ i_r \end{bmatrix} + \begin{bmatrix} \frac{d\psi_s}{dt} \\ \frac{d\psi_r}{dt} \end{bmatrix} + \begin{bmatrix} j\omega_s & 0 \\ 0 & j\omega_{s-r} \end{bmatrix} \begin{bmatrix} \psi_s \\ \psi_r \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} \psi_s \\ \psi_r \end{bmatrix} = \begin{bmatrix} L_s & L_m \\ L_m & L_r \end{bmatrix} \begin{bmatrix} i_s \\ i_r \end{bmatrix} \quad (2)$$

Where, u_s, u_r, i_s, i_r, ψ_s, ψ_r are the stator voltage, rotor voltage, current and flux, which are converted to the stator side, respectively. L_s, L_r, L_m are the stator and rotor inductance, mutual inductance, respectively. $L_{s\sigma}$, $L_{r\sigma}$ are the stator leakage, rotor leakage, R_s, R_r, R_{cb} are the rotor resistance and the rotor Crowbar resistance, ω_s is the synchronous speed, and $\omega_{s\text{-}r}$ is the slip speed.

During normal operation, the DFIG is excited by a rotor converter. When the power grid fails, the voltage of the DFIG drops suddenly, and a large transient voltage and current are induced in the rotor windings. The rotor winding should input the Crowbar to suppress the transient current and protect the converter from damage.

When a three-phase metallic short-circuit fault occurs in the power grid, assuming that the line reactance of the DFIG to the short-circuit point is X_e, the terminal voltage is $u_s = i_s X_e$. According to **Eqs 1**, **2**, the fault of the DFIG equivalent circuit can be obtained as shown in **Figure 2**.

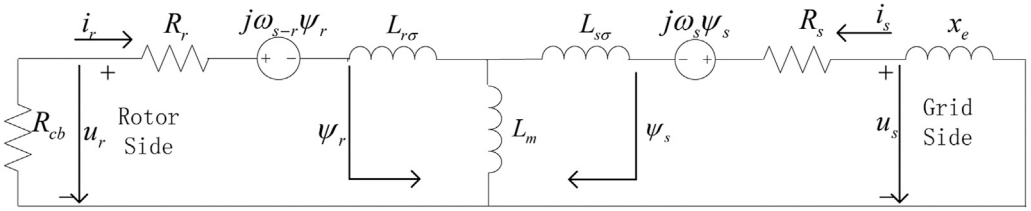


FIGURE 2 | Fault equivalent circuit of DFIG wind power generator.

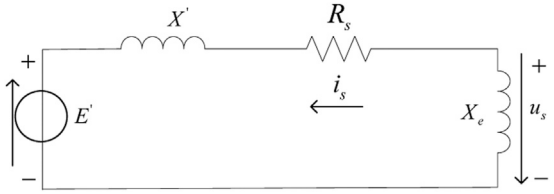


FIGURE 3 | Simplified equivalent circuit of DFIG.

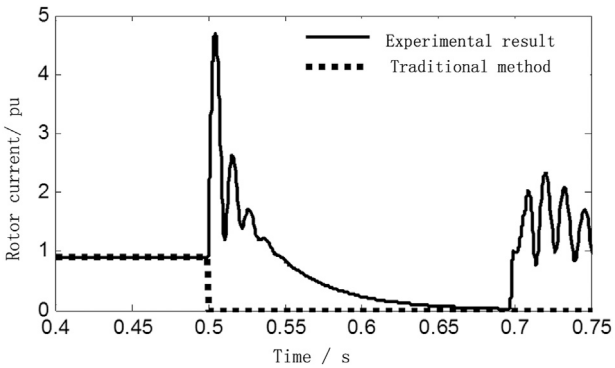


FIGURE 4 | The rotor current dynamic process of DFIG under three phase short circuit.

By eliminating the rotor current in **Eq. 2**, the stator flux linkage ψ_s can be obtained, then ψ_s is put into the **Eq. 1**:

$$u_s = R_s i_s + j\omega_s \frac{L_m}{L_r}\psi_r + j\omega_s\left(\frac{L_s L_r - L_m^2}{L_r}\right)i_s + \frac{d\psi_s}{dt} \tag{3}$$

Assuming that the DFIG's equivalent potential is $E' = j\omega_s L_m L_r^{-1}\psi_r - u_0 e^{-t/\tau' s}$. Transient reactance is $L's = L_s - L_m L_r^{-1} L_m$, $X' = j\omega_s L's$.The stator decay time constant is $\tau_s = R_s/L's$. A simplified diagram of the transient equivalent circuit of DFIG after the fault can be obtained, as shown in **Figure 3**. Stator flux linkage is $d\psi_s/dt = u_0 e^{-t/\tau_s}$, which is the attenuated DC flux linkage component. When calculating the short-circuit current AC component, it can be ignored (Swain and Ray, 2017; Ma and Liu, 2018). Simplify **Eq. 3** to obtain the relationship among the short-circuit current, transient reactance and equivalent potential.

$$i_s = \frac{E'}{-(R_s + X' + X_e)} \tag{4}$$

From the above analysis, the short-circuit current of the DFIG is determined by E', R_s, X', and X_e. Among them, R_s, X', and X_e are known quantities. Therefore, to calculate the short-circuit current of the DFIG accurately, it is required to solve the equivalent potential E' after the DFIG fault occurred.

CALCULATION METHOD OF DFIG SHORT-CIRCUIT CURRENT CONSIDERING THE DYNAMIC PROCESS OF THE ROTOR CURRENT

The equivalent potential of the DFIG is determined by its rotor flux linkage. Therefore, the key issue in studying the change law of the DFIG equivalent potential *E*', is how to accurately solve the rotor flux linkage during the fault.

Figure 4 shows the dynamic process of the DFIG rotor current after a three-phase short circuit. When the fault occurs, the rotor Crowbar is input and the rotor excitation circuit is shorted. The rotor current will increase to three to five times its rated value at first, and then decay to zero gradually after 30–50 milliseconds, which is similar to what is shown in **Figure 4**. However, compared with the traditional short-circuit calculation methods, the traditional methods ignore the rotor current dynamic, and assume the rotor current decays to zero directly after the fault occurs. Although the traditional method is convenient for calculation, it cannot accurately reflect the rotor current dynamic in the actual physical process and will cause certain errors in the calculation of the short-circuit current.

In this section, in order to ensure the accuracy of the rotor flux linkage solution, the dynamic process of the rotor current is taken into account, which was ignored before. And the equivalent potential is calculated. Finally, it is input to **Eq. 4**, and the accurate short-circuit current is obtained.

From **Eq. 2**, the stator rotor current is represented by flux linkage:

$$\begin{cases} i_s = \dfrac{L_r\psi_s - L_m\psi_r}{L_s L_r - L_m^2} = \dfrac{\psi_s}{L's} - \dfrac{L_m}{L_r}\dfrac{\psi_r}{L's} \\ i_r = \dfrac{-L_m\psi_r + L_r\psi_s}{L_s L_r - L_m^2} = -\dfrac{L_m}{L_s}\dfrac{\psi_s}{L'r} + \dfrac{\psi_r}{L'r} \end{cases} \tag{5}$$

Among them: $L'_s = L_r - L_m L_r^{-1} L_m$.

Bring **Eq. 5** into **Eq. 1**, to get detailed stator rotor flux linkage, which considering the rotor dynamics:

$$\frac{d}{dt}\begin{bmatrix}\psi_s\\ \psi_r\end{bmatrix} = \begin{bmatrix}\frac{-R_s}{L'_s} + j\omega_s & \frac{R_s L_m}{L'_s L_r}\\ \frac{R_r L_m}{L'_r L_s} & \frac{-R_r}{L'_r} + j\omega_{s-r}\end{bmatrix}\begin{bmatrix}\psi_s\\ \psi_r\end{bmatrix} + \begin{bmatrix}u_s\\ u_r\end{bmatrix} \quad (6)$$

In the process of solving **Eq. 6**, u_s, u_r, L_s, L_r, L_m, R_s, R_r are all known quantities. Therefore, **Eq. 6** is a first-order ordinary differential equation group about the stator rotor flux linkage. The Laplace transform method is used to solve the **Eq. 6**:

$$\begin{cases} u_s = \left(\frac{L'_s}{R_s} + j\omega_s + s\right)\psi_{s(s)} - \psi_{s(0)} - \frac{1}{\tau_s}\frac{L_m}{L_r}\psi_{r(s)} \\ u_r = \left(\frac{L'_r}{R_r} + j\omega_{s-r} + s\right)\psi_{r(s)} - \psi_{r(0)} - \frac{1}{\tau_r}\frac{L_m}{L_s}\psi_{r(s)} \end{cases} \quad (7)$$

Among them, $\tau_r = R_r/L'_r$, $\psi_{s(s)}$, $\psi_{r(s)}$ are the initial value of the stator and rotor flux linkage respectively, $\psi_{s(s)}$, $\psi_{r(s)}$are the stator and rotor flux linkage, which are Laplace transformed.

When a three-phase metal fault occurs at the machine terminal of the DFIG, the rotor Crowbar is input. Due to the short delay time of the converter control, if the delay time is ignored, that is u_r is 0 after the Crowbar is input, so it can be obtained from **Eq. 7** that:

$$\psi_{r(s)} = \frac{\left(\frac{1}{\tau_s} + j\omega_s + s\right)\psi_{r(0)} + \left(\frac{L_m}{\tau_r L_s}\right)\psi_{s(0)}}{(s+\alpha)(s+\beta)} \quad (8)$$

Among them: $\alpha = 1/\tau_s + j\omega_s - \eta$, $\beta = 1/\tau_s + j\omega_{s-r} + \eta$,
$\eta = L_m^2/L_s L_r/\tau_s\tau_r\left(\tau_r^{-1} - \tau_s^{-1} + j\omega_s + j\omega_r\right)$

.

Solve the time domain solution of **Eq. 8**:

$$\psi_{r(t)} = Ae^{\frac{-t}{\tau_s}}e^{-j\delta t} + \left[\psi_{r(0)} - A\right]e^{\frac{-t}{\tau_s}}e^{-j(\omega_r-\delta)t} \quad (9)$$

Among them: $A = -(\delta + j\kappa)\psi_{r(0)} + j(L_m/L_s\tau_r)\psi_{s(0)}/(\omega_r - \delta) + j(1/\tau_r - 1/\tau_s)$, $\omega_r = \omega_s - \omega_{s-r}$ is the rotor speed, $\kappa = \mathrm{Re}(\eta)$, $\delta = \mathrm{Im}(\eta)$ are the real part and the imaginary part of η.

In the synchronous coordinate system, the two parts of the rotor flux linkage attenuated at frequencies close to the DC and rotor speed, respectively. The initial stator rotor flux $\psi_{s(0)}$, $\psi_{r(0)}$ can be obtained by the pre-fault voltage and current through the voltage **Eq. 1**.

From **Eq. 4**, the amplitude of the fundamental frequency AC component in the equivalent potential is:

$$E'_f = j\omega_s L_m L_r^{-1}\left[\psi_{r(0)} - A\right]e^{\frac{-t}{\tau_r}} \quad (10)$$

The fundamental frequency RMS calculation method for the short-circuit current of the DFIG is in **Eq. 11**:

$$i_{sf} = \frac{E'_f}{-(R_s + X' + X_e)} \quad (11)$$

The formula of the rotor flux linkage after faults in traditional studies is in **Eq. 12**:

$$\psi_{r(t)} = \psi_{r(0)}e^{\frac{-t}{\tau_r}}e^{j\omega_r t} \quad (12)$$

Comparing **Eq. 9** with **Eq. 12**, the traditional studies used a simplified calculation method, which ignored the rotor current I_r and the stator resistance R_s. Although the obtained rotor flux linkage also includes the rotor speed attenuation part in the traditional method, the proposed method introduces η and A, which are constants related to the system parameters, to modify the calculated value of the rotor flux linkage in this paper. From the above analysis, The proposed method in this paper, taking into account the rotor current dynamic process, the calculated rotor flux linkage is modified, and the calculation accuracy of the short-circuit current is improved.

EXPERIMENTAL RESULTS AND ANALYSIS

Based on RTDS (real time digital simulation equipment of power system), an experimental platform with a DFIG converter actual control unit is established. The IGBT module is used to build the converter control unit. In the experimental platform, the parallel communication interface is used to realize the real-time transmission of the control data. The IGBT module is used to real-time control the rotor converter. The system structure diagram of the experimental platform built in this paper and the experimental scenario are shown in **Figure 5**.

Take the example of an actual DFIG farm connected to the grid in **Figure 6**. DFIGs are connected to 35 kv collection line through a transformer at the machine terminal. The lengths of lines AB and CD are 20 and 10 km respectively. The main relevant experiment parameters are as follows: the transformation ratios of the main transformer and machine terminal transformer are 220/35 kV, 35/0.69 kV respectively; The DFIG rated capacity is 2.0 MW; Stator resistance and leakage inductance are 0.016 p.u., 0.169 p.u. respectively; rotor resistance and leakage inductance are 0.009 p.u, 0.153 p.u respectively; excitation mutual inductance is 3.49 p.u.; the Crowbar resistance is 0.1 p.u. There are 10 identical DFIGs on each collection line in the DFIG farm.

Because the circuit between the wind turbines is short, on the same collection line, its influence can be ignored. On each collection line, the transient characteristics of wind turbines are almost the same, so the wind turbines on each collection line can be instead replaced by a single DFIG with equal capacity.

Assuming the DFIG is working under rated operating conditions, a three-phase metal short-circuit fault occurs at the C terminal of the CD line at t = 0.5 s, and the fault lasts for 0.2 s. **Figure 7** shows the instantaneous value of the three-phase short-circuit current of DFIG at terminals C, which is obtained during the experimental test. The RMS of the short-circuit current is extracted *via* the full-cycle Fourier algorithm, which is the solid line shown in **Figure 8**.

Under the same conditions, MATLAB was used to calculate the short-circuit current RMS of the traditional method and the

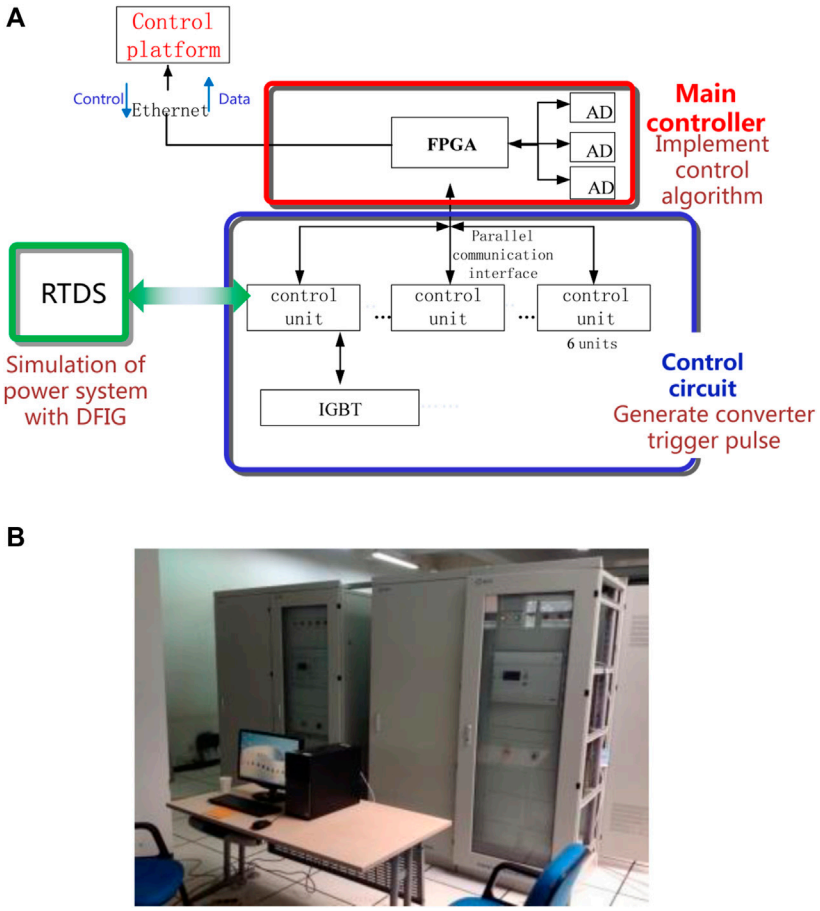


FIGURE 5 | Physical experiment equipment of DFIG based on RTDS. **(A)** System structure of experiment equipment. **(B)** Physical experiment equipment figure.

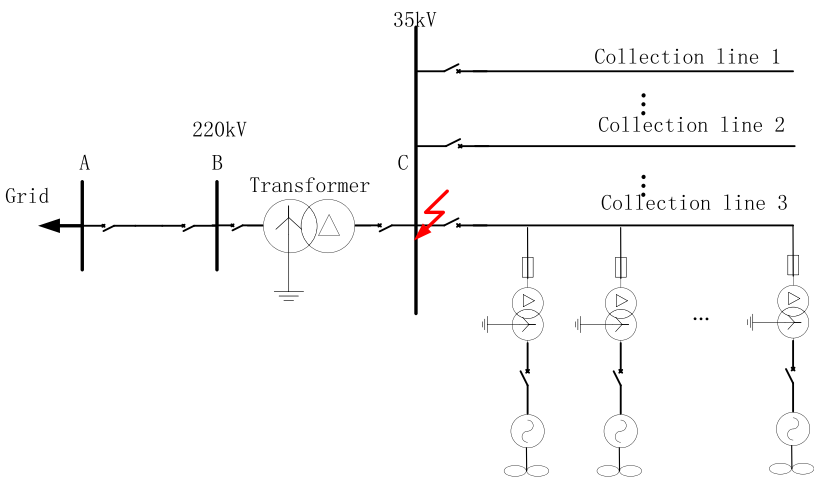


FIGURE 6 | Fault testing circuit structure of DFIG wind farm.

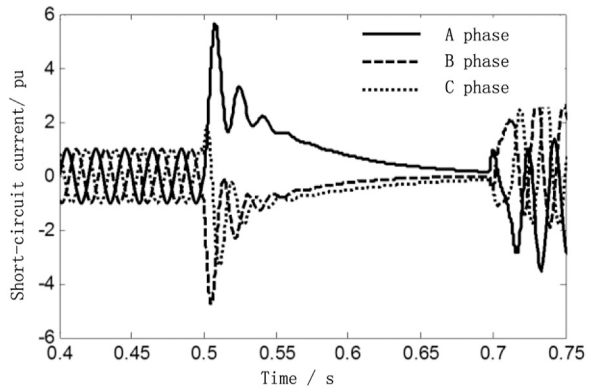


FIGURE 7 | DFIG short circuit current value of three-phase short circuit at bus C under rated operating conditions.

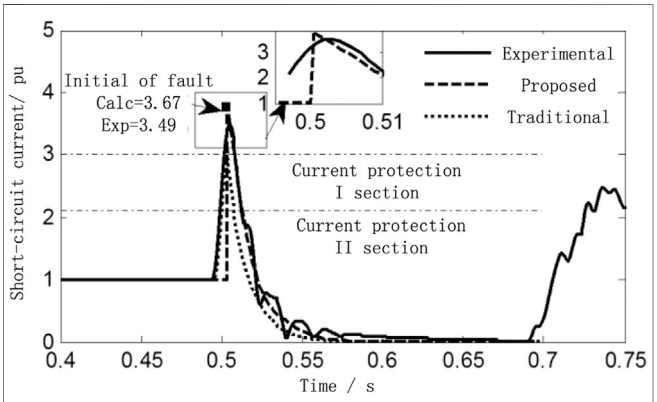


FIGURE 8 | Comparison figure between calculation result and experimental result of DFIG three-phase short circuit at bus C.

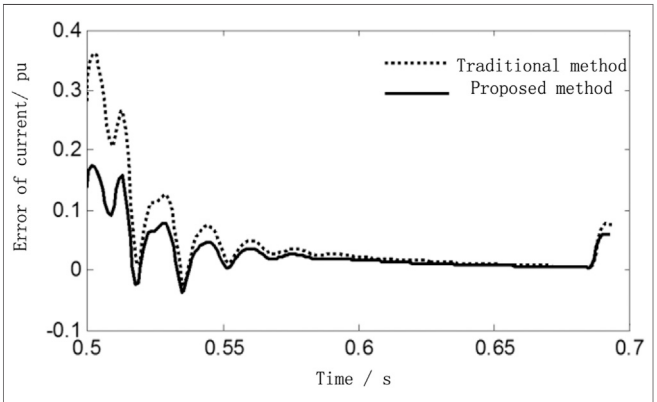


FIGURE 9 | Short circuit current error comparison between the traditional method and proposed method.

proposed method. Compared with the traditional calculation method, the proposed method and the short-circuit current RMS were measured by the experimental platform, and the short-circuit current calculation errors under the two methods were analyzed. **Figure 8** shows the trajectory of the DFIG short-circuit current RMS at terminal C, obtained by the traditional method, the proposed method, and the result of the experimental platform. **Figure 9** is the comparison of the short circuit current calculation errors between the traditional method and the proposed method in this paper.

It can be seen from the experimental results in **Figure 8**, when the fault occurs at 0.5 s, the RMS of the DFIG short-circuit current suddenly increases to 3.49 times of the rated value. The calculated result of the proposed method is 3.67 p.u., and the error is 5.1%, compared with the experimental test result. The result of the traditional method is 3.15 p.u., which ignores the rotor dynamic current. And the error of the traditional method is 9.7%, compared with the experimental test result. From the above analysis, the accuracy of the proposed method in this paper is improved by 4.6%, especially from the 0.5 s to the 0.55 s(short circuit current attenuation process). The curves fitting the degree between the proposed method calculation result and the experimental test result is extremely approach.

Comparing the fixed value of the I section current protection of the collector circuit in **Figure 8**, the calculation results of the traditional method may wrongly analyze the characteristics of the protection operation, resulting in the increase of protection rejection and mis-operation. However, the proposed method in this paper can be more accurate to analyze the effect of the DFIG short-circuit current, on the current protection operating characteristics of the collector circuit.

It can be seen from **Figure 9**, the error curve of proposed method is always below the error curve of the traditional method, and the error of proposed method in this paper is less than 0.19 p.u. (the calculation error is 6%) in the whole process.

After 20 sets of experiments, a large amount of experimental data was obtained, and the short-circuit current calculation results and experimental test result data are shown in **Table 1**. The average error and the maximum error of the short-circuit current are calculated, at 0, 10, 20, and 50 ms after the fault.

From **Table 1**, it can be seen that the accuracy of the method proposed in this paper is improved by 2–5% compared with the traditional calculation method, and during the fault, the error between the proposed method calculation result and experimental result does not exceed 6%. It can be concluded that the calculation method proposed in this paper can not only calculate the initial value of the short-circuit current more accurately, but also reveal the variation law of the whole attenuation process of the short-circuit current more accurately.

CONCLUSION

In order to correctly evaluate the protective action characteristics of the large-scale DFIG accessing to the grid, the influence of the dynamic process of rotor current is taken into account in short-circuit current calculation. The rotor flux of DFIG in the event of a short circuit is calculated accurately, an improved RMS calculation method of DFIG short-circuit current is proposed, and an experimental platform is

TABLE 1 | Comparison between proposed method calculation result and experimental result with 20 sets of experiments.

	Traditional method		Proposed method	
	Maximum error (%)	**Average error (%)**	**Maximum error (%)**	**Average error (%)**
Initial fault	9.9	8.4	5.1	3.7
After fault 10 ms	6.2	4.4	2.4	2.0
After fault 20 ms	7.6	5.3	4.2	3.8
After fault 50 ms	0.4	0.2	0.3	0.2

established. Based on the experimental platform, the accuracy of the short-circuit current calculation result is compared, between the traditional calculation method and the method proposed in this paper.

The experimental results show that, compared with the previous researches, the influence of the rotor current dynamic process on the flux linkage calculation is taken into account, in the RMS calculation method of the short-circuit current proposed in this paper. The method proposed in this paper can more accurately calculate the equivalent potential, which is composed of flux linkage. The calculated initial value of the short-circuit current and the dynamic path of the short-circuit current have higher accuracy, in the proposed method. This study lays the foundation for further study of the influence of the DFIG short-circuit current on the protection operation characteristics.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

JY: data curation, writing, and software.

FUNDING

This work is supported by the Key scientific research projects of colleges and universities in Henan Province (19A470003).

REFERENCES

Firouzi, M., and Gharehpetian, G. (2017). LVRT Performance Enhancement of DFIG-Based Wind Farms by Capacitive Bridge-type Fault Current Limiter. *IEEE Trans. Sustain. Energ.* 9 (3), 1118–1125. doi:10.1109/TSTE.2017.2771321

Fu, Y., Li, Y., He, X., and Han, P. (2017). Corrected Transient Analysis Model of Doubly Fed Induction Generator with Crowbar Protection under Grid Fault. *Proc. CSEE* 37 (16), 4501–4602. doi:10.13334/j.0258-8013.pcsee.1612

Howard, D. F., Habetler, T. G., and Harley, R. G. (2012). Improved Sequence Network Model of Wind Turbine Generators for Short-Circuit Studies. *IEEE Trans. Energ. Convers.* 27 (4), 968–977. doi:10.1109/tec.2012.2213255

Liu, M., Pan, W., and Yang, G. (2018). Practical Calculation Method of Short-Circuit Currents Contributed by Doubly-Fed Wind Turbine Cluster. *Power Syst. Technology* 42 (5), 1475–1481.

Ma, J., and Liu, J. (2018). Adaptive Directional Current protection Scheme Based on Steady State Component in Distribution Network with DG. *Electric Power Automation Equipment* 38 (1), 1–9.

Muljadi, E., Samaan, N., Gevorgian, V., Li, J., and Pasupulati, S. (2013). Different Factors Affecting Short Circuit Behavior of a Wind Power Plant. *IEEE Trans. Ind. Applicat.* 49 (1), 284–292. doi:10.1109/tia.2012.2228831

Okedu, K., and Barghash, H. (2020). Enhancing the Performance of DFIG Wind Turbines Considering Excitation Parameters of the Insulated Gate Bipolar Transistors and a New PLL Scheme. *Front. Energ. Res.* 8, 373. doi:10.3389/fenrg.2020.620277

Rahimi, M., and Azizi, A. (2019). Transient Behavior Representation, Contribution to Fault Current Assessment, and Transient Response Improvement in DFIG-Based Wind Turbines Assisted with Crowbar Hardware. *Int. Trans. Electr. Energ Syst.* 29 (1), e2698. doi:10.1002/etep.2698

Sulla, F., Svensson, J., and Samuelsson, O. (2011). Symmetrical and Unsymmetrical Short-Circuit Current of Squirrel Cage and Doubly Fed Induction Generators. *Electric Power Syst. Res.* 8 (1), 1610–1618. doi:10.1016/j.epsr.2011.03.016

Sun, L., and Wang, Y. (2018). LV Ride through Control Strategy of Doubly Fed Induction Generator Based on Crowbar Series Capacitor. *Power Syst. Technology* 42 (7), 2089–2096.

Swain, S., and Ray, P. K. (2017). Short Circuit Fault Analysis in a Grid Connected DFIG Based Wind Energy System with Active Crowbar protection Circuit for Ridethrough Capability and Power Quality Improvement. *Int. J. Electr. Power Energ. Syst.* 84, 64–75. doi:10.1016/j.ijepes.2016.05.006

Tamaarat, A., and Benakcha, A. (2014). Performance of PI Controller for Control of Active and Reactive Power in DFIG Operating in a Grid-Connected Variable Speed Wind Energy Conversion System. *Front. Energ.* 8 (3), 371–378. doi:10.1007/s11708-014-0318-6

Telukunta, V., Pradhan, J., Agrawal, A., Singh, M., and Srivani, S. (2018). Protection Challenges under Bulk Penetration of Renewable Energy Resources in Power Systems: A Review. *CSEE J. Power Energ. Syst.* 3 (4), 365–379. doi:10.17775/CSEEJPES.2017.00030

Wang, Z., Liu, Y., Lei, M., Bian, S., and Shi, Y. (2015). Doubly-Fed Induction Generator Wind Farm Aggregated Model Based on Crowbar and Integration Simulation Analysis. *Trans. China Electrotechnical Soc.* 30 (4), 45–52. doi:10.19595/j.cnki.1000-6753.tces.2015.04.006

Zhai, J., Zhang, B., Xie, G., Mao, C., and Wang, K. (2013). Three Phase Symmetrical Short Circuit Current Characteristics Analysis of Wind Turbine Driven DFIG with Crowbar. *Automation Electric Power Syst.* 37 (3), 18–22. doi:10.7500/AEPS201111217

Conflict of Interest: The 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.

frontiers in Energy Research

ORIGINAL RESEARCH
published: 25 June 2021
doi: 10.3389/fenrg.2021.701873

Multi-Field Collaborative of Oil-Immersed Transformer for Distributed Energy Resources Temperature Rise Considering the Influence of Heat Transfer Oil

*Kun Huang[1], Hua Li[1], Wei Wang[1], Lifeng Zhang[2], Anhui Feng[2] and Xiaohua Li[2]**

[1]Guangzhou Power Supply Bureau Limited, Guangzhou, China, [2]School of Electrical Engineering, Shanghai University of Electric Power, Shanghai, China

The current calculation method of fluid–structure interaction (FSI) has a defect for oilimmersed transformers for distributed energy resources (DERs). In order to solve this problem, a calculation method is proposed for the temperature rise of oil-immersed transformers in this article. The vibration of insulating oil is considered in this method. Different from the temperature field model established by FSI, the effect of insulating oil vibration on the temperature field is considered. The structure field is introduced to establish the insulating oil vibration model. The temperature field correction coefficient is introduced by coupling the insulating oil vibration and the natural convection of the insulating oil. The result shows that compared with FSI, the results of the calculation method in this study are consistent with the experiment, and the temperature field distribution in the oil-immersed transformer can be calculated more accurately.

Keywords: DERs, correction coefficient, insulating oil, multiphysical interaction, oil-immersed transformer, temperature field

OPEN ACCESS

Edited by:
Yang Li,
Northeast Electric Power University, China

Reviewed by:
Shaoyan Li,
North China Electric Power University, China
Chen Liang,
Nanjing University of Information Science and Technology, China

****Correspondence:***
Xiaohua Li
lixiaohua96@126.com

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

Received: *28 April 2021*
Accepted: *17 May 2021*
Published: *25 June 2021*

Citation:
Huang K, Li H, Wang W, Zhang L, Feng A and Li X (2021) Multi-Field Collaborative of Oil-Immersed Transformer for Distributed Energy Resources Temperature Rise Considering the Influence of Heat Transfer Oil. Front. Energy Res. 9:701873. doi: 10.3389/fenrg.2021.701873

INTRODUCTION

With the increasing demand for energy in industrial production and people's lives, as well as the contradiction between environmental problems and energy development, exploring how to improve the energy utilization rate on the premise of environmental friendliness has become a topic of common concern all over the world. With the development of microgrids and distributed generation technology, more and more clean energy is used in power systems, which puts forward new requirements for the stability of power grids. The operation state of a power transformer is directly related to the stability of the power grid, and in actual operation, insulation and thermal problems are the key factors affecting the operation state of the transformer. The loss will generate heat during the operation of the transformer, making the temperature rise. The presence of insulating oil will cause the heat accumulation above the core and winding, and the core and winding will locally overheat. Operating at high temperatures for a long time will destroy the insulation layer above the transformer, leading to thermal aging of the transformer, causing insulation breakdown and affecting the stability of the power system. Therefore, it is of great importance to obtain the accurate temperature field distribution in the transformer.

In order to balance the interests of IEO and users, a novel Stackelberg game-based optimization framework is proposed for the optimal scheduling of integrated demand response (IDR). IDR-

enabled integrated energy systems with uncertain renewable generations are proposed in Li et al. (2021). A novel optimal scheduling model based on chance-constrained programming (CCP), aimed at minimizing the generation cost, is proposed for a small-scale integrated energy system (IES) with CHP units, thermal power units, renewable generations, and representative auxiliary equipment (Li et al., 2020).The 2D simulation model was established in Liu et al. (2017) for the simulation analysis of the overall temperature rise of oil-immersed transformers by means of mixed numerical calculation. A method for transformer temperature field calculation was proposed in Liao et al. (2015) on the basis of the finite element method and finite volume method, in which a three-phase three-column transformer is calculated and analyzed by the establishment of the electromagnetic–fluid–temperature field coupling model. The effectiveness of the short-circuit method and regression analysis method was verified by Yang et al. (2011). Investigation of the effect of sulfur corrosion on the characteristics of transformer cellulosic insulation was proposed in Gao et al. (2020), and the results were compared with those of fluent software to verify the effectiveness of the method. Xia et al. (2017) proposed a method for calculating the temperature rise of transformer winding, which can effectively tackle the problem of low experimental efficiency of winding temperature rise.

At present, the loss of transformer and the hot spot temperature rise of winding are analyzed in most articles. However, the temperature field of an oil-immersed transformer, especially the calculation of the overall temperature rise, is seldom mentioned (Wang et al., 2017; Lin et al., 2004). Due to the existence of insulating oil, convective heat dissipation is the main heat dissipation mode of an oil-immersed transformer. For coupled heat transfer, the thermal boundary conditions are dynamically determined by the heat exchange process, and there is no need to specify them in advance (Wu et al., 2012; Zhang et al., 2014). However, the FSI calculation is the distribution of the convective heat transfer temperature field in the case of bulk rest of the transformer, tending to focus on the flow velocity in the case of natural convection or forced convection. Generally, it is suitable enough to calculate the heat dissipation of dry transformer air. Given that the insulating oil is much denser than air, the effect of forced vibration of the insulating oil on the temperature field caused by the vibration of the transformer body needs to be considered, making the calculation result closer to the real value.

In order to improve the stability of distributed energy grid operation and maintain the operation safety of the power transformer, considering that the density of insulating oil is much higher than that of air, the influence of forced vibration of insulating oil on the temperature field caused by the transformer body's vibration needs to be considered. Therefore, this article proposes a calculation method of transformer temperature rise considering the influence of vibration of insulating oil. An oil-immersed transformer can be an example: A multiphysical coupling model of an electromagnetic–structural–fluid–temperature field is established. The distribution of the temperature field in an oil-immersed transformer is analyzed by means of multiphysical field coupling simulation and verified by experiments. The experimental results show that the error of the model is reduced by 3.35%.

FIGURE 1 | Oil-immersed transformer model.

TABLE 1 | Transformer technical parameters.

Parameter	Value	Parameter	Value
High voltage (kV)	5	No-load current rating (%)	2.8
Low voltage (V)	220	Short impedance (%)	3
Cooling mode	ONAN	Capacity (KVA)	20

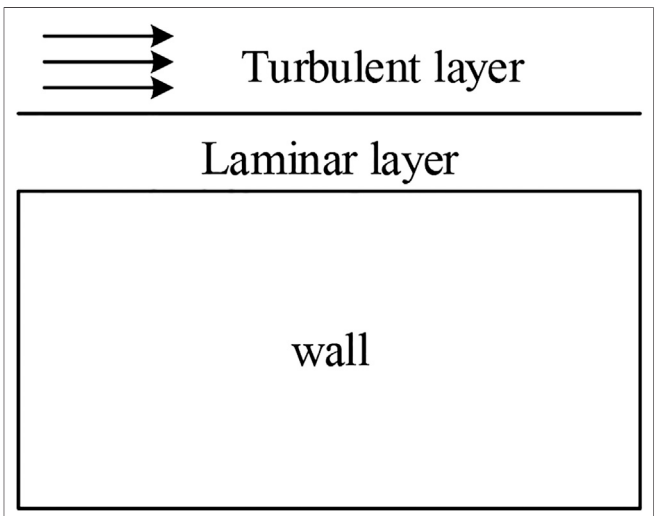


FIGURE 2 | Schematic diagram of insulating oil vibration.

VIBRATION MODEL OF INSULATING OIL

The oil-immersed transformer model and technical parameters adopted in this study are shown in **Figure 1** and **Table 1**.

As shown in **Figure 2**, we only consider the convective heat dissipation of insulating oil based on FSI, namely, the turbulent layer in **Figure 3**. In the turbulent layer, the insulating oil carries

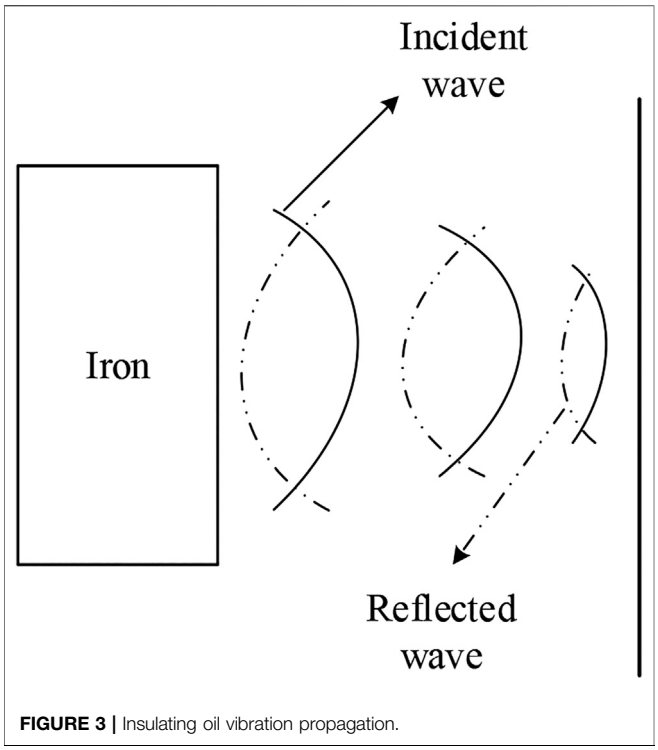


FIGURE 3 | Insulating oil vibration propagation.

away heat by flowing, playing the role of heat dissipation. However, the laminar flow layer still exists on the turbulent layer and the transformer's surface. Due to the viscous effect of the fluid, the flow velocity of the laminar flow layer is close to zero, and the temperature changes rapidly in the normal vector direction of the laminar flow layer. The laminar flow layer will greatly hinder the heat dissipation effect of the turbulent layer. The influence of the vibration of the insulating oil on the temperature field is considered, which is caused by the vibration of the wall surface, improving the heat dissipation effect by disturbing the laminar flow layer while increasing the flow velocity of the turbulent layer. Therefore, the vibration of the insulating oil is first calculated and then explained in this article.

Vibration Analysis of Insulating Oil

Considering the influence of vibration on the fluid field, the fluid travels from the transformer body to the tank wall as a pressure wave, which is manifested as fluid vibration. The surface pressure generated here is the key to coupling the fluid field of insulating oil vibration with the structural field of the transformer body's vibration. The propagation formula is as follows (Zhang et al., 2018):

$$\nabla^2 p = \frac{1}{c_0^2}\frac{\partial^2 p}{\partial t^2}, \tag{1}$$

where p is the fluctuating pressure, c_0 is the propagation velocity of the vibration wave in oil, t is the time variable, and ∇ is the Laplace operator.

When the vibration wave in the insulating oil is transferred to the surface of the oil tank, the wave impedance of the insulating oil is smaller than that of the tank wall, so the transfer wave will reflect as shown in **Figure 3**. At the same time, the hydraulic pressure formed by the gravity of the insulating oil itself will also suppress the vibration of the transformer.

It can be seen that the reflected pressure wave is still a longitudinal wave, but this time, the phase difference is 180 degrees, and the subsequently transmitted vibration wave interacts with the disturbance. The core and winding through which the vibration wave passes can no longer remain in phase of zero, indicating that the insulating oil vibration wave phase is no longer consistent with the vibration source phase. Phase difference of the $\Delta\varphi$ makes insulating oil vibration wave produce a reaction force for the surface of the core and winding, and vibration damping effect inhibition of the transformer.

Table 2 shows the physical property parameters of the transformer's insulating oil. In order to analyze the influence of insulating oil vibration on the suppression effect of the transformer body's vibration, in line with the analysis of transformer vibration, vibration fluid–solid interaction is adopted to analyze the vibration of the transformer's insulating oil, as shown in **Figure 4**. As can be seen, the bottom of the transformer core is based on the fixed constraint, and the vibration of the insulating oil is dominated by the core column and the top one.

As shown in **Figure 5**, the transformer is separated and the vibration of the winding and the core is measured separately by the acceleration sensor. In this experiment, the self-designed distributed feedback laser (DFB) fiber Bragg grating sensor is used for vibration measurement to avoid electromagnetic interference of the traditional piezoelectric sensor and overvoltage damage of the sensor since the sensor needs to have direct contact with the surface of the transformer.

As shown in **Figure 6**, the measured result of winding vibration acceleration is smaller than that of the iron core, indicating that the vibration of the transformer is dominated by the magnetostriction of the iron core, and then the vibration calculation is dominated by the iron core under normal circumstances, which is consistent with the results shown in **Figure 4**. Under normal circumstances, the vibration is dominated by iron core.

Vibration Suppression Model of Insulating Oil

The reaction force will be produced by the vibration of the insulating oil, which can restrain the vibration of the transformer and reduce the amplitude of transformer

TABLE 2 | Insulating oil property parameters.

Characteristic parameter	Fitting formula
Density (kg/m^3)	895
Heat conductivity (W/mK)	0.13215–0.00024 T
Dynamic viscosity (mm/s^2)	18.4–0.16 T
Dilatation coefficient (K)	0.00087

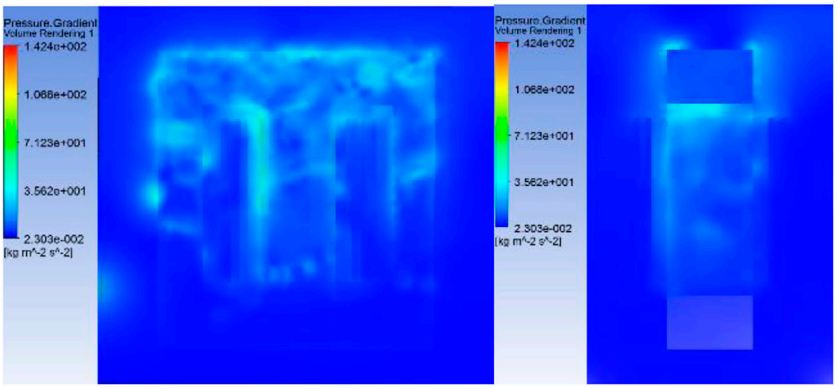


FIGURE 4 | Insulating oil vibration.

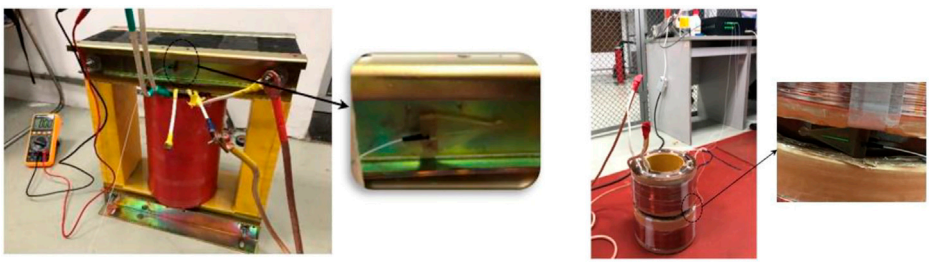


FIGURE 5 | Vibration experiment of DFB.

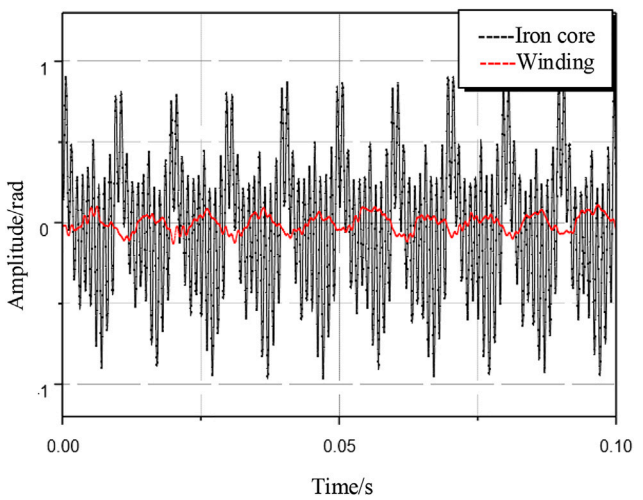


FIGURE 6 | DFB vibration experiment waveform.

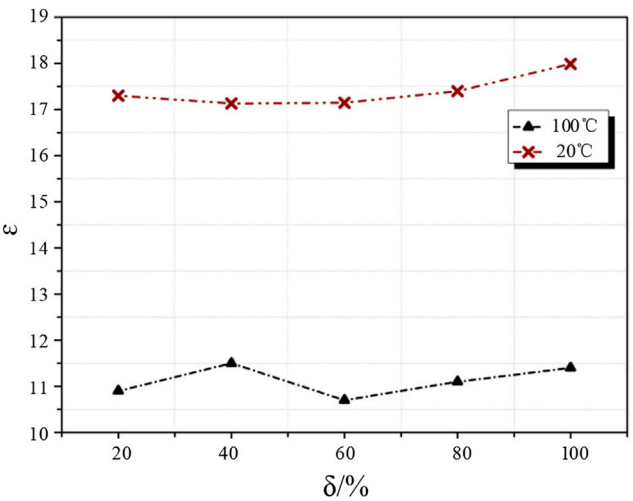


FIGURE 7 | Correction coefficient vibration influence diagram.

vibration. As the vibration source of the insulating oil, the vibration of the transformer is essential for its accurate calculation. Therefore, the effect of oil damping on vibration is calculated by introducing a percentage coefficient of vibration suppression:

$$S_{soil} = (1 - \varepsilon)S, \tag{2}$$

where S represents the vibration displacement of the transformer without considering the vibration suppression of the insulating oil, S_{soil} represents the vibration displacement of the transformer

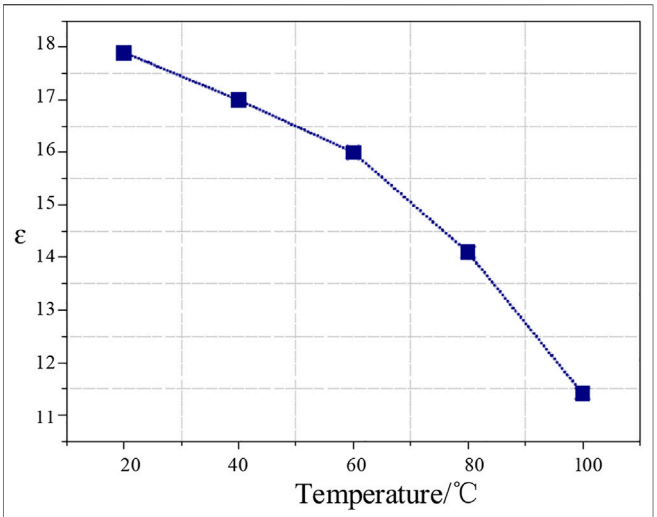


FIGURE 8 | Correction coefficient ε temperature influence diagram.

while considering the vibration suppression of the insulating oil, and ε (%) represents the ratio coefficient while considering the damping effect of the insulating oil. The correction coefficient ε can be affected by many factors. Since the vibration and temperature of the transformer vary significantly under different working conditions, it is assumed that the vibration intensity and temperature of the transformer with the correction coefficient ε are the influencing factors, which will be discussed, respectively, through finite element simulation.

As shown in **Figure 7**, the percentage coefficient δ is used to indicate the intensity of the vibration. It is worth mentioning that when the temperature remains unchanged, the value of the coefficient ε varies little with the change in δ, with the relative error of no more than 5%. It can be assumed to be a constant value by default, which is the case for both temperatures, indicating that the influence of the vibration intensity is weak in vibration suppression.

It should be noted that the ε is significantly affected by temperature. It can be seen in **Figure 7** that its value at different temperatures varies greatly because the wave pressure and viscous impedance in the fluid are closely related to the hydrodynamic viscosity. Temperature has a significant effect on the dynamic viscosity of insulating oil and the modulus of silicon steel. Only considering the influence of these factors can help accurately calculate the vibration displacement of the oil-immersed transformer. Therefore, the influence of temperature on vibration model modification should be further considered.

As shown in **Figure 8**, by simulating the correction coefficient at different temperatures and keeping the vibration intensity constant, it can be noted that there is an approximately linear relationship between the temperature and the correction coefficient ε at 100°C.

Through **formula (2)**, the value of ε can be calculated, and the influence of insulating oil on vibration displacement can be corrected. The ε parameter is set as a function of temperature T:

$$\varepsilon = \varepsilon_0 + K_{soil} e^{(T/\lambda)}, \tag{3}$$

where K_{soil}, α, and λ are the influence coefficients ($K_{soil} = -0.97$; $\varepsilon_0 = 19.33$; and $\lambda = 47.56$).

It can be found that considering the insulating oil vibration for the transformer vibration suppression model is also a function of temperature. With the increase in temperature, the vibration suppression effect of the insulating oil on the transformer becomes weaker and weaker due to the change in the physical parameters of the insulating oil with the temperature change. However, based on the calculations, the vibration of the insulating oil itself becomes more intense as the temperature increases. The vibration of the insulating oil can be accurately calculated with the suppression coefficient.

CALCULATION OF THE INFLUENCE OF INSULATING OIL VIBRATION ON TEMPERATURE FIELD

The influence of insulating oil vibration on the transformer temperature field is reflected in the influence on the surface heat dissipation coefficient (Yang and Tao, 2006). In general, when calculating convective heat dissipation by fluid–structure interaction, only the natural convective heat transfer generated by viscous volume expansion force is considered, while the fluid vibration acceleration generated by pressure resultant force is ignored. The formula is as follows (Tian et al., 2016):

$$\begin{aligned}
&\frac{\partial \rho}{\partial t} + v\frac{\partial \rho}{\partial z} + u\frac{\partial \rho}{\partial r} + \rho\left(\frac{\partial \rho}{\partial z} + \frac{1}{r}\frac{\partial}{\partial z}(ru)\right) = 0,\\
&\rho\left(\frac{\partial \rho}{\partial t} + v\frac{\partial \rho}{\partial z} + u\frac{\partial \rho}{\partial r}\right) = \left(\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + u\frac{\partial^2 u}{\partial z^2}\right) - \frac{\partial \rho}{\partial r},\\
&\rho\left(\frac{\partial v}{\partial t} + v\frac{\partial v}{\partial z} + u\frac{\partial v}{\partial r}\right) = \left(\frac{\partial^2 v}{\partial r^2} + \frac{1}{r}\frac{\partial v}{\partial r} + u\frac{\partial^2 v}{\partial z^2}\right) - \rho g - \frac{\partial \rho}{\partial z},\\
&\rho\left(\frac{\partial T}{\partial t} + v\frac{\partial T}{\partial z} + u\frac{\partial T}{\partial r}\right) = \frac{\lambda}{c_p}\left(\frac{\partial^2 T}{\partial r^2} + \frac{1}{r}\frac{\partial T}{\partial r} + \frac{\partial^2 T}{\partial z^2}\right),
\end{aligned} \tag{4}$$

where t is time; z and r are the axial and radial directions of cylindrical coordinates, respectively; u and v are the velocity components in the z direction and the r direction, respectively; λ is thermal conductivity; ρ is the insulating oil density; T is temperature; P is liquid pressure; and C_P is the specific heat capacity of the insulating oil.

In this study, the vibration velocity of the insulating oil is coupled with the natural convection velocity, and then the vibration velocity of the insulating oil is coupled with the temperature field on the basis of the natural expansion heat convection. The convective heat transfer coefficient of the transformer surface considering the vibration of the insulating oil can be expressed as follows:

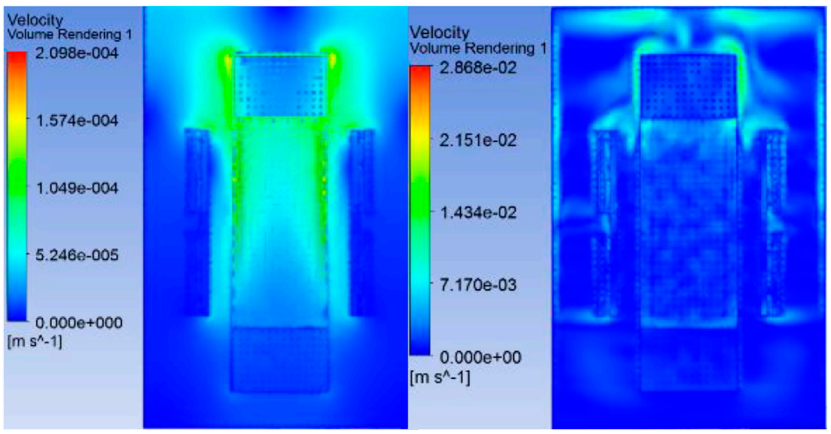


FIGURE 9 | Comparison of heat transfer oil disturbance and natural convection oil flow rate.

$$h_T = \frac{k_f k_p (q_t + q_s)^m C^m L^{(m-1)} Pr^n}{[\nu(C + q_t \cdot q_s)]^m}, \tag{5}$$

where k_f is the thermal conductivity of the insulating oil; P_r is the Prandtl number; L is the characteristic length; q_t and q_s are the buoyancy velocity and vibration velocity of the insulating oil microaggregate flowing through the transformer's surface; C is the cosine of the angle between the two velocities; and k_p, m, and n are constants.

According to **formula (4)**, the heat dissipation coefficient is directly proportional to the temperature rise and fluid flow velocity. The change of fluid velocity will change the temperature field. The fluid flow velocity of the insulating oil is not only composed of the natural buoyancy caused by thermal expansion, but also composed of the vibration velocity caused by the vibration of the insulating oil. According to the above calculation results of vibration of insulating oil, the comparison of two oil flow velocities when the temperature rise reaches 60°C is shown in **Figure 9**. The velocity of oil disturbed by the insulating oil is obviously lower than that of oil under natural convection.

The vibration of the insulating oil exerts a direct impact on the temperature field, without an impact on the transformer loss directly. However, it is difficult to calculate the overall distribution of the temperature field of the transformer after the coupling structure field, so the loss can be modified by **formula (6)** to achieve the equivalent correction of the influence of vibration on the temperature field:

$$P_T = (1 - \beta(V, T))[(k_h(f, B_m, T, V) f B_m^{\alpha} + k_e(f, B_m, T, V) f^2 B_m^2 + k_{ex}(f, B_m, T, V)(f B_m)^{3/2}) + P_{Cu}] \tag{6}$$

Considering the influence of the vibration of the insulating oil on the temperature field, the temperature correction coefficient β is directly related to the vibration inhibition coefficient ε. The correlation coefficient is obtained by fitting through simulation calculation, and the ambient temperature is 20°C, as shown in **Figure 10**. It can be seen that the trend of the correction coefficient ε is opposite to that of the temperature field correction coefficient β. With the increase in temperature, the inhibition of the vibration will be weakened by the insulating oil, while the vibration speed of the insulating oil will increase and the influence on the temperature field will be enhanced. It is shown that the temperature field model of insulating oil vibration is still a function of temperature.

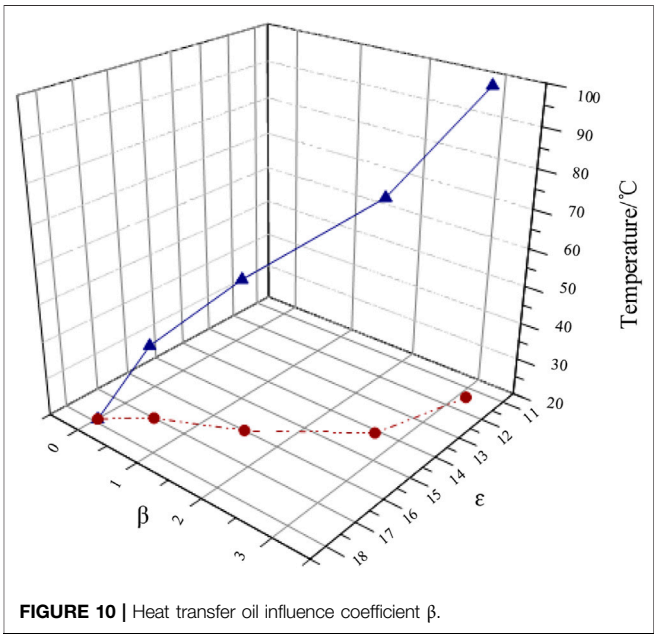


FIGURE 10 | Heat transfer oil influence coefficient β.

TRANSFORMER LOSS CALCULATION

Calculation Model of Iron Loss and Copper Loss

The iron loss is divided into three basic loss components: hysteresis, eddy current, and abnormal loss (Yang and Tao,

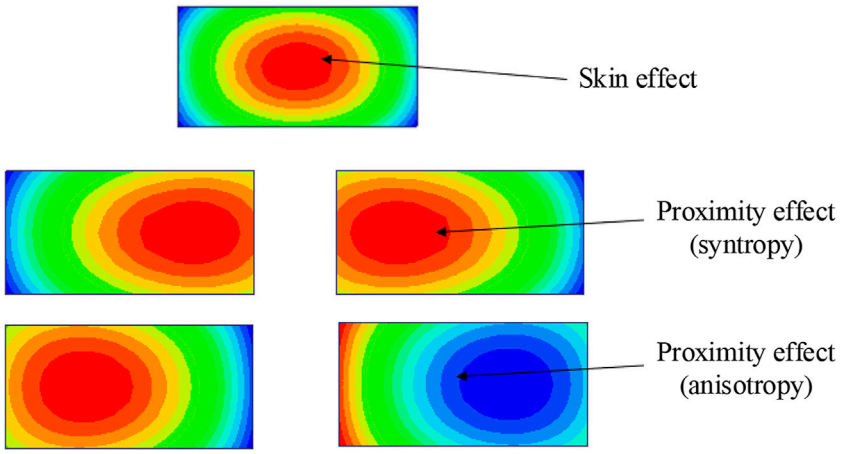


FIGURE 11 | Skin effect and proximity effect.

2006; Zou et al., 2010). In most cases, the extra loss is ignored. The classical trinomial formula is usually used to calculate

$$P_{Fe} = k_h f B_m^{\alpha} + k_e f^2 B_m^2 + k_{ex} (f B_m)^{3/2}, \tag{7}$$

where f is the frequency and B is the peak magnetic density.

AC copper loss can be equivalent to the composition of DC winding copper loss and eddy current loss (Xue et al., 2018):

$$P_{ac} = P_{dc} + P_{eddy} \tag{8}$$

Transformer windings are mostly flat copper wires. The skin effect and proximity effect of the flat copper wire windings are shown in **Figure 11**.

By ignoring the influence of eddy current redistribution on the electric field distribution, the average loss in conductors per unit length can be estimated (Lou and Li, 2016; Tian et al., 2016):

$$P = \frac{m^2 n^2 \omega^2 B_n^2}{8\pi\rho_c}, \tag{9}$$

where ρ_c is the resistivity of the conductor; B_n is the peak of flux density; and for angular frequency, m and n are the length and width of the flat copper wire, respectively. The resistance is equivalent to the skin depth of the windings, and then the eddy current loss is corrected:

$$R_S = R\left[1 + \frac{1}{48\pi^2}\left(\frac{mn}{\eta}\right)^2\right] \tag{10}$$

The high voltage side is a round copper wire winding, while the low voltage side is a flat copper wire winding. Although the distribution law of current density is the same, the difference in the shape makes the degree of skin effect and proximity effect different, resulting in different increase rates of equivalent resistance. **Figure 12** shows the measurement of the winding equivalent resistance using LCR.

The results indicate that there is almost no difference in the growth rate of equivalent resistance between the two wires at low and medium frequencies, and the value slightly increases at high frequencies, but the increase is limited. Therefore, a unified model is adopted to calculate the equivalent resistance of the two wires.

FIGURE 12 | Winding AC equivalent resistance experiment.

Temperature Correction Model

There exists a big error of iron loss calculated by the ordinary trinomial formula. For the trinomial formula, temperature will directly affect the coefficient values of core loss. Notably, there are few studies on the influence of temperature on the loss of silicon steel at home and abroad, and the loss data of silicon steel with the influence of temperature are not perfect. The missing material data of the silicon steel must be obtained directly from experiments.

The temperature model of a silicon steel sheet is introduced to correct the coefficient of the trinomial formula, and the temperature feedback model is used for two-way interaction calculation, thus reducing the loss error (Tian et al., 2016):

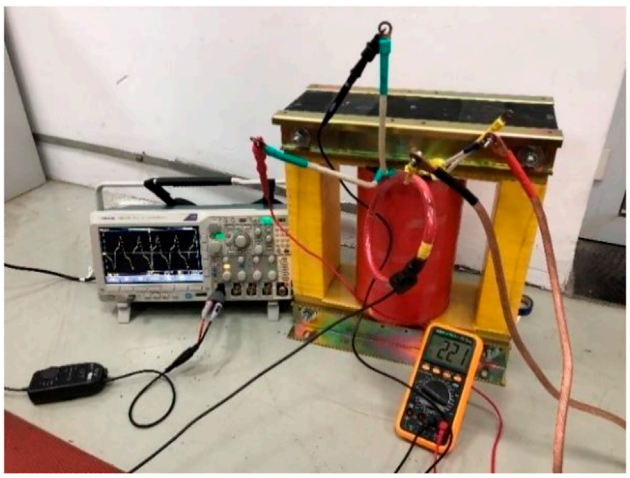


FIGURE 13 | Transformer no-load experiment.

$$k_h(T,f,B_m) = k_{Th}k_{h,T_0}(f,B_m), \tag{11}$$

$$k_e(T,f,B_m) = k_{Te}k_{e,T_0}(f,B_m), \tag{12}$$

where k_{Th} and k_{Te} are the correction coefficients considering temperature and k_{h,T_0} and k_{e,T_0} are the coefficients at time T_0.

The influence of temperature on iron loss can be obtained by measuring two sets of data, which can effectively solve the problem of imperfect material parameters of the silicon steel sheet. The correlation coefficient is obtained by using the no-load test of iron core as shown in **Figure 13** (k_{Th} = 1.012 and k_{Te} = 1.004).

TEMPERATURE RISE CALCULATION

The temperature simulation results of the electromagnetic–structural–fluid–temperature field of the multiphysical field interaction transformer filled with insulating oil established by considering the vibration correction model of the insulating oil are shown in **Figure 14**, and the oil temperature at the upper end of the transformer is obviously higher than that at the lower end.

In order to verify the influence of insulating oil vibration on transformer temperature, the basic temperature of the transformer is changed by changing the filling amount of the insulating oil. When the filling amount of the insulating oil is 80 and 100%, the experimental data and simulation data are compared, as shown in **Table 3**. Experiments for temperature rise to the top of the oil temperature in rated conditions have been measured. The upper insulating oil temperature of 55.2 and 54.75 ˚C in the model of the electromagnetic–structural–fluid–temperature field interaction correction, only considering the 57.3 and 56.41˚C in the model of three physical field electromagnetic–fluid–temperature field of fluid–solid interaction, there is a difference. The temperature distribution does not change, but the difference between the upper insulating oil temperature rises by about 2.1 and 1.6˚C. It can be seen that the error of model construction in this study is lower, with a temperature rise difference of 9.2 and 9.5%, which is reduced by 3.5 and 3.35%.

The multiphysical interaction model of the electromagnetic–structural–fluid–temperature field considering the vibration of the insulating oil is closer to the real value than the traditional FSI model. The multiphysical coupling method considering the influence of the vibration of the insulating oil is more reliable. It can be extended to all types of oil-immersed power transformers for the reason proposed in this article: The calculation method considers the influence of insulating oil vibration on the internal temperature field of the oil-immersed transformer as a function of temperature.

CONCLUSION

In this article, a more accurate method is proposed to calculate the temperature field of the transformer considering the vibration of the insulating oil. Taking an oil-immersed transformer as an example, the multiphysical coupling model of the electromagnetic–structural–fluid–temperature field considering the vibration of the insulating oil can be used to calculate the

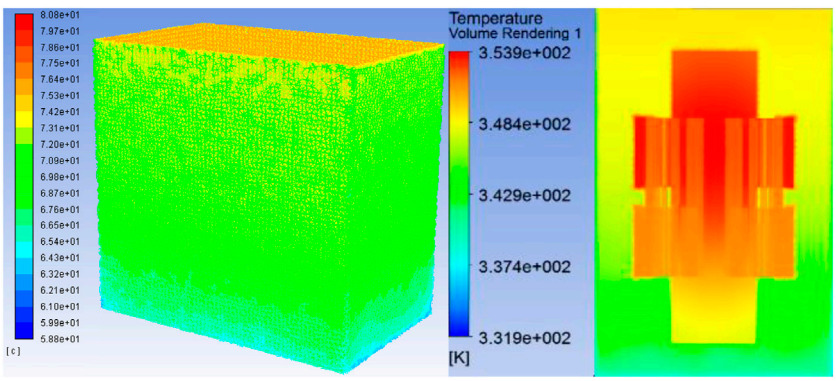


FIGURE 14 | Transformer temperature field (insulating oil vibration model).

TABLE 3 | Corrected results comparison.

Oil-filling capacity (%)	Model	Temperature rise (°C)	Absolute error (°C)	Relative error (%)
100	Convention	56.41	6.41	12.85
	Correction	54.75	4.75	9.5
	Experiment	50.0	—	—
80	Convention	57.3	7.27	12.7
	Correction	55.2	5.07	9.2
	Experiment	50.8	—	—

temperature field of the oil-immersed transformer more accurately. The following conclusions are drawn:

1) A modified model of vibration suppression of insulating oil is established, and a temperature-dependent vibration suppression function of insulating oil is established by introducing the correction coefficient ε.
2) A correction model for the influence of insulating oil vibration on the temperature field is established. As the temperature increases, the influence of heat conduction oil on the temperature field increases. At the same time, the influence of insulating oil vibration on temperature field considered in this study is still a function of temperature through correction coefficients ε and β.
3) The multiphysical coupling model of the electromagnetic–structural–fluid–temperature field is established to calculate the temperature field of an oil-immersed transformer.

The model of vibration suppression function and the calculation method of temperature field considering the vibration of insulating oil are presented in this article, which is of great significance to prolong the service life of power transformers and maintain the stable operation of distributed energy resource systems.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

The individual contributions of the authors are as follows: data curation, KH; formal analysis, HL; methodology, WW; supervision, XL; validation, LZ; writing (original draft), AF. All authors have read and agreed to the published version of the manuscript.

REFERENCES

Gao, S., Yang, L., and Ke, T. (2020). Failure Mechanism of Transformer Oil-Immersed Cellulosic Insulation Induced by Sulfur Corrosion. *Cellulose* 27 (12), 7157–7174. doi:10.1007/s10570-020-03271-x

Li, Y., Wang, C., Li, G., and Chen, C. (2021). Optimal Scheduling of Integrated Demand Response-Enabled Integrated Energy Systems with Uncertain Renewable Generations: A Stackelberg Game Approach. *Energ. Convers. Manage.* 235, 113996. doi:10.1016/j.enconman.2021.113996

Li, Y., Wang, C., Li, G., Wang, J., Zhao, D., and Chen, C. (2020). Improving Operational Flexibility of Integrated Energy System with Uncertain Renewable Generations Considering thermal Inertia of Buildings. *Energ. Convers. Manage.* 207, 112526. doi:10.1016/j.enconman.2020.112526

Liao, C., R, J., Liu, C., et al. (2015). Comprehensive Analysis of 3-D Electromagnetic-Fluid-thermal fields of Oil-Immersed Transformer. *Electr. Power Autom. Equip.* 35 (9), 150–155. doi:10.16081/j.issn.1006-6047.2015.09.024

Lin, D., Zhou, P., Fu, W. N., Badics, Z., and Cendes, Z. J. (2004). A Dynamic Core Loss Model for Soft Ferromagnetic and Power Ferrite Materials in Transient Finite Element Analysis. *IEEE Trans. Magn.* 40 (2), 1318–1321. doi:10.1109/TMAG.2004.825025

Liu, G., Jin, Y., Ma, Y., et al. (2017). Numerical Analysis of Fluid Field and Temperature Field of Oil-Immersed Transformer. *Transformer* 54 (5), 22–26. doi:10.19487/j.cnki.1001-8425.2017.05.007

Lou, Y., and Li, B. (2016). Methods for Reducing the Loss Caused by the Eddy Current in Electromagnetic Railgun Shielding. *High Voltage Eng.* 42 (6), 1935–1941. doi:10.13336/j.1003-6520.hve.20160401007

Tian, M., Zhu, J., and Song, C. (2016). Temperature Field Simulation of Coal Dry-type Transformer Based on Fluid-Solid Coupling Analysis. *High Voltage Eng.* 42 (12), 3972–3981. doi:10.13336/j.1003-6520.hve.20161128035

Wahyudi, M., Negara, I. M. Y., Asfani, D. A., Hernanda, I. G. N. S., and Fahmi, D. (2017). "Application of Wavelet Cumulative Energy and Artificial Neural Network for Classification of Ferroresonance Signal during Symmetrical and Unsymmetrical Switching of Three-Phases Distribution Transformer." in 2017 International Conference on High Voltage Engineering and Power Systems, Bali, Indonesia, Oct 02–05, 2017, (Indonesia: Pendidikan Tnggi Teknik Indonesia; Inst Teknologi Bandung; Inst Teknologi Bandung, STEI; PLN), 394–399. doi:10.1109/ICHVEPS.2017.82258770

Wang, W., et al. (2017). The Calculation Method of Temperature Field of Oil Immersed Transformer with FEM and FVM Combined Method. *Electr. Meas. Instrument.* (18), 42–47.

Wang, X., Lai, C., Yu, D., Xu, Y., and He, Y. (2019). "Application of Energy Performance Contracting in Distribution Transformer Based on Neural Network Prediction," in 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Con-ference (ITNEC), 42–46. doi:10.1109/ITNEC.2019.8729195

Wu, X., Shu, N., and Li, H. (2012). Thermal Field Calculation and Analysis of Gas Insulated Busbars Based on Fluid Multiple Species Transport. *Proc. CSEE* 32 (33), 141–147. doi:10.13334/j.0258-8013.pcsee.2012.33.020

Xia, Y., Zhang, J., and Xie, Q. (2017). Engineering Calculation Method for Winding Temperature Rise of Oil Immersed Power Transformer. *High Voltage Apparatus* 53 (9), 176–180. doi:10.13296/j.1001-1609.hva.2017.09.030

Xue, S., Feng, J., Guo, S., Peng, J., Chu, W. Q., and Zhu, Z. Q. (2018). A New Iron Loss Model for Temperature Dependencies of Hysteresis and Eddy Current Losses in Electrical Machines. *IEEE Trans. Magn.* 54 (1), 1–10. doi:10.1109/TMAG.2017.2755593

Yang, S., and Tao, W. (2006). *Heat Transfer*. 4th Edition. Beijing: Higher Education Press, 33–76.

Yang, Z., Wang, L., and Du, J. (2011). Temperature Rise Test of Oil -Immersed Power Transformer and its Calculating Method. *Transformer* 38 (6), 15–20. doi:10.19487/j.cnki.1001-8425.2001.06.006

Zhang, F., Ji, S., Shi, Y., et al. (2018). Research on Transfoermer Winding Vibration and Propagation Characteristics. *Proc. CSEE* 38 (9), 2790–2798. doi:10.13334/j.0258-8013.pcsee.170497

Zhang, Q., Lu, X., and Huang, S. (2014). Temperature Rise Calculations of High Density Permanent Magnet Motors Based on Multi-Domain Co-simulation. *Proc. CSEE* 34 (12), 1874–1881. doi:10.13334/j.0258-8013.pcsee.2014.12.006

Zou, J., Jiang, S., and Liang, W. (2010). AC Loss in a High Speed BLPM Motor Considering Proximity Effect. *Electric Machines and Control* 14 (5), 131–135. doi:10.15938/j.emc.2010.05.011

Conflict of Interest: KH, HL, and WW were employed by Guangzhou Power Supply Bureau Limited.

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

frontiers in Energy Research

BRIEF RESEARCH REPORT
published: 30 June 2021
doi: 10.3389/fenrg.2021.681244

Estimating Smart Grid's Carbon Emission Reduction Potential in China's Manufacturing Industry Based on Decomposition Analysis

Hua Fu[1], Yingying Shi[2] and Yongchao Zeng[3]*

[1] School of Government Management, Peking University, Beijing, China, [2] Evergrande School of Management, Wuhan University of Science and Technology, Wuhan, China, [3] School of Management and Economics, Beijing Institute of Technology, Beijing, China

China has a large manufacturing industry and shoulders the responsibility of reducing carbon emissions. Smart grid technologies can integrate multiple renewable energy technologies, which possess significant potential in reducing carbon emissions. To estimate the carbon emission reduction potential of the smart grid in China's manufacturing industry, this paper applies a temporal logarithmic mean Divisia index (LMDI) method to analyze the driving forces of carbon emission changes in the whole manufacturing industry and 28 sub-industries from 2000 to 2017, respectively. The results reveal that industrial activity and energy intensity are the key factors leading to the increase and mitigation of carbon emissions, respectively. Sub-industries with high emission intensity are crucial for the reduction of carbon emissions in the manufacturing industry. By applying a smart grid, the carbon emissions could be reduced by 27.51% in the optimistic scenario.

Keywords: smart grid, distributed energy resources, CO_2 emission, manufacturing industry, electricity, LMDI

OPEN ACCESS

Edited by:
Liang Chen,
Nanjing University of Information Science and Technology, China

Reviewed by:
Xiaoying Wang,
University of Calgary, Canada
Xue Wang,
Shanghai University of Finance and Economics, China
Yong Wang,
Chinese Academy of Social Sciences (CASS), China

****Correspondence:***
Yingying Shi
shiyingying06@126.com

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

Received: *16 March 2021*
Accepted: *19 April 2021*
Published: *30 June 2021*

Citation:
Fu H, Shi Y and Zeng Y (2021) Estimating Smart Grid's Carbon Emission Reduction Potential in China's Manufacturing Industry Based on Decomposition Analysis. Front. Energy Res. 9:681244. doi: 10.3389/fenrg.2021.681244

INTRODUCTION

Due to its continued economic growth, industrialization, and urbanization, China surpassed the United States as the world's largest emitter of carbon dioxide in 2007. According to the BP world energy statistics review, China's carbon emissions in 2016 reached 9.12 billion tons, accounting for 27.3% of global carbon emissions. The extreme climate phenomenon caused by the sharp increase in greenhouse gas has brought great challenges for human survival and development. In 2015, to undertake international responsibilities and obligations, the Chinese government proposed reducing carbon emissions per unit of GDP by 60–65% by 2030 from the 2005 level. As the pillar of China's economy, industry consumes 70% of the domestic energy and contributes more than 50% of the domestic carbon emissions. To achieve the goal of reducing carbon emissions while maintaining high-quality economic development, the manufacturing industry has made the reduction of carbon emissions its top priority. Considering China's significant coal consumption, the reduction of carbon emissions in the manufacturing industry is not only related to the industrial structure but also related to the energy structure. This unreasonable energy structure is an important factor that leads to the industry's high carbon emission intensity.

There is rich literature focusing on carbon emission reduction regarding the manufacturing industry (e.g., Li and Cheng, 2020). It has also been demonstrated that smart-grid technologies are helpful for reducing carbon emissions in power and transportation industries (Fu et al., 2012); however, the potential of smart-grid in terms of reducing carbon emissions in the manufacturing industry is rarely studied. Since a variety of energy resources are converted into electric power for

serving manufacturing industries, it is of great practical significance to estimate the carbon emission reduction potential of the smart grid.

In this paper, we use the logarithmic mean Divisia index (LMDI) decomposition method to analyze the influencing factors of carbon emissions in the manufacturing industry and attempt to clarify the impacts of industrial activities and energy structure on the sub-industry's carbon emissions. Then we estimate the potential of emission reduction by optimizing the energy structure through a smart grid.

LITERATURE REVIEW

Carbon emission is an important aspect of environmental economics and energy economics. As for the relationship between carbon emissions and energy structure, many studies have been carried out from the industry level and regional level, mainly focusing on manufacturing, the power industry, energy structure, coal consumption, coal power, and renewable energy power generation.

Some scholars have analyzed the relationship between China's energy consumption, especially coal consumption, and carbon emissions from a regional perspective. Sun et al. (2016) used Moran's I index and a geographical weighted regression model to illustrate the spatial features of provincial carbon emission transfer and its economic spillover effects. They found that 18 provinces reduced coal consumption through carbon emission transfer, indicating that the transfers occurred due to the presence of carbon emissions. Because these regions reduced coal consumption in the production processes, they had cleaner modes of production. Some scholars have studied the differences and reasons of carbon emissions in eastern, central, and Western China (Xu and Lin, 2017). Cai and Guo (2018) identified the driving forces of carbon emissions of 286 Chinese cities, finding that industry composition, development stage, and geographical location are important driving forces to increase the carbon emissions of Chinese cities.

Some scholars have analyzed the impacts of economic activities, energy intensity, and technology on carbon emissions from the perspective of industry. Since the decomposition analysis method can better quantify the potential of carbon emissions reduction, the LMDI method is widely used in research about the region and industry. By using a decomposition analysis of carbon emissions from China's manufacturing industry, Chinese scholars found that economic output effect and energy intensity effect are the most significant of carbon emissions, and the output effect is far greater than the inhibition effect of energy intensity and energy structure on carbon emissions (Xu et al., 2012; Tan and Lin, 2018). Energy intensity can be further decomposed into technology, intermediate input, and industrial output structure; however, the effect of technological change on carbon emissions is not obvious in recent years (Zhou and Ang, 2008). Lin and Du (2014) used a comprehensive decomposition model to study the impact of technological progress on the decline of China's energy intensity. In the long run, technological progress will be key to achieving the sustainable growth of energy productivity. Ang et al. (2015) expanded the LMDI decomposition method and studied carbon emissions of the manufacturing industries in 30 provinces of China. Based on this information, domestic scholars further analyzed the influencing factors of carbon emissions in the sub-industries of China's manufacturing industry (Shi et al., 2019).

Additionally, scholars have carried out research on the emissions reduction effects of a smart grid. Lamiaa and Tarek (2013) studied Egypt to explore the development of a smart grid and its contribution to carbon emissions reduction in the electricity sector. Some scholars have studied how to optimize the smart grid to encourage the use of more renewable energy for power generation (Li et al., 2021). Some scholars have designed emissions reduction schemes for smart grids in six representative EU countries and have found that by adopting new technologies and effective regulatory frameworks, we can realize the emissions reduction potential and achieve the 2020 emissions reduction target (Darby et al., 2013).

Although many scholars are concerned about the carbon emissions emitted by the manufacturing industry and believe that the unreasonable energy consumption structure is an important driving force, the contribution of renewable energy use to carbon emissions reduction is less significant than one might think. In this paper, smart grids and the manufacturing industry's carbon emissions are combined to bridge the gap in emissions reduction strategies.

RESEARCH METHODS AND DATA SOURCES

Calculation Methods

According to Kaya's equation (Kaya, 1989), the carbon emissions of the energy-related manufacturing industry is mainly driven by five forces: carbon emissions coefficient, energy structure, energy intensity, economic activities, and industrial scale. This can be expressed as follows:

$$C_{ij} = \sum_{ij} \frac{C_{ij}}{E_{ij}} \times \frac{E_{ij}}{E_i} \times \frac{E_i}{Y_i} \times \frac{Y_i}{P_i} \times P_i \quad (1)$$

On the right hand side of Equation 1, the first component could be interpreted as the CO_2 emissions coefficient since the CO_2 emissions is deflated by the energy consumed in sub-industry i of the manufacturing industry, where j repents the energy types (j = 1,2,. . .,18). The second component measures the energy structure since the CO_2 emissions is deflated by the of energy type consumed in the sub-industry. The third component could be interpreted as the energy intensity. An increase in energy usage technical efficiency will lead to a larger energy intensity change and therefore more of the change in E/Y. The fourth and fifth components measure output per capita and the number of employees in sub-industry i which represent industry activity and industry scale of the sub-industry, respectively.

Using the notations described above, we can rewrite Equation 1 as

$$C_i = \sum_j EC_{ij} + ES_{ij} + EI_i + EA_i + IP_i \quad (2)$$

According to the LMDI model given by Ang (2005), the total change of carbon dioxide emissions related to energy between the base year 0 and target year t of China's manufacturing industry can be divided into five driving forces: carbon emissions coefficient (EC) effect, energy structure (ES) effect, energy intensity (EI) effect, economic activity (EA) effect, and industry population (IP) effect. Now that the change of CO_2 emissions from the period 0 to t is as follows:

$$\Delta C_i^{(t-0)} = C_i^t - C_i^0 = \Delta C_{i,EC}^{(t-0)} + \Delta C_{i,ES}^{(t-0)} + \Delta C_{i,EI}^{(t-0)} + \Delta C_{i,EA}^{(t-0)} + \Delta C_{i,IP}^{(t-0)} \quad (3)$$

According to the LMDI model given by Ang (2005), the total change of carbon dioxide emissions related to energy between the base year 0 and target year t of China's manufacturing industry can be divided into five driving forces: EC effect, ES effect, EI effect, EA effect, and IP effect. The specific calculation is as follows:

$$\Delta EC = \sum_{ij} L\left(C_{ij}^t, C_{ij}^0\right) \ln\left(\frac{EC_i^t}{EC_i^0}\right) \quad (4)$$

$$\Delta ES = \sum_{ij} L(C_{ij}^t, C_{ij}^0) \ln\left(\frac{ES_i^t}{ES_i^0}\right) \quad (5)$$

$$\Delta EI = \sum_{ij} L\left(C_{ij}^t, C_{ij}^0\right) \ln\left(\frac{EI_i^t}{EI_i^0}\right) \quad (6)$$

$$\Delta EA = \sum_{ij} L(C_{ij}^t, C_{ij}^0) \ln\left(\frac{EA_i^t}{EA_i^0}\right) \quad (7)$$

$$\Delta IP = \sum_{ij} L(C_{ij}^t, C_{ij}^0) \ln\left(\frac{IP_i^t}{IP_i^0}\right) \quad (8)$$

where

$$L(C_{ij}^t, C_{ij}^0) = \begin{cases} \frac{C_{ij}^t - C_{ij}^0}{lnC_{ij}^t - lnC_{ij}^0} & C_{ij}^t \neq C_{ij}^0 \\ C_{ij}^t & C_{ij}^t = C_{ij}^0 \end{cases} \quad (9)$$

Data Source

According to the industrial classification of national economic activities (GB/T 4754-2011), the manufacturing industry is divided into 31 sub sectors. Due to the data discontinuity of C42, C43, and C37, this study covers 28 sub-industries of the manufacturing industry. The output value and employment population of manufacturing sub-industries are from China Industrial Statistics Yearbook (2001–2019). The output value is calculated using the constant price in 2000. Energy consumption and carbon dioxide emissions data of the manufacturing industry is from carbon emissions inventory and energy inventory of CEADs. According to the revised China's carbon emissions factors published in Nature by Liu et al. (2015), the list was compiled with China's energy statistical yearbook data and widely used by scholars at home and abroad.

ANALYSIS OF ENERGY AND CARBON EMISSION

Relationship Between Energy Consumption and Carbon Emissions

In recent years, with the growth of manufacturing output value, the energy consumption and carbon emissions of manufacturing industry are also increasing. In 2000, 28 sub-industries in the manufacturing industry consumed 6.05 million tons of standard coal, which increased to 1977 million tons of standard coal in 2017, an increase of 2.3 times in 18 years. In the same period, CO_2 emissions increased from 1365.1 tons to 4608.3 tons, an increase of 2.4 times. Thus, there is a positive correlation between energy consumption and CO_2 emissions. Although electricity is the second largest energy source for the manufacturing industry, the ratio of electricity to energy consumption remained low at less than 20%.

From 2000 to 2017, the carbon emissions and energy consumption of the manufacturing industry can be roughly divided into four stages. In the first stage (2000–2005), growth rate of energy consumption and CO_2 emissions increased sharply from less than 9% to about 20% while the proportion of electricity in energy consumption decreased. In the second stage (2005–2010), the growth rates were basically controlled within 10%, although they once declined to 4% in 2008. The proportion of electricity continued to shrink. In the third stage (2010–2014), the growth rates were both within 10%, and the ratio of electricity began to increase. In the last stage (2014–2017), the growth rates were both negative, and the proportion of electricity reached 23.8% in 2017.

Carbon Emissions for Different Energy Sources

The 18 types of energy in the energy emissions list fit into six categories: raw coal, coal products, oil, natural gas, and electric and thermal energy. As demonstrated in **Figure 1**, coal has always been the main source of energy consumed by the manufacturing industry and the largest source of carbon emissions; however, since 2010, the proportion of coal-related carbon emissions has dropped to less than half. **Figure 1** indicates that carbon emissions from raw coal reached a peak in 2004 (37%) and then gradually declined, and carbon emissions from coal products have gradually increased from 27 to 37%. It is obvious that electrical power is the second largest source of carbon emissions in the manufacturing industry. Electricity-related carbon emissions accounted for 17–24% of the total emissions from the manufacturing industry between 2000 and 2017, reaching a peak of 24% in 2017. The proportion of oil-related carbon emissions declined after

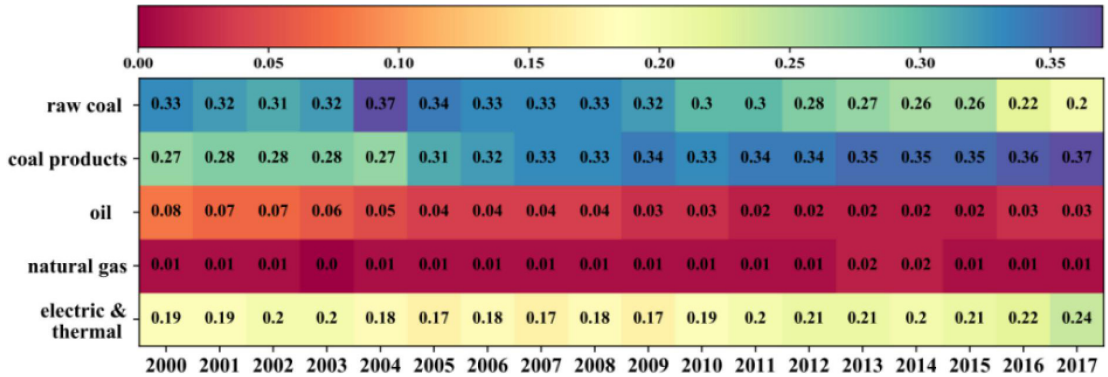


FIGURE 1 | Energy consumption of manufacturing industry in 2000–2017.

2000 and then increased again after 2013. The scale of natural gas used by the manufacturing industry is low compared to proportion of other energies used and tends to rise as a whole.

China's energy resource structure is different from developed countries, as it uses a larger amount of coal and lower amounts of low carbon energy resources. In the view of the average level of the world, the proportion of oil, natural gas, and coal is much more equal, that is 34, 24, and 27% respectively. Fossil energy use in America and Europe is dominated by oil and natural gas, while the proportion of coal is only 14 and 13% respectively. However, coal is the most important resource in the electric power industry of China, while the share of nuclear energy, photo-voltaic, wind power, and other non-fossil energy is pretty low.

CALCULATION PROCESS AND RESULTS ANALYSIS

Carbon Emissions From Manufacturing Sub-Industries

The CO_2 emissions' intensity of 28 manufacturing sub-industries shows great diversity in 2017. The industry with the highest CO_2 emissions intensity is C31 (447 tons/10000CNY), while the industry with the lowest CO_2 emissions intensity is C39 (0.04 tons/10000CNY). It is apparent that there are significant differences in CO_2 emissions across the 28 sub-industries. Therefore, according to the CO_2 emissions intensity, we divide 28 sub-industries into three categories: high emissions intensity industry (HEIs), medium emissions intensity industry (MEIs), and low emissions intensity industry (LEIs). HEIs include C31, C30, C25, C26, and C32. The CO_2 emissions of these five industries alone account for 82% of the total emissions of 28 industries. MEIs mainly include C22, C17, C28, and other traditional industries with high energy consumption. LEIs are mainly light industry, i.e., C39 and other high-tech industries. The average emissions intensity of HEIs is 1.99 tons/10000CNY, 5.9 times of MEIs (0.34 tons/10000CNY), and 20.6 times of LEIs (0.1 tons/10000CNY). Therefore, HEIs are the focus of our following research.

Time Series Decomposition Analysis of Manufacturing Industry

In this section, we use the LDMI decomposition model to analyze the main drivers of manufacturing CO_2 emission changes in 2000–2017. According to the trend of CO_2 emissions from 2000 to 2017, we have conducted research in four stages, namely, 2000–2005, 2005–2010, 2010–2014, and 2014–2017.

Decomposition Results for the Whole Manufacturing Industry

During 2000–2017, the CO_2 emissions of China's manufacturing industry increased from 1.617 billion tons to 5.032 billion tons, with an increase of 211%. **Figure 2** shows that EA effect and EI effect are the main driving forces for the increase and decrease of CO_2 emissions, respectively. Industrialization and urbanization are inferred to for the driving factor behind the expansion of the manufacturing industry and related CO_2 emissions since 2000. In terms of CO_2 emission reduction, the energy intensity decreased rapidly during 2000–2017. EI effect led to a reduction of 582.83 billion tons of CO_2 emissions, accounting for 170.7% of the reduction of CO_2 emissions in the manufacturing industry. IP effect accounts for 12.8% of the change of CO_2 emissions in the manufacturing industry. The CO_2 emissions coefficient of fossil energy is basically constant, and the change mainly comes from the technological progress and the improvement of production efficiency of electric heating production. In the four stages, CO_2 emissions increased much more rapidly than in the following years. Among the five influencing factors, EA effect is the most important force behind CO_2 emissions.

Decomposition of Manufacturing Sub-Industries

This section decomposes the 28 sub-industries of the manufacturing industry to study their energy-related carbon emissions and explore the main forces driving carbon emissions changes in sub-industries. Although the cumulative contribution

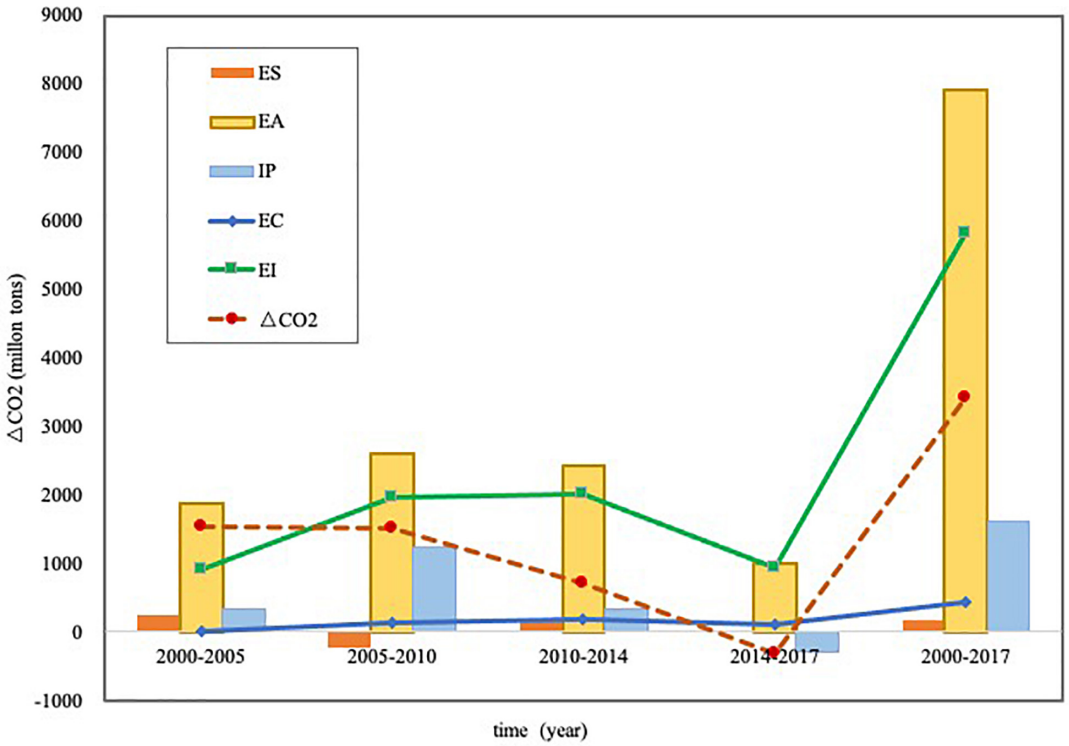


FIGURE 2 | Decomposition result of carbon emission changes in manufacturing industry.

of the influencing factors of carbon emissions in each sub-industry varied significantly between 2000 and 2017, the decomposition results from the 28 sub-industries show that the EA effect is the main force increasing carbon emissions, while the EI effect is the main contributor to carbon emissions reduction. As analyzed in the previous section 5.1, HEIs contribute a lot to carbon emissions, hence they are the focus of our following research.

For HEIs, as the rate of urbanization continues to accelerate in China, the demand for raw materials for infrastructure construction keeps increasing in 2000–2017, leading to the rapid development of HEIs and the increase of carbon emissions. As shown in **Figure 3**, the EC effect had a positive impact on the carbon emissions of HEIs from 2000 to 2017. The ES effect had a negative impact on the carbon emissions of C31 and a positive impact on other HEIs, but the impact was not significant. The EI effect is mainly driven by technological progress and remains relatively limited compared with the EA effect in 2000–2017. The EI effect of C31 reached −2417 million tons, followed by C30, C26, C25, and C32, which accounted for −161, −261.1, −127.9, −175.5, and −14.8 of the total changes in carbon emissions, respectively. The cumulative contribution of the EA effect in C31 is the largest of the 28 sub-industries, accounting for 226.6% of carbon emissions changes. Compared to the impact caused by EA effect, the expansion of the IP effect resulted in 738 million tons of carbon emissions from C31, followed by C30, C26, C25, and C32, which contributed 49.2, 49.9, 33, 44, and 13.3%, respectively.

Carbon Emissions Reduction Potential of Smart Grid

The EC effect is closely related to the energy consumption structure. By using smart power to produce clean energy, the proportion of thermal power generation can be significantly reduced. According to China's carbon emissions reduction commitment, the country will strive to achieve carbon neutrality in 2050. The report "2050 World and China Energy Outlook" issued by CNPC research institute points out that China's energy structure will continue to develop toward clean and low-carbon, clean energy will gradually replace coal after 2030, and the proportion of coal will drop to 47.1 and 32.4%, respectively, in 2030 and 2050. Therefore, based on the energy consumption structure and the carbon emissions produced in 2017 with the calculation formula (9) and (10), we set three scenarios: conservative estimation, normal estimation, and optimistic estimation. In the first scenarios, the raw coal power generation will be reduced by 30% on an existing basis; in the second and third, coal-fired power will be reduced by 50% and 70%, as shown in **Table 1**.

The use of a smart grid can help achieve a better emissions reduction effect. It is conservatively estimated that by applying a smart grid, the carbon emissions will be reduced by 543.25 Mt, which is 12% compared with the carbon emission of 4608.28 Mt in 2017. Under normal and optimistic conditions, carbon emissions will be reduced by 905.41 and 1267.58 Mt, which accounts for 19.65 and 27.51% in the manufacturing industry of 2017.

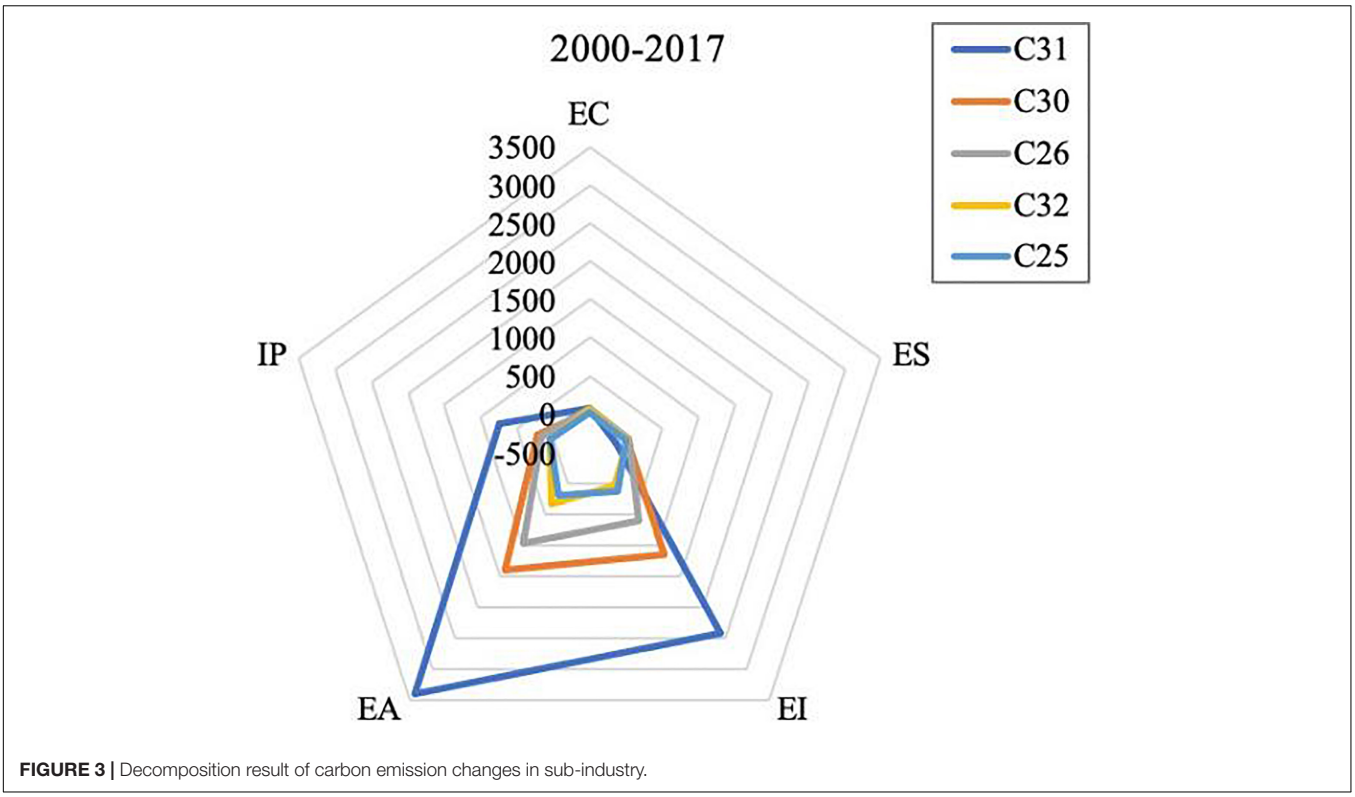


FIGURE 3 | Decomposition result of carbon emission changes in sub-industry.

TABLE 1 | Potential of Carbon Emission Reduction of Smart Grid in 2017.

Scenarios	The reduction of coal-fired power	The reduction of CO_2 emissions (Mt)	Change rate of carbon emission
Conservative estimation	30%	543.25	11.79%
Normal estimation	50%	905.41	19.65%
optimistic estimation	70%	1267.58	27.51%

Therefore, using a smart grid has great potential for carbon emissions reduction.

CONCLUSION AND POLICY IMPLICATIONS

Conclusions

In 2000–2017, the energy structure in which coal was dominant was the main reason for the increase of CO_2 emissions. In 2017, CO_2 emissions of HEIs accounted for 82% of total CO_2 emissions of the manufacturing industry, which are also major coal consuming industries. Economic activity (EA) effect and energy intensity (EI) effect are the main driving forces for the increase and decrease of CO_2 emissions in China's manufacturing industry respectively. Energy structure (ES) effect and industrial population (IP) effect promote the increase of CO_2 emissions, while CO_2 emissions coefficient (EC) effect alleviates CO_2 emissions. As our goal is to maintain green and sustainable development, a feasible way to reduce carbon emissions is to reduce energy intensity and optimize energy structure, rather than reduce economic activities.

Policy Implications

Firstly, restrict the development of high energy-consuming industries to reduce carbon emission intensity. It is necessary to implement more stringent threshold access and environmental impact assessment policies for high energy consumption enterprises and to restrict the expansion of the non-metallic mineral products industry and other high emission industries. It is also important to promote high energy consumption industries to continuously improve energy utilization efficiency.

Secondly, make efforts to optimize the energy structure. Gradually reduce the proportion of coal consumption in energy consumption, and control coal, oil, and natural gas in fossil energy consumption to an appropriate level. We should encourage the use of renewable energy sources, such as hydro-power and wind power, and appropriately encourage the development of nuclear power.

Finally, we should develop and use smart grid technology. By using smart grids, we can provide a strong power grid foundation and technical support for power transportation, enhance the strength of power grids, and access large-scale clean and renewable energy. This would then reduce the proportion of electric power, optimize the energy structure, and reduce carbon emissions.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

HF is the post-Ph.D. scholar who is mainly responsible for data processing and analysis, using decomposition model for decomposition analysis and thesis writing. YS proposes the research topic and helps with the methodology. YZ contributes to the analysis framework design, data visualization and analysis, and manuscript revision. All authors have read and agreed to publish the current version of the manuscript.

ACKNOWLEDGMENTS

The authors would like to thank the experts who attended the IEIS 2020 meeting for their valuable opinions on this manuscript. The draft of this paper was preprinted in the Springer Science and Business Media LLC, 2021 (Fu et al., 2021).

REFERENCES

Ang, B. (2005). The LMDI approach to decomposition analysis: a practical guide. *Energy Policy* 33, 867–871. doi: 10.1016/j.enpol.2003.10.010

Ang, B. W., Xu, X. Y., and Su, B. (2015). Multi-country comparisons of energy performance: the index decomposition analysis approach. *Energy Econ.* 47, 68–76. doi: 10.1016/j.eneco.2014.10.011

Cai, B., and Guo, H. (2018). Local strategies for China's carbon mitigation: an investigation of Chinese city-level CO2 emissions. *J. Clean. Prod.* 178, 890–902. doi: 10.1016/j.jclepro.2018.01.054

Darby, S., Strömbäck, J., and Wilks, M. (2013). Potential carbon impacts of smart grid development in six European countries. *Energy Effic.* 6, 725–739. doi: 10.1007/s12053-013-9208-8

Fu, H., Shi, Y., and Liu, J. (2021). "Research on the driving forces of carbon emissions in China's manufacturing industry: a multi-sector decomposition analysis," in *IEIS 2020 Proceedings of the 7th International Conference on Industrial Economics Systems and Industrial Security Engineering*, eds M. Li, G. Bohács, G. Hua, D. Gong, and X. Shang (Singapore, FL: Springer Singapore), 35–46.

Fu, L., Zhao, H., and Guo, S. (2012). An analysis on the low-carbon benefits of smart grid of China. *Physics Procedia* 24, 328–336. doi: 10.1016/j.phpro.2012.02.049

Kaya, Y. C. (1989). *Impact of Carbon Dioxide Emission on GNP Growth: Interpretation of Proposed Scenarios, Presentation to the Energy and Industry Subgroup, Response Strategies Working Group*. Paris: IPCC.

Lamiaa, A., and Tarek, E. (2013). Reducing carbon dioxide emissions from electricity sector using smart electric grid applications. *J. Eng.* 2013, 845–853. doi: 10.1155/2013/845051

Li, J., and Cheng, Z. (2020). Study on total-factor carbon emission efficiency of China's manufacturing industry when considering technology heterogeneity. *J. Clean. Prod.* 260:121021. doi: 10.1016/j.jclepro.2020.121021

Li, Y., Wang, C., Li, G., and Chen, C. (2021). Optimal scheduling of integrated demand response-enabled integrated energy systems with uncertain renewable generations: a Stackelberg game approach. *Energy Convers. Manag.* 235:113996. doi: 10.1016/j.enconman.2021.113996

Lin, B., and Du, K. (2014). Understanding the change of energy intensity in China: a comprehensive decomposition framework. *World Econ.* 4, 69–87.

Liu, Z., Guan, D., Wei, W., Davis, S. J., Ciais, P., Bai, J., et al. (2015). Reduced carbon emission estimates from fossil fuel combustion and cement production in china. *Nature* 524, 335H–338H. doi: 10.1038/nature14677

Shi, Y., Han, B., Zafar, M. W., and Wei, Z. (2019). Uncovering the driving forces of carbon dioxide emissions in Chinese manufacturing industry: an intersectoral analysis. *Environ. Sci. Pollut. Res.* 26, 31434–31448. doi: 10.1007/s11356-019-06303-7

Sun, L., Wang, Q., Zhou, P., and Cheng, F. (2016). Effects of carbon emission transfer on economic spillover and carbon emission reduction in China. *J. Clean. Prod.* 112, 1432–1442. doi: 10.1016/j.jclepro.2014.12.083

Tan, R., and Lin, B. (2018). What factors lead to the decline of energy intensity in China's energy intensive industries? *Energy Econ.* 71, 213–221. doi: 10.1016/j.eneco.2018.02.019

Xu, J., Fleiter, T., Eichhammer, W., and Fan, Y. (2012). Energy consumption and CO2 emissions in China's cement industry: a perspective from LMDI decomposition analysis. *Energy Pol.* 50, 821–832. doi: 10.1016/j.enpol.2012.08.038

Xu, R., and Lin, B. (2017). Why are there large regional differences in carbon emissions? Evidence from China's manufacturing industry. *J. Clean Prod.* 140, 1330–1343. doi: 10.1016/j.jclepro.2016.10.019

Zhou, P., and Ang, B. W. (2008). Decomposition of aggregate CO2 emissions: a production-theoretical approach. *Energy Econ.* 30, 1054–1067. doi: 10.1016/j.eneco.2007.10.005

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

frontiers in Energy Research

BRIEF RESEARCH REPORT
published: 04 August 2021
doi: 10.3389/fenrg.2021.672999

Research on Improved Fault Current Analysis Method for Flexible Direct Current Power Grid Considering Alternating Current Feed

Yinfeng Sun[1]*, *Xin Xiong*[1], *Zhenhao Wang*[1], *Guoqing Li*[1], *Xueguang Wu*[2,3] and *Weiru Wang*[1]

[1] Key Laboratory of Modern Power System Simulation and Control & Renewable Energy Technology, Ministry of Education (Northeast Electric Power University), Jilin, China, [2] Global Energy Interconnection Research Institute, Beijing, China, [3] Beijing Key Laboratory of DC Power Grid Technologies and Simulation, Beijing, China

Flexible direct current (DC) grid can realize large-scale renewable energy, wide-area coordinated complementation, and reliable power transmission. It is an important development that can be used to support high-voltage and large-capacity flexible DC transmission in the future. The short-circuit current of the DC line is one of the important bases for the selection of key main equipment parameters such as converter valves, DC circuit breakers, and reactors in the flexible DC grid. In this paper, a flexible DC grid equivalent circuit network model with alternating current (AC) feed-in is established. Aiming at the monopolar ground fault of the flexible DC grid grounded through the metal loop, an optimized traditional matrix calculation method is proposed to obtain the accurate line fault current value. On this basis, with an actual engineering background, the equivalent circuit model of the four-terminal bipolar flexible DC power grid is established, and the influence of grounding position, grounding parameters, and current-limiting reactor on the fault current of the DC line is analyzed. Finally, simulation using the PSCAD software verifies the effectiveness and accuracy of the proposed method. The method proposed in this paper can provide the necessary bases and references for the selection of flexible DC grid equipment.

Keywords: flexible DC grid, short-circuit current, AC feed, monopolar ground fault, matrix calculation

OPEN ACCESS

Edited by:
Liang Chen,
Nanjing University of Information Science and Technology, China

Reviewed by:
Shaoyan Li,
North China Electric Power University, China
Jia Cui,
Shenyang University of Technology, China

***Correspondence:**
Yinfeng Sun
273524083@qq.com

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

Received: *26 February 2021*
Accepted: *15 March 2021*
Published: *04 August 2021*

Citation:
Sun Y, Xiong X, Wang Z, Li G, Wu X and Wang W (2021) Research on Improved Fault Current Analysis Method for Flexible Direct Current Power Grid Considering Alternating Current Feed. Front. Energy Res. 9:672999. doi: 10.3389/fenrg.2021.672999

INTRODUCTION

The many advantages of voltage source converter-high-voltage, direct current (VSC-HVDC) make it suitable for new large-scale energy grid connection, grid interconnection (Allebrod et al., 2008; Perez et al., 2015; Guo et al., 2017; Li et al., 2021) and power quality control (Li et al., 2016), DC grid power flow control (Li et al., 2018), and other aspects. With the development of modular multilevel converter (MMC) in the direction of high-voltage, large-capacity, and long-distance power transmission, the use of overhead power transmission with obvious economic and technical advantages has become an inevitable choice. For flexible DC grids with overhead lines, the calculation and the protection of fault currents for DC lines have become an important research concern. Compared with the alternating current (AC) transmission system, the DC transmission system has a lower damping coefficient and response time constant. Therefore, it is hard to imagine

that once a DC fault occurs, the fault current will rise fast and have a large amplitude, which will seriously affect the safety and stability of the equipment and even the whole system (Franck, 2011; Liu et al., 2017). At present, the topological structure of half-bridge converters is mostly used in engineering, and the DC circuit breaker is used to remove the fault before converter station blocking, so as to reduce the adverse impact caused by converter station blocking, control the capacity, and reduce the cost. From the point of view of design, in order to verify the braking ability of the DC circuit breaker, it is necessary to accurately calculate the short-circuit current of the lines before the blocking of the converter station.

On the basis of the DC system, Shuai et al. carried out equivalent calculations for the capacitor discharge of the faulty module in a bipolar short-circuit fault and obtained some useful conclusions (Li et al., 2017). On the basis of the double-ended flexible DC system, they took into account the influence of the AC-side feed and converter control. Weiru et al. adopt the recursive equation method to calculate the short-circuit current from the perspective of the converter energy, making the calculation results of short-circuit current more accurate (Wang et al., 2019). However, the shortcoming of this method is that it tends to qualitative analysis and does not consider the fault current calculation in multi-terminal flexible DC grid scenarios. Using a general method for calculating short-circuit current of a DC grid, Li et al. constructed a bipolar flexible DC grid equivalent circuit network (Li et al., 2017). By establishing the initial matrix equations of the flexible DC grid before the fault and the modified matrix equations of the flexible DC grid after the fault, the transient characteristics of the short-circuit current of the flexible DC grid under different operating conditions are quantitatively analyzed. However, the shortcoming of this method is that only the bipolar short-circuit fault is considered and the influence of the AC grid is not involved. On the basis of the flexible DC power grid, the system model after the short-circuit fault was simplified in Tang and Dong (2019). However, the shortcoming of this method is that the influence of the AC feed is ignored. To sum up, in the current academic research on DC fault current, only a few researchers take the influence of the AC system into account; even if considered, there is no clear analytical expression. However, it is undeniable that this neglect and approximation will inevitably produce certain errors, which is not conducive to the accurate selection, precise control, and setting calculation of the DC equipment. The method of considering the influence of the AC feed into the calculation of DC short-circuit current cannot obtain the accurate analytical expression of DC short-circuit current, and the research on the quantitative analysis of the influence of injected power of the converter on fault current before DC grid failure is still blank. When volatile large-scale wind power, photovoltaic power, and other renewable energy sources are connected to the flexible DC grid, especially when the renewable energy adopts low voltage ride-through or fault ride-through strategies, fault current may be continuously fed to the fault point under the fault condition. Because of the higher probability of monopolar ground faults and the need for the influence of the grounding method to be considered, the calculation is more complicated. Therefore, this article focuses on a detailed and in-depth analysis of the grounding method through metal loops and monopolar ground faults used in actual projects.

A mathematical model of the MMC converter considering AC feed is established firstly on the basis of traditional AC and DC equivalent channels. The AC feed as a controlled DC source is incorporated into the calculation of DC capacitance and discharge, and the analysis method of single-pole grounding fault of metal loop grounding flexible DC network is improved. At the same time, the equivalent circuit model of bipolar flexible DC network is constructed. Secondly, the initial matrix equations of the flexible DC grid before the failure are established, and the initial matrix equations before the failure are revised twice to obtain the fault current correction matrix equations of the flexible DC grid after the failure. Finally, a simulation analysis of a typical example is performed to verify the correctness and effectiveness of the proposed method.

THE EQUIVALENT MODEL FOR FAULT CURRENT CALCULATION WITH INCLUSION OF AC FEED

When a pole bus or a converter fails in a bipolar symmetrical system structure, it only affects the fault pole and the converter connected with the fault pole and has little effect on the other pole, so as to effectively improve the reliability and safety of the system operation. This paper takes the positive electrode short-circuit fault as an example to illustrate. The detailed topology of an MMC converter station positive pole is shown in **Figure 1**.

In **Figure 1**, the voltage and current of each bridge arm are, respectively, denoted as u_{rj} and i_{rj} (subscript $r = p, n$, respectively, represent the upper and lower bridge arms; subscript $j = a, b, c$, respectively, represent the three phases, namely, a, b, and c, similarly hereinafter). In addition, m_{rj} and $u\Sigma\ rj$ represent the modulation signal and voltage of each bridge arm, respectively. i_{vj} represents the output AC current of the MMC. u_{gj} represents the voltage at the PCC point of the AC grid. L_t and R_t represent the leakage inductance and resistance of the connected transformer, respectively, while U_{dc} and i_{dc} represent the DC positive voltage and DC current, respectively. As shown in the figure, the directions of all electrical quantities are marked.

Under the condition of system equilibrium, according to the Kirchhoff Voltage Law (KVL), the voltage equations of the upper and lower bridge arms in a single-phase circuit can be obtained as follows (Harnefors et al., 2013):

$$\begin{cases} u_{gj} - L_t\frac{di_{vj}}{dt} - R_t i_{vj} - L_{arm}\frac{di_{pj}}{dt} - R_{arm}i_{pj} + u_{pj} = u_{dcp} \\ u_{gj} - L_t\frac{di_{vj}}{dt} - R_t i_{vj} + L_{arm}\frac{di_{nj}}{dt} + R_{arm}i_{nj} - u_{nj} = 0 \end{cases}. \quad (1)$$

The input ratios of the sub-modules in the upper and lower bridge arms of MMC are, respectively,

$$\begin{cases} m_{pj} = \frac{1-M\cos(\omega_g t+\varphi_j)}{2} \\ m_{nj} = \frac{1+M\cos(\omega_g t+\varphi_j)}{2} \end{cases}, \quad (2)$$

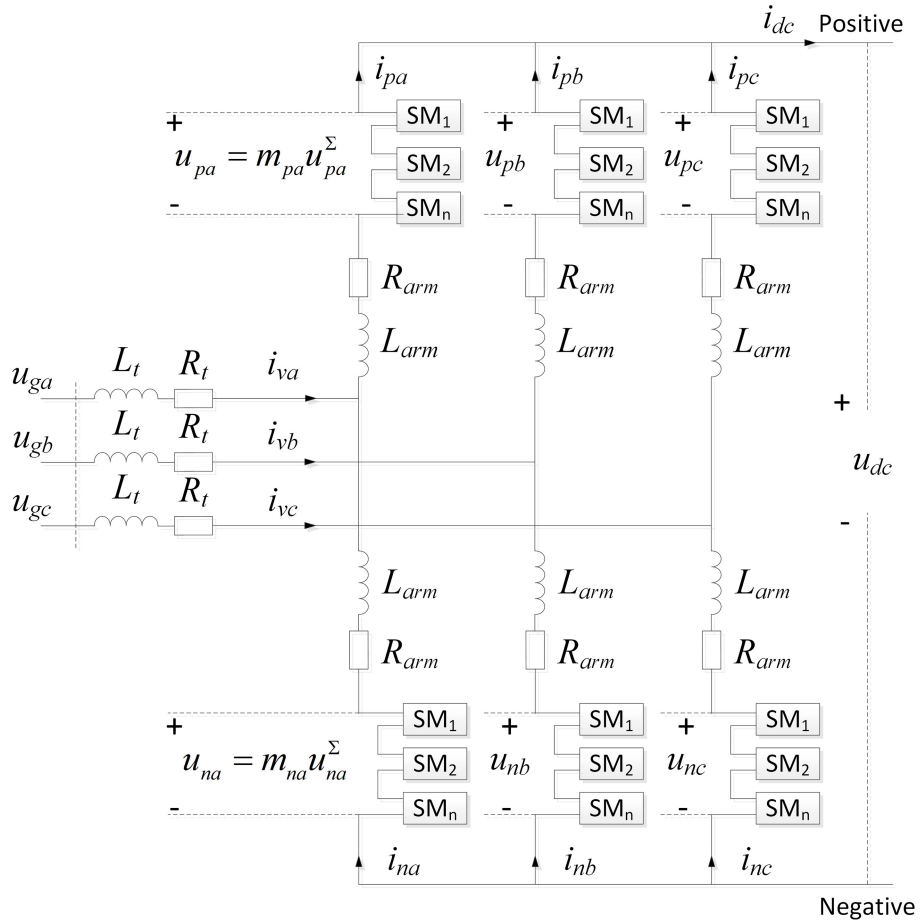


FIGURE 1 | Detailed topology of an MMC converter station positive pole.

where M is the modulation ratio; ω_g is the fundamental angular frequency; ϕ_j is the initial phase.

According to the current-voltage relationship in **Figure 1**,

$$\begin{cases} i_{vj} = i_{pj} - i_{nj} \\ u_{vj} = \frac{u_{nj} - u_{pj}}{2} \\ u_{comj} = \frac{u_{pj} + u_{nj}}{2} \\ i_{cirj} = \frac{i_{pj} + i_{nj}}{2} \end{cases} \tag{3}$$

where u_{vj} represents the output voltage of phase j of the MMC, also known as the internal potential of the MMC; u_{comj} is defined as the common-mode voltage of the upper and lower arms of phase j; and i_{cirj} represents the circulating current of phase j.

The mathematical equation of the dynamic characteristics of the equivalent circuit of the DC side of the MMC can be obtained by subtracting the Equation (1) as

$$L_{arm}\frac{di_{cirj}}{dt} + R_{arm}i_{cirj} = u_{comj} - \frac{u_{dcp}}{2}. \tag{4}$$

First, the three-phase superposition on the DC side dynamic Equation (4) of MMC is performed to obtain

$$L_{arm}\frac{d\sum_{j=a,b,c} i_{cirj}}{dt} + R_{arm}\sum_{j=a,b,c} i_{cirj} = \sum_{j=a,b,c} u_{comj} - \frac{3u_{dcp}}{2}. \tag{5}$$

In addition, as can be seen from the figure the relation between the sum of the three-phase circulation and the DC current is

$$\sum_{j=a,b,c} i_{cirj} = i_{dc}. \tag{6}$$

The equivalent capacitor voltage u_{ce} of the MMC is defined as

$$u_{ce} = \frac{2}{3}\sum_{j=a,b,c} u_{comj}. \tag{7}$$

Equation (6) and (7) are then substituted into Equation (5). Then, the equivalent circuit model of the DC side of the MMC can be initially obtained:

$$\frac{2}{3}L_{arm}\frac{di_{dc}}{dt} + \frac{2}{3}R_{arm}i_{dc} = u_{ce} - u_{dcp}. \tag{8}$$

The following two pairs of equations are introduced here to further establish a DC equivalent model taking into account the AC effects:

$$\begin{cases} u_{pj} = m_{pj} u_{pj}^{\sum} \\ u_{nj} = m_{nj} u_{nj}^{\sum} \end{cases} \tag{9}$$

$$\begin{cases} \frac{du_{pj}^{\sum}}{dt} = -\frac{N m_{pj}}{C_0} i_{pj} \\ \frac{du_{nj}^{\sum}}{dt} = -\frac{N m_{nj}}{C_0} i_{nj} \end{cases}. \tag{10}$$

The above equation describes the internal relationship between the output voltage of the bridge arm of the MMC and the sum of the capacitive voltages of all the sub-modules of the bridge arm (Harnefors et al., 2013).

Assuming that the capacitor voltages of the sub-modules are completely balanced, $u\Sigma\ pj \approx u\Sigma\ nj$; Equations (2), (3), and (9) are combined to obtain

$$\begin{aligned} u_{comj} &= \frac{1}{4}\left(u_{pj}^{\sum} + u_{nj}^{\sum}\right) + \frac{1}{4} M \cos\left(\omega_g t + \varphi_j\right)\left(u_{nj}^{\sum} - u_{pj}^{\sum}\right) \\ &= \frac{1}{4}\left(u_{pj}^{\sum} + u_{nj}^{\sum}\right). \end{aligned} \tag{11}$$

Upon substitution of Equation (11) back into Equation (7), expressing the equivalent capacitance voltage u_{ce} of MMC,

$$u_{ce} = \frac{2}{3}\sum_{j=a,b,c} u_{comj} = \sum_{j=a,b,c} \frac{1}{6}\left(u_{pj}^{\sum} + u_{nj}^{\sum}\right). \tag{12}$$

Derivation of the above equation and combination with Equation (10) can be obtained

$$\frac{du_{ce}}{dt} = -\frac{N}{6C_0}\sum_{j=a,b,c}\left(m_{pj} i_{pj} + m_{nj} i_{nj}\right). \tag{13}$$

Further, Equation (2) is substituted into the above equation for simplification. The obtained equation contains a more intuitive physical meaning.

$$\begin{aligned} C_e \frac{du_{ce}}{dt} &= \sum_{j=a,b,c} \frac{M\cos(\omega_g t + \varphi_j)}{2}\left(i_{pj} - i_{nj}\right) - \\ \sum_{j=a,b,c} \frac{i_{pj} + i_{nj}}{2} &= \frac{\sum_{j=a,b,c} u_{vj} i_{vj}}{u_{ce}} - i_{dc} = \frac{p_{con}}{2u_{ce}} - i_{dc}. \end{aligned} \tag{14}$$

At this time, $C_e = 6C_0/N$ is defined as the equivalent capacitance of MMC; C_0 represents the capacitance value of the converter sub-module; P_{con} represents the injected power of the converter station; i_{dcs} represents the controlled DC current source of the equivalent circuit of the DC side of the MMC.

According to Equation (14), the controlled DC current source can be calculated as

$$i_{dcs} = p_{con}/2u_{ce}. \tag{15}$$

In the steady-state operation of the system, the charge and discharge of the MMC capacitor maintain a dynamic balance, and the capacitor voltage fluctuation is relatively small, which causes the capacitor current i_{ce} to be approximately zero, that is

$$i_{ce} = -C_e \frac{du_{ce}}{dt} \approx 0 \tag{16}$$

By combining Equations (14) and (16), the following can be obtained:

$$i_{dcs} = p_{con}/2u_{ce} = i_{dc} = p_{dc}/u_{dc}, \tag{17}$$

where p_{dc} represents the active power on the DC side of the converter station, while u_{dc} represents the voltage between the DC poles.

At the same time, due to the minimal resistance of the MMC bridge arm, the internal loss of the MMC can be ignored. According to the power balance principle of the AC and DC sides, the following can be obtained:

$$\begin{cases} u_{ce} \approx u_{dcp} = u_{dc}/2 \\ p_{con} \approx p_{dc} \\ p_{con}/2u_{ce} \approx p_{dc}/u_{dc} \end{cases} \tag{18}$$

The theoretical analysis result of Equation (18) is the same as the mathematical deduction result of Equation (17), which verifies the correctness of the formula of the controlled DC current source.

IMPROVED ANALYTICAL METHOD OF FAULT CURRENT IN FLEXIBLE DC GRID

Improved Initial Matrix Analysis Method

A four-terminal DC equivalent circuit network model is provided in **Figure 2**, where L_e represents the equivalent inductance of the discharge circuit, while R_e represents the equivalent resistance of the discharge circuit; n_1, n_2, n_3, and n_4 represent the positive node of the converter station, n_1', n_2', n_3,' and n_4' represent the negative node of the converter station, and n_{10}-n_{20}-n_{30}-n_{40} represent the metal loop. The equivalent capacitance voltage matrix of the converter station of the flexible DC grid can be obtained by listing the KVL equations among the various discharge circuits:

$$A_0 U = R_0 I_0 + L_0 I_0', \tag{19}$$

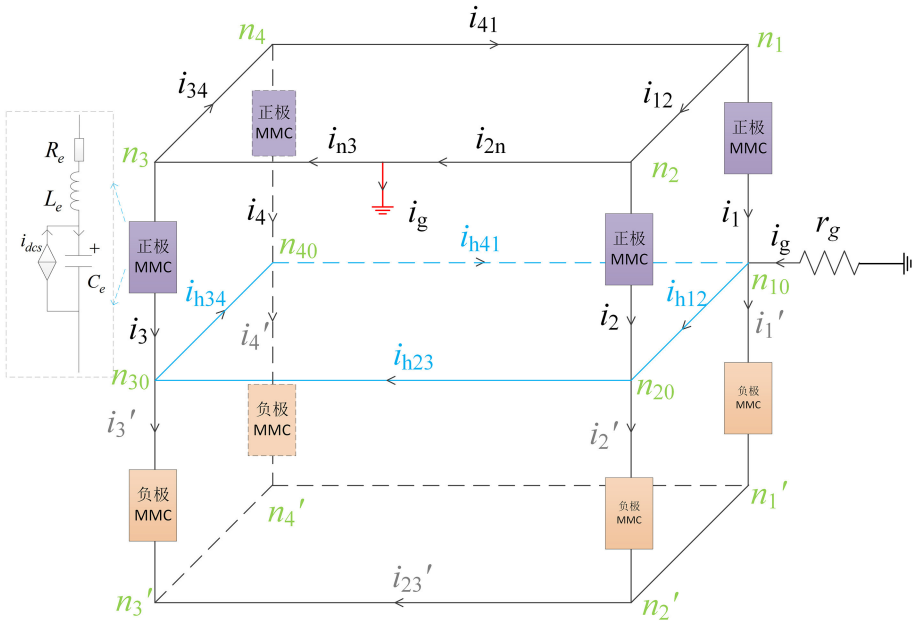


FIGURE 2 | A four-terminal DC equivalent circuit network model.

where $U = [u_{c1},u_{c2},u_{c3},u_{c4}]^{\mathrm{T}}$ represents the equivalent capacitance voltage matrix of the converter station; $I_0 = [i_{12},i_{23},i_{34},i_{41}]\mathrm{T}$ represents the line current matrix between the stations; A_0 represents the initial correlation matrix; R_0 represents the initial resistance matrix; and L_0 represents the initial inductance matrix. The parameters A_0, R_0, and L_0 can be represented in the matrix as follows:

$$\begin{cases} A_0 = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ -1 & 0 & 0 & 1 \end{bmatrix} \\ R_0 = \begin{bmatrix} R_{e1}+R_{12}+R_{e2} & -R_{e2} & 0 & -R_{e1} \\ -R_{e2} & R_{e2}+R_{23}+R_{e3} & -R_{e3} & 0 \\ 0 & -R_{e3} & R_{e3}+R_{34}+R_{e4} & -R_{e4} \\ -R_{e1} & 0 & -R_{e4} & R_{e4}+R_{41}+R_{e1} \end{bmatrix} \\ L_0 = \begin{bmatrix} L_{e1}+L_{12}+L_{e2} & -L_{e2} & 0 & -L_{e1} \\ -L_{e2} & L_{e2}+L_{23}+L_{e3} & -L_{e3} & 0 \\ 0 & -L_{e3} & L_{e3}+L_{34}+L_{e4} & -L_{e4} \\ -L_{e1} & 0 & -L_{e4} & L_{e4}+L_{41}+L_{e1} \end{bmatrix} \end{cases}, \tag{20}$$

where R_{ei} and L_{ei} represent the equivalent resistance and equivalent inductance of the i converter and R_{ij} and L_{ij} represent the line resistance and the line inductance between the i and j converters.

The relationship between equivalent capacitance current and DC line current in the converter station is as follows:

$$I_c = -A_0{}^T I_0 - I_{dcs}, \tag{21}$$

where $I_c = [i_{c1},i_{c2},i_{c3},i_{c4}]^{\mathrm{T}}$ represents the equivalent capacitance current matrix of the converter station, while $I_{dcs} = [i_{ac1},i_{ac2},i_{ac3},i_{ac4}]^{\mathrm{T}}$ represents the equivalent AC feed current matrix of the converter station.

The relationship between the equivalent capacitance voltage matrix of each converter station and the line current matrix is given as

$$U' = C \cdot I_c = C \cdot \left(-A_0{}^T I_0 - I_{dcs}\right). \tag{22}$$

The C matrix is defined as

$$C = \begin{bmatrix} 1/C_{e1} & 0 & 0 & 0 \\ 0 & 1/C_{e2} & 0 & 0 \\ 0 & 0 & 1/C_{e3} & 0 \\ 0 & 0 & 0 & 1/C_{e4} \end{bmatrix}. \tag{23}$$

Substituting Equation (15) into the above equation yields

$$U' = C \cdot \left(-A_0{}^T I_0 - P_0 U^{-1}\right), \tag{24}$$

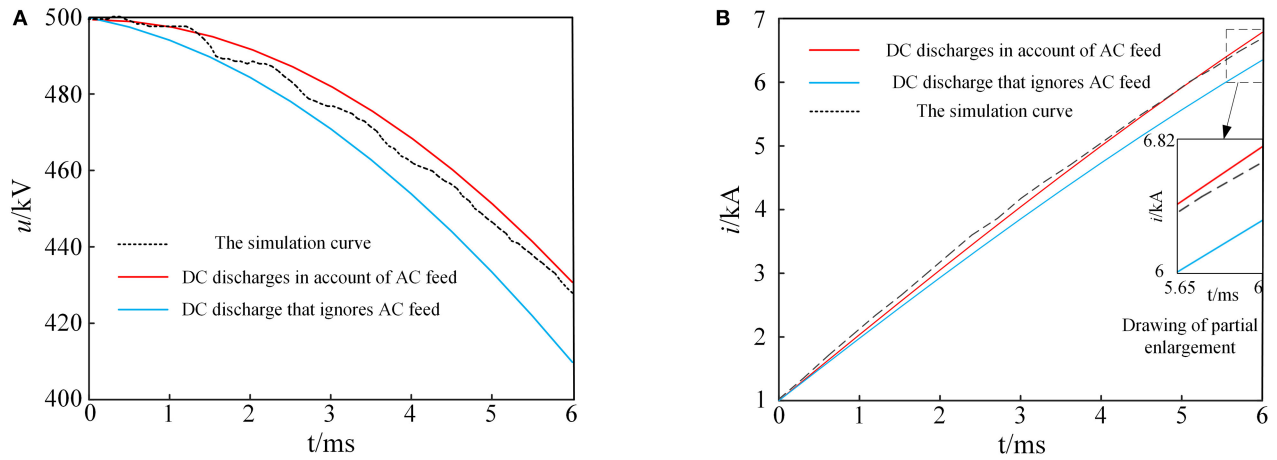


FIGURE 3 | DC capacitor voltage and DC fault current during monopolar ground fault. **(A)** Sub-module capacitance voltage. **(B)** DC fault current.

where U^{-1} represents the matrix of the reciprocal of the capacitor voltage of each converter station, i.e., $U^{-1} = [1/u_{c1},1/u_{c2},1/u_{c3},1/u_{c4}]^T$; P_0 represents the initial output power of each converter station matrix, and the parameters are as follows:

$$P_0 = \begin{bmatrix} p_{10} & 0 & 0 & 0 \\ 0 & p_{20} & 0 & 0 \\ 0 & 0 & p_{30} & 0 \\ 0 & 0 & 0 & p_{40} \end{bmatrix}_{4\times4}. \tag{25}$$

The initial matrix equations of the flexible DC grid are

$$\begin{cases} A_0 U = R_0 I_0 + L_0 I_0' \\ U' = C \cdot \left(-A_0{}^T I_0 - P_0 U^{-1}\right) \end{cases} \tag{26}$$

Improved Modified Matrix Analytical Method

Assuming that the short-circuit fault occurs on the branch b_{ij}, the distance between the fault location and the node n_i, denoted by x, accounts for the percentage of the total length of the faulted line. After a fault, the flexible DC grid will add a node, n_0, and the number of branches will increase by one, accordingly. The branch resistance matrix R and the inductance matrix L are increased by one row and one column, respectively; the branch resistance r_{ij} is updated to r_{i0} and r_{0j}, and the branch inductance L_{ij} is updated to L_{i0} and L_{0j}, as shown in Equation (27).

$$\begin{cases} r_{i0} = x r_{ij} \\ l_{i0} = x l_{ij} + L_l \\ r_{j0} = (1-x)\, r_{ij} \\ l_{j0} = (1-x)\, l_{ij} + L_l \end{cases} \tag{27}$$

The branch current i_{ij} is corrected to i_{i0} and i_{0j}, and the line current matrix between stations $I_0 = [\cdots,i_{ij}, \cdots]^T$ is corrected to $I_t = [\cdots,i_{i0},i_{0j}, \cdots]^T$. The U dimension of the equivalent capacitance voltage matrix of the flow station remains unchanged. Correspondingly, one row is added to the initial incidence matrix A_0.

The corrected equivalent capacitance voltage matrix of the converter station in the flexible DC grid is expressed as

$$A_t U = R_t I_t + L_t I_t. \tag{28}$$

If the metal loop is included in the fault circuit, the voltage drop of the metal loop needs to be taken into account, and the correction equation becomes

$$\begin{cases} A_t U = R_t I_t + L_t I_t + B U_h \\ U_h = R_h I_h \end{cases} \tag{29}$$

where B is the incidence matrix with a size of 5×4; $b_{ki} = 1$ is the kth loop that passes through the metal loop; $b_{ki} = 0$ is the kth loop that does not pass through the metal loop. R_h is the resistance matrix of each branch of the metal return line, which has a size of 4×4; I_h represents the current matrix of each branch of the metal return line, i.e., $I_h = [i_{h12},i_{h23},i_{h34},i_{h41}]^T$. The parameters of B and R_h are as follows:

$$\begin{cases} B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ R_h = \begin{bmatrix} r_{h12} & 0 & 0 & 0 \\ 0 & r_{h23} & 0 & 0 \\ 0 & 0 & r_{h34} & 0 \\ 0 & 0 & 0 & r_{h41} \end{bmatrix} \end{cases} \tag{30}$$

As shown in **Figure 2**, by listing the KCL relation of n-1 neutral points and the KVL relation of the metal return line, it can be

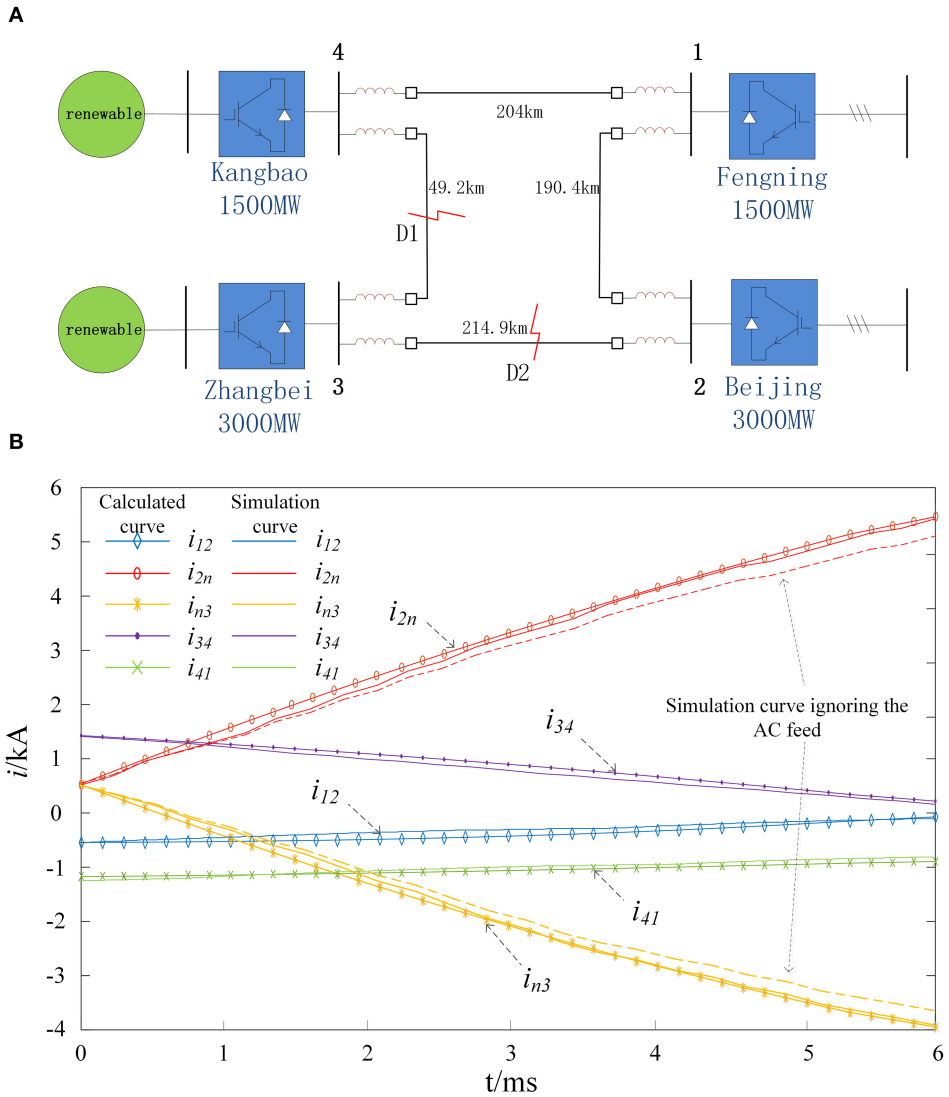


FIGURE 4 | Analysis of DC fault current in MMC flexible DC grid. **(A)** Four- terminal DC power grid topology. **(B)** DC line current during the monopolar ground fault.

concluded that

$$\begin{cases} i_g + i_{h23} - i_{h34} = i_{30} - (i_{n3} - i_{34}) \\ i_{h34} - i_{h41} = i_{40} - (i_{34} - i_{41}) \\ i_{h41} - i_{h12} = i_{10} - (i_{41} - i_{12}) \\ i_{h12} - i_{h23} = i_{20} - (i_{12} - i_{2n}) \\ i_{h12} \cdot r_{12} + i_{h23} \cdot r_{23} + i_{h34} \cdot r_{34} + i_{h41} \cdot r_{41} = 0 \end{cases} \tag{31}$$

The matrix expression of the metal loop current can be obtained from the above equation

$$I_h = R_1 I_t + R_2 I_{n0}, \tag{32}$$

where R_1 is the resistance matrix of size 4×5; R_2 is the resistance matrix of size 4×4; and I_{n0} is the current matrix of each converter station of sound pole, i.e., $I_{n0} = [i_{10}, i_{20}, i_{30}, i_{40}]^{\mathrm{T}}$

The equivalent capacitance voltage matrix equation of the converter station of the flexible DC grid can be obtained by substituting Equation (32) into Equation (29). By combining the line current matrix equation obtained before, the current of each line and the capacitor voltage of each converter station can be obtained, as given in the equation below.

$$\begin{cases} A_t U = (R_t + B R_h R_1) I_t + L_t I_t^{'} + B R_h R_2 I_{n0} \\ U^{'} = C \cdot \left(-A_t^T I_t - P_0 U^{-1}\right) \end{cases} \tag{33}$$

SIMULATION RESULTS

The model parameters correlate to the Zhangbei DC project in China as follows: rated DC voltage is ±500 kV, rated DC current is 3 kA, rated power are 1,500/3,000/3,000/1,500 MW, sub-module capacitances are 8/15/15/11.2 mF, bridge arm

inductances are 100/50/50/100 mH, the inductance of the DC bus smoothing reactor is 150 mH, and the inductance of the neutral smoothing reactor is 300 mH.

Modular Multilevel Converter-High-Voltage, Direct Current Fault Current Analysis

As shown in **Figure 3**, the calculation results of DC fault current not considering the influence of AC-side feed and the calculation results of DC fault current taking into account the influence of AC-side feed are compared and analyzed in this paper.

As shown in **Figure 3**, the difference between the DC discharge voltage which does not consider the influence of AC feed and the DC discharge voltage which takes the influence of AC feed into account is 20.20 kV, accounting for 4.04% of the rated voltage of the converter station. Meanwhile, the difference between the DC discharge current, not taking into account the influence of the AC feed, and the DC discharge current considering the influence of the AC feed is 547.8 A, which accounts for 8.12% of the line fault current. The results show that the capacitor voltage and the fault current curves of the sub-module while considering the AC influence are more consistent with the simulation curves.

Analysis of DC Fault Current in MMC Flexible DC Grid

As shown in **Figure 4**, a monopolar ground fault at the midpoint of the DC lines of converter station 2 and converter station 3 should be set, and the transient current waveforms of each DC line within 6 ms after the fault should be recorded. As shown in **Figure 4**, the line current calculation waveform of the improved method, the line current simulation waveform with the AC feed, and the line current simulation waveform without the AC feed are presented. The AC feed line current simulation waveform should be ignored and the circuit breaker should be set on the AC side. When the fault occurs, the AC circuit breaker acts to remove the influence of the AC feed on the fault line current. The dotted line is the simulation current of the line on both sides of the fault point, ignoring the AC feed. **Figure 4** shows that, the results obtained by the improved method in this paper are in good agreement with the simulation results, and the error of 6 ms current accounts for 0.6% of the fault current, which verifies the rationality and accuracy of the improved method. In addition ,it can be seen that, compared with the calculated value and simulation value of fault current considering the influence of AC feed, the error of ignoring AC feed is about 300A, accounting for about 6% of the peak fault current.

CONCLUSIONS

In this paper, research on an improved fault current analysis method of the DC grid with the AC feed taken into account is proposed, and the effectiveness of the method is verified by applying the method to an actual project. The main work performed on this paper and conclusions are summarized as follows:

(1) The mathematical model of the DC discharge equivalent circuit with the inclusion of the AC feed is established, and the current source on the AC side of the model is equivalent.
(2) The results obtained by the improved method in this paper are in good agreement with the simulation results, which not only verifies the influence of the AC feed on the fault current of the DC line but also proves the rationality of the improved method.
(3) As can be seen, omitting the influence of the AC feed will reduce the calculation accuracy of the DC line short-circuit current, which will inevitably affect the accuracy of the DC grid equipment selection and control parameter setting.

To sum up, the fault current analysis method of a flexible DC grid proposed in this paper can be used to provide necessary bases and references for the equipment selection and the parameter setting of a flexible DC grid.

With the further-increasing new energy capacity access to the DC grid, its volatility and uncertainty can easily lead to a wide range of fluctuations in AC feed power and DC power flow, and the resulting changes in the DC fault current cannot be ignored, especially when considering the large-scale off-grid or grid-connected conditions of wind power, photovoltaic power, etc. Of course, at this time, the adaptability of the method in this paper should be discussed in depth based on the respective fault crossing strategies of power supply and DC grid, which is also the research topic that the authors are carrying out at present.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

YS, ZW, and GL contributed to conception and design of the study. XX organized the database. YS and XW performed the statistical analysis. YS wrote the first draft of the manuscript. YS, XX, XW, and WW wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This work was supported by the National Key R&D Program of China (No. 2018YFB0904600), Jilin Provincial Department of Education Research Project (JJKH20190704KJ), and Key Project of Jilin Science and Technology Bureau (2019301163).

REFERENCES

Allebrod, S., Hamerski, R., and Marquardt, R. (2008). *New Transformerless,Scalable Modular Multilevel Converters for HVDC-Transmission[C]//Power Electronics Specialists Conference(PESC)* (Rhodes: Institute of Electrical and Electronics Engineers), 174–179. doi: 10.1109/PESC.2008.4591920

Franck, C. M. (2011). HVDC circuit breakers: a review identifying future research needs. *IEEE Trans. Power Delivery* 26, 998–1007. doi: 10.1109/TPWRD.2010.2095889

Guo, C., Liu, W., Zhao, C., and Iravani, R. (2017). A frequency-based synchronization approach for the VSC-HVDC station connected to a weak AC grid. *IEEE Trans. Power Delivery* 32, 1460–1470. doi: 10.1109/TPWRD.2016.2606495

Harnefors, L., Antonopoulos, A., Norrga, S., Angquist, L., and Nee, H.-P. (2013). Dynamic analysis of modular multilevel converters. *IEEE Trans. Indust. Electron.* 60, 2526–2537. doi: 10.1109/TIE.2012.2194974

Li, B., Zhou, S., Xu, D., Yang, R., Xu, D., Buccella, C., et al. (2016). An improved circulating current injection method for modular multilevel converters in variable-speed drives. *IEEE Trans. Indust. Electron*. 63, 7215–7225. doi: 10.1109/TIE.2016.2547899

Li, C., Zhao, C., Xu, J., Ji, Y., Zhang, F., and An, T. (2017). A pole-to-pole short-circuit fault current calculation method for DC grids. *IEEE Trans. Power Syst.* 32, 4943–4953. doi: 10.1109/TPWR.S.2017.2682110

Li, S., Guo, C., Zhao, C., and Xu, J. (2017). A novel MMC topology with lower power loss and DC fault ride-through capability. *Proce. CSEE* 37, 1–9. doi: 10.13334/j.0258-8013.pcsee.162109

Li, Y., Li, Y., Li, G., Zhoa, D., and Chen, C. (2018). Two-stage multi-objective OPF for AC/DC grids with VSC-HVDC: incorporating decisions analysis into optimization process. *Energy* 147, 286–296. doi: 10.1016/j.energy.2018.01.036

Li, Y., Wang, C., Li, G., and Chen, C. (2021). Optimal scheduling of integrated demand response-enabled integrated energy systems with uncertain renewable generations: a Stackelberg game approach. *Energy Conver. Manage.* 235:113996. doi: 10.1016/j.enconman.2021.113996

Liu, G., Xu, F., Xu, Z., Zhang, Z., and Tang, G. (2017). Assembly HVDC breaker for HVDC grids with modular multilevel converters. *IEEE Trans. Indust. Electron.* 32, 931–941. doi: 10.1109/TPEL.2016.2540808

Perez, M. A., Bernet, S., Rodriguez, J., Kouro, S., and Lizana, R. (2015). Circuit topologies,modeling,control schemes,and applications of modular multilevel converters. *IEEE Trans. Indust. Electron.* 30, 4–17. doi: 10.1109/TPEL.2014.2310127

Tang, L., and Dong, X. (2019). An approximate method for the calculation of transmission line fault current in MMC-HVDC grid. *Proc. Chin. Soc. Electr. Eng.* 39:181327. doi: 10.13334/j.0258-8013.pcsee.181327

Wang, W., He, Z., Li, G., Xin, Y., and Gu, H. (2019). Recursive calculation method of MMC-HVDC DC fault current with AC effect. *Chinese J. Electric. Eng.* 39, 313–320. doi: 10.13334/j.0258-8013.pcsee.191045

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

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership