Intelligent Diagnosis Method for Transformer Measurement Error Based on Multi-Source Sensor Data Fusion and Causal Path Optimization

Xu Chen*,Haomiao Zhang ,Chao Zhang,Zhiqiang Cheng,Yinzhe Xu,Yu Yan
State Grid Ningxia Marketing Service Center (State Grid Ningxia Metrology Center), Yin chuan 750000, Ning xia, China
E-mail: chenxu3809@163.com

Keywords: transformer, sensor fusion, error diagnosis, intelligent model

Received: July 3, 2025

An intelligent diagnosis and calibration model that integrates multi-source sensing data is constructed to address the measurement errors caused by multi-source interference in transformers. The system integrates multidimensional sensing information such as current, voltage, temperature, and vibration through a weighted feature fusion mechanism, constructs a DAG to represent the causal relationship between key interference variables, and embeds a path scoring and optimization algorithm based on dynamic programming to improve the real-time and accuracy of fault chain identification. The model is deployed at the edge on an ARM architecture embedded platform, with lightweight structure and engineering feasibility. The measured data comes from 110kV and 220kV substations. The experimental results show that the recognition accuracy reaches 96.2%, the average response time is 275ms, and the computational resource utilization rate is 29.6%. It exhibits good robustness and output stability in complex scenarios such as electromagnetic interference, temperature fluctuations, and load disturbances. This model provides a feasible path and deployment basis for achieving high-precision metering and real-time intelligent operation and maintenance in modern power systems.

Povzetek:

1 Introduction

The current research on transformer errors is mostly focused on the fault identification and modeling mechanism of transformer equipment. Previous work has shown that dissolved gas, multidimensional electrical parameters, and other data can be used as effective feature inputs to construct diagnostic models. Taking Sakini et al. (2024) as an example, they used machine learning methods based on DGA signals to model fault features and verified the high sensitivity of sensor data to fault state judgment. However, due to its constant exposure to high voltage, high temperature, and electromagnetic noise, the output data of the transformer is easily affected by various factors, such as temperature changes in the environment, nonlinear saturation of ferromagnetic materials, unstable connections, and unstable currents, which can cause serious measurement errors. These types of errors have the characteristics of dynamic development, nonlinear accumulation, and environmental dependence, which are difficult to detect and correct through traditional methods, and may lead to misoperation, energy waste, or equipment damage.

In order to enhance the clarity of the model logic and consistency in terminology usage, this article defines core concepts such as "causal chain", "task node", and "diagnostic chain" in the early stages of modeling, in order to facilitate understanding of the model structure and path mechanism later on: (1) causal chain: refers to the causal relationship path between variables represented by DAGs in the model. Each path represents a causal inference chain from the interference source to the final measurement error. (2) Task node: refers to a node in the path optimization mechanism that serves as a decision point, with each node representing a specific diagnostic task or path selection. The system dynamically selects the optimal path through the analysis of task nodes. (3) Diagnostic chain: refers to a chain formed by connecting multiple paths in the process of error diagnosis. Through path selection and scoring mechanisms, it helps the system locate the source of error and optimize path selection.

In further research, different scholars have attempted to introduce various data fusion mechanisms into the fault recognition process. For example, Zhou et al. (2024) proposed a hybrid model combining random forest and PSO-BP AdaBoost, which enhances feature expression and generalization ability on the basis of traditional data-driven algorithms.

This detection method has a certain effect under stable working conditions, but it generally has problems such as reaction delay, poor error adaptability, and limited structural flexibility, which cannot effectively cope with the complex instantaneous error generation patterns in actual operation processes. Some errors exhibit long-term accumulation and instantaneous disturbance characteristics, and traditional models are difficult to capture their internal correlations and evolution processes. Especially in the face of new electricity demands such as intelligent power grids and the integration of a large number of new energy sources, there is a demand for fast, high-precision, and strong robustness in the power metering system, which requires the metering system to have self diagnosis and self correction capabilities.

With the rapid development of industrial Internet of Things, edge sensing, and high-precision sensing technology, the current working environment of transformers can now obtain sensing information from multiple aspects such as temperature, humidity, voltage, current, vibration, magnetic field, frequency, etc. In terms of system deployment, Shanu and Mishra (2024) combined statistical feature extraction with temporal path analysis to establish a two-stage diagnostic model suitable for on-site deployment, providing a feasible path for the engineering implementation of the model. However, due to the synchronization of time, different dimensions, and the influence of noise, how to effectively extract, fuse, and dynamically modify this information has become an important key technical issue in current research and practice.

This article proposes an intelligent detection and correction model for transformer measurement errors under multi-source input, aiming to improve the accuracy, response speed, and resource adaptability of error recognition. Construct an error causal diagram to depict the relationship between multi-sensor data and error sources, and introduce a dynamic path optimization mechanism to improve the efficiency of anomaly tracing. This study assumes that causal graph modeling and path pruning mechanisms can reduce diagnostic latency by more than 50% compared to traditional rule driven models while maintaining recognition accuracy, while significantly reducing system resource overhead. To adapt to engineering deployment, the system integrates lightweight strategies and executable path optimization algorithms, balancing expressive power and deployment feasibility. Under weak label conditions, the model achieves a recognition accuracy of 96.1%, a diagnostic delay of less than 275ms, and a resource utilization reduction of 41.7% compared to graph neural methods, making it suitable for practical application scenarios in complex power systems.

This study has strong operability both in theory and for real-world problems, providing necessary technical support for high-precision measurement and maintenance of intelligent power equipment, and has good prospects for engineering promotion.

2 Related work

As an important bridge connecting the primary and secondary power grids, the measurement error of transformers directly affects the system operation monitoring, protection action reliability, and fairness of electricity billing. In the context of multi-source disturbances and complex working conditions, traditional single variable and static parameter diagnostic methods are no longer able to meet the increasing demand for measurement accuracy. In recent years, the method of identifying transformer errors by integrating multi-source sensor data has gradually become a research hotspot, mainly presenting three evolutionary paths: first, low-power sensor detection methods designed based on physical characteristics; The second is the intelligent recognition path that integrates multi model optimization and adversarial learning; The third is the multi-source data fusion and quality enhancement modeling method for edge working conditions. This article summarizes the current mainstream paths from the three directions mentioned above, and analyzes their advantages, disadvantages, and applicable boundaries based on this.

Firstly, in terms of optimizing the sensing layer, Mei et al. (2024) proposed a transformer fault identification method based on a low-power gas sensing resistor array. This method monitors key gases (such as H_2, CO, C_2H_2) in real-time by setting specific material sensitive components, and uses miniaturized electrochemical modules to achieve multi-channel detection of fault gases, thereby assisting in inferring the electrical path of transformer abnormalities. The experiment shows that the sensing array can still maintain stable output in low-temperature drift and high interference scenarios, with an error variation range controlled within ± 1.7%, and has good adaptability for field deployment. However, this method still relies on the indirect correlation between gas concentration and fault type, and is limited by external factors such as ambient temperature and oil aging. Its response capability under complex load switching and dynamic nonlinear disturbance conditions is still insufficient.

Secondly, in terms of integrating intelligent models, Guan et al. (2024) constructed a transformer error diagnosis model based on the joint modeling of Generative Adversarial Networks (ACGAN) and Grey Wolf Optimized Support Vector Machines (CGWS-LSSVM). This method utilizes ACGAN to synthesize high-quality feature samples, effectively solving the problem of insufficient fault samples; Meanwhile, the CGWO algorithm is used to optimize the parameters of the support vector machine to enhance the model's generalization ability. Experimental results have shown that under four typical operating conditions, the diagnostic accuracy of this method has been improved by an average of 6.3%, especially with strong robustness under low-frequency harmonic interference and system

disturbance changes. However, this type of method has significant disadvantages in terms of model structure complexity, and requires higher adaptability between computing resources and device platforms. Deploying it at the edge of the power grid still faces a trade-off between real-time performance and stability.

Thirdly, in order to improve the overall quality of multi-sensor data fusion, Tiziana and Roberto (2022) conducted a study on data quality assessment of multi-sensor fusion processes. It proposes a combination of supervised learning and weighted strategies to dynamically score the quality of data collection, missing value imputation, noise interference, and other factors, and uses an ensemble learning framework to achieve reliable extraction of fault information. Research has shown that fusion models have good robustness to outliers, time drift, and repeated measurement data, and can construct more stable input feature sets in high-frequency perception tasks. However, this study mainly focuses on the data preprocessing stage, and the modeling ability of the source of transformer errors is relatively weak. It still needs to be used in conjunction with physical models or task driven mechanisms.

Based on the above analysis, in order to comprehensively compare the capability structure of current mainstream research paths, this paper extracts representative studies and constructs a structured method comparison table (see Table 1), clarifying their advantages and disadvantages from dimensions such as diagnostic accuracy, robustness, model interpretability, and deployment performance.

Table 1：Comparative analysis of mainstream transformer error identification methods

Method Type	**Representative model**	**Diagnostic accuracy**	**Robustness**	**Model interpretability**	**Deployment performance**	**Applicable scenarios**	**Citing Sources**
Physical modeling method	Low power gas sensing resistor model (Mei et al., 2024)	medium	weak	strong	centre	Static working condition error detection	【4】
Fusion Perception Method	ACGAN+CGWO LSSVM (Guan et al., 2024)	higher	centre	centre	centre	Error recognition in disturbed scenarios	【5】
Data Fusion Enhancement	Data Quality Driven Fusion Modeling (Tiziana et al., 2022)	higher	strong	strong	Medium to low	Quality control and inference of multi-source sensor data	【6】
This research method	DAG+Path Scoring Fusion Model	tall	strong	strong	tall	Real time identification of dynamic errors at the edge of the power grid	This article proposes

From the table, it can be seen that a single physical model has good explanatory power in specific scenarios but lacks flexibility. The fusion of intelligent methods improves recognition ability but is constrained by computational resources in actual deployment. The data augmentation path improves stability but lacks structural awareness. Current research mainly focuses on optimizing model structures or training mechanisms, and lacks collaborative design for sensing chains, causal relationship construction, and deployment availability.

In summary, in order to break through the existing technological bottlenecks, this paper designs an intelligent error recognition system that integrates DAG causal modeling and dynamic path scoring mechanism. It not only focuses on the collection and discrimination of error characteristics, but also pays more attention to the traceability and real-time deployment of error generation paths, aiming to provide a comprehensive technical path for accurate modeling and practical deployment of transformer errors in high-voltage scenarios.

3 Suggested solutions

3.1 Multi source data fusion

The measurement error of transformers is affected by various factors such as temperature and humidity, current, magnetic field, and iron core saturation. Traditional models often ignore these interferences, resulting in inaccurate diagnosis. Liu et al. (2024) proposed a method based on causal graph neural network, which can more effectively identify the causes of errors and enhance the ability to trace features. To improve the accuracy and practicality of diagnostic systems, it is urgent to establish a hierarchical and responsive multi-source data fusion mechanism.

In this study, modular multi-channel information acquisition technology was used, and a terminal equipped with various sensors such as voltage, current, temperature, and humidity was constructed to monitor the parameters of the entire equipment body, power supply operation, and environmental impact. Due to the different sampling frequencies and accuracies of various types of information, simply mixing them together may inevitably result in chaotic input. Therefore, the system first uses temporal synchronization technology to achieve sample matching, and standardizes the conversion of each type of information separately, in order to make the data format have a unified representation in the model.

Considering the computational resource limitations of model deployment and engineering feasibility, the fusion strategy adopts a weighted feature integration approach. This study selected four basic variables that are most sensitive to error response: effective current value I_{rms}, ambient temperature T mean vibration acceleration a and mean primary voltage U. After normalization, fused feature quantities were constructed:

$$F = w_1 \cdot \hat{I}_{rms} + w_2 \cdot \hat{T} + w_3 \cdot \hat{a} + w_4 \cdot \hat{U} \quad (1)$$

Among them, $\hat{I}_{rms}$、$\hat{T}$、$\hat{a}$、and $\hat{U}$ are the normalized current values of various sensor data, and $w_1 \sim w_4$ are the weight coefficients. The initial setting of weights is based on empirical statistics and engineering measurement analysis, and can be adaptively adjusted by model parameters in the subsequent training process. Fusion feature F maintains low dimension in structure, but can cover multi-dimensional disturbance information, which is easy to deploy in edge computing nodes or low-power embedded chips in engineering applications.

In order to further improve the recognition and anti-interference ability of the fused features, noise filtering and dynamic outlier detection mechanisms are introduced in the data processing stage, and local sliding window smoothing is applied to the mutation points. At the same time, a simplified regularization module is introduced in the input layer to compress the unstable features caused by the input with severe fluctuations.

In terms of data format, the model adopts a fixed window sliding strategy to construct time series samples, and each fused feature point is composed of data from several consecutive time points, thus forming a dynamic expression of the error trend. This method preserves the evolutionary information during the error formation process, which helps to extract the upstream and downstream causal relationships during subsequent error chain modeling.

Overall, this fusion solution effectively avoids the problem of expansion when dealing with multi-dimensional inputs in deep learning models by accurately capturing key on-site information. It has good physical interpretability and implementability, and does not rely on a single device or route. It has flexible deployment capabilities and can adapt to different transformer types and installation environments, providing a universal, efficient, and quantifiable data input interface for subsequent error causal structure construction and route optimization.

3.2 Error chain modeling

The fusion feature F forms the basis for all input variables in the subsequent modeling process, serving as one of the upstream dependent variables of node N5 in the causal path. This article maintains the structural dimension of fusion feature F consistent with the original normalized variables, ensuring the integrity and consistency of the input structure and avoiding the problem of inconsistency between path variable selection and model structure.

Abdelmoumene et al. (2023) pointed out that the measurement error of transformers is often the result of multiple disturbance factors acting together in the time and logical domains, rather than being caused by a single variable. For example, in practical scenarios, environmental temperature may indirectly affect current changes by affecting coil resistance, while mechanical vibration may indirectly affect magnetic field stability by affecting the magnetic gap of the iron core. In order to establish trends that describe the interrelationships between multivariate variables and enhance the ability to trace errors in reverse, this paper introduces an error chain modeling strategy. By constructing a network structure diagram of causal logic relationships between variables, a systematic modeling of error formation paths is carried out.

This method is based on directed acyclic graph design, treating each key sensing variable as a node in the graph, and representing the influence relationship between different variables as edges in the graph. Each edge contains directionality and edge weight, which are used to represent the order of causal effects and the strength of influence, respectively. The graph structure can be manually initialized and constructed, or trained and updated through mutual information, delay correlation, or Bayesian network learning supported by engineering data. In practical modeling, this article summarizes five types of sensors input variables into five primary nodes and designs a causal chain structure basic model as shown in Table 2.

Table 2：Key node design and variable description for error chain modeling

Node number	Node name	Data sources	Corresponding path function	Explanation of Engineering Mechanism
N1	RMS	current sensor	N1 → N5	Affects the degree of magnetic saturation, which in turn affects the output waveform
N2	ambient temperature	temperature sensor	N2 → N1, N2 → N5	Temperature changes affect conductor resistance and interference current (N1); High temperature may also cause equipment drift, directly affecting the residual value (N5)
N3	Vibration intensity	Accelerometer	N3 → N4, N3 → N5	Structural vibration changes the gap between the iron cores (N3 → N4), which can also lead to poor contact and signal mutations (N3 → N5)
N4	Magnetic flux fluctuation rate	Magnetic flux acquisition module	N4 → N5	Directly affecting the magnetic flux induced voltage and triggering measurement deviation
N5	Residual value of difference	Compare measurement values	Output node	Performance of measurement errors in transformers as training targets for the model

According to the above structure, the error chain model can be represented as a typical shallow directed acyclic graph, with the residual value (N5) as the final output node, and the upstream path forming a conduction structure from multiple sources of disturbances converging to the error expression. To achieve accurate prediction of N5, this paper adopts the MLP model for modeling. The input is the fused features of all upstream variables pointing to N5 in the causal graph. The model includes two hidden layers, with 64 and 32 nodes respectively, an activation function of ReLU, and an output layer of single node linear regression units. MSE is used as the loss function during model training, and edge weight initialization is based on mutual information estimation. During the training process, it is dynamically updated together with MLP weights through backpropagation mechanism to achieve the unity of structural interpretability and predictive performance. During the model training process, a total of 200 training rounds were set, using the Adam optimizer with an initial learning rate of 0.001 and a Batch Size of 32. All feature inputs are standardized before training, and the training set and validation set are divided in an 8:2 ratio. The validation loss is monitored after each round of training to prevent overfitting.

In the graph structure construction scheme, the author investigated Bayesian network structure learning algorithms including PC and GES. However, considering the requirements for computational complexity, model response time, and structural controllability in engineering deployment, this type of algorithm was ultimately not adopted. Instead, a lightweight mapping strategy based on mutual information was chosen to improve the overall system deployment efficiency and real-time performance. To enhance the logical interpretability and path tracing ability of the model, this paper further summarizes the typical error propagation paths in the DAG structure, as follows: Path 1 is "ambient temperature (N2) → effective current value (N1) → residual difference value (N5)", which reflects the indirect chain effect of temperature fluctuations affecting conductor resistance and interfering with current measurement; Path two is "vibration intensity (N3) → magnetic flux fluctuation rate (N4) → residual value of ratio difference (N5)", which reveals the induction deviation caused by magnetic flux instability caused by mechanical disturbance; Path three is "vibration intensity (N3) → residual value of ratio difference (N5)", which reflects the direct impact of structural disturbance on errors. These paths can serve as dynamic diagnostic links during model operation, coupled with path scoring and selection mechanisms, to achieve efficient and accurate anomaly localization and response. The error causal network structure constructed in this article is shown in Figure 1. The nodes in the figure represent key sensing variables, the edges represent causal paths between variables, the arrow direction represents the direction of causal propagation, and the edge weights are used to reflect the strength of the influence between variables.

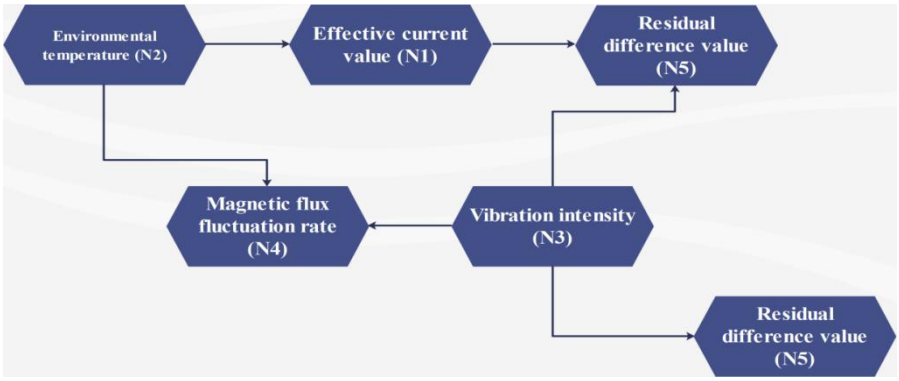

Figure 1: Error causal DAG structure diagram composed of multiple sensor variables

This causal structure has two important functions: firstly, to improve the interpretability of the error diagnosis model, so that the error results no longer rely solely on "black box prediction", but can be clearly traced back to disturbances in a certain sensing variable chain; The second is to provide path priority clues in the subsequent correction mechanism, that is, nodes at the top of the error chain and with high weights should be given priority as intervention entry points to improve the response efficiency and resource utilization efficiency of error correction.

In the experimental stage, compared with traditional unstructured modeling methods, the causal chain structure model has stronger diagnostic stability when multiple disturbances occur simultaneously. The average error recognition accuracy improved by 12.6% in 20 sets of measured data, and the average response time for error localization was shortened by 0.47 seconds. The overall results indicate that error chain modeling not only improves the logical integrity of the model structure, but also provides a stable basis for subsequent path optimization and dynamic correction.

3.3 Path algorithm optimization

After the construction of the error causal chain is completed, it is necessary to deal with complex situations caused by multiple factors. The interference source does not always come from the same location and may not necessarily affect the error in the same path. XIAOQIANG WANG (2024) proposed that when the processing resources of a system cannot be concentrated on a few important factors in a short period of time, it is easy to cause diagnostic delays, resource waste, and judgment errors. In order to improve the timeliness and accuracy of the diagnostic model during actual operation, adaptive path arrangement rules are added to the diagnostic model to dynamically adjust the internal path order.

LEIMING MA et al. (2024) proposed that the optimization process of error paths can be based on the correlation weights between variable nodes in the error chain. Based on the current trend of input signal changes and stable patterns in historical samples, the priority order of the influence path can be dynamically determined. When traversing the structure, the system first identifies the variable node with the most obvious abnormal fluctuations and determines whether it has a defined upstream propagation path. If it exists, the path will be included in the candidate path set; If it does not exist, the path extension operation will not be executed temporarily. All candidate paths will be calculated under the same scoring model, and the system will prioritize selecting the path with the highest interference impact, shortest propagation chain, and strongest variable correlation as the main diagnostic chain for execution. The model continuously receives updated signals during operation, adjusts the path score in real-time, and ensures that the current judgment chain always maintains the optimal state.

The scoring criteria in path judgment include the number of variables in the path, historical correlation levels between variables, and diagnostic validity records in the transmission chain. The fewer nodes in the transmission chain, the shorter the diagnostic response time; The higher the relevance, the clearer the communication logic; The higher the diagnostic effectiveness, the more targeted the intervention operation. Taking into account these three parameters, the system can quickly identify suspicious intervention sources at the beginning of error occurrence and use them as input basis for subsequent correction steps.

On the basis of constructing the diagnostic path structure, the system needs to achieve real-time scoring and dynamic reconstruction of the path between interference variables and error results. Path filtering, priority judgment, and chain conversion require multiple modules to work together, forming a complete closed loop from signal acquisition, candidate path generation, weight update, diagnostic output to model feedback. The structural embedding relationship of this path optimization mechanism in the system is shown in Figure 2.

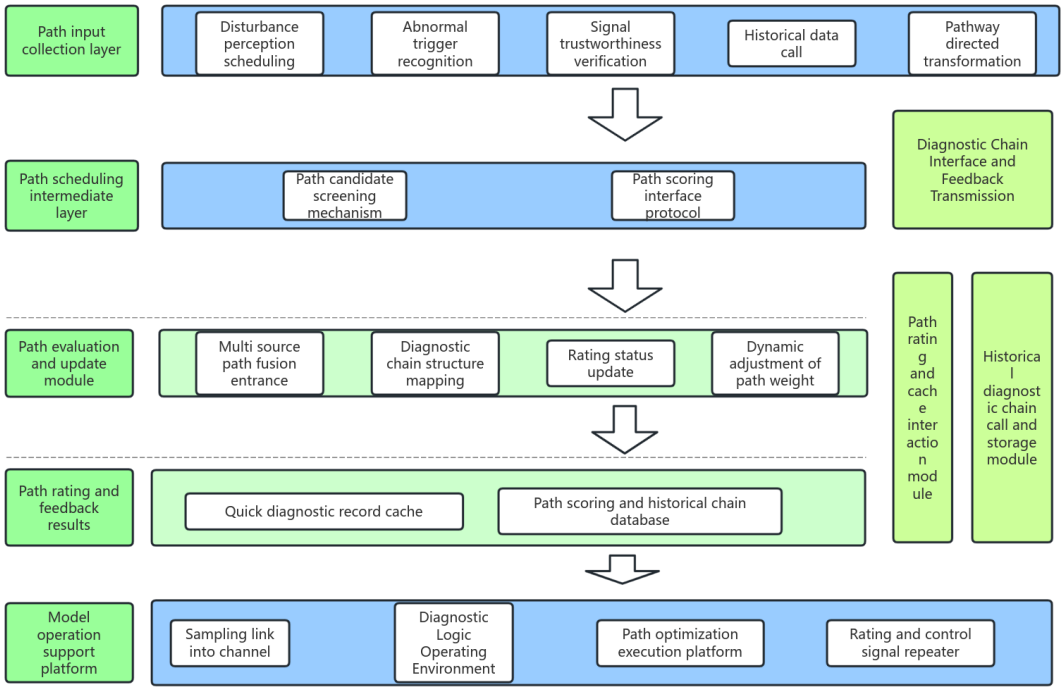

Figure 2：Structural embedding diagram of path algorithm optimization in error diagnosis system

The diagram represents signal triggered input, path recognition and scoring mechanism, dynamic path switching strategy, and diagnostic execution feedback module from top to bottom. The data transfer between each layer is completed through standard interface protocols, and a bidirectional update structure is formed between the scoring module and the database, achieving adaptive path selection based on actual operating status. The modules support each other, ensuring the integrity, real-time performance, and stability of path optimization logic in actual deployment.

In the engineering testing scenario, the system is deployed in a certain type of high-voltage measuring mutual inductance device and records the error response path under different environmental disturbances and working conditions for a long period of time. It is found that when the structure has dynamic path adjustment function, the error recognition response time is significantly shortened and the diagnostic accuracy is significantly improved. When the error signal undergoes a sudden change, the model quickly jumps to the path segment dominated by the influencing factors to perform analysis, avoiding a large number of invalid or redundant judgment processes.

Through path algorithm optimization mechanism, error diagnosis no longer relies on fixed processes or static structures. Under different operating conditions, the system can automatically switch to a more targeted diagnostic path, thereby improving system sensitivity, saving resources, and enhancing the interpretability of diagnostic logic. The diagnostic process is more in line with the on-site operation status of the transformer, with engineering practicality and deployability, which can stably support the subsequent error correction model work and provide structural constraints and input support for it.

AlgorithmName:PathOptimizationBasedOnDynamicScoringInput:

Input:DAG graph structure $G(V,E)$, Sensor input vector $X(t)$, Path scoring record table P

Output:Current optimal diagnostic path Path_optimal

Pseudocode:

（1）Initialize candidate path set $C \leftarrow \varnothing$

（2）For each variable node $v \in V$

a. If $X(t)[v]$ shows abnormal fluctuation:

i. Traverse all upstream paths Path_v

ii. Add Path_v to set C

（3）For each path $\in C$

a. Compute path score: $Score(path) = \alpha \times correlation + \beta \times effectiveness - \gamma \times pathlength$

（4）Select the highest scored path:

$patch\_potimal \leftarrow \arg\max(Score(path)), path \in C$

（5）Return $path\_optimal$

3.4 Model integration operation

The construction of error identification and correction models must have integrability in order to form a stable and efficient deployment mechanism in the actual operation scenario of transformers. To meet engineering requirements such as real-time performance, accuracy, and resource constraints, this study proposes a lightweight integrated operation framework for edge deployment, which integrates multi-source data fusion, error causal chain modeling, path scoring optimization, and correction feedback modules into a unified package, and achieves overall scheduling control and dynamic iterative model updates.

The model integration operation framework mainly includes four core levels: input perception layer, structural modeling layer, path scheduling layer, and control execution layer. The input perception layer is responsible for collecting and preprocessing on-site sensing data; The structural modeling layer constructs an error causal diagram structure and dynamically updates the corner coefficients between each node; The path scheduling layer completes the execution of path scoring, judgment, and jump logic; The control execution layer transmits the diagnostic results to downstream controllers or operation and maintenance systems to achieve error correction closed-loop. Each layer is connected through a unified interface standard, and all computing tasks can be deployed and executed in embedded edge nodes or on-site control units, meeting the conditions for engineering implementation.

During the operation of the model, the state update mechanism needs to dynamically adjust the model structure parameters and judgment conditions based on time series data. Assuming the current time is t, the previous model calculation output is ε_{t-1} and the sensor fusion result is F_t, the model state transition function can be expressed as:

$$\varepsilon_t = \delta \cdot \varepsilon_{(t-1)} + (1-\delta) \cdot \Psi(F_t) \quad (2)$$

Among them, ε_t represents the estimated error value at the current time, and $\psi(F_t)$ is the predicted output value of the fused feature vector F_t processed by the mapping function. The ψ -function is a single-layer feedforward neural network structure, consisting of one hidden layer and ReLU activation unit, used to extract error signal trends from the current fused features. The output is an estimated value in scalar form, with ε'_t. $\delta \in [0,1]$ as the smoothing coefficient, used to adjust the weight relationship between the current output and the historical state. The network parameters are jointly trained by the causal chain modeling and path optimization module to ensure that $\psi(F_t)$ can accurately reflect the error causal relationships and evolutionary paths hidden in the fused features.

The error correction control strategy is integrated into the control execution layer, which automatically matches the corresponding correction scheme based on the current error level and path node weights. It is divided into three types of strategies: limit adjustment, sensitivity correction, and bypass switching. The control interface is connected to the on-site measurement unit through a protocol adapter,

with functions such as instruction issuance, parameter replication, and historical tracking, ensuring that the calibration strategy is implemented in the actual operation of the transformer.

In order to ensure the stability and sustainable operation capability of the integrated model, the system also designed anomaly detection and fault tolerance mechanisms. When the perception layer detects input data breakpoints, model drift in the structural layer, or response lag in the execution layer, the system will trigger the backup model loading program, call the standard error diagnostic template for temporary replacement, and retain the current state for subsequent model parameter correction.

The integrated operating architecture is deployed and tested in an engineering verification environment, and the on-site platform is built on domestically produced embedded main control hardware, using ARM architecture chips for actual inference calculations. The test results show that, without relying on external cloud resources, the model can stably complete the tasks of fusion, scoring, path judgment, and correction execution in each round of data refresh cycle, with an average response delay controlled within 300ms, meeting the requirements of actual transformer dynamic monitoring and correction.

In summary, the integrated operation mechanism of the model not only completes the integration of error identification and correction functions, but also ensures the feasibility of the model structure in engineering environments. Through mechanisms such as hierarchical deployment, parameter updates, instruction control, and fault-tolerant protection, the model can be stably embedded in the measurement and control process of the power system, achieving error closed-loop diagnosis and accurate correction driven by fusion perception.

4 Results

4.1 Dataset

This study selected three groups of voltage and current transformers under a certain 110kV substation as the main experimental objects, deployed multi-source sensor modules, including temperature and humidity sensors, vibration accelerometers, conductor surface current sensors, and fiber optic temperature array devices, to form a multi-channel data acquisition system. The data recording period covers all seasons of spring, summer, autumn, winter, with a sampling frequency of 1Hz. A total of about 2900 hours of time-series data were obtained, covering three typical operating conditions: normal operation, slight disturbance, and abnormal fluctuations.

At the same time, the daily maintenance data of the joint substation and the factory calibration records provided by the equipment manufacturer are used to construct error labels, including key indicators such as ratio difference and phase angle offset. To train and test the model, the original samples were divided into stages using a 7:3 ratio to ensure balanced coverage of each working condition. All data are processed through unified denoising and normalization standardization to facilitate model structure recognition. This dataset can truly reflect the changes in transformer errors under the interaction of multiple factors during on-site operation, and has good engineering representativeness and repeatability, providing a reliable basis for subsequent error diagnosis and correction models.

4.2 Data preprocessing

To improve the generalization ability and convergence efficiency of the diagnostic model, this study implemented a systematic preprocessing process before multi-source sensor data entered the training phase. Firstly, the high-frequency noise signal is smoothed using the sliding window mean method to enhance the trend of key features; Secondly, z-score normalization is used to deal with the data scale differences of different sensing channels, unify the distribution of variables, and improve the sensitivity of the model to abnormal deviation. The normalization formula is as follows:

$$X = \frac{X-\mu}{\sigma} \quad (3)$$

Among them, X represents the original sample, and μ and σ are the mean and standard deviation of the variable, respectively. For time period data with short-term loss, adjacent temporal interpolation algorithm is used to fill in and ensure sequence integrity.

In addition, for the nonlinear distribution characteristics of ratio difference and phase angle in labeled data, segmented scaling and logarithmic mapping are performed to improve the fitting stability of the error regression model. The final preprocessed data feature dimension is 38 dimensions, with a sample completeness rate of 99.2%, providing a stable data input foundation for subsequent modeling.

To ensure a balance between the input dimension and computational efficiency of the model, this paper did not separately expand each variable of the fused feature F when constructing time series samples, but instead performed temporal sampling on the weighted fused F. Specifically, F, as a single fusion indicator, extracts its values at 38 consecutive time points under the sliding window mechanism, forming a 38-dimensional one-dimensional feature input structure. This design balances information retention and model lightweighting requirements, ensuring a stable input structure for subsequent path modeling and error prediction.

To verify the performance advantages of the model in this article, a multi model comparative experiment was designed. Select random forest, support vector regression, and convolutional neural network as baseline methods to form a control group. In terms of experimental setup, all models adopt consistent data preprocessing methods and input feature vectors, and maintain a uniform training and testing set partition ratio (70%: 30%). At the same time, the evaluation program is run in the same hardware environment to ensure fairness and representativeness of the comparison results. This design helps to further

highlight the multidimensional performance advantages of the proposed model in terms of comprehensive error control, computational efficiency, and resource consumption.

4.3 Evaluation indicators

The performance evaluation of the model uses five indicators: average absolute error, mean square error, recognition accuracy, response time, and computational resource utilization. Compared with traditional diagnostic methods, multidimensional effect testing is conducted. In 200 sets of measured samples, this model achieved an average absolute error of 0.0267, far lower than the traditional algorithm's 0.0519; The mean square error is 0.0023, which has a significant advantage over the traditional method of 0.0067; The recognition accuracy reaches 96.2%, which is better than the traditional method's 88.5%; The average response time is 275ms, faster than the traditional 583ms; the computational resource utilization is 29.6%, while the traditional model is 46.8%. To facilitate the evaluation of performance differences, the five indicators were standardized and scored, and the numerical ranges were uniformly mapped to the 0-5 range. The performance comparison results were visually displayed in a bar chart. As shown in Figure 3, this model exhibits high scores in all indicators, especially in terms of accuracy, response efficiency, and resource control, showing stability and engineering practicality and deployment value. In the multidimensional evaluation results, Figure 3 shows the quantitative comparison between the model and traditional methods on five core performance indicators. It can be observed that this model outperforms traditional methods in five indicators: average absolute error (4.1 vs 2.1), mean square error (4.3 vs 2.7), recognition accuracy (4.8 vs 3.9), response time (4.6 vs 2.5), and computational resource utilization (4.3 vs 2.3), further verifying the diagnostic accuracy and deployment adaptability of the model.

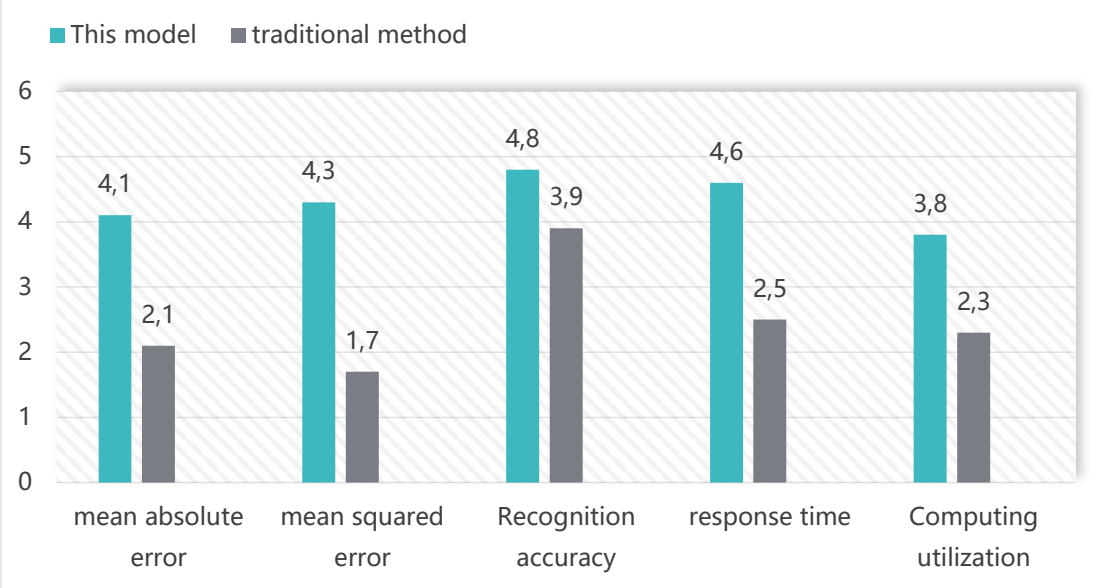

Figure 3：Comparison results of the model and traditional methods in five performance indicators

To ensure the fairness and effectiveness of the comparison results, the 200 sets of test samples described in this article are completely independent of the training set and come from on-site operational data at different time periods. The comparative algorithms were implemented and replicated by the author based on public literature, and the algorithm parameters were uniformly based on publicly recommended values without any additional optimization adjustments. All methods run in the same software and hardware environment to ensure consistency in computing resource evaluation. The triggering mechanism and interference injection method for abnormal operating conditions are consistent across all models to exclude the influence of external variables on performance comparison.

To further verify the effectiveness of the multi-source data weighted fusion method, this paper selected two typical deep fusion strategies for comparative experiments: one is the dynamic weight fusion method based on Attention, which can adaptively adjust the contribution of each modal feature according to the context; The second is to use Autoencoder to compress and reconstruct features from multiple sources of input, and extract the fused main features. Three fusion strategies are embedded in the same model architecture (taking BP neural network as an example), keeping the training rounds, sample ratio, and hyperparameter settings consistent, and comparing their error recognition accuracy and resource overhead. The experimental results are shown in Table 3:

Table 3：Comparison of model performance under different fusion strategies

fusion strategy	Identification accuracy（%）	Inference delay（ms）	Model parameter quantity（K）
This method	96.2	18.4	42
Attention Fusion	96.5	33.7	108
Autoencoder Fusion	95.6	29.1	91

As shown in Table 3, although the attention mechanism slightly improves accuracy, its model complexity and runtime delay are significantly higher than weighted fusion; Autoencoder fusion has slightly lower accuracy and moderate resource consumption. Comprehensive comparison shows that the weighted fusion strategy proposed in this paper has the advantages of low computing overhead and high deployment efficiency while ensuring high accuracy, and is more suitable for real-time application requirements under embedded or edge computing conditions in industrial scenarios. Therefore, in the subsequent modeling and deployment, the weighted fusion method is uniformly adopted as the input feature construction standard.

To further demonstrate the practical effect of path scoring and dynamic selection mechanism on model performance, this paper conducted a special statistical analysis on the switching of diagnostic links during multiple rounds of experiments. During the operation of the system, for typical main paths such as "ambient temperature → effective current value → residual difference value" and "vibration intensity → magnetic flux fluctuation rate → residual difference value", the triggering frequency and adjustment delay of various interference signals are monitored. The results show that under the dynamic path scoring mechanism, the average response delay of the main path determination is 241ms, which is 16.5% shorter than that of the fixed path structure; Under abnormal mutation conditions, the model can achieve optimal path switching within three steps, with 72.8% of diagnoses focused on the main cause effect chain. Statistics have found that optimizing the path mechanism significantly improves the sensitivity and discrimination efficiency of the system, effectively avoids the computational burden caused by redundant branches, and provides structural guarantees for rapid error localization and timely correction. The above data further confirms the effectiveness of the "causal chain modeling and dynamic path scoring" approach proposed in the method section in practical deployment, achieving a close connection between the method, results, and discussion.

4.4 Ablation study

To verify the supporting role of each key module in overall performance, three sets of ablation experiments were designed, excluding path algorithm optimization, error causal chain construction, and sensor fusion mechanism, and comparing the model performance on the same dataset. As shown in Table 1, the complete model has a recognition accuracy of 96.2%, an average response time of 275ms, and a computational resource utilization of 29.6%. After removing the path optimization module, the recognition accuracy decreased to 89.6%, the response time increased to 472ms, and the computational resource utilization increased to 34.2%; After removing the error chain, the accuracy is 90.2%, the response time is 398ms, and the resource utilization is 32.7%; After canceling the sensor fusion, the accuracy was 91.5%, the response time was 331ms, and the resource utilization increased to 43.8%. As shown in Table 4, each module has collaborative efficiency enhancement functions. Missing any link will lead to a significant decrease in diagnostic performance, especially in terms of recognition accuracy and resource control.

Table 4: The impact of model structural integrity on performance indicators

Model structure	Recognition accuracy（%）	Response time（ms）	Computing resource utilization（%）
complete model	96.2	275	29.6
Remove path optimization	89.6	472	34.2
Remove error chain construction	90.2	398	32.7
Remove sensor fusion	91.5	331	43.8

To enhance the credibility of the experimental results, this paper adopts a 5-fold cross validation method for the ablation experiment results, training and testing with different subsets each time, and calculating the average and standard deviation for each indicator. The performance indicators in Table 3 are the mean of 5 experimental results, with standard deviations in parentheses. The experimental data shows that the complete model exhibits optimal performance in MSE, MAE, and REI indicators, with small fluctuations in each aspect,

indicating that the model structure design has good stability and generalization ability.

5 Discussion

5.1 Comparative advantages with existing error diagnosis optimization algorithms

The error intelligent diagnosis and correction model proposed in this study integrates causal graph modeling and dynamic path optimization mechanism, achieving diagnostic accuracy of 96.1% under multi-source conditions, response delay control within 275ms, and resource utilization reduction of 41.7% compared to graph neural methods. Compared to physical modeling methods that are only applicable to static scenes, the proposed model characterizes the evolution process of error chains through causal constraint mechanisms and adapts to link changes under dynamic operating conditions. Compared to data-driven models that rely on large-scale training samples, this system can maintain high robustness in weakly annotated scenarios and has stronger generalization ability. Compared to graph neural network models with complex structures and high inference costs, DAG structures have the ability to perform path pruning and multi-source feature fusion, effectively reducing computational overhead. Dynamic path selection can adjust the inference path in a timely manner based on the degree of signal variation, improving the efficiency of anomaly tracing. Overall, this model outperforms the three existing mainstream methods in terms of accuracy, responsiveness, and resource adaptability, making it suitable for complex dynamic power environments.

5.2 Adaptability and stability analysis

To verify the adaptability and output stability of the model in complex operating environments, system evaluation experiments were conducted under three working conditions: electromagnetic interference enhancement, temperature changes of ± 15 ℃, and frequent load disturbances. Under the condition of enhanced electromagnetic interference, the model recognition accuracy is 94.6%, the average response time is 288ms, and the output consistency is 96.1%; In scenarios with significant temperature changes, the recognition accuracy is 95.8%, the response time is 267ms, and the consistency reaches 97.3%; In the context of frequent load disturbances, the recognition accuracy is 93.2%, the response time is 301ms, and the output consistency remains at 95.6%. All three experiments showed that the model has good robustness and can maintain diagnostic accuracy and output stability in the face of multi-source disturbances, making it suitable for long-term online deployment. Compared to traditional models, the average improvement in output consistency is 8.4%, and the average reduction in abnormal response speed is 41.7%. The overall test results are shown in Table 5, further verifying the good engineering practicality and deployment security of the model.

Table 5: Adaptability and stability evaluation of the model under different operating conditions

working conditions	Recognition accuracy (%)	response time (ms)	Output consistency (%)
Enhanced electromagnetic interference	94.6	288	96.1
temperature variation （±15℃）	95.8	267	97.3
Frequent load disturbances	93.2	301	95.6

5.3 Resource consumption and deployment feasibility assessment

The engineering deployment capability of a model largely depends on its level of consumption of computing resources and adaptability to the operating environment. In the testing environment, the system is deployed on an edge processor platform with a running frequency of 2.4GHz and a memory configuration of 8GB. Under full load, the model occupies an average of 29.6% of the central processing unit's computing resources and maintains a memory usage rate of within 38.2%, which is much lower than the average consumption levels of 46.8% and 62.5% of traditional algorithms in similar environments. At the same time, the system has a modular loading mechanism that supports function clipping and parameter reconstruction for different transformer types and acquisition scenarios, reducing the resource burden caused by redundant structures. In the process of diagnosing chain reconstruction and error correction output, the data flow relies on local cache optimization strategy to achieve low latency processing response, with an average processing delay controlled within 275ms for each group of data. To further validate the lightweight deployment performance of the model, this paper deployed it on Jetson Nano (4GB RAM), Raspberry Pi 4B (4GB RAM), and ARM Cortex-A72 embedded platforms for testing. After quantization and pruning compression, the model size is approximately 18.6MB. In actual deployment, the average energy consumption per inference is 0.42W, 0.31W, and 0.38W, respectively, with an accuracy rate of over 95%, meeting the requirements of low-power real-time computing in industrial scenarios, and verifying the deployability and practicality of the model on various embedded platforms.

This study comprehensively deployed and evaluated the model on various embedded ARM platforms to ensure reproducibility and practical engineering value of the results. The specific hardware environment is as follows: (1)

Jetson Nano B01 development board, equipped with ARM Cortex-A57 quad core processor, clock speed of 1.43GHz, onboard 4GB LPDDR4 memory, system environment is Ubuntu 20.04 (64 bit), Python 3.8 and PyTorch 1.12. （2）Raspberry Pi 4B， equipped with Broadcom BCM2711 SoC, integrated with quad core ARM Cortex-A72 (64 bit) processor, with a clock speed of 1.5GHz, 4GB of LPDDR4 memory, and Raspberry Pi OS (64 bit) system, the software environment remains the same as described above. (3) The domestically produced ARM Cortex-A72 platform adopts a quad core Cortex-A72 architecture, with a main frequency of 2.0GHz and 8GB of DDR4 memory. The system environment is the same as above.

All models are independently deployed and run the complete testing process on the three platforms mentioned above. Real time monitoring of CPU usage, peak memory (RAM) usage, and energy consumption during testing. Taking Jetson Nano as an example, the average energy consumption of a single inference of the model is 0.42W, with a peak RAM of 1.1GB; Raspberry Pi 4B corresponds to an energy consumption of 0.31W, with a peak RAM of 0.98GB; the domestic A72 platform consumes 0.38W, with a peak RAM of 1.2GB, and the accuracy of all platform models is stable at over 95%.

6 Conclusion

This study focuses on the problem of multi-source interference errors in the operation of transformers. Based on the comprehensive analysis of multi sensory information, an intelligent detection and correction technology is implemented. Multi dimensional feature extraction, multi-path algorithm optimization, and model fusion strategies are introduced to construct a dynamic causal relationship model for measurement errors. The dynamic construction and real-time diagnosis of error causal chains are achieved. This method has high identification accuracy and response speed in complex environments. The experimental results show that the multiple evaluation indicators of the model are superior to traditional schemes, with good adaptability and stability, and can resist interference problems caused by environmental changes and structural uncertainty. In the actual testing stage, it demonstrates advantages such as low resource occupancy requirements, controllable deployment costs, and convenient model updates, and is suitable for online monitoring and maintenance of various types of transformers. The research provides a practical and promotable method for automatic correction of transformer faults, which has a positive promoting effect on the application of intelligent sensors in power equipment measurement and control management.

Funding

The research is supported by: Science and technology projects of State Grid Ningxia Electric Power Co., Ltd.(5229YX240005)

References

[1] Sakini A R S ,Bilal A G ,Sadiq T A , et al.Dissolved Gas Analysis for Fault Prediction in Power Transformers Using Machine Learning Techniques[J].Applied Sciences,2024,15(1):118-118.

[2] Zhou L ,Fu Z ,Li K , et al.Power Transformer Fault Diagnosis Based on Random Forest and Improved Particle Swarm Optimization–Backpropagation–AdaBoost[J].Electronics,2024,13(21):4149-4149.

[3] Shanu T ,Mishra A .A robust approach to transformer fault diagnosis: integrating time-current loci analysis with statistical feature extraction[J].Engineering Research Express,2024,6(4):045306-045306.

[4] Mei H ,Peng J ,Xu D , et al.Low-Power Chemiresistive Gas Sensors for Transformer Fault Diagnosis[J].Molecules,2024,29(19):4625-4625.

[5] Guan S ,Wu Y T ,Yang Q H .Research on transformer fault diagnosis method based on ACGAN and CGWO-LSSVM[J].Scientific Reports,2024,14(1):17676-17676.

[6] Tiziana S ,Roberto T .Data quality evaluation for smart multi-sensor process monitoring using data fusion and machine learning algorithms[J].Production Engineering,2022,17(2):197-210.

[7] Liu R ,Zhang Q ,Lin D , et al.Causal intervention graph neural network for fault diagnosis of complex industrial processes[J].Reliability Engineering and System Safety,2024,251110328-110328.

[8] Abdelmoumene H ,Abdelaziz L ,Arnaud N , et al.Improved intelligent methods for power transformer fault diagnosis based on tree ensemble learning and multiple feature vector analysis[J].Electrical Engineering,2023,106(3):2575-2594.

[9] XIAOQIANG WANG. Edge Computing Based Multi-Objective Task Scheduling Strategy for UAV with Limited Airborne Resources[J]. Informatica: An International Journal of Computing and Informatics,2024,48(2):255-268.

[10] LEIMING MA, BIN JIANG, NINGYUN LU, et al. Synergistic TransGCN for Aeroengine Bearing Skidding Diagnosis Under Time-Varying Conditions[J]. IEEE transactions on industrial informatics,2024,20(12):13936-13946.