frontiers | Frontiers in Energy Research

OPINION
published: 24 May 2022
doi: 10.3389/fenrg.2022.909520

Multisource Heterogenous Data Fusion for Fault Identification of Buried Substations

Yifang Jin[1], *Baling Fang*[1,2]*, *Ziqi Da*[1], *Shanhe Huang*[1] *and Chengxin Zhao*[1]

[1]*College of Electrical and Information Engineering, Hunan University, Changsha, China,* [2]*Huaxiang Xiangneng Technology Co., Ltd., Yiyang, China*

Keywords: acoustic signals, multi-source heterogeneous data fusion, buried transformer, fault identification, feature extraction, meta-learning

OPEN ACCESS

Edited by:
Bo Yang,
Kunming University of Science and Technology, China

Reviewed by:
Youwei Jia,
Southern University of Science and Technology, China
Xian Zhang,
Harbin Institute of Technology, China

***Correspondence:**
Baling Fang
5911866@qq.com

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

Received: *31 March 2022*
Accepted: *19 April 2022*
Published: *24 May 2022*

Citation:
Jin Y, Fang B, Da Z, Huang S and Zhao C (2022) Multisource Heterogenous Data Fusion for Fault Identification of Buried Substations. Front. Energy Res. 10:909520. doi: 10.3389/fenrg.2022.909520

INTRODUCTION

With the acceleration of urban and rural power grid construction and the increasing demand for residential electricity, the voltage level of power transmission and transformation equipment has been continuously improved. High-voltage, large-capacity power transformers and some substations have been located around urban areas with high load density, which leads to resident complaints and even disputes (Xiong et al., 2020; Zhang et al., 2022). Considering the urban planning and land-use restrictions, buried substations, with the reductions in space demand, noise emissions, and power line loss compared to conventional substations, are recognized to have a significant role in the sustainability and urbanization of power distribution systems (Li et al., 2016; Sadovskaia et al., 2019). However, for buried substations that are prefabricated and with smaller internal space, it can become much more inconvenient to operate and have routine maintenance of power supply and distribution equipment. Therefore, in order to improve the operational reliability and safety of buried substations, it is of great significance to accurately perceive and identify the comprehensive state of the buried substations.

Conventional condition perception and fault identification methods include dissolved gas analysis in oil and partial discharge monitoring technology. Previous work by Fan et al. (2019) proposed an analysis technology of dissolved gas in oil for fault analysis, but this technique has drawbacks such as absolute coding boundary and large diagnostic equipment. The partial discharge monitoring technology is also susceptible to electromagnetic interference (Hussain et al., 2013). Moreover, most of the existing transformer state perception and fault identification technologies only use a single type of data for identification, such as the concentration of a specific dissolved gas or the pulse current generated by partial discharge, with certain limitations. Therefore, the accuracy and reliability of the fault identification results of the buried substations need to be improved. Since most buried substations are prefabricated and adopt a fully sealed box design, the buried transformer is located in a closed and narrow buried space, and the maintenance entrance and exit areas are small. Consequently, the traditional transformer fault identification methods are no longer suitable for new buried substations.

In this study, a multisource heterogenous data fusion for fault identification of buried substations is proposed. First, the state information of the buried substation is collected from various aspects, such as acoustic signals, temperature, humidity, voltage, and current, and the collected acoustic signals of buried substations are decomposed and features extracted. Then, the extracted acoustic features and non-acoustic features are input into the fault identification network for multisource heterogenous data fusion. Finally, the fault identification results are obtained. This method can collect more operating status information of buried substations and provide cross-dataset verification, thereby enhancing fault identification accuracy and efficiency in buried substation analysis. The main work steps include three parts: signal processing, feature extraction, and fault identification.

SIGNAL PROCESSING AND FEATURE EXTRACTION

The acoustic signals generated by the transformer under various operating conditions, which are also called the frequency of voiceprint waves, are found to be different through models and experiments (Hui et al., 12014; Riaz et al., 2016). In addition, the acoustic signals can be easily collected by microphone sensors or recording equipment, which belongs to a non-contact measurement. This non-contact measurement method is more suitable for the buried substations where the indoor space is small and the contact measurement is inappropriate. In order to better recognize the operating state of the buried substations and identify typical fault conditions, feature analysis of the acoustic signals can be carried out, and the corresponding features of the acoustic signals under different operating states can be extracted. Considering that the acoustic signals collected by the sound sensor are the time-domain signals with few available features, the acoustic signals can be converted from time-domain signals to frequency-domain signals for research. Thereby, time-frequency analysis can be implemented on the acoustic signals and then the time-frequency features of the acoustic signals are extracted, which enriches the available features of the acoustic signals.

The methods that can be utilized for time-frequency analysis and signal decomposition on time-domain signals include Fourier transform (FT), short-time Fourier transform (STFT), wavelet transform (WT), empirical mode decomposition (EMD), and local mean decomposition (LMD) (Zhao et al., 2019). Among them, FT does not have time-frequency analysis but only converts the time-domain signals into frequency-domain signals to analyze. STFT does not consider adaptation, whose size and shape of each window are precisely given, and it is still not very effective for signals with large differences between high frequency and low frequency. As a time-frequency analysis method, WT has good localization characteristics in both time and frequency domains. However, WT is not adaptive, and each component obtained after decomposition loses its own physical meaning. EMD and LMD are improved based on WT, which both belong to recursion mode decomposition. It is seen that the error will gradually be accumulated in the decomposition process, and there are problems with modal mixing and endpoint effect, which may affect the subsequent fault identification results (Zuhaib et al., 2020). In order to overcome the shortcomings of the analysis method mentioned earlier, a novel variable-scale processing method was proposed, which was called variational mode decomposition (VMD). VMD is a new time-frequency analysis method of adaptive, entirely non-recursive modal variation (Dragomiretskiy and Zosso, 2014). VMD can adaptively match the optimal center frequency and limit the bandwidth of each mode. In addition, the method can achieve effective separation of intrinsic mode functions (IMFs) and frequency-domain division of signals. Then, the effective decomposition components of the given signals are obtained, and the optimal solution to the variational problem is obtained. Compared with EMD and LMD methods, VMD can avoid some problems such as modal mixing and endpoint effect and has high decomposition, which can improve the accuracy and efficiency of subsequent fault identification.

In order to facilitate the input of useful feature information contained in the acoustic signals to the subsequent fault identification network, it is essential to implement feature extraction on each modal component obtained after the signal decomposition. The methods that can be used to extract features include Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Local Preservation Map (LPM), and Deep Belief Network (DBN). Compared with the previous three methods, DBN has strong feature extraction capabilities and good compatibility with other algorithms, which can fully map the fault information hidden in the original vibration signals (Zhang et al., 2018). Moreover, there is no need to rely on experience to extract data features. After the automatic extraction of the underlying network, the results of applying the extracted features to other networks will be much better than the results of manually extracting features.

MULTISOURCE HETEROGENOUS DATA FUSION FRAMEWORK FOR FAULT IDENTIFICATION OF BURIED SUBSTATIONS

Algorithms that can be used for fault identification include probabilistic neural network (PNN), radial basis function (RBF) neural network, wavelet artificial neural network (ANN), support vector machine (SVM), and random forest (RF). However, most of the existing transformer fault identification methods use a single network to judge the fault type, which may lead to inaccurate fault identification results (Ke Meng et al., 2010; Tripathy et al., 2010; Musa et al., 2018; Jamali et al., 2020). The meta-ensemble fault identification model is applied to input the acoustic features and non-acoustic features into two networks and finally output the fault identification results by integrating the outputs of the two networks, which can implement multisource data fusion on various features of buried substations. Compared with the traditional methods of fault identification using a single data source, multisource heterogenous data fusion can analyze the operating state of the buried substations according to multiple data sources, which can improve the accuracy and efficiency of fault identification. The meta-ensemble fault identification model includes at least one sub-classification network, an ensemble weight dynamic iterative network, and an ensemble network. The steps of training the model are as follows.

First, the time-frequency features of the sample acoustic signal data are input into the sub-classification network, and the output data of each classification network are integers representing the state of the buried substations. The n sub-classification networks can be similar classifiers with different hyperparameters, such as XGBOOST classification networks with n different

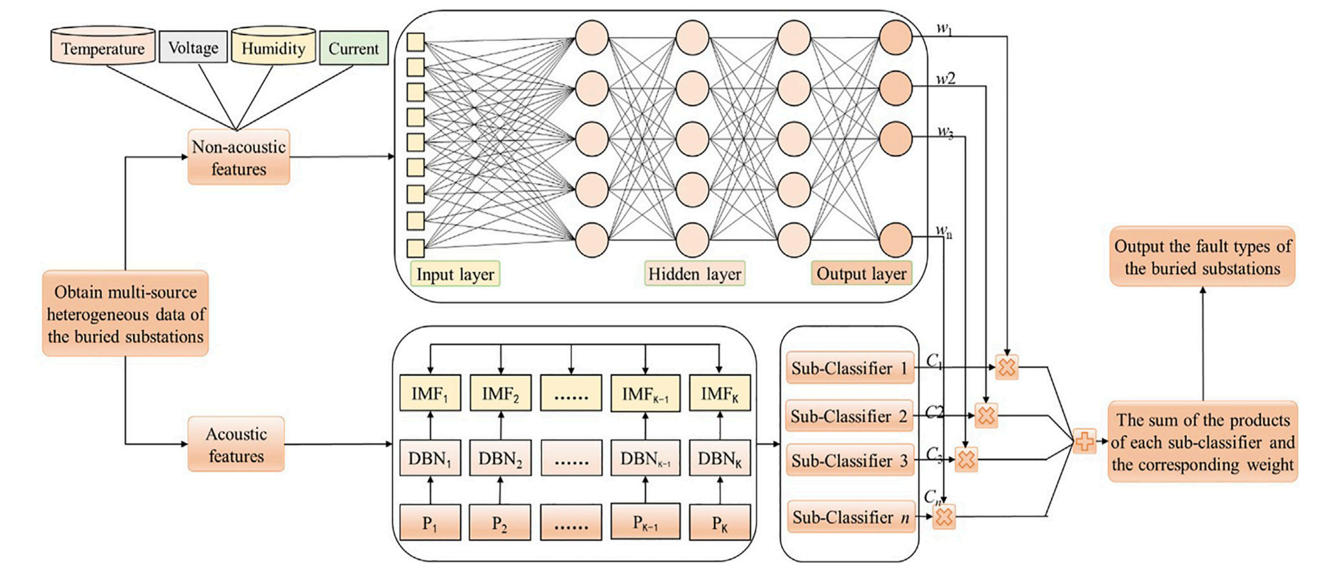

FIGURE 1 | Multisource heterogenous data fusion framework for fault identification of buried substations.

hyperparameters, and the n sub-classification networks can also be different types of classification networks, such as XGBOOST, SVM, and RF. Then, non-acoustic features of the sample are input into the ensemble weight dynamic iterative network. The ensemble weight dynamic iterative network includes an ensemble weight dynamic decision-making network, which can be achieved by the BP neural network, and the output layer has n neurons in total. Finally, the n outputs of the sub-classification networks and the outputs of the BP neural network are integrated through the ensemble network to obtain the forward propagation results. Based on the forward propagation results, the loss function of the meta ensemble fault identification model is calculated. Through iterative training, the loss function of the meta-ensemble fault identification model is minimized, and the trained meta-ensemble fault identification model is obtained. The meta-ensemble machine learning method in this study is adopted to train the model, which affects the output of the BP neural network by updating the weights and biases of each layer in the BP neural network, and thus affects the judgment of the buried state. The overall frame diagram is shown in **Figure 1**.

DISCUSSION AND CONCLUSION

Through the comprehensive analysis of the collected multiple data sources, the operating state information of the buried substations can be fully excavated. Therefore, research on multisource heterogenous data fusion for fault identification of buried substations is necessary, which is essential for deep learning of the various features corresponding to the various operating states of the buried substations and timely discovery of the fault type of the buried substations. The signal decomposition and feature extraction of the acoustic signals can make full use of all the information contained in the acoustic signals. Furthermore, the acoustic features and non-acoustic features are simultaneously input into the pretrained meta-ensemble fault identification model so as to implement multisource heterogenous data fusion on various features of the buried substations, and the fault identification results of the buried substations are obtained. The method can improve the effectiveness and the accuracy of fault identification in buried transformer analysis and further enhance the operational reliability and safety of the buried substations. Verifications of the method will be part of further research.

AUTHOR CONTRIBUTIONS

Writing the original draft and editing, YJ; conceptualization, BF; formal analysis, ZD; visualization and contribution to the discussion of the topic, SH and CZ.

FUNDING

This work was supported by the Hunan High-tech Industry Science and Technology Innovation Leading Foundation of China (2022GK4055).

REFERENCES

Dragomiretskiy, K., and Zosso, D. (2014). Variational Mode Decomposition. *IEEE Trans. Signal Process.* 62 (3), 531–544. Feb.1. doi:10.1109/TSP.2013.2288675

Fan, J., Liu, Z., Meng, A., Yin, H., Sun, Q., Bin, F., et al. (2019). Characteristics of Tin Oxide Chromatographic Detector for Dissolved Gases Analysis of Transformer Oil. *IEEE Access* 7, 94012–94020. doi:10.1109/ACCESS.2019.2928869

Hui, X., Zheng, S., Zhou, J., Chi, H., Jin, X., and Zhang, X. (1201). Hilbert-huang Transform Time-Frequency Analysis in ϕ -OTDR Distributed Sensor. *IEEE Phot. Technol. Lett.* 26 (23), 2403–2406. doi:10.1109/LPT.2014.2358262

Hussain, G. A., Kumpulainen, L., Kluss, J. V., Lehtonen, M., and Kay, J. A. (2013). The Smart Solution for the Prediction of Slowly Developing Electrical Faults in MV Switchgear Using Partial Discharge Measurements. *IEEE Trans. Power Deliv.* 28 (4), 2309–2316. doi:10.1109/TPWRD.2013.2266440

Jamali, S., Bahmanyar, A., and Ranjbar, S. (2020). Hybrid Classifier for Fault Location in Active Distribution Networks. *Prot. Control Mod. Power Syst.* 5 (1), 1–9. doi:10.1186/s41601-020-00162-y

Ke Meng, K., Zhao Yang Dong, Z. Y., Kit Po Wong, D. H., and Wong, K. P. (2010). A Self-Adaptive RBF Neural Network Classifier for Transformer Fault Analysis. *IEEE Trans. Power Syst.* 25 (3), 1350–1360. doi:10.1109/tpwrs.2010.2040491

Li, C., Liu, X., Zhang, W., Cao, Y., Dong, X., Wang, F., et al. (2016). Assessment Method and Indexes of Operating States Classification for Distribution System with Distributed Generations. *IEEE Trans. Smart Grid* 7 (1), 481–490. Jan. 2016. doi:10.1109/TSG.2015.2402157

Musa, M. H. H., He, Z., Fu, L., and Deng, Y. (2018). A Cumulative Standard Deviation Sum Based Method for High Resistance Fault Identification and Classification in Power Transmission Lines. *Prot. Control Mod. Power Syst.* 3 (3), 291–302. doi:10.1186/s41601-018-0102-4

Riaz, F., Hassan, A., Rehman, S., Niazi, I. K., and Dremstrup, K. (2016). EMD-based Temporal and Spectral Features for the Classification of EEG Signals Using Supervised Learning. *IEEE Trans. Neural Syst. Rehabil. Eng.* 24 (1), 28–35. doi:10.1109/TNSRE.2015.2441835

Sadovskaia, K., Bogdanov, D., Honkapuro, S., and Breyer, C. (2019). Power Transmission and Distribution Losses - A Model Based on Available Empirical Data and Future Trends for All Countries Globally. *Int. J. Electr. Power & Energy Syst.* 107, 98–109. doi:10.1016/j.ijepes.2018.11.012

Tripathy, M., Maheshwari, R. P., and Verma, H. K. (2010). Power Transformer Differential Protection Based on Optimal Probabilistic Neural Network. *IEEE Trans. Power Deliv.* 25 (1), 102–112. doi:10.1109/TPWRD.2009.2028800

Xiong, L., Liu, X., Liu, Y., and Zhuo, F. (2020). Modeling and Stability Issues of Voltage-Source Converter Dominated Power Systems: A Review. *Csee Jpes*, 1–18. doi:10.17775/CSEEJPES.2020.03590

Zhang, C., He, Y., Yuan, L., and Xiang, S. (2018). Analog Circuit Incipient Fault Diagnosis Method Using DBN Based Features Extraction. *IEEE Access* 6, 23053–23064. doi:10.1109/ACCESS.2018.2823765

Zhang, K., Zhou, B., Or, S. W., Li, C., Chung, C. Y., and Voropai, N. (2022). Optimal Coordinated Control of Multi-Renewable-To-Hydrogen Production System for Hydrogen Fueling Stations. *IEEE Trans. Ind. Appl.* 58 (2), 2728–2739. March-April 2022. doi:10.1109/TIA.2021.3093841

Zhao, W., Shang, L., and Sun, J. (2019). Power Quality Disturbance Classification Based on Time-Frequency Domain Multi-Feature and Decision Tree. *Prot. Control Mod. Power Syst.* 4 (1), 1–6. doi:10.1186/s41601-019-0139-z

Zuhaib, M., Rihan, M., and Saeed, M. T. (2020). A Novel Method for Locating the Source of Sustained Oscillation in Power System Using Synchrophasors Data. *Prot. Control Mod. Power Syst.* 5 (1), 1–12. doi:10.1186/s41601-020-00178-4

Conflict of Interest: Author BF was employed by Huaxiang Xiangneng Technology Co., Ltd, and currently is also a part-time postdoctoral fellow at Hunan University.

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.