Chapter 7 Power Transformer Fault Location

AI based power transformer fault diagnosis and condition assessment have been extensively covered in previous chapters. As a reasonable extension, this chapter deals with fault location techniques – another important issue in power equipment test and maintenance field.

7.1 Why Fault Location

Fault location can provide critical information for power transformer maintenance, failure investigation, and restoration. According to the knowledge of possible faulty areas, a plan of detailed fault investigation can be made long before the scheduled shut-down, or the necessary spare part/equipment and repairing materials/resources can be prepared adequately in case of an unavoidable failure. The resultant could be a considerably reduced transformer down time and restoration effort, which is essential in today's competitive electric power market.

Fault investigation is a major issue of power transformer maintenance. The results are the basis of a trade-off decision to continue operation, re-energize the protection-tripped transformer after treatment, partly/fully repair or replace the failed transformer. For instance, the fault could be an unintentional core ground that can be eliminated using the "sledge hammer" method [Danny91], or a through-fault (fault occurs outside the transformer but induces very large currents through the transformer windings) that may or may not harm the transformer too much. Fault location can help identifying these kinds of fault by discriminating between "core fault" and "winding fault", thus provide the decision-making support for maintenance engineers.

7.2 The Basis of Fault Location

The intrinsic relationships between fault-related materials and fault types are the basis of transformer fault location. For example, winding fault will most probably related to solid insulation overheating and discharges, therefore the present of these two types of fault could be an indication of fault in winding areas. However, 100% rule-based fault location is not possible.

AI technology provides a possible way to locate fault in power transformers. Upon the collection of a large amount of real data cases, some AI based methods may be able to give a rough estimation of where inside the transformer the fault is.

A total of 174 data samples were collected for this study, which were detailed in Appendix 8. Each sample includes the seven popular gas-in-oil concentrations (H_2, CH_4, C_2H_6, C_2H_4, C_2H_2, CO and CO_2), fault location descriptions, and a group of fault location representing numbers. The numbers were determined according to the fault location descriptions, and correspond to the following five fault location categories. Since the descriptions are natural languages, sometimes they cannot identify the location clearly. Knowledge of transformer structure and fault phenomena is a crucial factor in the determination process.

The fault location categories in this study is defined as:

- LTC

  Include LTC tap board terminals, in-tank LTC components, and surrounding areas. Conventionally, LTC inside the tank is one of the most troublesome components of power transformers. Its mechanical parts become loose due to frequent operations. Its contacts experience overheating and/or arcing when changing taps under load conditions. All these factors contribute to the fault development. Although modern transformers usually have a separate LTC compartment, the tap board is still inside the tank and causing problems, because the mechanical and thermal stresses can transfer from the separate compartment to it.

- TANK

  Include the oil tank case, core laminations and assembly bolts, etc. The major problem is overheating caused by close loop current. Careless assemblage of core or tank parts may leave conducting loops in the magnetic circuit. Vibration of the core may damage the core bolt insulation and generate conducting loops also. These loops will allow large current to flow and the core/tank could be seriously overheated.

- LEADS

  Include leads between winding coils, between windings and bushings, between windings and LTC tap board, between neutral point and ground, etc. The high voltage leads are more likely to have problems than the low voltage ones. These leads are often related to high-energy discharges because they are usually close to ground potential. They can also experience electrical overheating because they usually have joints, whose resistance will generate a large amount of heat under normal and especially overload condition.

- WNDG

  Refers to windings. Problems with windings are mostly insulation degradation or breakdown. The degradation is caused by high temperature of the core or the conductor, and electric field applied to the insulation. The symptoms include the development of partial discharges and arcing. Overload and through fault cause excessive electrical overheating. Abnormal eddy current in the core causes excessive magnetic overheating. The overheating accelerates the degradation process.

- OTHER

  Refers to areas other than the previously defined ones, such as forgotten tools in the tank, static shielding, cooling system (fans, oil pumps), etc.

Other categorization methods may exist but the current approach is believed to be more practical from a maintenance point of view.

Of the 174 data samples, 110 were selected to train AI classifiers, which formed the training set (LOC_TRN), the other 64 were used to test the performance of the trained AI classifiers, which formed the testing set (LOC_TST).

7.3 Logistic Regression Based Fault Location

The principle of logistic regression based fault location is simple compare to neural network based or other AI techniques. Assume p is the probability that an event occurs for a given $N\times1$ data vector $\boldsymbol{X}$, the logistic function is defined as:

$$\mathrm{LOGIT}(p) = \log\left(\frac{p}{1-p}\right) = \boldsymbol{W}_0 + \boldsymbol{W}^T\boldsymbol{X} \qquad (7\text{-}1)$$

Where $\boldsymbol{W}_0$ is the intercept of the regression function, $\boldsymbol{W}$ is the $N\times1$ slope parameter vector, and T denotes a vector transform. W_0 and $\boldsymbol{W}$ can be estimated using a regression algorithm from a training data set like the above LOC_TRN. For each data sample in the training data set, p is either 1 or 0, representing whether or not the event occurred.

For the data samples in the testing data set like the above LOC_TST, the following equation can be used to derive the probability that the event will occur.

$$p = \frac{\exp(p)}{1+\exp(p)} = \frac{\exp\left(\boldsymbol{W}_0 + \boldsymbol{W}^T\boldsymbol{X}\right)}{1+\exp\left(\boldsymbol{W}_0 + \boldsymbol{W}^T\boldsymbol{X}\right)} \tag{7-2}$$

This is very similar to the sigmoid activation of neural networks and it can be approximate by a two-layer multiplayer perceptron (MLP) neural network.

For diagnostic purposes, a probability level P_{lvl} is defined to indicate if the data sample corresponds to an event or not. When p is larger than P_{lvl}, the event is said to have occurred. P_{lvl} is so selected that the testing accuracy of the training data set is high.

Corresponding to the 5 fault location categories, there are 5 logistic regression-based classifiers and therefore 5 outputs for each input vector. The meaning of these outputs is not explicit because P_{lvl} could be higher or lower than 0.5 depending on the fault location category. To better understand and compare the performance of the method, the output p of each classifier is mapped to $y \in [0,1]$ using the following equation.

$$y(p) = ap^2 + bp + c \tag{7-3}$$

Where a, b and c are constants derived from

$$\begin{cases} y(0) = 0 \\ y(P_{lvl}) = 0.5 \\ y(1) = 1 \end{cases} \tag{7-4}$$

If the y value according to Equation (7-3) is larger than 1, make it to be 1. This will give a fuzzy representation of the diagnosis.

7.4 Pattern Representation For Fault Location

It is well known that pattern representation, by which we mean the formation of the input vector of a pattern recognition classifier, affects the training and performance of the classifier greatly. An optimal pattern representation should facilitate the training and ensure high accuracy in the classifier's diagnostic performance, but may never be reached in practice. In this study, logistic regression analysis was used not only to obtain fault location classifiers but also to seek a quasi-optimal pattern representation for AI based fault location.

Three pattern representation methods (PRM) were studied and compared:

PRM1: H_2/TDCG, CH_4/TDCG, C_2H_2/TDCG, C_2H_4/TDCG, C_2H_6/TDCG, CO/TDCG, CO/CO_2

PRM2: H_2/TDHG, CH_4/TDHG, C_2H_2/TDHG, C_2H_4/TDHG, C_2H_6/TDHG, CO/CO_2

PRM3: H_2, CH_4, C_2H_2, C_2H_4, C_2H_6, CO, CO_2

Here PRM1 and PRM2 recognized the importance of CO/CO_2 ratio that may be useful in paper degradation diagnosis.

The three methods yielded different diagnostic accuracies on the training data set LOC_TRN and testing data set LOC_TST, as shown in Table 7-1. Here the test accuracies (TA) are listed for two schedules, denoted by TA1 and TA2, respectively. TA1 applies when the actual fault location is clearly indicated, i.e. the corresponding classifier output is the highest among the 5 outputs for a specific pattern representation. TA2 includes TA1 and also applies when the corresponding output is the second highest.

The confidence of TA2 diagnosis is somewhat lower than TA1 but it has practical importance. In a real fault investigation, it is not wise to look into just one particular area and not pay attention to others. For example, in a tear-down investigation case a large amount of work is in the open-tank process, once the tank is open, coil, leads and other areas will be equally accessible. Under this situation, having two suspect areas in mind will avoid missing potential fault areas.

Table 7-1 Test Accuracy (%) of Logistic Regression Based Fault Location Classifiers

Data Set	LOC_TRN		LOC_TST		LOC_TRN + LOC_TST	
	TA1	TA2	TA1	TA2	TA1	TA2
PRM1	51.8	69.1	36.9	53.8	46.3	63.4
PRM2	46.4	70.0	40.0	56.9	44.0	65.1
PRM3	52.7	78.2	44.6	64.6	49.7	73.1
Fuzzy	50.0	72.7	36.9	60.0	45.1	68.0

By comparison we can see that PRM3 is better than PRM1 and PRM2, regardless of the test data set. On the other hand, even if PRM3 is used, the overall accuracy of the logistic regression based fault location is still less than 80%, which is considered unacceptable.

7.5 Fuzzy Combination of Logistic Regression Based Fault Location

Based on the experience of improved performance with combined fault diagnosis, it is hoped that some kind of combination of the three pattern representation methods may yield better performance than the individuals.

The first combination method was tried using a fuzzy calculation equation:

$$y' = \min\left(y^{\mathrm{PRM1}}, \quad y^{\mathrm{PRM2}}, \quad y^{\mathrm{PRM3}}\right) \tag{7-5}$$

Where y^{PRM1}, y^{PRM2} and y^{PRM3} are the outputs from the three pattern representation methods PRM1, PRM2 and PRM3 of Table 7-1. Unfortunately the test accuracies are extremely low and are not worth given here. The reason is probably that this is not a strictly defined fuzzy logic problem, i.e. the combination is not a strict AND logic function.

Another combination method is based on Equation (7-6) and (7-7). First the product of the three individual outputs was obtained using Equation (7-6). Then the product was mapped to $d \in [0, \quad 1]$ using Equation (7-7). The results are listed in the last row of Table 7-1.

$$y' = y^{\mathrm{PRM1}} \times y^{\mathrm{PRM2}} \times y^{\mathrm{PRM3}} \tag{7-6}$$

$$d(y') = a(y')^2 + by' + c \tag{7-7}$$

Parameters in (7-7) were derived from

$$\begin{cases} y(0) = 0 \\ y(0.125) = 0.5 \\ y(1) = 1 \end{cases} \tag{7-8}$$

The second condition implies that when all the three outputs are equal to 0.5, the mapping should also yield 0.5.

However, from Table 7-1 we could not see any improvement for this second fuzzy combination method either.

7.6 ANN based Fault Location

Since the performance of the logistic regression based fault location is not satisfactory, a three-layer 7×21×5 MLP network was selected to do the job. Because pattern representation PRM3 was proved to have better performance in the logistic regression analysis, the seven inputs were selected to be the scaled gas-in-oil concentration of the seven gases (H_2, CH_4, C_2H_6, C_2H_4, C_2H_2, CO and CO_2). The scaling method is same as for transformer fault diagnosis, i.e. ppm value divided by 1000. The five outputs of the 7×21×5 MLP network correspond to the 5 fault location categories: LTC, TANK, LEADS, WNDG and OTHER.

The 7×21×5 MLP network was trained using data set LOC_TRN. The convergence of the training is very slow lately and the training is terminated when the average training error is below 0.1. Table 7-2 shows the test accuracy after the training. For the training data set LOC_TRN, the test accuracies at Schedule TA1 and TA2 are both above 90%. The minor difference between the two shows the advantage of using TA2 instead of TA1 but it is not a big deal. For the testing data set LOC_TST however, the test accuracies at Schedule TA2 is much higher than that at TA1, clearly shows the advantage of using TA2.

Table 7-2 Test Accuracy (%) of the ANN Based Fault Location Classifier

Data Set	LOC_TRN		LOC_TST		LOC_TRN + LOC_TST	
	TA1	TA2	TA1	TA2	TA1	TA2
PRM3	90.0	94.5	52.3	75.4	76.0	87.4

Comparison between Table 7-1 and Table 7-2 shows that the 7×21×5 MLP network is much better than the logistic regression based classifiers. This is reasonable because the input-output equation of the 7×21×5 MLP network is much more complicated than the logistic function and can represent more nonlinear relationships. This also reveals that the fault location task is not a simple logistic problem, the boundaries of the feature space is highly irregular and requires highly nonlinear mapping techniques such as the neural networks. Because of this, trying to setup a rule-based expert system to replace the neural network will be very difficult, if not completely impossible.

7.7 Summary and discussion

This chapter studies AI techniques for power transformer fault location. A total of 174 data samples were collected and used in the study. Two kinds of AI techniques were studied, including the logistic regression and the multiplayer perceptron (MLP) neural network.

Logistic regression is basically a statistical method but it can also be considered as a special form of neural networks – the simplest two-layer MLP. It may be effective when the feature space of the problem is regularly clustered, i.e. data point of the same category scatter around a center. But it will not work properly when the cluster is too loose, such as the fault location case we have here.

The potential of logistic regression lies in its capability of auditing the feature selection process. It can tell in general what pattern representation is good and what is not, making the feature selection of the neural network application much easier.

MLP network is capable of approximating highly nonlinear input-output relationships and is the final solution of power transformer fault location problems. With the help of logistic regression analysis, its input selection is easier than fault diagnosis.

The better performance of the 7×21×5 MLP network with respect to the logistic regression method suggests the complexity of the fault location problem and the difficulty of setting up fault location expert systems. However, this should not discourage the effort of rule-based fault location studies.