Electrical Power and Energy Systems 93 (2017) 308–315

Contents lists available at ScienceDirect

Electrical Power and Energy Systems

journal homepage: www.elsevier.com/locate/ijepes

Application of a general regression neural network for health index calculation of power transformers

Md Mominul Islam [a,*], Gareth Lee Dr. [a], Sujeewa Nilendra Hettiwatte Dr. [b]

[a] School of Engineering and Information Technology, Murdoch University, Murdoch, Australia
[b] School of Engineering, National School of Business Management, Homagama, Sri Lanka

ARTICLE INFO

Article history:
Received 22 March 2017
Received in revised form 5 June 2017
Accepted 8 June 2017

Keywords:
Power transformer
Health index
General Regression Neural Network (GRNN)
Artificial intelligence
Condition monitoring

ABSTRACT

A power transformer is one of the most important components in a transmission network. To assess the overall condition of this valuable asset, health index calculations are recently gaining more attention from the utility companies that operate networks. Only limited research has been conducted on health index calculations of transformers. Most of the past approaches are based on the linear combination of weighted scores of measurements following the industry standards such as IEEE, IEC and CIGRE. A few previous methods based on artificial intelligence and statistical approaches such as fuzzy logic, multivariate analysis and binary logistic regression have been published in recent years. In this paper, a General Regression Neural Network (GRNN) which has a nice nonlinear property and can work with measurements without quantization has been evaluated. The GRNN allows multi-dimensional measurements to be combined through an optimal weighting and scoring system to compute a quantitative health index of power transformers. The weighting of each test was assigned based on a smoothly interpolated continuous function. The efficacy of the model has been validated against expert classifications and data sets published in the literature. The comparative results demonstrate that, the proposed method is reliable and very effective for condition assessment of transformers through an automated health index calculation.

1. Introduction

Power transformers are the most valuable assets in a power grid and comprise a significant fraction of the total investment in a power delivery system. The failure of a transformer can have a profound impact on the end users or utilities along with the cost implications of replacement, lost revenue and customer impacts [1,2]. Moreover, a sudden failure can damage the environment through oil leakages and can pose a risk to utility personal by causing fire and explosions. Therefore, to maintain an optimal balance among maintenance costs, safety and capital investments, proper condition assessment of these valuable assets is essential.

Over recent decades, health index (HI) calculations have been used as a powerful tool to assess the condition of transformers. In these HI calculations, a transformer's current information such as various test results, expert observations and field inspection data are combined into a single quantitate index to reflect its overall condition. Although the HI cannot reflect the status of any particular part of a transformer, it measures the level of long-term degrading that cannot be easily determined by one-off inspections. The calculated HI score is a result of interaction between different routine and diagnostic tests that are not considered by the classical condition monitoring techniques [3]. Thus the calculated HI can identify the transformers that are close to their end-of-life and differentiate the transformers that have a higher probability of failure [4–6]. This information helps utilities to manage their assets appropriately, through clearly identifying the transformers that need more attention or major capital expenditure.

The challenging task in any HI calculation is to identify the most significant measurements and incorporate them through justified weightings. One of the established practices of many utilities is to use the recommended conditional score and weighting factors supplied by industry standards organisations such as IEEE, IEC and CIGRE and combine the test results in a linear way. Mathematically, the linear approach can be expressed by the following equation.

$$h(\boldsymbol{x}) = \frac{w_1x_1 + w_2x_2 + \cdots + w_nx_n}{w_1 + w_2 + \cdots + w_n} \tag{1}$$

* Corresponding author.
E-mail addresses: md.islam@murdoch.edu.au (M.M. Islam), Gareth.Lee@murdoch.edu.au (G. Lee), sujeewa@nsbm.lk (S.N. Hettiwatte).

http://dx.doi.org/10.1016/j.ijepes.2017.06.008
0142-0615/

where h is a health index metric, w and x represent the weight and conditional score of each test respectively and n represents the number of tests included in a HI calculation.

After computing the HI, the conditional categories of transformers are determined based on the standard deviation within a known set of calculated scores. A similar approach has been used by Jahromi [4] where the HI scores of each transformer have been calculated by summing the product of scores and weights based on 24 tests. One of the limitations of this method is that a large number of tests need to be conducted simultaneous to apply the method. As some of the tests are conducted irregularly, at most points in time the set $\{x_1, x_2, \ldots, x_n\}$ is incomplete or contains outdated measurements. Moreover, a wide range of test results are treated equally and the method is insensitive to a single but critical result that could potentially lead a transformer to a catastrophic failure. Ahmed in [5], has used a Fuzzy Logic approach where the membership of a measurement in the five health condition categories was calculated using a number of Fuzzy sets (membership functions). To calculate the degree of membership, a certain portion of each set was overlapped with neighbouring sets using linear boundaries. Therefore, a measurement could be a member of single or multiple sets. The percentage of the overlapping zone is completely subjective and dependent on the experts' knowledge. Moreover, the derivation of the expert rules is also critical, since the tests are applied as a decision tree; hence the order of tests is important. Ahmed gave most importance to the furan concentration over other tests. However, the actual estimation of furan from oil sampling in some cases is difficult as furan can disappear due to evaporation through the open breathing system or become disassociated under oxidizing and increased temperature conditions [7,8]. Oil treatment in the form of reclamation, replacement or top-up also has an impact on the furan concentration. Therefore, the calculated HI based on Fuzzy rules may have a certain degree of inaccuracy. Zuo in [9] also proposed another nonlinear approach where the conditional HI scores of transformers were calculated by using a binary logistic regression (BLR). The regression model can be expressed by (2).

$$H(x;\beta) = \frac{1}{1 + e^{-(\beta_0 + \sum_{i=1}^{n} \beta_i x_i)}} \tag{2}$$

where β_0 is a constant and β_i is a coefficient that reflects the contribution of each independent test x_i.

It is apparent from (2) that the accuracy of the method is dependent on the estimation of a set of coefficients $\{\beta_i\}$. Although, the BLR model can learn a sigmoidal surface that is similar to a single layer perceptron, it has much more limited interpolation than can be achieved by a General Regression Neural Network (GRNN) as used here.

In order to overcome the shortcomings of the currently available methods, a novel GRNN technique has been adopted for this work. The GRNN tackles the HI calculation as a function interpolation problem consisting of a nonlinear mapping from a set of six input variables (containing information about the operational condition of transformers) onto a single output variable representing the predicted overall health condition of transformers. The GRNN is a non-parametric method as opposed to Fuzzy Logic and Binary Logistic Regression that require a manual choice of model complexity and estimation of the resulting parameters (or coefficients). This means the GRNN will naturally form a model whose complexity is justified by the availability of training examples, complying with Risannen's minimum description length principle [10]. The GRNN has the ability to control model complexity when the set of transformer data is constrained and this is a significant contribution arising from this work. This property also makes the GRNN well suited to multivariate interpolation of high-dimensional training examples. The GRNN is from the family of radial basis function networks and its sole parameter controls the smoothness of interpolation offered. It can learn quickly the underlying function of high dimensional measurements from limited number of training examples. The performance of the model is evaluated by making a comparison between predicted and measured values based on the mean square error. The arrangement of the paper is as follows. Section 2 describes the overview of General Regression Neural Network. Section 3 presents the methodology followed in this research. Section 4 describes results that were achieved with the method. Section 5 presents the comparative analysis with other methods and Section 6 presents a summary of the results and conclusions.

2. Overview of General Regression Neural Network

The General Regression Neural Network (GRNN) is a single-pass learning algorithm that uses radial basis functions as a form of standard statistical kernel regression [11,12]. It is a kind of probabilistic neural network that has the capability to converge to the underlying function of measurements by induction from a limited number of training examples. The GRNN possesses a simple network structure that is fast when learning and quickly converges to the optimal regression surface. It exhibits excellent approximation to arbitrary functions having inputs and outputs from sparse and noisy sources, and can achieve optimal performance by adjusting a single smoothing parameter. Therefore, it has numerous applications in engineering and scientific data analysis.

In the GRNN, the value of a dependent variable $Y = \{Y_n\}$ is predicted from a number of given D-dimensional independent measurements $\boldsymbol{X} = \{X_n\}$ where $X_n \in \mathscr{R}^{\mathrm{D}}$. The prediction is the most probable value of Y for each value of X based on a finite set of n measurements and their associated Y values. Thus the GRNN learns a mapping from an input domain containing X to an output codomain containing Y, where either space can be multidimensional, but here Y is assumed as a scalar. If a GRNN is trained with a finite number of available measurements, it can estimate a linear or nonlinear regression surface to predict the most probable value of Y for any new measurement of X. If the joint continuous probability distribution function (pdf) $f(\boldsymbol{X}, Y)$ of a random variable vector $\boldsymbol{X}$ and a scalar random variable Y is known, the conditional expectation [11] of Y for a given value of X can be expressed as:

$$E[Y|\boldsymbol{X}] = \frac{\int_{-\infty}^{+\infty} Y f(\boldsymbol{X}, Y) dy}{\int_{-\infty}^{+\infty} f(\boldsymbol{X}, Y) dy} \tag{3}$$

In practice, for most cases the joint probability distribution function of $\boldsymbol{X}$and Y is unknown. However, the joint pdf can be empirically approximated using nonparametric Parzen window estimation from a finite set of measurements (a form of training set) [13]. This approach is free from any assumption of a particular functional form, as the parameters of the model are determined directly from the training set. Thus the GRNN learns and is able to generalize immediately. Cacoullos in [14] has shown the suitability of Parzen window estimation for multidimensional cases. The probability density of X at any sample point x in a hypercube region $\mathscr{R}^{\mathrm{D}}$ can be estimated [15] as:

$$f(\boldsymbol{x}) = \frac{1}{N\sigma^D} \sum_{n=1}^{N} k\left(\frac{x - X_n}{\sigma}\right) \tag{4}$$

where σ is the side length of a hypercube, σ^D is the volume of the hypercube and $k(x)$ is a statistical kernel or window function [15].

To estimate the joint pdf of $\boldsymbol{X}$and Y, different kernel functions $k(x)$ such as rectangular or Gaussian kernels can be used. As the Gaussian function is simple to implement and has continuous