ISSN 2250-3153

Optimization of LS-SVM Parameters Using Genetic Algorithm to Improve DGA Based Fault Classification of Transformer- A Review

Ms. Aparna R. Gupta, Prof (Mr). V. R. Ingle

Abstract- This paper gives the comparison of different methodologies for the fault classification in transformer based on Dissolved Gas Analysis. This paper describes the Dissolved Gas Analysis based fault classification of Transformer using Least Square Support Vector Machine. The parameters of Support Vector Machine are optimized using Genetic Algorithm. Failure of a large power transformer not only results in the loss of very expensive equipment but it can cause significant collateral damage as well.

Index Terms- dissolved gas analysis, least square support vector machine, genetic algorithm, transformer

I. Introduction

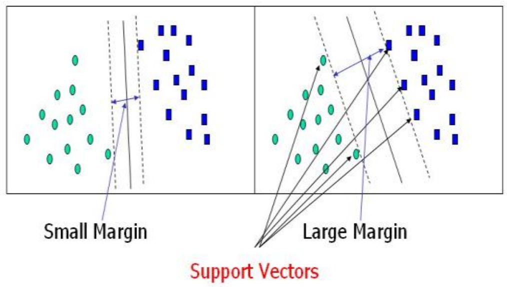
Dissolved Gas Analysis:
The transformer oil provides both cooling and electrical insulation. It baths every internal component and contains a lot of diagnostic information in the form of dissolved gases. Since these gases reveal the faults of a transformer, they are known as Fault Gases. They are formed in transformer oil, due to natural ageing and as a result of faults inside the transformer. The detection method is called Dissolved Gas Analysis (DGA). **DGA** is the study of dissolved gases in transformer oil. It is the most sensitive and reliable technique which gives an early indication of abnormal behavior of a transformer. Formation of fault gases is due to oxidation, vaporization, insulation decomposition, oil breakdown and electrolytic action.

The different types of gases in transformer oil are Hydrogen(H_2), Methane(CH_4), Ethane (C_2H_6) , Ethylene (C_2H_4), Acetylene(C_2H_2), Carbon Monoxide (CO), Carbon Dioxide (CO_2), oxygen, nitrogen. Also, gases such as propane, butane, butene and others can be formed as well, but their use for diagnostic purposes is not widespread. The concentration of the different gases provides information about the type of incipient-fault condition present as well as the severity. Different methods Rogers, fuzzy, neural, key gas method, duval, dornenburg ratio etc. are available for fault detection using DGA data.

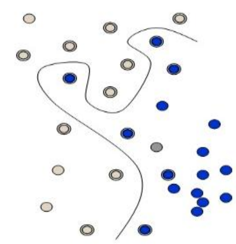
B. Support Vector Machine

A support vector machine (SVM) is a concept in computer science for a set of related supervised learning methods that analyze data and recognize patterns, used for classification and regression analysis. The standard SVM takes a set of input data and predicts, for each given input, which of two possible classes the input is a member of, which makes the SVM a non-probabilistic binary linear classifier.

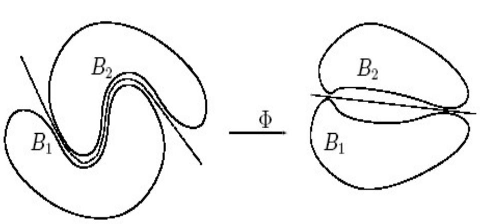
A SVM performs classification by constructing an *N*-dimensional hyperplane that optimally separates the data into two categories. One reasonable choice as the best hyperplane is the one that represents the largest separation, or margin, between the two classes.



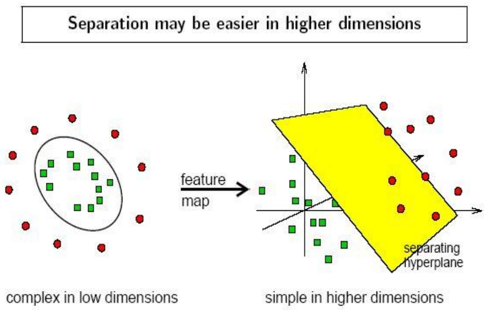
The simplest way to divide two groups is with a straight line, flat plane or an N-dimensional hyperplane. But what if the points are separated by a nonlinear region.



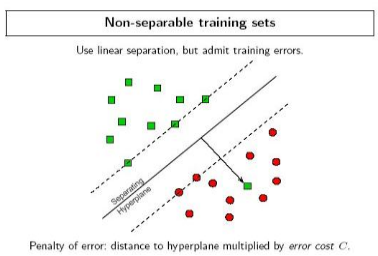
In 1992, Bernhard Boser, Isabelle Guyon and Vapnik suggested a way to create nonlinear classifiers by applying the kernel trick. The resulting algorithm is formally similar, except that every dot product is replaced by a nonlinear kernel function. The transformation may be nonlinear and the transformed space high dimensional; thus though the classifier is a hyperplane in the high-dimensional feature space, it may be nonlinear in the original input space.



ISSN 2250-3153



Ideally an SVM analysis should produce a hyperplane that completely separates the feature vectors into two non-overlapping groups. However, perfect separation may not be possible, or it may result in a model with so many feature vector dimensions that the model does not generalize well to other data; this is known as *over fitting*.



To allow some flexibility in separating the categories, SVM models have a cost parameter, *C*, that controls the trade off between allowing training errors and forcing rigid margins. It creates a *soft margin* that permits some misclassifications. Increasing the value of *C* increases the cost of misclassifying points and forces the creation of a more accurate model that may not generalize well. The effectiveness of SVM depends on the selection of kernel, the kernel's parameters, and soft margin parameter C.

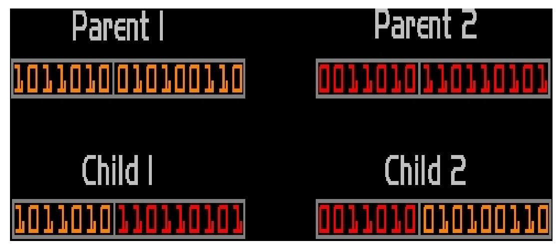
C. Genetic Algorithm

A genetic algorithm (GA) is a search heuristic that mimics the process of natural evolution. This heuristic is routinely used to generate useful solutions to optimization and search problems. Genetic algorithms belong to the larger class of evolutionary algorithms (EA), which generate solutions to optimization problems using techniques inspired by natural evolution, such as inheritance, mutation, selection, and crossover.

In a genetic algorithm, a population of strings (called chromosomes or the genotype of the genome), which encode candidate solutions (called individuals, creatures, or phenotypes) to an optimization problem, evolves toward better solutions. Traditionally, solutions are represented in binary as strings of 0s and 1s, but other encodings are also possible. The evolution usually starts from a population of randomly generated individuals and happens in generations. In each generation, the fitness of every individual in the population is evaluated, multiple individuals are stochastically selected from the current population (based on their fitness), and modified (recombined and possibly randomly mutated) to form a new population. The new population is then used in the next iteration of the algorithm. Commonly, the algorithm terminates when either a maximum number of generations has been produced, or a satisfactory fitness level has been reached for the population.

Selection: While there are many different types of selection. During each successive generation, a proportion of the existing population is selected to breed a new generation. Individual solutions are selected through a *fitness-based* process, where fitter solutions (as measured by a fitness function) are typically more likely to be selected. Certain selection methods rate the fitness of each solution and preferentially select the best solutions. Other methods rate only a random sample of the population, as this process may be very time-consuming.

Cross over: The next step is to generate a second generation population of solutions from those selected through genetic operators: crossover (also called recombination), and/or mutation. The most common solution is something called crossover, and while there are many different kinds of crossover, the most common type is single point crossover. In single point crossover, you choose a locus at which you swap the remaining alleles from one parent to the other. This is complex and is best understood visually.



As it can be seen, the children take one section of the chromosome from each parent. The point at which the chromosome is broken depends on the randomly selected crossover point. This particular method is called single point crossover because only one crossover point exists. Sometimes only child 1 or child 2 is created, but oftentimes both offspring are created and put into the new population. Crossover does not always occur, however. Sometimes, based on a set probability, no crossover occurs and the parents are copied directly to the new population. The probability of crossover occurring is usually 60% to 70%.

Mutation: After selection and crossover, a new population full of individuals is created. Some are directly copied, and others are produced by crossover. In order to ensure that the individuals are not all exactly the same, we allow for a small chance of mutation. You loop through all the alleles of all the individuals, and if that allele is selected for mutation, you can either change it by a small amount or replace it with a new value. The probability of mutation is usually between 1 and 2 tenths of a percent. A visual for mutation is shown below.

As you can easily see, mutation is fairly simple. You just change the selected alleles based on what you feel is necessary and move on. Mutation is, however, vital to ensuring genetic diversity within the population.

D. 1.4 Fault Classification procedure using GA and SVM

The SVM (support vector machine) uses the principle of minimization of structural risk to enhance the generalization ability, solves the problems of less sample, non-linear, local minima, it is fit for the gas forecasting of the transformer oil chromatography faults. The LS-SVM is an extension of the standard SVM The LS-SVM is an extension of the standard SVM, the quadratic term is used as the optimization index entries, and it also use the equality constraints instead of inequality constraints of the standard SVM, namely, the quadratic programming problem is transformed into a linear equation groups, it reduces the computational complexity, increases the speed of the solving.

The selection of the parameters of the optimal model influences the algorithm performance greatly. Whether the parameters are appropriate or not, could bring great impact on the training error and the punishment of the ample data, so it is necessary to optimize the parameters. GA is applied to choose the parameters of LS-SVM. The GA uses the encoding mechanism; it generates the initial population randomly, expends the search space fast, realizes the LS-SVM parameters optimization, and improves the fault diagnosis accuracy.

II. DGA TRANSFORMER REVIEW

LS-SVM (least square support vector machines) is applied to solve the practical problems of small samples and non-linear prediction better and it is suitable for the DGA in power transformers [1]. But in this model, the selecting of the parameters, impact on the result of the diagnosis greatly, so it is necessary to optimize these parameters. The IGA (improved genetic algorithm) is applied in this paper to make an optimization of these parameters about LS-SVM. In this 5 types of gases for fault types were used.

LS-SVM (least square support vector machines) is applied to solve the practical problems of small samples and non-linear prediction better and it is suitable for the DGA in power transformers [2]. But in this model, the selecting of the parameters, impact on the result of the diagnosis greatly, so it is necessary to optimize these parameters. The IGA (improved genetic algorithm) is applied in this paper to make an optimization of these parameters about LS-SVM. In this 7 types of gases for forecasting results of the gas in oil were used.

III. COMPARATIVE STUDY

Data mining (DM) technology based on Fuzzy Rough Set (FRS) and Support Vector Machine (SVM) are presented to classify the Fault of power transformer [3]. In order to make full use of the classification ability of SVM and improve the fault classification accuracy, FRS is used to pre-classify the transformer fault and the multi-level power transformer fault diagnosis model based on FRS and SVM.

A fuzzy model based on genetic programming (GPFM) is proposed to diagnose the fault types of insulation of power transformers [4]. The proposed GPFM algorithm constructs the fuzzy relationship between input and output fuzzy variables by genetic programming algorithms.

A new prediction method combined variable weight Gray Verhulst model and gray integrated relation grade was proposed to solve the problem of power transformer fault prediction [5]. Variable weight Gray Verhulst model was proposed based on 2 improved Gray Verhulst models with 2 difference select rules of parameter ρ in background function. Genetic algorithm chose parameter ρ for variable weight Gray Verhulst model.

Genetic-based neural networks (GNNs) for the assessment of the condition of power transformers [6]. The GNNs automatically tune the network parameters, connection weights and bias terms of the neural networks, to yield the best model according to the proposed genetic algorithm. Due to the global search capabilities of the genetic algorithm and the highly nonlinear mapping nature of the neural networks, the GNNs can identify complicated relationships among the dissolved gas contents in the transformers insulation oil and hence the corresponding fault types.

Reference papers	**[1]**	**[2]**	**[3]**	**[4]**	**[5]**	**[6]**
Methodology	IGA and LS-SVM	IGA and LS-SVM	FRS and SVM	GP and fuzzy	Grey Verhulst model and GA	GA and Neural Network
Samples required	Small samples	Small samples	large	large	Historical data (small samples, prediction)	Large
complexity	Reduced	Reduced	-	-	Two models(two GA each for posterior and prediction error)	-
accuracy	91.7%	-	91.7%	80.9%	93.7%	89.05% 92.22%

IV. PROPOSED METHODOLOGY

To increase the fault classification ability, LS-SVM parameters will be optimized using Genetic Algorithm. The Probable steps will be –

- Collection of DGA data
- Normalization of Data
- Classification of fault using LS-SVM
- Determine LS-SVM Parameters
- LS-SVM parameters will be Optimized using GA

Simulate the model to diagnose faults of the samples.

REFERENCES

[1] "*The Application of the IGA in Transformer Fault Diagnosis Based on LS-SVM* " Li Yanqing , Huang Huaping, Li Yanqing , Huang Huaping, Li Ningyuan, Xie Qing, Lu Fangcheng ; IEEE 2010.

[2] "Transformer Fault Gas Forecasting Based on the Combination of Improved Genetic Algorithm and LS-SVM" Li Yanqing , Huang Huaping, Li Ningyuan,Xie Qing, Lu Fangcheng ; IEEE 2010.

[3] "Application of Data Mining Technology Based on FRS and SVM for Fault Identification of Power Transformer" Zhihong Xue, Xiaoyun Sun, Yongchun Liang, in IEEE 2009, International Conference on Artificial Intelligence and Computational Intelligence.

[4] "A Genetic Programming Based Fuzzy Model for Fault Diagnosis of Power Transformers". Zheng Zhang, Kangling Fang and Weihua Huang in IEEE 2010, Third International Conference on Intelligent Networks and Intelligent Systems.

[5] "Transformer Fault Prediction based on GA and Variable Weight Gray Verhulst Model" Rui-rui Zheng, Ji-yin Zhao, Bao-chun Wu; IEEE 2009.

[6] "Condition assessment of power transformers using genetic-based neural networks", Y.-C. Huang; *IEE Proc.-Sci. Meus. Techiol, Vol. 1.50, No. 1. Junuary 2003.*

[7] "Fault Diagnosis of Transformer Based on Quantum behaved Particle Swarm Optimization-based Least Squares Support Vector Machines". Zhi-biao Shi, Yang Li, Yun-feng Song, Tao Yu; IEEE 2009.

AUTHORS

First Author – Aparna Gupta, B.D.C.O.E,Sevagram,Wardha, Addr-C/o Dr. R. K. Gupta, 37/2, Baid Layout, Gopuri Road, Wardha-442001,(MH) India, Mobile no.- 8600290176
apigupta01@gmail.com

Second Author – Prof (Mr).V. R. Ingle, Asst. Prof. P. G. Department of Engineering,B.D.C.O.E. Sevagram, Wardha-442001, (MH), India, Mobile no.- 9881489025