Ms. Aparna R. Gupta, Prof (Mr). V. R. Ingle, Dr. M. A. Gaikwad / Nternational Journal Of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com
Vol. 2, Issue4, July-August 2012, pp.1806-1809

LS-SVM Parameter Optimization Using Genetic Algorithm To Improve Fault Classification Of Power Transformer

Ms. Aparna R. Gupta*, Prof (Mr). V. R. Ingle, Dr. M. A. Gaikwad****

*(Research Scholar, B. D. College of Engg., Sevagram, Wardha(M.S.), India

** (P. G. Department of Electronics Engg. B. D. College of Engg., Sevagram, Wardha(M.S.), India

ABSTRACT

The 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. The selection of the parameters, impact on the result of the diagnosis greatly, so it is necessary to optimize these parameters. The parameters of Support Vector Machine are optimized using GA (Genetic Algorithm). The GA generates the initial population randomly, expands the search space fast and improves the global search ability and convergence speed. Finally, the optimized LS-SVM is used for analysis of multiple sets of DGA (Dissolved gas analysis) data of transformers, the results show that the accuracy of fault diagnosis has been effectively improved.

Keywords - Dissolved Gas Analysis, Least Square Support Vector Machine, Genetic Algorithm

I. INTRODUCTION

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. The DGA is the study of dissolved gases in transformer oil.

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. The soft computing techniques like Fuzzy, Neural and Neuro-fuzzy utilizes limited parameters where as the parameters are not compressive, hence resulted into inaccurate classification of it.

Support Vector Machines is a powerful methodology for solving problems in nonlinear classification. The LS-SVM is an extension of the standard SVM, the quadratic term is used as the optimization index entries, and it also uses the equality constraints instead of inequality constraints of the standard SVM, namely, the quadratic term programming problem is transformed into a linear equation groups, reducing the computational complexity, increasing the speed of the solving. It solves the problems of less sample and nonlinear data.

The accuracy of an SVM model is largely dependent on the selection of the model parameters. This paper uses genetic algorithm to optimize the parameters of LS-SVM. Genetic algorithm uses selection, crossover and mutation operation to search the model parameter. This paper compares the results obtained by traditional IEC Ratio method, LS-SVM and LS-SVM parameters improved by genetic algorithm.

II. Least Square 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 as shown in figure1.

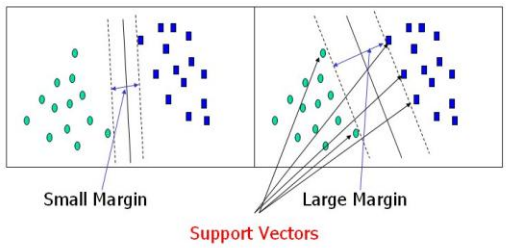

Figure: 1 Linear Separation

The simplest way to divide two groups is with a straight line, flat plane or an N-dimensional hyperplane.

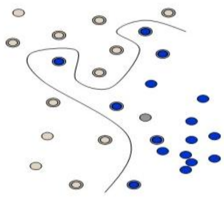

Figure: 2 Nonlinear Separation

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 as shown in Figure 3.

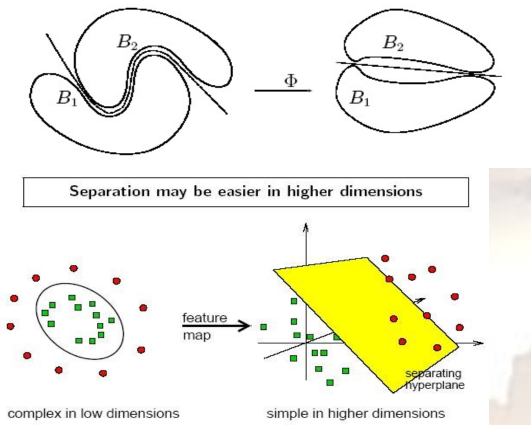

Figure 3: Higher Dimension Plane

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*.

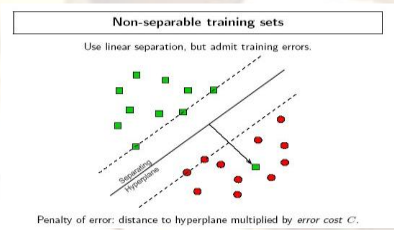

Figure: 4

To allow some flexibility in separating the categories, as shown in Figure 4 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.

III. 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. 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: A population of 756 different values of γ and σ^2 was created. The parameter γ controls the penalty degree, and the parameter σ^2 is the kernel function parameter. While there are many different types of selection. The selection is done randomly of 10 such values of γ and σ^2. 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.

Cross over: The next step is to generate a Child thro crossover (also called recombination), and/or mutation. The most common solution is something called crossover.

Thus, crossover was applied and such eight child were produced. Adaptability of each child was checked using fitness function accuracy. Thus best child was then selected. At last both the best parent and best child fitness level were compared, and finally that value of γ and σ^2 was selected which had more chance of adaptability.

IV. Fault Diagnosis procedure using IECRatio Method

Insulating oils under abnormal electrical or thermal stress breakdown to liberate small quantities of gases. The composition of these gases is dependent upon type of fault. By means of dissolved gas analysis (DGA), it is possible to distinguish fault such as partial discharge (corona), overheating, and arcing in a great variety of oil filled equipment. DGA can give early diagnosis and increase the chances of finding the appropriate cure.

Total 94 DGA samples where collected, which contained seven types of gases, CO_2 (Carbon

Dioxide), H_2 (Hydrogen), CH_4 (Methane), C_2H_6 (Ethane), C_2H_4 (Ethylene), C_2H_2 (Acetylene), and CO (Carbon Monoxide). As per the IEC 60599 firstly the safe limits of gas concentrations where considered. After analysis it was found that there where 28 Normal cases. Thus the remaining 66 suspicious cases where then applied with IEC Ratio method.

Diagnosis of faults is accomplished Via a simple coding scheme based on ranges of the ratios as shown in table 1.

X= C_2H_2 / C_2H_4
Y= CH4/ H_2
Z= C2H4/ C_2H_6

Table 1: IEC Ratio Codes

Ratio Code	Range	Code
X	<0.1	0
	0.1-1.0	1
	1.0-3.0	1
	>3.0	2
Y	<0.1	1
	0.1-1.0	0
	1.0-3.0	2
	>3.0	2
Z	<0.1	0
	0.1-1.0	
	1.0-3.0	0
	>3.0	1
		2

The combination of the coding gives 9 different types of transformer faults. The type of faults based on the code is shown in table 2 below.

X	Y	Z	Fault Type
0	0	0	Normal ageing
0	1	0	Partial discharge of low energy density
1	1	0	Partial discharge of high energy density
1-2	0	1-2	Discharge of low energy
1	0	2	Discharge of high energy
0	0	1	Thermal fault <150C
0	2	0	Thermal fault 150-300C
0	2	1	Thermal fault 300-700C
0	2	2	Thermal fault > 700C

Table2: Classification based on IECRatio Codes.

The right guess by IEC Ratio method was of 30 cases out of 94 cases, thus the accuracy was 31.91%. To improve the fault diagnosis accuracy the LS-SVM model with optimized parameters using genetic algorithm is then applied.

V. Fault Diagnosis Based on GA and LS-SVM

This paper uses GA to optimize two important parameters γ and σ^2 in LS-SVM model. The parameter γ controls the penalty degree, and the parameter σ^2 is the kernel function parameter. These parameters greatly affect the accuracy of fault diagnosis, so to have the best combinations about γ and σ^2 the optimization model parameter. The complete optimization flow chart is as in figure 5.

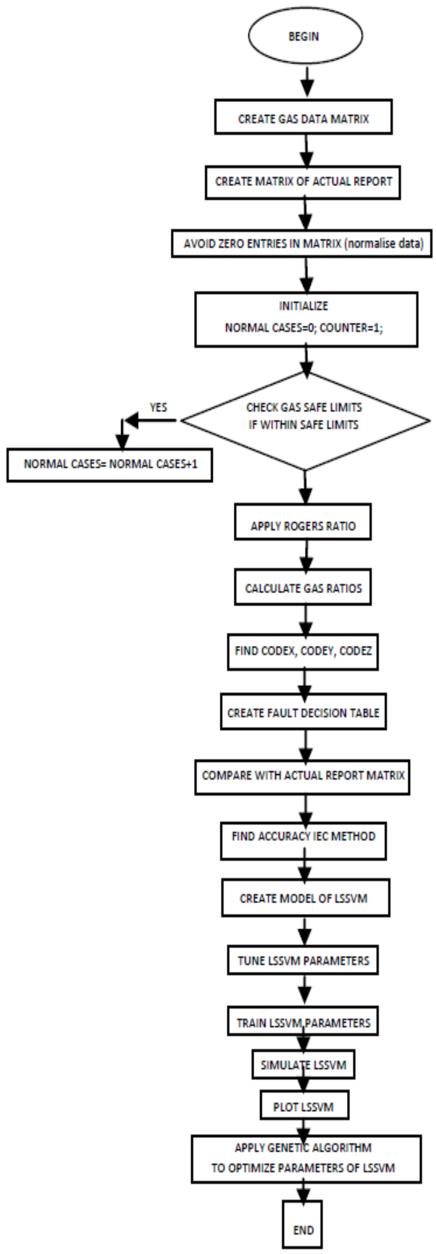

Fig 5: Flow chart of GA based LSSVM

VI. Simulated Results

The 94 DGA samples were collected in which 7 types of gases where present CO2, H_2, CH_4, C_2H_6, C_2H_4, C_2H_2, and CO. The table 3 below shows the correct fault diagnosis results of IEC code, LS-SVM and LS-SVM optimized by GA.

	Total Cases	Normal Cases	Suspicious Cases
	94	28	66
Right Guess IEC	30		
Accuracy IEC	31.91%		
Right Guess LS-SVM	52		
Accuracy LS-SVM	55.32%		
Right Guess GA-LS-SVM	64		
Accuracy GA-LS-SVM	68.09%		

VII. Conclusion

The Genetic Algorithm is applied to optimize the parameters of the LS-SVM, γ (gam) and σ (sig2). The LS-SVM solves the multi-classification problem of less samples and nonlinear data. Using genetic algorithm for optimizing LS-SVM parameters the fault diagnosis accuracy has been improved. The results show that it is much effectively than the IEC ratio method and traditional LS-SVM.

This method can improve the precision greatly, it is useful for fault diagnosis of power transformer.

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. January 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.

[8] "On the Kernel Widths in Radial-Basis Function Networks" NABIL BENOUDJIT and MICHEL VERLEYSEN, Neural ProcessingLetters 18: 139–154, 2003.

[9] "Interpretation of Gas-In-Oil Analysis Using New IEC Publication 60599 and IEC TC 10 Databases" Michel Duval, Alfonso dePablo in DEIS Feature article, March/April 2001 — Vol. 17, No. 2.

[10] "Article " Jashan Deep Singh, Yog Raj Sood and Piush Verma in The IUP Journal of Electrical and Electronics Engineering. October, 2008 .

[11] "Application of LS-SVM by GA for Dissolved Gas Concentration Forecasting in Power Transformer Oil " XIE Hong-ling, LI Nan, LU Fang-cheng, XIE Qing, in IEEE 2009.

[12] "Intelligent Fault Diagnosis in Power Transformer with Using Dissolved Gas Analysis in different Standards by Fuzzy Logic" AmirHoseyn Zaeri ; Rahmat Houshmand ; Mohammad Kafi in Majlesi Journal of Electrical Engineering. 2007. vol.1. Issue 2.

[13] "A Thermodynamic Approach to Evaluation of the Severity of Transformer Faults" Fredi Jakob, Member, IEEE, Paul Noble, and James J. Dukarm, Member, IEEE 2011.

[14] "A Review of Faults Detectable by Gas-in-Oil Analysis in Transformers" Michel Duval, in IEEE 2002

[15] "The Research on the Method of Condition Estimate of Power Transformer Base on Support Vector Machine" Zhang Weizheng, Yang Lanjun, Du limin in 2010 China International Conference on Electricity Distribution .