J. Basic. Appl. Sci. Res., 3(1s)125-129, 2013
© 2013, TextRoad Publication

ISSN 2090-4304
Journal of Basic and Applied Scientific Research
www.textroad.com

Introduction of Improved XCSR Algorithm using Limited Training Data: A Case Study for Fault Diagnosis in Analog Circuits

N. Moshtaghi Yazdani[1], A. Yazdani.Sequerloo[2], M. Shariat Panahi[3]

University of Tehran

Received: June 10 2013
Accepted: July 9 2013

ABSTRACT

Daily advancement of electronic science and analog/digital circuits has resulted in circuits with complicated tasks. Increased reliability of these systems, along with correct test, fault diagnosis and troubleshoot of these circuits have become very important and critical issue. Step-by-step examination of the circuits during manufacturing and before its delivery to user is mentioned as a technique to enhance reliability of the circuits. Therefore, the best approach would be generation of a template of faults. Since extended classification system (XCS) is known as one of the most successful learning agents for problem solving, XCS and other sample-based learning algorithms are utilized in this paper to diagnose the fault in analog circuits. For example, an analog to digital converter (ADC) is used in this regard. Efficiency of these methods is also evaluated through comparison of the results (about the sample problem).

KEYWORDS: analog to digital converter, extended classification system (XCS), neural network , reliability.

1-INTRODUCTION

Electronics is the science of studying transmission of electrical current through various materials (e.g. semiconductors, resistances, inductors and capacitors) and its associated effects. The electronic circuits are used to perform different tasks. The main applications of the electronic circuits are: control and process of data, conversion and distribution of electric power.

A typical electric system can be divided into three parts:

Input: electronic and mechanical sensors (or energy converters). These equipments receive signals or information from outside and convert them to current, voltage or digital signals.

Signal processor: These circuits are in fact responsible for management, interpretation and conversion of the input signals for being used in appropriate application. Digital signal processors are usually responsible for processing the mixed signals in this section.

Output: Activators or other equipments (like energy converters) that convert the voltage/current signals to proper output.

On the other hand, the electronic circuits are categorized under two groups: analog and digital.

Most of the analog electronic devices are comprised of several basic circuits. The analog circuits unlike the digital circuits use continuous amplitude of voltage. There a large number of different analog circuits, because an analog circuit may contain a circuit with only one element or thousands of elements. These circuits are also called linear circuits, although many nonlinear elements like detectors, compilers and etc. are used in them. Transistor amplifiers or vacuum lamps, operational amplifiers and oscillators are some good examples for the analog circuits. Today, some of the analog circuits benefit from digital elements or even microprocessors to improve performance of the circuit. These circuits are usually called “mixed signal”. It might be difficult to distinguish an analog circuit from a digital one, because both linear and nonlinear elements are used in some of the circuits. The digital circuits are those which are designed based on some separate levels of voltage. These circuits are the most common examples for introduction of Boolean algebra and form the basic principles of all digital computers. In most cases, there are two voltage states in a digital circuit which are shown with High and Low. For this particular case, the high voltage is a non-zero voltage which differs depending on the source type. Continuous developments in technology of manufacturing the digital circuits have led to produce high frequency switching elements in micrometer dimensions and even smaller [1]. It has become rather impossible to distinguish the digital parts behind them which are described with the analog behaviors. AT the same time, fault diagnosis in the digital circuits has been properly advanced, but it is still difficult to test the analog circuits. Parameters of the analog circuits can take any value from zero to infinity when a fault occurs. The former stands for a short circuit, while the latter is equivalent with an open circuit. No one can analyze the circuit of test in the whole range of possible values of the part. Connection between

*Corresponding Author: N. Moshtaghi Yazdani, University of Tehran, Arezoo.Yazdani@ut.ac.ir

changes of the elements and specifications of the circuits are usually nonlinear upon fault in the analog circuits. Thus, fault diagnosis is rather difficult and complicated in all working conditions. Analysis of the analog circuits is more complicated and difficult than the digital circuits. This is because the analog circuits must be analyzed sequentially which will cause impractical and incorrect analysis times [2]. Meanwhile, diagnosis of fault and its exact location are of great importance in the test systems. A great deal of research work is currently being done in this field. Verilog and VHDL software languages contribute to analyze the analog circuits in the digital environment. However, creation of an accurate and effective system of fault diagnosis in the analog circuits is one major challenge in testing these circuits [2]. There are two general methods for the fault detection problem, namely model-based method and data-based method. In the former, the user has access to the system model the behavior of which is monitored and analyzed the data by running mathematical model acquisition [7]. This model can be either analytical or based on personal information. Linear systems are one application of the model-based methods, since they can be simply introduced and examined. On the other hand, in the data-based method, the user directly adopts to analyze the data without acquisition of the mathematical model. Capability of this model is well understood when the system is monitored or it is nonlinear and complicated. Therefore, the system is rather complicated in spite of being linear, so it cannot be obtained from mathematical equations [6]. Two methods are mainly considered for learning rule from the samples, namely decision tree, and separation and solution. One rule is learnt at each step in the separation and solution approach. Then the samples covered by that rule are separated and this procedure is repeated on the remaining samples.

2- Sample-Based Learning Algorithms

K-NN or nearest neighborhood algorithm is one of the sample-based learning algorithms. This algorithm does nothing special during the learning and just stores the training samples. To show class tag of a data sample, the algorithm calculates the distance of this sample with other training samples. The most common way to calculate such distance is by Euclidean criterion, although there are some other criteria like Manhattan Minovsky which are used for this purpose [3]. Having calculated the distance, majority rule is conducted among k-nearest training sample to the current test sample, and majority tag of this sample is allocated to the test sample. K is a parameter which is judged by the user. These algorithms are also called lazy algorithms since they do nothing special in the learning phase and just store the test samples.

C4.5 is one of the most famous algorithms to make the decision tree [4]. This algorithm was first developed in 1993 from ID3 algorithm. Every middle node formed in the tree represents a test on the values of one property, while every branch stands for an allowed value of that property. The criterion used to select the appropriate property for a node is information gain which leads to create a bias in favor of the properties with various values. Gain Ratio is suggested to be used for solving this problem. ID3 algorithm just supports discrete properties, whereas C4.5 algorithm manages continuous properties in addition to the discrete ones. Moreover, management of the properties with unknown values is another advantage of C4.5 over ID3. To avoid overfitting in the developed classification model, tree pruning techniques are often recommended. The overfitting phenomenon occurs when accuracy of the classification model developed based on the training data is significantly high, but it does not yield a good accuracy on the test data set. In other words, the overfitted classification model is obtained from the training data and this overfitting will not necessarily lead to greater classification accuracy on the test data set. There are two major techniques for pruning a tree. Growth of the tree is stopped in some points before completion of the tree in the former technique which is called pre-pruning. On the other hand, in the second technique called post-pruning the tree grows completely and then some sub-trees are replaced with a leaf node then. The generated tree can be transformed into a set of equivalent classification rules. Pruning of the rules is implemented later by removing some of its preconditions.

Kstar algorithm is a sample-based learner which classifies every new record through comparing it with the available classification records in the data base. This algorithm presumes that the similar samples have the same classes. There are two basic sample-based learning components including distance function which determines similarity of the samples to each other, and classification function which defines how similarity of the samples would lead to a final classification for the new sample. Kstar algorithm is classified under lazy learners of the K-nearest neighborhood which uses fatigue to determine distance or similarity between the samples. The approach adopted by this algorithm to measure the distance between two samples is somehow derived from information theory. Based on this theory, the distance between two samples involves the complexity in transformation of one sample to another. Complexity is calculated in two steps. A constant set of transformations, which map some samples to some other ones, are defined first. A program for transformation of sample “a” to sample “b” is a constant series of transformations starting from “a” and ending to “b”. Such programs are made by adding a termination sign to each

sequence. Complexity of a program is generally defined as length of the shortest sequence that displays it. Based on this definition, the distance between two samples would be length of the shortest sequence which transforms two samples to each other. This approach concentrates on a transformation (the shortest one) among the possible transformations [5].

A wide range of training techniques have been introduced in the field of "machine learning" (either supervised or not) to make the machine needless of searching through a great volume of information and data. It also addresses a template which can be used to perform predictable (e.g. classification and regression) or descriptive (e.g. clustering) tasks. The methods which work based on a set of rules are known as the most famous machine learning methods. This is because they are more comprehensible thanks to the techniques the regularly adopt. They use a limited set of "action-condition" rules to demonstrate a small contribution of the entire solution space. Conditions address a part of the problem domain, while the actions indicate the decision based on sub-problems which have been highlighted by the conditions. In the base state, the classification systems include a set of rules, with each rule being a favorite solution for the target problem. These classifications gradually become effective by application of a supporting design in which genetic algorithm is formed on the dividers and causes

The first intelligent classification system developed by Holland (LCS) was designed to work for both discrete and continuous problems. This classification learning system is a sample machine learning which combines temporal differences and learning supervisions with genetic algorithm and adopts to solve both simple and complicated problems. Based on the supervision provided by Holland, LCS system uses a single property called power for each of the classifiers. Power of a classifier is indicative of its affectivity and is exclusively determined by the percentage of correlation the answer has with the expected answers. These criteria are known with principles which are discussed in the supervised training.

Some other types of LCS (e.g. improved XCS) have been developed after the first introduction of the main LCS. Before 1995, when the classification system was not introduced yet, ability of a classifier to find proper answers was effective on reinforcement system of them. Thereby, basic and simple classification systems were gradually transformed into more accurate decision making factors. It is currently believed that the improved XCS is able to solve even more complicated problems with no need to further modify the parameters. This system is thus accounted for the most successful learning system. However, based on the common approach proposed for training the XCS, fitness is increased only for rules which can answer the training data correctly. This means that the chance of each rule for survival and participation in reproduction process is directly dependent on how it responds to the training data, and realistic assessment of this chance requires application of a huge number of the training data. Moreover, neural network is comprised of a number of nodes and edges. Weight of each edge indicates how the node affects its adjacent node. A subset of the model nodes with no connection between them is called input nodes, while the other subset is denoted output nodes. Abnormality diagnosis is done in the following three steps by this technique: The first step includes determining type of the input and output data under study. The second step involves learning phase of the method which includes training the network and modulates the input-output based on weight of the edges. The input nodes acquire data from the system in the third step. Output of the network is observed at this step and according to the output value it can be judged whether

3. PROPOSED METHOD

In this method, the limited set of training data are generally used to modify specifications of the rules (i.e. "prediction", "prediction error" and "fitness"). This is done using the following equations:

Updating Prediction and Prediction Error

If $exp_i < 1/\beta$ then $P_i = P_i + (R - P_i) / exp_i$, $\varepsilon_i = \varepsilon_i + (|R - P_i| - \varepsilon_i) / exp_i$

If $exp_i \geq 1/\beta$ then $P_i = P_i + \beta (R - P_i)$, $\varepsilon_i = \varepsilon_i + \beta (|R - P_i| - \varepsilon_i)$

Updating Fitness

If $\varepsilon_i < \varepsilon_0$ then $k_i = 1$

If $\varepsilon_i \geq \varepsilon_0$ then $k_i = \beta (\varepsilon_i / \varepsilon_0) - \gamma$

$F_i = f_i + \beta [(k_i / \sum k_j) - f_i]$

Where, β represents the learning rate, γ represents the accuracy power of the rule, ε stands for the prediction error, and exp stands for the experience of rule. P is the prediction of rule, while R is the reward received from environment. k denotes the accuracy of rule, and f shows its fitness. Finally, index i represents the number of rule in the set of rules.

In the next step and for development of the variety in the data set, "stochastic selection with residual" was used among the sequences which indicate the condition phase of the existing data, in order to select several pairs as the parent. At the same time, the condition part of the new data is also generated using middle crossover which is applied on these parent sequences. The value of each conditional variables is given by the equation below:

$$a_i = \alpha(a_i F) + (1-\alpha)(a_i M) \qquad (1)$$

Where, a_i is the value of the ith conditional variable within the new data, $a_i F$ represents the value of the ith conditional variable in the first parent (father), while $a_i M$ represents the value of the ith conditional variable in the second parent (mother). α is the participation coefficient of the parents which is determined adaptively. The action part of the new data is also made using a nonlinear mapping from the space of conditional variables to that of actions, which is generated using the existing data.

Diversification to the existing data continues until the termination condition (e.g. when percentage of the correct answers to the test data reaches a predefined threshold) is met by the completed data.

4. Application of Sample-Based Learning Algorithms for Testing ADC Circuits

Taking into account the daily increasing technology of the electronic circuits and their complexity, safe application of these systems and circuits entails proper tests to ensure their reliability. Any deficiency in these circuits might even cause irreparable damages since in critically important applications which deal with human life.

In testing the circuit, some inputs are introduced to the circuit under examination which is also called test template. Having applied these templates, response of the circuit to these stimuli is compared with the expected outputs, so that one can judge whether it is correct or not. Therefore, generation of test templates and required stimuli is the basic step to achieve the test rules. Various techniques have been suggested to make the test templates, including universal template generation and quasi-random.

An analog to digital converter (ADC) is used in this paper (Figure 1) to create fault data bank by choosing different fault models. Thereby, the results obtained from diagnosis of the sample-based learning algorithms are addressed.

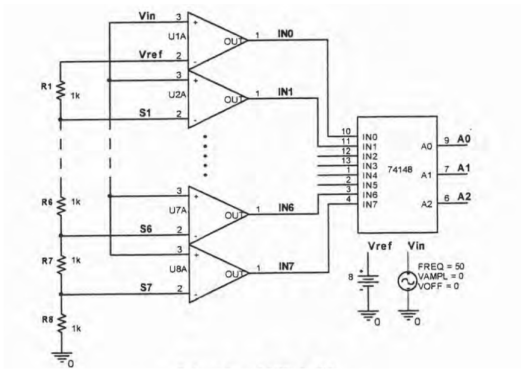

Fig.1. *ADC used in this research*

The faults which are introduced to the circuit are from the following models:

1) Bridge fault: it means that two points of the circuit are short circuited [8];
2) Closed tolerance fault: a given amount of tolerance is considered for one of the existing elements in the figure above, such that the changes beyond it would be inferred as fault [9,10];
3) Short circuit of a node of circuit with zero voltage;
4) Short circuit of a node of circuit with unit (source) voltage.

ORCAD 16.5 software was utilized to apply these faults in addition to obtain the samples (both training and test). It was observed after application of the faults that a bridge fault between 56 pairs of the circuit node, a tolerance fault in 8 resistances of the circuit, a short circuit fault for a node of the circuit to zero voltage in 18 nodes of the circuit, and a short circuit fault for a node of the circuit to unit voltage in 32 nodes of the circuit have occurred. The fault data bank was generated having applied these faults. 80% of these faults were considered for training the sample-based learning algorithms, while the other 20% were allocated for testing. The results of this evaluation have been summarized in Table 1 below.

Table1. *Evaluation of the proposed algorithm in comparison with other methods*

Algorithm	C4.5	LCS	K-NN	Kstar	XCS	neural network	proposed method
Accuracy	78.56%	80.47%	85.24%	85.62%	87.6%	92.5%	91.14%

5. CONCLUSION

As can be seen in the table above, the neural network provides the best performance and diagnoses a greater percentage of different types of faults in the ADC circuit among all artificial intelligent methods. Neural networks are successful method in fault diagnosis. This is because of their excellent learning capability. However, many neural networks which are used in fault diagnosis still suffer from some deficiencies like incomprehensiveness, being time consuming, stability and flexibility. The proposed method diagnoses smaller percentage of different faults in the ADC circuits, but it needs considerably shorter training time in comparison with the neural network and no longer incorporates stability issues of the neural network. Therefore, with respect to the great fault diagnosis of this method, it can be used for this purpose in analog circuits and particularly in ADC circuits.

REFERENCES

[1] Meijer, P. B. (2001). Neural networks for device and circuit modelling. In*Scientific Computing in Electrical Engineering* (pp. 251-258). Springer Berlin Heidelberg.

[2] Chakrabarti, S., Cherubal, S., & Chatterjee, A. (1999). Fault diagnosis for mixed-signal electronic systems. In *Aerospace Conference, 1999. Proceedings. 1999 IEEE* (Vol. 3, pp. 169-179). IEEE.

[3] Sharpe, P. K., & Glover, R. P. (1999). Efficient GA based techniques for classification. *Applied Intelligence*, *11*(3), 277-284.

[4] Quinlan, J. R. (1993). *C4. 5: programs for machine learning* (Vol. 1). Morgan kaufmann.

[5] Witten, I. H., & Frank, E. (2005). *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann.

[6] J.starzyk , J.pang , "Fault Diagnosis in Analog and Mixed Model Low Testability System 4th Multi-Conference on system cybernetics and information (SCI2000) Orlando, Florida, July 2000 ,pp,452-460 Best Paper award .

[7] Singh, M., & Koren, I. (2003). Fault-sensitivity analysis and reliability enhancement of analog-to-digital converters. *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, *11*(5), 839-852.

[8] Mourad, S., & Zorian, Y. (2000). *Principles of testing electronic systems*. Wiley-Interscience.

[9] Lundberg, K. H. (2002). Analog-to-Digital Converter Testing. *Unpublished Manuscript.*

[10] Arpaia, P., Cennamo, F., & Daponte, P. (1999). Metrological characterisation of analog-to-digital converters-a state of the art.