ANFIS-Based Prediction of Head Performance Downgrade at Quality Test

Rangson Hunkon[1], Pornchai Supnithi[2]
[1]College of Data Storage Innovation,
[2]Faculty of Engineering, [1,2]King Mongkut's Institute of Technology Ladkrabang
Bangkok, Thailand
[1]rangson.hunkon@wdc.com, [2]ksupornc@kmitl.ac.th

Piya Kovintavewat[3]
[3]Data Storage Technology Research Center,
Nakhon Pathom Rajabhat University
Nakhon Pathom, Thailand
[3]piya@npru.ac.th

***Abstract*—The hard disk drive (HDD) quality failure is an important parameter that crucially impacts factory efficiency and productivity. In general, a major failure is normally caused by head performance downgrade. Thus, this paper proposes a method to predict this failure, which in turn can help improve HDD reliability. Specifically, we apply an analytic tool called an adaptive neuro-fuzzy inference system (ANFIS) to predict the head performance downgrade, and find an opportunity to reduce the quality failure. As an initial study, we found from an experiment that the ANFIS model can be employed to predict a failure from head performance downgrade at quality test with an accuracy of about 80%.**

***Keywords*—Adaptive neuro-fuzzy inference system (ANFIS), head performance downgrade, quality test.**

I. INTRODUCTION

Hard disk drive (HDD) is a growing business due to worldwide customers' demand. In practice, HDDs are important for data storage in every company and personal use (e.g., desktop, notebook, tablet, game boxes, and so on). Hence, HDD manufacturer must improve the processes continuously to get the highest output with the highest quality to support customer requirements. Generally, the HDD manufacturing process can be divided into two processes. The first build is done in a class 100 clean room [1], where contamination is carefully controlled. Next, the second build is performed in a backend area, where drives are tested for performance and quality control before shipping them to customers, as illustrated in Fig. 1.

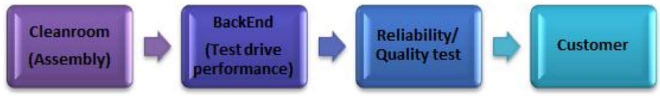

Fig. 1. A flow of hard disk drive manufacturing process.

Practically, HDD manufacturer must continuously improve the manufacturing processes and the quality performance so as to support customer's requirements. However, we still found many HDD failures returned from customers. In general, the failure at quality test is about 1.5%, whose major cause is head performance downgrade. This paper is aimed at investigating if there is a possible way to identify such potential failures and predict the failures at quality test so as to let them fail at the backend test process and also prevent failed drives shipping to the customers.

To study the prediction and reduce head performance downgrade at quality test, we found that the existing key head stack parameters are BLP (baseline popping), VGA (variable gain amplifier), SNR (signal-to-noise ratio), MEW (magnetic erasure width), OW (overwrite), and EM (error margin) from the backend test process. These parameters are required for an analytic tool called an adaptive neuro-fuzzy inference system (ANFIS) [2] to predict the failures.

Practically, ANFIS is a fuzzy inference system implemented in a framework of an adaptive network. By using a hybrid learning procedure, the ANFIS can construct an input-output mapping, based on human knowledge and stipulated input-output data pairs. In general, ANFIS is suitable for modeling a non-linear system and predicting a chaotic time series. For example, Roy [3] employed the ANFIS to predict the surface roughness in a turning operation for a set of given parameters with two different membership functions (MFs) and then compared the prediction accuracy. It was found that the bell-shaped MF has the prediction accuracy of 97.84%, whereas the triangular MF has the prediction accuracy of 96.13%. Altaher [4] studied a neural fuzzy classifier based on ANFIS for malware detection, which was found that this ANFIS classifier can detect the malware exe files effectively. Nazmy and Messiry [5] presented an intelligent diagnosis system using a hybrid approach of ANFIS for classification of the electrocardiogram signals, whose results indicated an accuracy level of more than 97%. Tepin [6] proposed a neural network rank level fusion applied on key parameters measured in the manufacturing process to predict customer failures resulted from head disk interaction (HDI). It was found that the result of rank level fusion classification model is able to achieve 86.61% accuracy for testing samples, and potentially to affect HDI failure. Then, Asawatongtip [7] introduced a method to predict the root causes of drive downgrade during the assembly and test processes by data mining using the Bayesian network, which a achieve an accuracy of 80.7%. Therefore, this paper proposes a new prediction method based on an ANFIS model by using the head stack input parameters to predict head performance downgrade at quality test.

The rest of this paper is organized as follows. Section II briefly summarizes an ANFIS model. Section III explains the experimental method. Simulation results are given in Section IV. Finally, Section V concludes this paper.

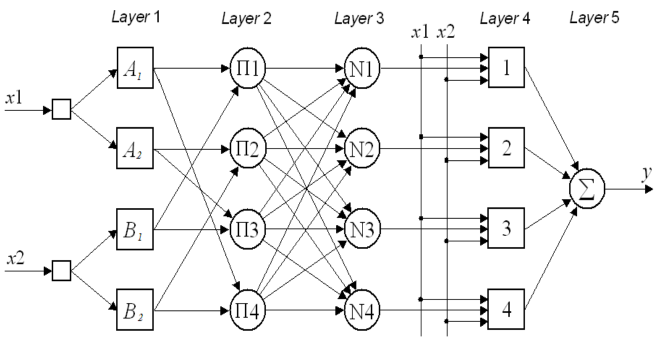

Fig. 2. Adaptive neuro-fuzzy inference system (ANFIS).

II. Adaptive Neuro-Fuzzy Inference System

Fuzzy inference systems are efficient techniques for studying the behavior of nonlinear systems by using fuzzy logic rules. ANFIS is an adaptive neuro-fuzzy inference system that uses the learning techniques of neural networks, which is usually employed in many applications in control and prediction. Practically, ANFIS utilizes a hybrid learning algorithm to specify parameters. Specifically, it uses the least-squares method with the back propagation gradient descent method to train the ANFIS membership function parameters based on a given training data set.

In general, the ANFIS structure is similar to a neural network structure. Fig. 2 illustrates the ANFIS model based on Takagi and Sugeno model [8]. Clearly, it consists of 5 layers connected through direction links, where the 1st is a fuzzy layer, the 2nd is a product layer, the 3rd is a normalized layer, the 4th is a de-fuzzy layer, the 5th is a total output layer, x_1 and x_2 are the inputs, and y is the output. Note that each layer is characterized by a node function with fixed adjustable parameters.

According to the Takagi and Sugeno model [8], the rule sets and the function of each layer are as follows

Rule Set:

If (x_1 is A_1) and (x_2 is B_1) then $f_1 = p_i x_1 + q_i x_2 + r_i$

If(x_1 is A_2) and (x_2 is B_2) then $f_2 = p_2 x_1 + q_2 x_2 + r_2$

where p_1, p_2, q_1, q_2, r_1 and r_2 are linear parameters and A_1, A_2, B_1 and B_2 are non-linear parameters.

Layer 1: It is an input fuzzy layer. Every node in this layer is an adaptive node that satisfies the following equations

$$O_{1,i} = \mu_{A_i}\left(x_1\right), \qquad i = 1, 2 \tag{1}$$

$$O_{1,i} = \mu_{B_{i-2}}\left(x_2\right), \qquad i = 3, 4 \tag{2}$$

where $O_{1,i}$ denote the output functions, μ_{A_i} and μ_{B_i} denote of the membership functions.

A membership function for a fuzzy set A on the universe of discourse x is defined as $\mu_A(x) \rightarrow [0, 1]$, where each element of x is mapped to a value between 0 and 1. This value quantifies the grade of membership of the element in x to the fuzzy set A. The membership function (MF) allows to graphically represent a fuzzy set, where the x axis represents the universe of discourse, and the y axis represents the degrees of membership in the [0,1] interval. For instance, if the triangular MF is employed, μ_{A_i} is given by

$$\mu_{A_i}\left(x\right) = \max\left[\min\left(\frac{x - a_i}{c_i - a_i}, \frac{c_i - x}{c_i - b_i}, 0\right)\right], \tag{3}$$

where a_i, b_i, and c_i are the MF parameters. On the other hand, if the generalized bell-shaped MF is used, μ_{A_i} will be given by

$$\mu_{A_i}\left(x\right) = \frac{1}{1+\left(\frac{x-c_i}{a_i}\right)^{2b_i}}, \tag{4}$$

where a_i, b_i, and c_i are the MF parameters.

Layer 2: It is a product layer, where each node in this layer computes the impact of each rule through the multiplication by

$$O_{2,i} = w_i = \mu_{A_i}\left(x_1\right)\mu_{B_i}\left(x_2\right), \tag{5}$$

where $O_{2,i}$ is the output of layer 2.

Layer 3: It is a normalization layer, which computes the normalized effect of a given rule by

$$O_{3,i} = \overline{w}_i = \frac{w_i}{w_1 + w_2}, \tag{6}$$

where $i \in \{1, 2\}$, and $O_{3,i}$ is the output of layer 3, which can be called "normalized firing" strength are normalized with a maximum equal to 1 and a minimum equal to 0.

Layer 4: It is a de-fuzzy layer, where the parameters in this layer are considered as consequent parameters that follow

$$O_{4,i} = \overline{w}_i\ f_i = \overline{w}_i\ \left(p_i x_1 + q_i x_2 + r_i\right), \tag{7}$$

where $O_{4,i}$ is the output of layer 4 and $\{p_i, q_i, r_i\}$ are called linear parameters or consequent parameters.

Layer 5: It is a total output layer to calculate the sum of the output for all incoming signals given by

$$O_{5,i} = \sum_i \overline{w}_i \ f_1 = \frac{\sum_i w_i f_i}{\sum_i w_i} \ , \tag{8}$$

where $O_{5,i}$ is the total output layer and $\sum$ computes the sum of all incoming signals.

III. Experimental Method

This study presents a prediction method of head stack performance downgrade at HDD quality test, which is based on ANFIS functions in MATLAB R2008b, where 6 input parameters (BLP, VGA, SNR, MEW, OW and EM) and 1 output parameter (Pass or Fail) are considered in our ANFIS model. Note that based on our observation from historical data statistics, we found that these 6 input parameters in the HDD backend test process are the most relevance to the head performance at the quality test.

First, we collect the unseen data used for testing about 10000 data points in the manufacturing test process. Then, we divide the data (passed and failed) into two groups randomly, with 80% for a training data set and 20% for a verifying data set. In the training process, we separate it into 2 steps, where the first step is to find the MF by using a "genfis2" function in MATLAB and to determine the membership boundary by radius between 0.4 and 1.0, and the second step is to adjust the ANFIS model to obtain the best settings for prediction. After the ANFIS model with proper parameters is obtained, we then verify the accuracy of the model by using the (unseen) verifying data set to verify the accuracy of the model.

Fig. 3 shows the data distribution of 6 input parameters, which is obtained by separating each input into 1 dimension and 2 dimensions for both passed (blue color) and failed (red color) drives. It is apparent that this distribution data is complicated to classify the data. For example, let us consider the two inputs BLP and EM (i.e., a lower-left corner figure). The scatter plot cannot tell us that the passed and the failed drives are not significant different. However, we can utilize the ANFIS model to classify and predict the failures as shown in simulation.

IV. Simulation Results

In this section, we will verify the sensitivity and the accuracy of the proposed ANFIS model that is used to predict the head performance downgrade, as depicted in Fig. 4 and Fig. 5, respectively, where the x-axis is a radius between 0.4 and 1.0, and the y-axis is the percentage of accuracy. In addition, the blue line represents 6 inputs (all features), the orange line represents 3 inputs (BLP, SNR, EM), the red line represents 2 inputs (BLP, EM) and the yellow line represents 2 inputs (SNR, EM). These inputs are the data used as the inputs to the ANFIS model. In general, we need to consider the results of both sensitivity and accuracy so as to explain the prediction performance.

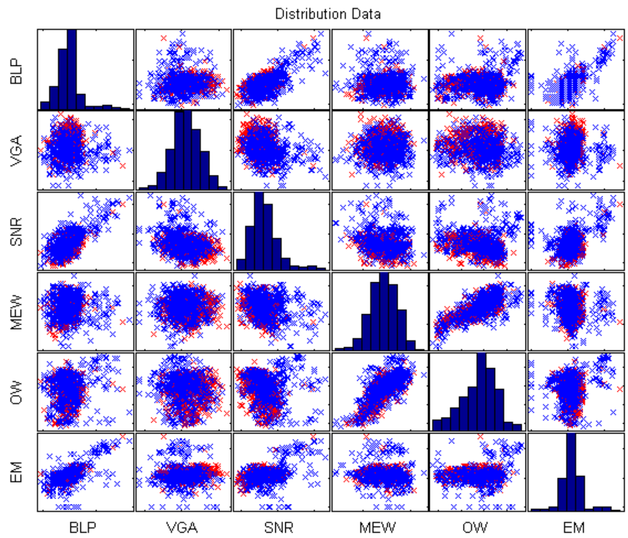

Fig. 3. The distribution data of 6 inputs parameter.

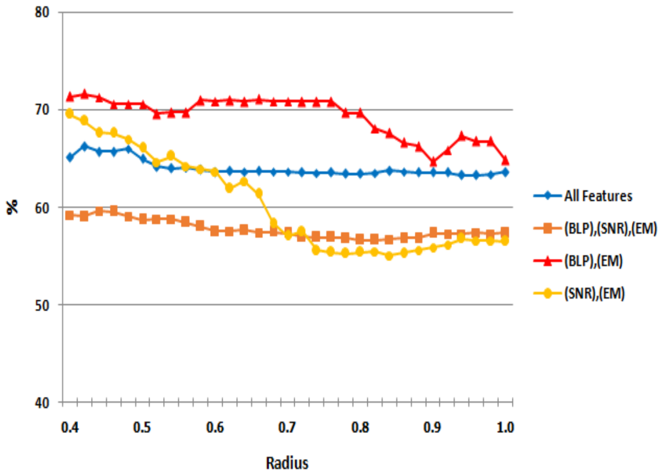

Fig. 4. Sensitivity of ANFIS prediction.

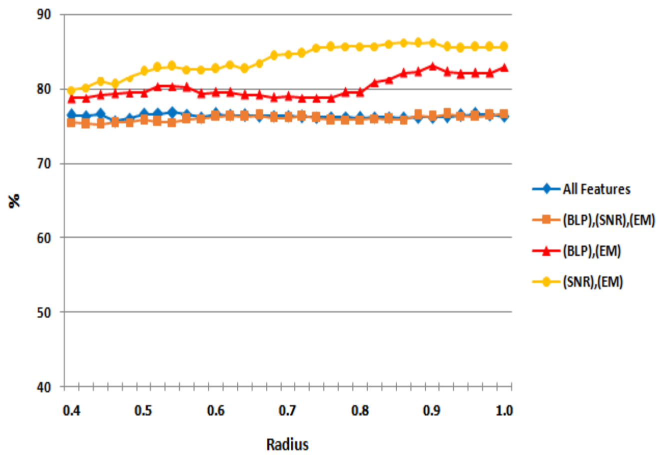

Fig. 5. Accuracy of ANFIS prediction.

In this work, the prediction accuracy and the sensitivity are evaluated by

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN}, \tag{9}$$

$$\text{Sensitivity} = \frac{TP}{TP+FN}, \tag{10}$$

where TP is True Positive (prediction of 1 when the sample test result has a 1), TN is True Negative (prediction of 0 when the sample test result has a 0), FP is False Positive (prediction of 1 when the sample test result has a 0), and FN is False Negative (prediction of 0 when the sample test result has a 1).

As shown in Fig. 4 and Fig. 5, if we use 6 inputs (all features) in the ANFIS model, it will approximately give the prediction error with sensitivity of 65% and accuracy of 77%. However, we found that using only 2 inputs (i.e., BLP and EM) as the inputs to the ANFIS model, we can obtain the best prediction result, i.e., the prediction error with sensitivity of 70% and accuracy of 80%.

Based on the result, we found that using two inputs will yield a better prediction result than using six inputs. This might be because six parameters may have more complexity and overfitting than two parameters. It should be noted that overfitting generally occurs when a model is excessively complex, such as having too many parameters relative to the number of observations. A model which has been overfitting will generally have poor predictive performance, as it can exaggerate minor fluctuations in the data.

V. CONCLUSIONS

This paper studies the prediction failure of head performance downgrade at quality test by using an ANFIS method and the data from the backend test process. As an initial study, we found that only two input parameters, namely BLP and EM, have a good relationship with quality failure. Specifically, we obtain the best prediction result at 80% accuracy. Therefore, it can be implied that the ANFIS model can be used to classify the complicated data and to predict the failures of head performance downgrade at quality test with good result. Nonetheless, it should be pointed out that the proposed ANFIS model can still be improved to achieve higher accuracy. This can be done by using a better input parameter that is closely related to the head performance downgrade, and optimizing the ANFIS parameters. Once we obtain a better ANFIS model, it will be useful to improve the quality performance of HDD manufacturing.

ACKNOWLEDGMENT

I would like to thank you Assoc.Prof.Dr.Pornchai Supnithi (King Mongkut's Institute of Technology Ladkrabang) and Assoc.Prof.Dr.Piya Kovintavewat (Data Storage Technology Research Center, Nakhon Pathom Rajabhat University), who provided me the guidance, advices and reviews my work. Also, I would like to thank you my WD Company to support everything during I study and do my thesis.

REFERENCES

[1] R. Duarte, "Environmental Monitoring of Clean Rooms", DDK Scientific, Corp., 2008.

[2] J. S. R. Jang, "ANFIS: Adaptive Neuro-Fuzzy Inference System", IEEE Transaction on System, vol. 23, no. 3, 1993.

[3] S. Roy, "Design of adaptive Neuro-Fuzzy Interface System for predicting surface roughness in turning operation", Journal of Scientific & Industrial research, vol. 64, 2005.

[4] A. Altaher, "Application of Adaptive Neuro-Fuzzy Inference System for Information Security", Journal of Computer Science 8(6): 983-986, 2012.

[5] T. M. Nazmy, H. EL-Messiry, "Adaptive Neuro-Fuzzy Inference System for Classification of ECG Signals", Journal of Theoretical and Applied Information Technology, 2009.

[6] W. Tepin, "Customer Failure Modes Prediction for Hard Disk Drive using Neural Networks Rank-Level Fusion", in Proc. of ECTI-CON, pp. 476 – 479, 2011.

[7] V. Asawatongtip, "Data Mining For Predict Hard Disk Drive Downgrade," Asian Institute of Technology, 2012.

[8] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control", IEEE Transactions on Systems, Man and Cybernetics, vol. 15, pp. 116-132, 1985.