Recycled integrated circuit detection using reliability analysis and machine learning algorithms

Udaya Shankar Santhana Krishnan | Kalpana Palanisamy

Department of Electronics & Communication Engineering, PSG College of Technology, Coimbatore, 641004, India

Correspondence
Udaya Shankar S., Department of Electronics & Communication Engineering, PSG College of Technology, Coimbatore, India. Email: rajkrish18392@gmail.com

Abstract
The use of counterfeit integrated circuits (ICs) in electronic products decreases its quality and lifetime. Recycled ICs can be detected by the method of aging analysis. Aging is carried out through reliability analysis with the effect of hot carrier injection and bias temperature instability (BTI). In this work, three machine learning methods, namely K-means clustering, back propagation neural network (BPNN) and support vector machines (SVMs), are used to detect the recycled IC aged for a shorter period (1 day) with minimum data size. This work also distinguishes the effects of degradation due to process variations and reliability effects. The reliability and Monte Carlo simulation are performed on benchmark circuits such as c17, s27, b02 and fully differential folded-cascode amplifier using the Cadence Virtuoso tool, and the parameters such as minimum voltage, delay value, supply current, gain, phase margin and bandwidth are measured. Machine learning methods are developed using MATLAB to train and classify the parameters. From the results obtained, it is observed that the classification rate for the benchmark circuits is 100%, and using BPNN, K-means clustering and SVM and the proposed method, recycled IC or used IC is detected even if it was used for 1 day.

1 | INTRODUCTION

Counterfeit integrated circuits (ICs) are a major problem in the electronic supply chain because of its reliability and security issues. This affects many devices used in the various application fields such as biomedical, telecommunications, automotive, consumer electronics and defence. Counterfeit parts are of various forms such as recycled, overproduced, remarked, cloned, out-of-spec and defective. The recycled ICs degrade the device performance due to aging. In addition, the reliability effects also further degrade the performance of the devices over a period of time.

Degradation in the performance of the IC is impacted by a variety of working conditions such as bias voltage, temperature and workload. There are various reliability issues such as bias temperature instability (BTI), hot carrier injection (HCI), electromigration and time-dependent dielectric breakdown (TDDB) in ICs. These conditions change over a period of time due to the aging of devices. Several methods have been discussed in the literature to detect a recycled IC based on different operating conditions and parameters.

A prominent method used to detect a recycled IC is on-chip aging sensor-based detection. Further, the aging sensors have been enhanced with pre-sampling and stability checking [1], electromigration-induced detection [2] and timing warning capabilities [3]. Recycled ICs are detected with reliability analysis (RA) and statistical methods [4, 5], LUT-based prediction methods [6], enhanced BIST architectures [7] and reflected EM signatures [8]. Ring oscillator (RO)-based aging detection sensor is used for the detection of aged devices by measuring the delay of the various paths in the circuit [9–11]. The on-chip aging sensor-based detection methods either add an additional logic element or require a high-expensive component to detect the recycled IC. Several machine learning (ML) techniques are developed in the process of detection of recycled ICs. Some of the ML methods used are support vector machines (SVMs) [12–21], data correlation [22], K-means clustering [17–19], multivariate adaptive regression splines analysis (MARS) [23, 24], linear regression [25], vector quantization [20], Naïve Bayes classifier [20], random forest method [26], extreme gradient boosting [26], K-nearest neighbours [26] and neural networks (NNs) [26, 27]. These ML techniques do not need...
additional logic elements; it only trains the set of data samples to detect the recycled IC.

Parameters are measured from the ICs provided by the trusted and untrustworthy vendors [12]. One-class SVM [12, 13] is used to classify the brand new and the recycled IC, based on the decision function. The degradation due to the effect of NBTI in devices at different time periods is used to train the SVM to detect the counterfeit IC. This method is not suited to detect the recycled devices for a short period of aging.

Early failure rate analysis is used to collect parameters from fresh and recycled devices [16]. One-class SVM classifier (OCC) and degradation curve sensitivity analysis (DCSA) are used to classify the parameters. OCC has an impact on process parametric variations. If the process variation is large, then the recycled IC will fall within the boundary space. Stress has to be applied in the IC for a long time to make it go beyond the boundary. The parametric measurement with a similar degradation rate of fresh and recycled IC is selected in the DCSA in order to get a better detection of recycled ICs. This ML classification is verified with industrial benchmark circuits and an analog folded-cascode amplifier circuit. The accuracy of classification using OCC in [16] is affected by the impact of process variations in the parameters or features, while in DCSA, the parameters are independent of process variations, but DCSA could detect the recycled ICs aged more than 1 month.

BPNN and Elman NN [27] are used to predict the aging of CMOS low-noise amplifier (LNA) due to HCI and NBTI. The prediction model is developed based on a multistep-ahead prediction model. The parameters are obtained from the LNA in measurement and stress stages. The S parameters are used for degradation analysis. Noise figure (NF), supply voltage and third-order intercept point (IIP3) are used for degradation analysis due to its very small variations due to aging. Three different prediction models such as 6-step ahead, 9-step ahead and 12-step ahead prediction models are developed for training using the NN model. Elman NNS consume more time compared to BPNN [27]. The parameters with minor variation due to aging are not preferred for prediction. This limits only certain parameters that are eligible to predict aging in the circuit.

In [24], MARS-based aging prediction model is developed in order to train the time-variant working conditional parameters. The parameters may be temperature and workload activity. The process parametric variation compensation factor is determined by testing the new manufactured device and from the corner simulation for worst and best cases. This factor compensates for the impact of aging due to HCI and NBTI on process variation for a certain period of time. The delay on the critical paths indicates the aging. Root-mean-square error (RMSE) evaluates the accuracy of the prediction model. MARS has lesser RMSE value compared to other prediction models like SVM and RNN [24].

In the proposed method, recycled IC is detected using RA with NBTI and HCI to predict the age of the device in the initial design stage, and they are classified using the ML techniques. The effect of NBTI and HCI is primarily considered in the proposed work. Three ML methods such as K-means clustering, BPNN and SVM classifier are used to predict the recycled or aged IC from the new ICs with the help of parametric measurements. ML approaches use low area overhead, low cost and low power detection method for recycled ICs. It requires only the input data samples to be trained. The parameters for the aged IC are differentiated from the fresh IC due to NBTI and HCI. The parameters measured at different time instants are used for training the ML methods. The ML methods used here provide a better classification between the fresh and aged ICs.

Section 2 describes the background of reliability effects in transistors like BTI and HCI. In Section 3, ML methods like K-means clustering, BPNN, and SVM are discussed. Section 4 explains the proposed method with the simulation and data sample collection along with the classification using K-means clustering, BPNN and two-class SVM models, and Section 5 provides the performance comparison of the three used models.

## 2 AGING EFFECTS IN TRANSISTORS

With the scaling down of CMOS technology into the nanometre, the reliability effect has become a significant issue. Issues like parameter variability, faults and soft errors make the device to be unreliable in different technology nodes. This parameter variability changes the fabricated devices to have characteristic features different from the intended designed circuit. The process variations make the transistors or gates in the designed circuit to operate with parametric variations after fabrication.

During the chip operation, the characteristics of the transistors or gates will degrade over a period of time. The variations in the parameters of the devices are due to different sources like temperature and voltage variations. The variations in temperature and voltage depend on the input workload activity, frequency and time of the operation. Hence, they cause dissimilarity on the properties of different gates of the IC at different time instants during the operational lifetime.

Wear-out effects cause transistor aging due to runtime variations. These wear-out effects increase the threshold voltage ($V_{th}$) of the transistors and the switching delays of the gates. This leads to a timing failure of the IC. In this study, aging analysis is performed through reliability issues, namely BTI and HCI.

### 2.1 Bias temperature instability

BTI [28, 29] shifts the value of the threshold voltage ($V_{th}$), mobility ($\mu$) of transistors, transconductance ($g_{m}$) and linear and saturation drain currents over a period of time under DC stress. This leads to an increase in circuit delay. BTI is classified as negative and positive BTI. NBTI [30, 31] affects PMOS transistors if the gate terminal of PMOS is connected to the negative terminal of the voltage source, thus creating an
interface trap at the boundary layer of the channel and the gate oxide layer. Similarly, PBTI affects NMOS transistors, if the gate terminal of NMOS is connected to the positive terminal of the voltage source. Depending on the bias conditions of the gate terminal of the PMOS and NMOS transistors, there are two phases for BTI, namely stress and recovery. The stress phase occurs when the gate terminal is under DC voltage stress. The recovery phase occurs when the transistors are in the OFF state to release the voltage stress. BTI models have been developed using trapping–detrapping (TD) and reaction–diffusion (RD) mechanisms [29]. The main part of BTI is hole trapping and interface trap generation.

The TD model in [29] captures the features of the stress and recovery phase physically. Based on the model, the threshold voltage \( V_{\text{th}} \) of the transistor increases logarithmically, and the overall dynamic BTI behaviour is shown in Figure 1. Suppose a single PMOS or NMOS transistor is turned on at time \( t = 0 \), the stress period starts and no DC voltage stress is applied before. The increase in the threshold voltage \( \Delta V_{\text{th}} \) until stress time \( t_{\text{st}} \) is given by

\[
\Delta V_{\text{th}}(t_{\text{st}}) = \Phi_{\text{st}}(A + \log(1 + C_{\text{st}})).
\]  

(1)

The total shift in the threshold voltage \( \Delta V_{\text{th}}(t_{\text{st}} + t_{\text{rec}}) \) for a recovery interval of \( t_{\text{rec}} \) after the stress phase is as

\[
\Delta V_{\text{th}}(t_{\text{st}} + t_{\text{rec}}) = \Phi_{\text{rec}}(A + \log(1 + C_{\text{rec}})) + \Delta V_{\text{th}}(t_{\text{st}}) \\
\times \left( 1 - \frac{A + \log(1 + C_{\text{rec}})}{A + \log(1 + C(t_{\text{st}} + t_{\text{rec}}))} \right),
\]  

(2)

\[
\Phi \sim K \exp \left( \frac{-E}{kT} \right) \exp \left( \frac{BV_{\text{th}}/t}{kT_{\text{ox}}} \right),
\]  

(3)

where \( A, B \) and \( C \) are constants in a particular technology node, \( K \) is the fitting parameter, \( T \) is the temperature, \( k \) is Boltzmann's constant, \( t_{\text{ox}} \) is the oxide thickness, \( E \) is activation energy, and \( V_{\text{th}} \) and \( V_{\text{dh}} \) are the supply voltages under stress and recovery.

2.2 | Hot carrier injection

When the transistor is in saturation mode, some of the carriers in the high electric field gain higher energy to overcome gate oxide and channel barrier and collide with the atoms in the pinch-off region and produce electron–hole pairs due to the impact ionization process. These carriers get injected and trapped into the gate oxide. Due to the interface traps in the gate oxide, the device characteristics such as threshold voltage and drain current are degraded.

HCI [32] effect occurs in NMOS transistors. The effect of HCI in PMOS transistors is negligible. Hot carriers are formed during signal transitions. Hence the threshold voltage degradation due to HCI depends on the switching activity of the input signal [32]:

\[
\Delta V_{\text{th}} = A_{\text{HCI}} \times SW \times f \times e^{\frac{E_{ox}}{kT}} \times t^{0.5},
\]  

(4)

\[
E_{\text{ox}} = \frac{V_{\text{GS}} - V_{\text{th}}}{t_{\text{ox}}},
\]  

(5)

where \( A_{\text{HCI}} \) is a technology-dependent constant, \( f \) is the clock frequency, \( SW \) is the switching activity factor, \( t_{\text{ox}} \) is the oxide thickness, \( V_{\text{th}} \) is the threshold voltage and \( V_{\text{GS}} \) is the gate-source voltage of the transistor, \( E_{l} \) is a constant equal to 0.8 V/\( \text{nm} \) [32] and \( t \) is the total time.

The effect of HCI depends on the temperature of the device. Hence the change in the threshold voltage is modified as follows:

\[
\Delta V_{\text{th}} = A_{\text{HCI}} \times SW \times f \times e^{\frac{E_{ox}}{kT}} \times e^{\frac{E_{ox}}{kT}} \times t^{0.5}.
\]  

(6)

3 | MACHINE LEARNING METHODS

ML systematically applies different algorithms to synthesize the essential relationships among the data and the information. ML systems can be trained to classify the changing process conditions, in order to model variations in the operating behaviour. The different forms of ML methods are supervised, unsupervised, semi-supervised and reinforcement learning. The ML process involves various steps like input data collection, data preparation, training the data, evaluation of the results and tuning. In the proposed work, three different ML methods are used, namely K-means clustering, BPNN and SVM.
Where \( \| \cdot \| \) defines the Euclidean distance

\[
\text{kmeans}(X \in \mathbb{R}^{n \times d}, kC)
\]

1: while any \( c_j \) change location do
2: for \( i \in \{1, \ldots, n\} \) do
3: \( \text{class}(x_i) \leftarrow \arg \min_j \| x_i - c_j \| \)
4: end for
5: for \( j \in \{1, \ldots, k\} \) do
6: \( c_j \leftarrow \frac{1}{|I_j|} \sum_{i \in I_j} (\text{class}(x_i) = j) \)
7: end for
8: end while
9: return \( C \)

**FIGURE 2** Steps of K-means clustering [36, 37]

### 3.1 K-means clustering

K-means clustering [33, 34] is one of the most widely used clustering algorithms. It starts with the random initialization of the centroids. Each data point is allocated to the closest centroid based on a measurement of squared Euclidean distance. After the formation of K clusters, the centroids of each K cluster are updated. The centroid updates or changes the location until all the data points are assigned to the closest centroid in the K clusters.

For a given dataset \( D = \{X_1, X_2, \ldots, X_n\} \) consisting of \( n \) data points, after applying K-means clustering algorithm \( k \) clusters are obtained by \( C = \{C_1, C_2, \ldots, C_k\} \). The objective function or sum of squared error (SSE) for the K-means clustering is shown in the Equation (7), where \( C_j \) is the centroid of cluster \( C_k \). The main goal of the K-means clustering is to form the K-clusters with the minimized SSE value [35]. The steps for the K-means clustering algorithm [36, 37] are shown in Figure 2. The K-means algorithm is validated by measuring the silhouette value (SV). SVs define the connection between the data in the same and different clusters. The SV should be large enough to make the good clustering of data:

\[
\text{SSE} = \sum_{i=1}^{K} \sum_{j=1}^{n_j} \| X_j - C_j \|^2, \quad (7)
\]

### 3.2 Back propagation neural network

BPNN [38–40] is a type of supervised NN where the output values are known earlier. The BPNN consists of input, hidden and output layers as shown in Figure 3. BPNN has two stages, namely forward pass (propagation) and a backward pass (propagation). In the forward propagation, the outputs are measured and compared with the preferred output values. Then the errors are calculated from the preferred and actual output values. In the backward propagation, the error is used to change the weights in the NN in order to minimize the size of the error.

Let \( X_j \) be the input given to the input layer, \( y_j \) be the output of the hidden layer and \( Z_i \) be the output from the output layer. Suppose \( w_{ji} \) is the weight of the NN between the input and hidden layers and \( \psi_{lj} \) is the weight of the NN between the hidden and output layers. The expected output value is \( t_l \) and \( f(\cdot) \) is the activation function. The computational formula of the model [40] is expressed as follows:

In the forward propagation, the output from the hidden layer is

\[
y_j = f \left( \sum_i w_{ji} x_i - \theta_j \right) = f(\text{net}_j), \quad (8)
\]

\[
\text{net}_j = \sum_i w_{ji} x_i - \theta_j, \quad (9)
\]

The output of the output layer is

\[
Z_l = f \left( \sum_j \psi_{lj} y_j - \theta_l \right) = f(\text{net}_l), \quad (10)
\]

\[
\text{net}_l = \sum_j \psi_{lj} y_j - \theta_l. \quad (11)
\]

In the back propagation, to reduce the error, the gradient descent method is adopted to control the weight value of all layers. The change in the weights \( w_{ji} \) and \( \psi_{lj} \) is given by

\[
\Delta w_{ji} = -\eta \frac{\partial E}{\partial w_{ji}} = \eta \delta_i y_j, \quad (12)
\]

\[
\Delta \psi_{lj} = -\eta \frac{\partial E}{\partial \psi_{lj}} = \eta \delta_j x_i, \quad (13)
\]

\[
\delta_j = f'(\text{net}_j) \cdot \sum_i \delta_i \psi_{lj}. \quad (14)
\]

The learning rate of the BPNN is given by \( \eta' \), where \( \sum \delta_i \psi_{lj} \) is the error in the hidden layer, \( \delta'_j \) defines the error \( \delta_j \) of output, \( Z_i \) is propagated back through \( \psi_{lj} - y_j \) to turn into an error of the hidden layer.
3.3 | Support vector machine classifier

SVM [41–43] is a type of classifier method that classifies the data points by creating hyperplanes in a multidimensional space that divides different class labels as shown in Figure 4. A hyperplane function \( g(y) = W^T y + b \) separates the two classes with margins.

For a given set of data points \( X_i \) which belongs to two separate classes \( \omega_1 \) and \( \omega_2 \), the distance of any data point from the hyperplane is equivalent to \( \frac{g(X)}{||W||} \). SVM [43] tries to determine the value of \( W \) and \( b \), such that \( g(X) \) is equivalent to 1 for the closest data belonging to \( \omega_1 \) and -1 for the closest data belonging to \( \omega_2 \):

\[
W^T X + b = \begin{cases} 
1; & X \in \omega_1 \\
-1; & X \in \omega_2 
\end{cases}.
\]  

The SVM training involves the minimization of the objective function, which is expressed in the following equation:

\[
J(w) = \frac{1}{2} ||w||^2.
\]  

Subject to the constraint,

\[
y_i(w^T x + b) \geq 1, \quad i = 1, 2, ..., N
\]  

The objective function is augmented by adding it to the weighted sum of constraints and multiplied by the Lagrange multipliers:

\[
L(w, b, \lambda) = \frac{1}{2} w^T w - \sum_{i=1}^{N} \lambda_i [y_i(w^T x_i + b) - 1],
\]  

where \( W \) and \( b \) are the primal variables and \( \lambda_i \) is the Lagrange multipliers.

4 | PROPOSED METHOD TO DETECT AGED INTEGRATED CIRCUITS

The manufactured ICs undergo process parametric variations. The parameter variations in recycled ICs need to be differentiated from process parametric variation of newly manufactured IC. In the proposed method, three classification algorithms are used to classify the aged IC from the new IC. The classification algorithms used are K-means clustering, BPNN and two-class SVM (TCSVN). The following sections deal with the overview and summary of the method of classifying the recycled and new devices using the three ML algorithms. The overall flow of the proposed work is shown in Figure 5.

Figure 5 shows that two-level detection is used in the proposed method. First, the Monte Carlo (MC) simulation \( m(t) \) is performed using a fresh statistical device model, and reliability simulation is performed using degraded device models in the circuits. Then the parameter data samples are gathered from MC and reliability simulations.

The MC simulation \( m(t) \) using a statistical device model results in a response \( O_i(t) \) that implies the effect of process variations on fresh IC. A reliability simulation \( m(t) \) is performed to obtain the data response \( O_i(t) \) under reliability effects such as HCI and NBTI in the Device Under Test (DUT) with a degraded device model containing reliability parameters. Let \( C_{kl} \) be the DUT with \( k \) inputs and \( l \) outputs. Then,
\[ O_i(t) = m\{C_{kl}(t)\}, i = 1, 2, \ldots, N \text{ and } t = 1, 2, \ldots, N. \]  
(19)

\[ O'_i(t) = m'\{C_{kl}(t)\}. \]  
(20)

The data samples collected are the parameter outputs from both simulations that change with respect to the aging of time. Then the collected parameters are trained and tested using the three ML algorithms like K-means clustering, BPNN and SVM.

### 4.1 Simulation and data sample collection

MC simulation is carried out by varying the process parameters defined in the Process Development Kit. Reliability simulation [44] with effect of BTI and HCI is carried out for an aging time period varying from 1 day to 10 years. The fresh device is affected by the process variations.

The parameter data samples are determined by the measurements from the MC and aging simulations. The parameters measured are minimum voltage \(v_p\) and \(v'_p\), delay value \(d_p\) and \(d'_p\) and supply current \(I_p\) and \(I'_p\) for ISCAS and ITC benchmark circuits. The ISCAS 85 benchmark circuit used is c17 (Six NAND Gate Circuit), ISCAS 89 benchmark circuit used is s27 and the ITC 99 benchmark circuit used is b02 (Finite state machine that recognizes Binary-coded decimal numbers). The number of primary inputs (PI), primary outputs (PO) and gates and transistors in the benchmark circuits are shown in Table 1. For benchmark from Texas Instruments, the parameters obtained are gain, phase margin, bandwidth and supply current (Iddq). All the necessary parameters are measured from the corresponding simulations. These parameters vary with respect to aging time and reliability effects. The variation in the parameters determined from the fresh and reliability simulation can be greatly classified by using the above-mentioned classifiers.

For MC simulation of circuit \(m\{C_{kl}(t)\}\), the parameters measured are \(v_p\), \(I_p\) and \(d_p\).

For reliability simulation of circuit \(m'\{C_{kl}(t)\}\), the parameters measured are \(v'_p\), \(I'_p\) and \(d'_p\). The various parameters measured are represented as follows:

\[ v_p = \min\{O'_i(t)\}; p = 1, 2, 3, \ldots, 50; \]  
(20)

\[ i = 1, 2, 3, \ldots, N \text{ and } N = 300, \]  

\[ v'_p = \min\{O'_i(t)\}; t = 1, 2, \ldots, N, \]  
(22)

\[ d_p = OP_i(t) - O_i(t). \]  
(23)

\[ d'_p = O'_i(t) - O_i(t). \]  
(24)

The minimum voltage and supply current are directly measured. The delay values from MC simulation are obtained.
by determining the difference between the response of optimal simulation without and with process variation. The delay values from reliability simulation are calculated by determining the difference between the responses from MC and reliability simulations.

Totally, 50 MC simulations are performed. In case of reliability simulation, the data is collected for different time periods. It means that the reliability DC stress simulations are performed for different 50 uniformly sampled aging time periods $t_\nu$, where $\nu = 0.003$ years (1 day), 0.04 years (15 days), 0.25, 0.5, 1, …, 10 years. Then for each simulation of $t_\nu$ aging time period like 0.5 years, the parameters are obtained.

As discussed in previous sections, for 50 MC simulations, single minimum voltage, single supply current and single delay value are calculated for each simulation. So, for the MC simulation, a total of 150 data sample values are measured. For the reliability simulation, a single minimum voltage, single supply current and single delay value are calculated for each aging time period simulation. So, from the reliability simulation, a total of 150 data sample values are measured.

The method of classification using output parameters from the simulation trains only a few numbers of samples. It greatly reduces the storage space and time.

An industrial benchmark fully differential folded-cascode amplifier (FDPCA) from Texas Instruments, shown in Figure 6, has been analysed by the proposed method of recycled IC detection designed in 90-nm technology. The parameters considered for the analysis of this circuit are gain, phase margin, bandwidth and supply current (Iddq). The MC and aging simulations are also performed for the amplifier circuit to determine the parameter outputs. The obtained parameters are given as input to the ML algorithms for classification. The degradation of gain, phase margin and Iddq of FDPCA for an aging of 10 years is shown in Figures 7–9.

### 4.2 Classification using K-means clustering

First, the parametric samples are collected and given as input to the K-means clustering algorithm. The algorithm clusters the data samples into K clusters and then computes the SV. SV evaluates the performance of the K-means clustering. Two clusters are formed: one is for fresh and the other is for the degraded device depending on the parameter values. The difference between the parameters obtained from the MC
simulation and the reliability simulation is due to reliability effects, HCI and NBTI, in the devices during the reliability simulation.

K-means clustering [33, 34, 36] is the fastest and efficient way of clustering data. It is used in the detection of counterfeit IC. K-means clustering algorithm groups parameters obtained from the MC simulation and reliability simulation into clusters and determines the centroid values of every cluster.

The steps of the K-means clustering algorithm are shown in Figure 10. The input parameter data samples are applied to the clustering in the proposed method. The number of clusters K to be formed is supplied to the algorithm. K centroids are randomly initialized for the data.

The squared Euclidean distance is computed between the initialized K centroids and each pair of the parameter data sample. The data is assigned to a particular cluster if the Euclidean distance between the data and the centroid is lesser. Then the mean of the data and centroid in the clusters is measured. Move the centroid and then compute the distance between the new data points and centroid values, and place the data in the minimum distance cluster. Centroid stops moving until all the data points are placed in a particular cluster. The distance between a data point \((A_1, B_1)\) to the centroid \((C_1, C_2)\) in a cluster should be minimal compared to the distance between that data point \((A_1, B_1)\) to the centroid \((C_3, C_4)\) of the other cluster. Thus two clusters are formed and the data is grouped into two clusters as cluster 1 and cluster 2. Test data is provided to the cluster model to evaluate the clustering.

The clustering of parameters of the new and degraded devices using the K-means algorithm results in high accuracy. After the formation of clusters, the SVs are estimated. The SV is measured by the distance between two data in the same cluster and between two data in the different clusters. The distance between the two data in the same cluster is \(D_i\) and the distance between the two data in different clusters is \(D_j\). The SV for K clusters \((S_K)\) is calculated as follows:
The accuracy of K-means clustering is determined by the SV. A maximum value of SV supports the used K-means clustering algorithm. The K-means clustering method is tested for the output parameters of benchmark circuit s27, c17, b01 and FDFCA. Using the output parameters obtained, K-means clusters were generated and the SVs are estimated for K clusters. In this study, two clusters are formed, namely aged/degraded and fresh. Figure 11 shows the two clusters of output parameters of benchmark b02 using K-means clustering after aging for a period of 10 years. Cluster 1 and cluster 2 define the fresh and degraded device responses.

Table 1 shows that the K-means cluster performance of different benchmark circuits for aging time instants \( t_v \) with the input of output parameters. Figure 12 shows the maximum average SVs for the K clusters of b02 after 10 years of aging. The aging periods considered are \( t_1 = 1 \) day, \( t_2 = 15 \) days, \( t_3 = 0.25 \) years, \( t_4 = 0.5 \) years, \( t_5 = 1 \) year and \( t_6 = 5 \) years, \( t_7 = 10 \) years. The detection rate (DR) of K-means clustering is measured by the value of the silhouette \( S_K \). Let \( Z_i \) be the number of input data in K clusters closer to \(-1\) and \( K_i \) denotes the total number of data in K clusters. It indicates that \( Z_i \) is misclassified in K clusters. Then the DR is given as follows:

\[
S_K = \frac{D_i - D_j}{\text{Max}\{D_i, D_j\}}; \quad -1 < S_K < 1. \tag{25}
\]
The maximum average SV of K-clusters is calculated and shown in Table 1. Since the SV is greater than the mean of 0.7 at all aging time instants, the clustered data has high accuracy. The first column in Table 1 defines the type of benchmark circuits and the second column defines the number of gates (G) and transistors (T) in the benchmark circuits. The remaining columns describe the DR of aged IC and the maximum average SV for K clusters at various aging time instants. The benchmark circuit c17, s27, b02 and FDFCA have a high possibility of classification between the new and degraded IC from 1 day to 10 years using the K-means clustering. The SV is also high in all aging time instants. The formation of two clusters produces high SVs. The number of data samples with negative values of silhouette reduces the clustering performance. But in all aging time instants, the silhouette analysis has no negative values.

Once the aging time increases, the SV decreases slightly indicating that the parametric features may be difficult to be clustered by K-means clustering. Thus it is difficult to detect the degradation in parameter of the circuit due to process variation or reliability effects after a particular limit of aging.

4.3 | Back propagation neural network (BPNN) classification

The performance of the BPNN network is analysed with the value of mean square error (MSE) and receiver-operating curve (ROC). The number of neurons and the number of hidden layers for the network are fixed at the initial stage. A large number of neurons for hidden layers are required to train a large amount of data. Hence it increases the cost of the model. The amount of input data used in the training and testing process for the used BPNN model is low; the dataset contains three features for c17, s27 and b02 benchmark circuits and four features for FDFCA benchmark circuit. The total size of data is 300 for all benchmark circuits, hence the number of hidden neurons is chosen as 2 and the number of hidden layers is chosen as 2 for the BPNN model. The BPNN model classifies the parameter data. The training and testing are carried out for three different mixtures of train and test samples. Here the amount of data considered for testing is 20% and data considered for training is 80%. The classification with a higher DR is extracted and tabulated.

The BPNN classification model shown in Figure 13 is applied to benchmark circuits like c17, s27, b02 and FDFCA. The MSE and accuracy of the prediction model are computed for the training (TR) and testing (TE) of data at aging time instants $t_1, t_2, t_3, t_4$ and $t_5$ (Tables 2 and 3). The MSE of the BPNN model during testing is higher than training. The area under curve (AUC) in the ROC from the proposed method for the BPNN classification method is higher. Table 4 shows the AUC in ROC for different benchmark circuits used in the BPNN classification of new and degraded devices. The MSE value must be close to zero and the AUC must be close to one; then the performance of the BPNN classification model is high. Values of MSE and AUC in ROC from the model validates that the BPNN is suitable for the classification of recycled IC from the new IC using output response as the input data samples.

In order to evaluate the best performance of the BPNN model, MSE is plotted in terms of epochs as shown in Figure 14. The data used for training, validation and testing are 70%, 10% and 20%, respectively. The best validation performance value is $4.4616e-09$ at epoch 30 of the BPNN model classification for data of circuit b02 after the aging of 10 years. The best validation performance MSE decreases as the number of epochs increases, but it may increase in the validation data samples as the NN starts to overfit the training data.

4.4 | Two-class support vector machine (TCSVM) classification

The parameter data is applied to the TCSVM model to classify the fresh and degraded data. The data $(v_p, v_p', I_p, I_p', d_p, d_p', I_p, I_p')$ are provided to the model for classification.

The SVM model $S$ classifies the data as follows:

$$S = \begin{cases} 0; \text{Fresh} \\ 1; \text{Aged/Recycled} \end{cases}$$
### Table 2
MSE for BPNN classification of benchmarks at different aging time instants

| BM  | $t_1 = 1$ day | $t_2 = 15$ days | $t_4 = 0.5$ years | $t_5 = 1$ year | $t_6 = 5$ years | $t_7 = 10$ years |
|-----|---------------|------------------|-------------------|--------------|--------------|----------------|
|     | TR | TE | TR | TE | TR | TE | TR | TE | TR | TE | TR | TE |
| c17 | 1.074e-04 | 4.988e-04 | 3.243e-09 | 2.643e-04 | 3.008e-05 | 4.801e-05 | 1.737e-04 | 0.007 | 5.939e-04 | 7.779e-05 | 0.003 | 2.2982e-04 |
| s27 | 1.448e-09 | 1.628e-08 | 4.697e-05 | 6.403e-05 | 0.007 | 0.005 | 1.308e-06 | 9.234e-06 | 1.555e-06 | 3.449e-07 | 0.006 | 1.926e-05 |
| b02 | 7.357e-05 | 0.0271 | 3.243e-04 | 2.643e-04 | 0.010 | 0.038 | 0.001 | 0.012 | 1.463e-04 | 9.185e-05 | 2.498e-05 | 2.219e-05 |
| FDFCA | 7.357e-05 | 0.02 | 7.36e-05 | 0.028 | 7.365e-05 | 0.026 | 7.370e-05 | 0.03 | 7.300e-05 | 0.031 | 7.400e-05 | 0.032 |

Abbreviations: BPNN, back propagation neural network; FDFCA, fully differential folded cascode amplifier; MSE, mean square error; TR, training; TE, testing.

### Table 3
Accuracy (%) for BPNN classification of benchmarks at different aging time instants

| BM  | $t_1 = 1$ day | $t_2 = 15$ days | $t_4 = 0.5$ years | $t_5 = 1$ year | $t_6 = 5$ years | $t_7 = 10$ years |
|-----|---------------|------------------|-------------------|--------------|--------------|----------------|
|     | Accuracy (%) | Accuracy (%) | Accuracy (%) | Accuracy (%) | Accuracy (%) | Accuracy (%) |
|     | TR | TE | TR | TE | TR | TE | TR | TE | TR | TE | TR | TE |
| c17 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| s27 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| b02 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| FDFCA | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

Abbreviations: BPNN, back propagation neural network; FDFCA, fully differential folded cascode amplifier; TR, training; TE, testing.
Two-class SVM is used as the classification algorithm. It has a high level of accuracy compared to one-class SVM [45]. The disadvantages of TCSVM are its cost and time [45]. In this classification, the size of the input sample is not high and it consumes a reasonable training time (TT) and model size. The main goal of TCSVM is to classify the input data with higher accuracy within a short period of TT. The kernel used in TCSVM is linear. The TCSVM classified data for the s27 benchmark circuit after 10 years of an aging time period is shown in Figure 15. X-axis and Y-axis represent the parameters of fresh and degraded devices. The TCSVM model is evaluated by ROC and TT. The AUC in ROC must be of higher value to achieve good classification efficiency. Benchmark circuits s27, c17, b02 and FDFCA are tested with the used TCSVM classification. The MC and
TABLE 5 Performance of evaluation of TCSVM for benchmarks at different aging time

| BM  | $t_1 = 1$ day | $t_2 = 15$ days | $t_3 = 0.25$ years | $t_4 = 0.5$ years | $t_5 = 1$ year | $t_6 = 5$ years | $t_7 = 10$ years |
|-----|--------------|----------------|-------------------|------------------|---------------|----------------|----------------|
| c17 | 1            | 0.492          | 1                 | 0.727            | 1             | 0.647          | 1              | 0.761          | 1              | 0.6714         | 1              | 3.9275         |
| s27 | 1            | 1.036          | 1                 | 0.416            | 1             | 1.192          | 1              | 0.530          | 1              | 0.4653         | 1              | 0.4359         |
| b02 | 1            | 0.550          | 1                 | 4.043            | 1             | 3.931          | 1              | 0.525          | 1              | 0.4615         | 1              | 2.9875         |
| FDFCA | 1         | 0.600          | 1                 | 0.619            | 1             | 0.624          | 1              | 0.550          | 1              | 0.547          | 1              | 0.6207         | 1              | 0.539          |

Note: Obtained Area Under Curve (AUC) maximum value of 1 confirms the good Two-Class Support Vector Machine (TCSVM) based classification of fresh and aged Benchmark (BM) circuits like c17, s27, b02 and Fully Differential Folded-Cascode Amplifier parameters collected at various aging time. Training Time (TT) describes that the used TCSVM model can be trained within a few seconds to classify and detect the aged circuit from fresh circuit.

Abbreviations: BM, Benchmark Circuit; AUC, area-under curve; FDFCA, fully differential folded cascode amplifier; TCSVM, two-class support vector machine; TT, training time.

FIGURE 16 (a) Confusion matrix and (b) receiver-operating characteristics from TCSVM for s27 after 10 years of aging

reliability simulation are performed and output parameter values are collected. The output parameter values are applied to the TCSVM model. The results shown in Table 5 obtained from TCSVM are AUC in ROC and the model TT for the benchmark circuits at the aging time of $t_1$, $t_2$, $t_3$, $t_4$, $t_5$, $t_6$ and $t_7$. The confusion matrix and ROC curve of the benchmark circuit s27 for an aging time period of 10 years are shown in Figure 16. The TT of used TCSVM is higher in terms of several seconds.

5 | COMPARISON OF USED MACHINE LEARNING METHODS IN THE DETECTION OF COUNTERFEIT INTEGRATED CIRCUITS

The performance of used ML methods for classification is compared in terms of its DR of the aged or degraded IC from the new IC. The DR of K-means clustering defines the performance of identifying the aged device. The AUC in ROC obtained in
BPNN and TCSVM describes its performance. The best detection or classification rate for all benchmark circuits on applying the ML methods is tabulated in Tables 6–9. The three algorithms used in this work produce a 100% classification rate to classify the new and recycled devices using the output parameters obtained from the device aged from 1 day to 10 years.

The proposed methods are compared with the ML methods in [16] for an aging period of 1 month as shown in Table 10. The ML methods in our work achieve a 100% detection or classification rate similar to DCSA [16]. When compared to one-class SVM in [16], the K-means, BPNN and TCSVM produce higher classification rates. DCSA could produce a 100% classification rate only at an aging period of 1 month, but the ML algorithms used in our work produces the same classification rate even for 1 day of usage.

6 | CONCLUSION

In this study, reliability effects like HCI and NBTI are used for aging analysis. Benchmark circuits c17, s27 and b02 and FDFCA are tested at a different time period of aging starting from 1 day to 10 years. The output parameters like delay, supply current, minimum output voltage and gain and phase margin are used to train and test the used ML methods. BPNN, K-means clustering and TCSVMs are used to detect recycled or aged ICs. The results obtained show that the BPNN, K-means clustering and TCSVM perform well to give a better classification rate. These three algorithms result in a high DR of recycled ICs even at the age of 1 day with minimum data size.
ACKNOWLEDGEMENT

The author(s) received no financial support for the research, authorship, and/or publication of this article in the journal.

ORCID

Udaya Shankar Santhana Krishnan 0000-0001-6417-6345

REFERENCES

1. Agarwal, M., et al.: Circuit failure prediction and its application to transistor aging. Paper presented at Proceedings of 25th IEEE VLSI Test Symposium, Berkeley, CA, May (2007).

2. He, K., Huang, X., Tan, S.X.D.: EM-based on-chip aging sensor for detection and prevention of counterfeit and recycled ICs. Paper presented at Proceedings of IEEE/ACM International Conference on Computer-Aided Design, Austin, pp. 146-151 (2015).

3. Jang, B., et al.: On-chip AGING prediction circuit in nanometer digital circuits. Paper presented at Proceedings of International SoC Design Conference, Jeju, South Korea, pp. 68-69 (2015).

4. Lu, Y., et al.: Statistical reliability analysis under process variation and aging effects. Paper presented at Proceedings of ACM/IEEE Design Automation Conference, San Francisco, CA, pp. 514-519, July (2009).

5. Zhang, X., Xiao, K., Tehranipoor, M.: Path-delay fingerprinting for identification of recovered ICs. Paper presented at Proceedings of IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems, Austin, TX, pp. 13-18, October (2012).

6. Yang, Z., et al.: Workload-aware failure prediction method for VLSI devices using an LUT based approach. Paper presented at Proceedings of IEEE International Instrumentation and Measurement Technology Conference, Houston, TX, May (2018).

7. Savanur, P.R., Alladi, P., Tragoudas, S.: A BIST approach for counterfeit circuit detection based on NBTI degradation. Paper presented at Proceedings of IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems, Amherst, pp. 123-126, October (2015).

8. Shinde, S., et al.: Wideband microwave reflectometry for rapid detection of dissimilar and aged ICs. IEEE Trans. Instrum. Meas. 66, (8), pp. 1-10, (2017).

9. Sengupta, D., Sachin, S.: Predicting circuit aging using ring oscillators. Paper presented at Proceedings of 19th Asia and South Pacific Design Automation Conference (ASP-DAC), Singapore, pp. 430-435 (2014).

10. Sengupta, D., Sachin, S.: ReSCALE: recalibrating sensor circuits for aging and lifetime estimation under BTI. Paper presented at Proceedings of IEEE/ACM Int. Conf. on Computer-Aided Design (ICCAD), San Jose, CA, pp. 402-407 (2014).

11. Sengupta, D., Sachin, S.: Estimating circuit aging due to BTI and HCl using ring-oscillator-based sensors. IEEE Trans. Comput. Aided Des. Integr. Circ. Syst., 36, (10), pp. 1688-1701 (2017).

12. Huang, K., Carulli Jr, J.M., Malikis, Y.: Parametric counterfeit IC detection via support vector machines. Paper presented at Proceedings of IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems, Austin, pp. 7-12, October (2012).

13. Huang, K., Carulli, Jr Jr, J.M., Malikis, Y.: Counterfeit electronics: a rising threat in the semiconductor manufacturing industry. Paper presented at IEEE International Test Conference, Anaheim, pp. 1-4 (2013).

14. Dogan, H., Forte, D., Tehranipoor, M.M.: Aging analysis for recycled FPGA detection. Paper presented at IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems, Amsterdam, pp. 171-176, October (2014).

15. Konur, A., et al.: Fine-grained aging prediction based on the monitoring of run-time stress using DFT infrastructure. Paper presented at Proceedings of IEEE/ACM Int. Conf. on Computer-Aided Design (ICCAD), Austin, pp. 51-58 (2015).

16. Huang, K., et al.: Recycled IC detection based on statistical methods, IEEE Trans. Comput. Aided Des. Integr Circ. Syst., 34, (6), pp. 973-980 (2015).

17. Alam, M.M., Tehranipoor, M., Forte, D.: ‘Recycled FPGA detection using exhaustive LUT path delay characterization’. Paper presented at Proceedings of IEEE International Test Conf (TC), Fort Worth, pp. 1-10. (2016).

18. Vijayan, A., et al.: ‘Fine-grained aging-induced delay prediction based on the monitoring of run-time stress, IEEE Trans. Comput. Aided Des. Integr. Circ. Syst., 37, (5), pp. 1064-1075 (2018).

19. Vijayan, A., et al.: ‘Machine-learning-based aging analysis’ in ELFADEL I, In: Boning D., Li, X. (eds.) Machine Learning in VLSI Computer-Aided Design, pp. 265-289. Springer (2019).

20. Frazier, P.D., et al.: A novel counterfeit detection approach for integrated circuit supply chain assurance. J. Hardw. Syst. Secur., 2, pp. 240-250 (2018).

21. Wang, Y.T., et al.: Aging-aware chip Health prediction adopting an innovative monitoring strategy. Paper presented at Proceedings of 24th Asia and South Pacific Design Automation Conference, Tokyo, pp. 179-184, January (2019).

22. Baranowski, R., et al.: On-line prediction of NBTI-induced aging rates. Paper presented at Proceedings of the Design, Automation & Test in Europe Conference & Exhibition (DATE), Grenoble, pp. 589-592, March (2015).

23. Karimi, N., Huang, K.: Prognosis of NBTI aging using a machine learning scheme. Paper presented at Proceedings of the IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems, Storrs, September (2016).

24. Huang, K., Zhang, X., Karimi, N.: Real-time prediction for IC aging based on machine learning. IEEE Trans. Instrum. Meas. 68(12): 4756-4764 (2019).

25. Torres, E.S., Leite Jr, P.F., Drechsler, R.: Unintrusive aging analysis based on offline learning. Paper presented at Proceedings of IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems, Cambridge, October (2017).

26. Chang, N., et al.: Applying machine learning to design for reliability coverage. Paper presented at Proceedings of IEEE International Reliability Physics Symposium, Monterey, CA, April (2019).

27. Yang, C., Feng, F.: ‘Multi-step-ahead prediction for a CMOS low Noise amplifier aging due to NBTI and HCI using neural networks, J. Electron Test. 35, pp. 797-808, (2019).

28. Mahapatra, S. (ed.): Fundamentals of Bias Temperature Instability in MOS Transistors Characterization Methods, Process and Materials Impact, DC and AC Modeling, Springer Series in Advanced Microelectronics, 1st ed, Springer India (2016).

29. Guo, X., Stan, M.R.: Circadian Rhythms for future Resilient Electronic Systems Accelerated Active Self-Healing for Integrated Circuits, 1st ed. Springer International Publishing (2020).

30. Schroeder, D.K.: Negative bias temperature instability: What do we understand? Microelectron. Reliab. 47, pp. 841-852 (2007).

31. Stathis, J.H., Zafar, S.: The negative bias temperature instability in MOS devices: a review, Microelectron. Reliab. 47, pp. 841-852 (2006).

32. Tan, S., et al.: Long-Term Reliability of Nanometer VLSI Systems Modeling, Analysis and Optimization, 1st ed, Springer International Publishing (2019).

33. Li, Y., Wu, H.: A clustering method based on K-means algorithm, Phys Proc. 25, pp. 1104-1109 (2012).

34. Kanungo, T., et al.: An efficient k-means clustering algorithm: analysis and implementation, IEEE Trans. Pattern Anal. Mach. Intell. 24(7), pp. 881-892 (2002).

35. Peter, J.R.: Silhouettes: a graphical aid to the interpretation and validation of cluster analysis, J. Comput. Appl. Math. 20, pp. 53-65 (1987).

36. Gregory, J.H.: Learning Structure and Concepts in Data Through Data Clustering (PhD Thesis), University of California (2003).

37. Hudail, A., Khanafesh, M., Surikadi, O.: An improved version of K-medoids algorithm using CBO, Mod. Appl. Sci., 12(2), pp. 116-127 (2018).
38. Neural networks and deep learning, http://neuralnetworksanddeeplearning.com/chap2.html

39. Hecht-Nielsen, R.: Theory of the back propagation neural network. Paper presented at Proceedings of International Joint Conference on Neural Networks, Washington, pp. 593-605, (1989)

40. Li, J., et al.: Brief introduction of back propagation (BP) neural network algorithm and its improvement. In: Jin, D., Lin, S. (eds.) Advances in Computer Science and Information Engineering. Advances in Intelligent and Soft Computing, 1st ed., Springer, Berlin, pp. 553-558 (2012)

41. Elshinawy, M.Y., et al.: Comparing one-class and two-class SVM classifiers for normal mammogram detection. Paper presented at Proceedings of IEEE 39th Applied Imagery Pattern Recognition Workshop, Washington, October (2010)

42. Hsu, C.-W., Lin, C.-J.: A comparison of methods for multiclass support vector machines. IEEE Trans. Neural Netw. 13(2), pp. 415-425 (2002)

43. Awad, M., Khanna, R.: Efficient Learning Machines Theories, Concepts, and Applications for Engineers and System Designers, Apress, Berkeley (2015)

44. Maricau, E., Gielen, G.: Analog IC Reliability in Nanometer CMOS, Analog Circuits and Signal Processing, 1st ed., Springer-Verlag New York (2013)

45. Senf, A., Chen, X., Zhang, A.: Comparison of one-class SVM and two-class SVM for fold recognition. In: King, I., et al. (eds.) Neural Information Processing ICONIP 2006 Lecture Notes in Computer Science, Vol. 4233, Springer, Berlin, pp. 140-149 (2006)

---

**How to cite this article:** Santhana Krishnan US, Palanisamy K. Recycled integrated circuit detection using reliability analysis and machine learning algorithms. *IET Comput. Digit. Tech.* 2021;15:20–35. [https://doi.org/10.1049/cdt.12005](https://doi.org/10.1049/cdt.12005)