Identification of Characteristics After Soft Breakdown with GA-Based Neural Networks

Hsing-Wen Wang

National Changhua University of Education, College of Management,
Department of Business Administration, Changhua 500, Taiwan, R.O.C.
shinwen@cc.ncue.edu.tw

Abstract. In this research, we analyze the low-frequency noise power spectrum of drain current (S_{id}) in electrically stressed S_iO_2 film, and then propose the evolutionary neural networks-based model named ENN-SBD to identify the highly nonlinear degraded characteristics of low frequency noise around the soft breakdown (SBD). The S_{id} data follow the $1/f^{\gamma}$ relationship with different value of power exponent γ. The spatial oxide traps distribution is proposed to account for the different γ value. It is found that the S_{id} correlates closely with the gate fluctuations via the trapping and detrapping processes and hence it is feasible to build the model represents the behavior of soft breakdown. The results also indicate that ENN-SBD has more precisely identification capability than typical Lorentzian spectrum method. Besides, it is superior to the backpropagation neural networks-based model (BNN-SBD) while the system identification is proceeding. This paper is helpful for breakdown detection and saving the cost of testing from quality assurance in the process of advanced CMOS technology.

Keywords: Low-frequency noise, evolutionary neural networks, CMOS, soft breakdown, degraded characteristics.

1 Introduction

System modeling or system identification based on the conventional mathematical skills (e.g. differential equations) isn't suited for dealing with uncertain, non-structure or ill-define systems [1], [2]. The degraded characteristics of low frequency noise around the soft breakdown are typical problems known as the highly nonlinear and data-oriented while modeling its degradation characteristics. In fact, it costs much in the testing processes of quality assurance, especially for advanced CMOS technology. The reliability of thin S_iO_2 film used as gate dielectrics is one of the most concerned issues for advanced CMOS technology. When the gate oxide thickness is less than 5nm, there is new anomalous degradation and breakdown characteristics, called quasi-breakdown [3] or soft breakdown [4]. After the occurrence of SBD, the fluctuation phenomena are observed in the time evolution of the gate voltage or the gate current. It has been reported that the gate fluctuations are due to the trapping and detrapping of electrons in gate oxides [5]. Moreover, in thinner oxides the so-called carrier-number fluctuation model is that the trapping and detrapping processes within the gate oxide can modulate the underlying channel potential [6].

M. Ali and R. Dapoigny (Eds.): IEA/AIE 2006, LNAI 4031, pp. 564 – 572, 2006.

However, neither the experimental evidence in support of the carrier-number fluctuation model found yet nor the SBD model built. In this paper, during high-field stress we explore the noise power spectrum of drain current (S_{id}) characterization. In other words, the S_{id} data in fresh, stress-induced leakage current (SILC) and SBD mode are presented. We also discuss that for thinner gate oxide the S_{id} correlates with the gate current fluctuations and present the evolutionary neural networks-based SBD model.

Neural networks have been played an important role in attacking bottlenecks of advanced semiconductor technologies because of its distinguishing characteristic [7], [8], [9], [10], [11]. Ling et al (2003) demonstrates that the self-evolving neural networks contain the capability of fault tolerance to filter out the noise and provide high degree of accuracy of nonlinear function used to produce the model identification [12]. In view of this, this paper has adopted the ENN model to perform the non-parametric estimation with evolutionary process which systematical determines the optimal networks structure to obtain the global optimal solution instead of local optimal solution using traditional neural networks, and hence to obtain soft breakdown model with degraded characteristics of low frequency noise.

2 Evolutionary Neural Networks: Architectures and Algorithms

As to the type of ENN model adopted, it should be one where the users can easily define its exogenous variables or can add new elements to the ENN for system identification. In addition, the user should not have to worry about the problems of setting initial values to the weights of neurons, the number of neurons, the number of hidden layers and the neuron connections, etc., as attention can be devoted to optimizing the structure of networks on their own through genetic algorithms in the neural network learning process. Genetic algorithms are very effective at finding near global optimal solutions to a highly nonlinear function and a wide variety of problems [13], [14].

The ENN model used in this research is briefly described as follows. Chromosomes are used to express the neural network architecture and the parameters of the structure as presented in Fig. 1.

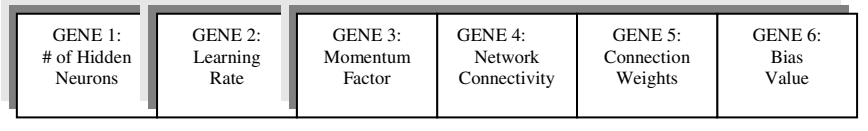

Fig. 1. Structure of chromosomes

The formation of each gene within a chromosome and the length of the chromosome can be determined depending on the cases. The network's evolutionary process is demonstrated in Fig. 2, and includes a training cycle and an evolutionary cycle. The steps for each stage are summarized below: (1) Initiate networks: randomly produce the initial networks' structure. (2) Training cycle: networks are built through genetic rules and a combination of weighted tuning. Training time is utilized in exchange for

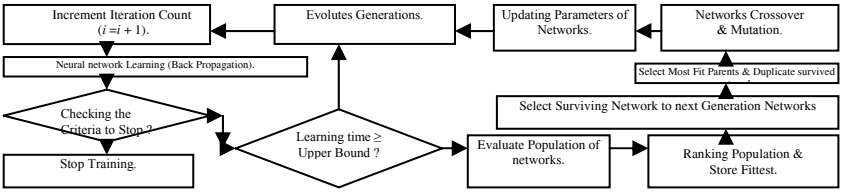

Fig. 2. The architecture of evolutionary cycle with nested training cycle for the ENN

the quality of an approximate optimal solution until the upper bound of the learning numbers can be reached. (3) Evolutionary cycle: the level of suitability of various networks for evaluating the fitness function is based on the mean square error, and the evolution of the networks then commences. In addition, based on the surviving networks which are decided by the suitability of the various networks, the reproduction, crossover and mutation of the surviving networks can be performed so as to generate a new generation of networks. (4) Return to Step (2) to conduct new generation network training until a satisfactory learning result or a pre-set termination condition is reached. We conclude the four stage operations of ENN as a pseudo code in Fig. 3:

```
Procedure GeNe
    Begin
        e = 0;
        initial population Pc(e);
        fitness Pc(e);
    While (termination criterion not
          reached)
        e= e + 1;
        Select Pc(e) from Pc(e-1);
        Crossover Pc(e);
        Mutate Pc(e);
        Fitness Pc(t);
        End.
```

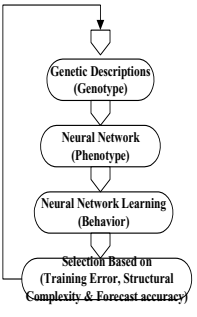

Fig. 3. The pseudo code of ENN

The parameters of neural networks is automatic fine tuned consist of two major adaptive methods, including the genetic search through moving window from observations, forecast horizon, network architecture space and control parameters to select the best performers. Finally, the back-propagation learning in each network evaluates the selected architectures to complete the evolution cycle.

We make good use of three criteria to measure the learning, testing and estimation accuracy in the ENN model so called performance indices, including R^2, *NRMSE* and *MAE*. The R^2 measures the interpretive capability of the SBD model to the real observations which is given as *Eq*. (1). It's more and more correlative to the real observations while R^2 is closer to 1. The NRMSE (Normalized Root Mean Square Errors) and MAE (Mean Absolute Errors) here is set to be the objective function for minimizing the tracking errors between the model outcomes and observations for SBD using ENN, which is given as *Eq*. (2) and *Eq*. (3).

$$R^2 = (n\sum_{i=1}^{n} x_i y_i - (\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i))^2 \Big/ (n\sum_{i=1}^{n} x_i^{\,2} - (\sum_{i=1}^{n} x_i)^2 (n\sum_{i=1}^{n} y_i^{\,2} - (\sum_{i=1}^{n} y_i)^2) \quad (1)$$

$$RMSE = (\sum_{i=1}^{n} (y_i - x_i)^2 \Big/ \sum_{i=1}^{n} (y_i - (mean(y_i)))^{0.5} \quad (2)$$

$$MAE = \sum_{i=1}^{n} |y_i - x_i| \Big/ n \quad (3)$$

Where x_i: input variables; y_i: estimation results by ENN model (output variable); n: # of observations.

3 Experiments and System Identification

In the experiment part, the n-channel MOSFET used in the study was fabricated in a 0.18 um process. In this process, the physical gate oxide thickness was determined to be 3.3 nm by using a C-V method. A constant gate voltage of 5.5V was adopted to stress the gate oxide, with source, drain, and substrate tied to ground. The high-field stress was interrupted several times for characterization of S_{id}. The S_{id} measurement set-up comprised a HP 35665A dynamic signal analyzer, BTA 9603 FET noise analyzer and HP4156B semiconductor parameter analyzer, as illustrated in the Fig. 4. The transistor with gate oxide area of 10×10 um^2 was measured at the inversion condition (V_{DS}=0.1V and V_{GS}=1V). The frequency range of S_{id} was from 1 Hz to 1K Hz. Three noise filters were used to eliminate the residual noise in all bias sources.

In the modeling part, when considering the behavior of soft breakdown, the factors are induced into the ENN model. On account of the significant correlation on drain

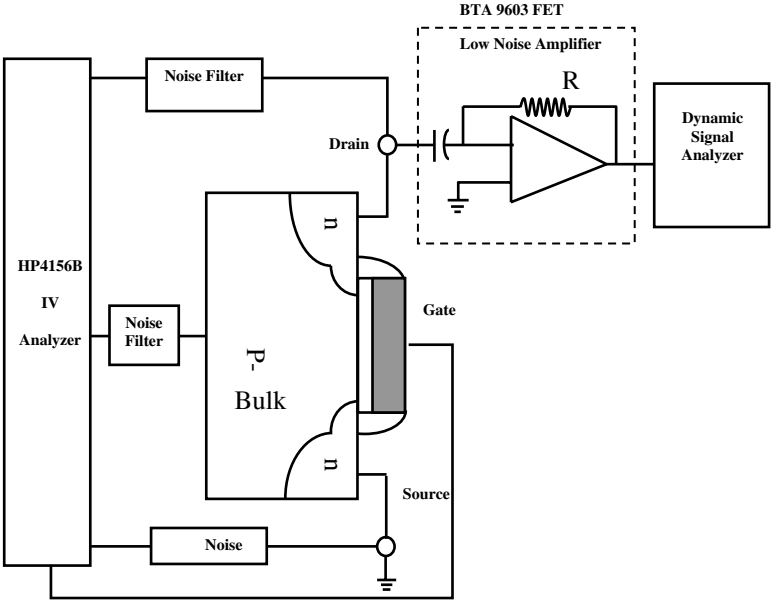

Fig. 4. Schematic diagram of the drain current noise measurement system

current in electrically stressed S_iO_2 film between various frequencies for several stress times before hard breakdown during from 1 Hz to 16.5 Hz (see Panel A, Table 1) which is also can be verified as exhibiting in Fig. 5. Fig. 5 shows that a significant correlation plot, that is, the normalization of noise power ($=S_{id}/I_D^2$) versus the gate current I_G, can serve as supporting evidence of the model since the S_{id} can be adequately traced to the gate current. A striking phenomenon of the gate current reduction in the SILC mode can be drawn herein. We can attribute it to the accumulation of the tunneling electrons trapped in the oxide.

Table 1. Correlation coefficient matrix for several stress times

Panel A. 1 Hz ≤ Q ≤ 16.5 Hz, Q: frequency			
	Fresh	10 sec	30 sec
Fresh	1	0.5885*	0.7301*
10 sec	0.5885*	1	0.7934*
30 sec	0.7301*	0.7934*	1
Panel B. 17 Hz≤ Q ≤1K Hz, Q: frequency			
	Fresh	10 sec	30 sec
Fresh	1	0.3316	0.1428
10 sec	0.3316	1	0.3652
30 sec	0.1428	0.3652	1

On the contrary, they trap insignificant correlation while the frequency increases into 1K Hz (see Panel B, Table 1). The auxiliary variables, I_G(A), I_D(A) and V_G(V) is necessary to induced jointly in our model as *Eq*. (4)

$$S_{id}(ST_y) = F(Q, I_G, I_D, V_G, ST, S_{id}(ST_x)),\ x \neq y. \tag{4}$$

where, I_G(A):gate current; I_D(A): drain current; V_G(V): gate voltage; Q: frequency; ST: stress time; $S_{id}(ST_x)$: S_{id} for x second stress time; $S_{id}(ST_y)$: S_{id} for y second stress time.

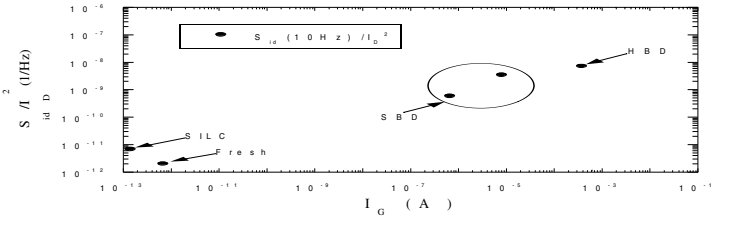

Fig. 5. A correlation plot between normalized drain current noise at 10 Hz frequency and gate current, all from the same bias condition of V_{DS}=0.1V and V_{GS}=1V

Hence, the factors of premise part of ENN-SBD include I_G(A), I_D(A), V_G(V), Q, ST, $S_{id}(ST_x)$ and the consequence part is $S_{id}(ST_y)$. In this study, we employ three criteria to stop training and testing. One is training epochs is equal to 50,000 times, the other is the training error doesn't change in late 1,000 times, error tolerance is less than 10^{-24}

or divergence happened. ENN would stop evolution and learning process automatically provided either of these criteria stands.

4 Experimental Results and Discussions

The time evolution of the gate current IG during constant voltage stress depicted in Fig. 6, showing two distinct events: SILC and SBD. After the occurrence of SBD, the gate current exhibits fluctuation phenomena as partially magnified in the insert of Fig. 6. The current fluctuations arise from the trapping and detrapping processes in and around SBD damaged region. Moreover, the high-field stress is interrupted several times to observe the characterization of S_{id}.

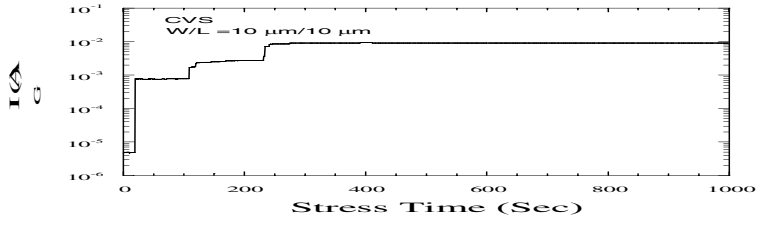

Fig. 6. Measured gate current versus stress time under a constant voltage stress

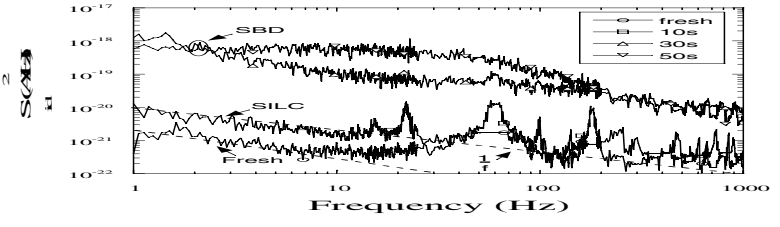

Fig. 7. Measured drain current noise power spectrum (S_{id}) for several stress times

Fig. 7 shows the measured S_{id} for different stress times. The S_{id} data follow the $1/f^{\gamma}$relationship with different value of power exponent γ. The spatial traps distribution within the gate oxide can account for different γvalue. The interpretation is similar to a literature model report [15]. Firstly, for S_{id} in the SILC mode the γ of 1 indicates that the traps are spatially uniform distribution within the gate oxide. Also, the oxide traps number increases with the time. The amount of oxide traps reflects the S_{id} magnitude increase. Subsequently, we observe a typical Lorentzian spectrum component after the occurrence of SBD. This component mainly arises from the trapping and detrapping processes in the localized oxide damage region, that is, a non-uniform oxide traps distribution. The inference is analogous to a Lorentzian spectrum observation on the small area device [16]. The area of the localized oxide damage region of around 0.3 nm^2 is proposed to account for this [17].

In addition, the time constant of the Lorentzian spectrum is estimated to be 3.5 *ms*, as shown in the Fig. 8. The time constant in S_{id} is associated with the net effect of the interaction between the oxide traps within the localized SBD damaged region and the underlying channel carrier, which is accounting for the carrier-number fluctuation model.

Fig. 8 shows ten fitting models for soft breakdown identification for 50 sec tress time, including linear, logarithmic, power, interpolate, polynomial, moving average, exponential, cubic-spline, $1/f^{\gamma}$ and ENN-SBD. This is consistent with the observation in Table 1. For most of the fitting models, they are all lose efficacy during the damage higher low-frequency region occurs area except the ENN. Besides, ENN has more precisely identification capability than typical Lorentzian spectrum method ($1/f^{\gamma}$).

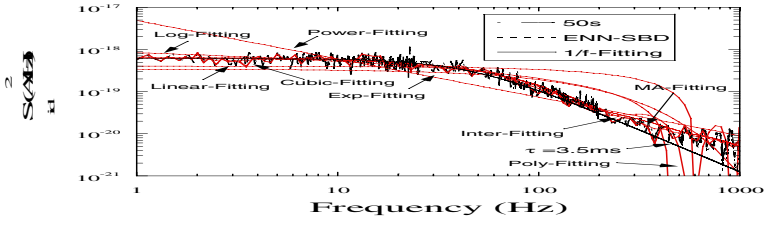

Fig. 8. Extracted the time constant of the Lorentzian spectrum from the data in Fig. 7

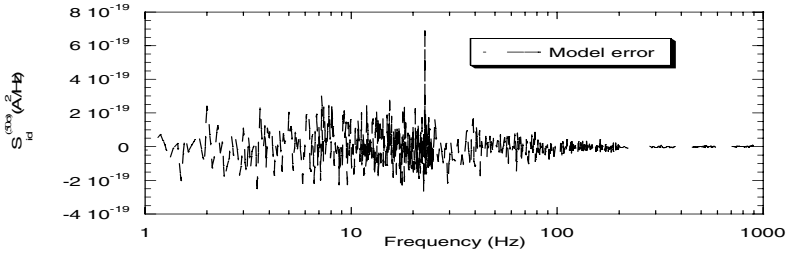

Fig. 9. ENN-SBD estimation error versus frequency

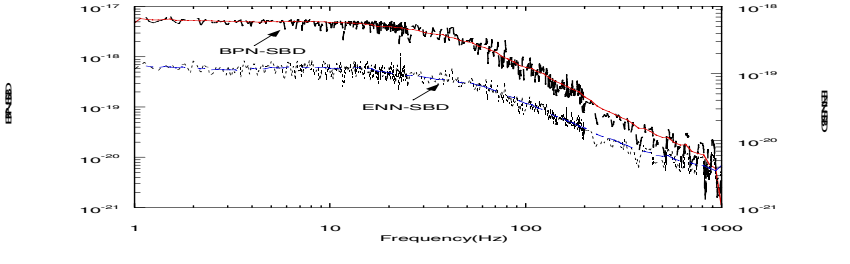

Fig. 10. Model comparisons between ENN-SBD and BPN-SBD for different frequency

In Fig. 9, there are 735 observations for each input/output variable. The moving window technique is employed to rolling the observations for batch learning in case of stress time is 50 sec. The first 485 pairs (training set) was used for training the ENN while the remaining 200 pairs for testing (used for validating the identified model), and 50 pairs for estimating. The resulting 111 neurons are generated after stop learning. It is quit unusual to observe the phenomenon that $RMSE_{\text{train}} \leq RMSE_{\text{test}}$ during the training process. Considering both the *RMSE*'s are very small, we conclude that: (1) the ENN has captured the essential components of the underlying dynamics; (2) the training data contains the effects of the initial conditions which might not be easy accounted for by the essential components identified by the ENN. As a comparison, we performed the same estimation by using the BPN with the same number of observations.

Fig. 10 shows the model comparisons between ENN-SBD and BPN-SBD with various frequencies. Obviously, ENN-SBD has a better performance than BPN-SBD in spite of committing to convergence at the end. It might indicate that the ENN can easily capture the spirit of "rule of thumb" used by humans, from another angle, it's more adaptive than BPN through genetic algorithms in the architecture.

5 Conclusions and Extensions of Current Work

First of all, during high-field stress, the origins of the S_{id} data in the SILC and SBD mode have been presented. The S_{id} data follow the $1/f^{\gamma}$ relationship with different value of power exponent γ. The spatial oxide traps distribution has been proposed to account for the different γ value. In addition, the Lorentzian spectrum is associated with the net effect of the interaction between the oxide traps within the localized SBD damaged region and the underlying channel carrier, that is, accounting for the carrier-number fluctuation model. It has pointed out that the S_{id} correlates closely with the gate fluctuations via the trapping and detrapping processes. Secondly, we have described the ENN-SBD mechanisms with case studies for the behavior inference of the degraded characteristics of low frequency noise after soft breakdown. By employing a hybrid learning procedure, the proposed model can refine the architecture to describe a complex CMOS system and has better performance than original Lorentzian spectrum method and BPN mechanism. Fabs build our model could thereby perform the simulation work instead of hardware operation and testing to reduce the reverent testing time and the costs.

Another important issue in the breakdown identification is how to modeling the hard breakdown which is even more difficult to estimate because it is tends to be an uncertainty problem. However, it would be more valuable to prevent the catastrophic hard breakdown occurs.

Acknowledgements

This paper was supported by the National Science Council under contract number NSC-94-2516-S-018-017. In addition, the author is very grateful to Dr. Kang for his valuable supports and the anonymous reviewers for their suggestions and comments.

References

1. Rietman, Edward A., Whitlock, Stephen A., Andrew Roy, Milton Beachy and Willingham, Timothy L. A System Model for Feedback Control and Analysis of Yield: A Multistep Process Model of Effective Gate Length, Poly Line Width, and IV Parameters. IEEE Transactions on Semiconductor Manufacturing. 14:1 (2001) 32-47.
2. Newnes, Linda B., Mileham, Tony R. and Doniavi, Ali. A Systems Approach to Semiconductor Optimization. IEEE Transactions on Electronics Packaging Manufacturing. 24:3 (2001) 171-177.
3. Lee, S. H., Cho, B. J., Kim, J. C., and Choi, S. H. Quasi-breakdown of ultrathin gate oxide under high field stress. IEDM Tech. Dig. (1994) 605-608.
4. Depas, M., Nigam, T., and Heyns, M. M. Soft breakdown of ultra-thin gate oxide layers. IEEE Trans. Electron Devices. 43 (1996) 1499-1504.
5. Houssa, M., Vandewalle, N., Nigam, T., Ausloos, M., Mertens, P. W., and Heyns, M. M. Analysis of the gate voltage fluctuations in ultra-thin gate oxides after soft breakdown. IEDM Tech. Dig. (1998) 909-912.
6. Weir, B. E., Silverman, P. J., Monroe, D., Krisch, K. S., Alam, M. A., Alers, G. B., Sorsch, T. W., Timp, G. L., Baumann, F., Liu, C. T., Ma, Y., and Hwang, D. Ultra-thin gate dielectrics : they break down, but do they fail? IEDM Tech. Dig. (1997) 73-76.
7. Braha, Dan and Shmilovici, Armin. Data Mining for Improving a Cleaning Process in the Semiconductor Industry. IEEE Transactions on Semiconductor Manufacturing, 15:1 (2002) 91-101.
8. Baker, M. D., Himmel, C. D. and May, G. S. Time series modeling of reactive ion etching using neural networks. IEEE Transactions on Semiconductor Manufacturing. 8:1 (1995) 62–71.
9. Rietman, E. A. and Lory, E. R. Use of neural networks in modeling semiconductormanufacturing processes: An example for plasma etch modeling. IEEE Transactions on Semiconductor Manufacturing. 6 (1993) 343–347.
10. Rietman, E. A. A neural network model of a contact plasma etch process for VLSI production. IEEE Transactions on Semiconductor Manufacturing. 9 (1996) 95–100.
11. Han, S. S. and May, G. Using neural networks process models to perform PECVD silicon dioxide recipe synthesis via genetic algorithms. IEEE Transactions on Semiconductor Manufacturing. 10 (1997) 279–287.
12. Ling, S.H., Leung, F.H.F., Lam, H.K., Yim-Shu Lee and Tam, P.K.S. A novel genetic-algorithm-based neural network for short-term load forecasting. Industrial Electronics, IEEE Transactions. 50: 4 (2003) 793 – 799.
13. Wong, F. and Lee, D. A Hybrid Neural Network for Selection. Proc. of the 2nd Annual Int. Conf. on Neural Networks (1993).
14. Wong, F., Tan, P. and Zhang, X. Neural Networks, Genetic Algorithms and Fuzzy Logic for Forecasting. Proceedings of the 3rd Int. Conf. on Advanced Applications and Worldwide (1992).
15. Surya, C. and Hsiang, T. Y. Theory and experiment on the $1/f^{\gamma}$ noise in p-channel metal-oxide-semiconductor field-effect transistors at low drain bias', Phys. Rev. B, 33 (1986), 4898-4905.
16. Uren, M. J., Day, D. J., and Kirton, M. J. $1/f$ and random telegraph noise in silicon metal-oxide-semiconductor field-effect transistors. Appl. Phys. Lett., 47 (1985) 1195-1197.
17. Chen, M. J., Kang, T. K., Liu, C. H., Chang, Y. J., and Fu, K. Y. Oxide thinning percolation statistical model for soft breakdown in ultrathin gate oxides. Appl. Phys. Lett. 77 (2000) 555-557.