Modeling Attack Tests and Security Enhancement of the Sub-threshold Voltage Divider Array PUF

Shengjie Zhou[1, †], Yongliang Chen[1,†], Xiaole Cui[1,2,*], Yun Liu[1]
1) Key Lab of Integrated Microsystems, Peking University Shenzhen Graduate School, Shenzhen 518055, China
2) Peng Cheng Laboratory, Shenzhen 518066, China
cuixl@pkusz.edu.cn

***Abstract*— Physical unclonable function (PUF) is widely used as the root of trust in the IoT systems. The sub-threshold voltage divider array PUF was reported as an anti-modeling-attack PUF. It utilizes the nonlinear I-V relationship in the sub-threshold region of MOS transistors to improve the security. However, the security of this PUF has not been soundly analyzed. This work presents the simulation results of modeling attack tests which reveal the vulnerability of this PUF. In the attack, the sub-threshold voltage divider array PUF is modeled by a dedicated artificial neural network (ANN). The nonlinearity of the PUF is simplified based on its working principle. The simulation results show that the prediction accuracy achieves 97% when the number of training CRPs is 350 for the single-stage PUF, and it achieves 90% when the number of training CRPs is 300 for the multi-stage PUF. Furthermore, this work improves the original structure of the sub-threshold voltage divider array PUF, to enhance its anti-modeling-attack capability. The simulation results of modeling attack tests show that the prediction accuracy of the improved multi-stage PUF is reduced to about 50%, which implies that the improved PUF has a strong anti-modeling-attack capability.**

Keywords—The sub-threshold voltage divider array PUF, modeling attack test, ANN.

I. INTRODUCTION

In recent years, the physical unclonable function (PUF) is adopted as the root of trust in the resource constrained applications [1]. The input and output of PUF are referred as the challenge and the response respectively. The challenge and its corresponding response form a challenge-response pair (CRP) [2]. The CRPs of PUF are generated based on the fluctuation of physical parameters, instead of being stored in memories. This fluctuation cannot be read out directly, as in the traditional non-volatile memories. Thus the PUFs are less vulnerable to the reverse engineering and the probing attacks. However, many PUFs are proved to be vulnerable to the modeling attacks. Their CRPs can be predicted by the mathematical model constructed based on a sub-set of CRPs. The modeling attacks become a real threat to the PUF-based system. Therefore the anti-modeling-attack capability becomes an important quality of PUF.

Researchers have proposed some anti-modeling-attack PUF schemes [4-6]. Recently, the nonlinearity of the I-V curve in the sub-threshold region of MOS transistors is utilized to build the PUF with higher security. Abilash Venkatesh et al. [7] proposed a PUF using the sub-threshold voltage divider array, as shown in Fig. 1. It is referred as the sub-threshold PUF in this paper for simplicity. The structure of the sub-threshold PUF is presented in Fig. 1. It consists of the upper array and the lower array with N PUF units. A PUF unit comprises of a CMOS inverter with the shorted gate and drain, along with an NMOS tail current source biased at the threshold voltage. The working process of the sub-threshold PUF is as follows. The same challenge vector, noted as [c_1 , c_2 ,... , c_N], is applied to the two arrays to determine which PUF units are connected. Then the output voltage $Vout$ is generated by the parallel-connected units in the upper array, as shown in Fig. 2. Similarly, the lower array also generates another $Vout$. These two $Vout$s are compared to determine the final PUF response. The test results indicated that the sub-threshold PUF is able to resist the modeling attacks based on the logistic regression (LR), the support vector machine (SVM) and the multi-layer perceptron (MLP) methods [7,8].

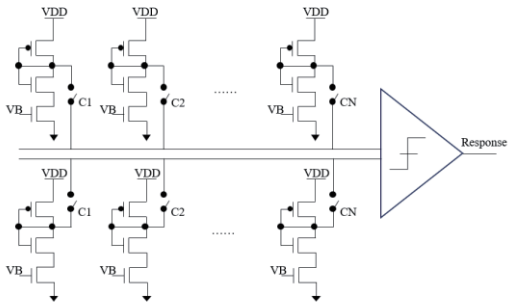

Fig. 1. The sub-threshold voltage divider array PUF [7].

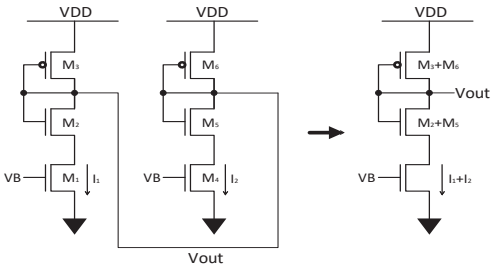

Fig. 2. The nonlinearity of two connected PUF units in the sub-threshold voltage divider array PUF [7].

This work proposes a specialized attack method to further test the security of the sub-threshold voltage divider array PUF.

* This author is the corresponding author.
† These authors contributed equally to this work.

979-8-3503-4859-0/DATE24/© 2024 EDAA

A dedicated artificial neural network (ANN) is constructed for the attack test based on the mathematical analysis. The specialized attack method helps to deliver a more realistic test result since the ANN is built according to the working principle of the target PUF. The prediction accuracies are higher than 90% with 350 training CRPs for both the single-stage and the multi-stage PUFs. The simulation test results show that the sub-threshold PUF is also vulnerable to the modeling attacks. Furthermore, an improved scheme is proposed to enhance the security of sub-threshold PUF. The effectiveness of the proposed security improvement for the sub-threshold PUF is also verified by the simulation results of the attack tests.

II. The Attack Tests on the Sub-threshold Voltage Divider Array PUF

A. The Attack Test on the PUF with Single-stage Sub-threshold Voltage Divider Array

The workflow of the sub-threshold PUF is analyzed as follows. For the PUF with single-stage sub-threshold voltage divider array in Fig. 2, two selected PUF units are connected in parallel to generate V_{out}. The currents I_1 and I_2 passing through the two PUF units can be expressed as (1) and (2).

$$I_1 = I_S \cdot exp\left(\frac{V_B - V_{th1}}{mV_T}\right) \tag{1}$$

$$I_2 = I_S \cdot exp\left(\frac{V_B - V_{th4}}{mV_T}\right) \tag{2}$$

In (1) and (2), V_{th1} and V_{th4} are the threshold voltages of transistors M_1 and M_4 respectively; V_T is the thermal voltage kT/q; V_B is the bias voltage of NMOS tail current source. It is assumed that V_{ds} of M_1 and M_4 are greater than 100mV. When these PUF units are connected together, the output voltage can be computed as (3) and (4).

$$I_1 + I_2 = I_{S1} \cdot exp\left(\frac{V_{DD} - V_{out} - |V_{thp}|}{mV_T}\right) \tag{3}$$

$$V_{out} = V_{DD} - |V_{thp}| - ln\left(\frac{I_S^{mV_T}}{I_{S1}}\right) - mV_T \cdot ln\left[exp\left(\frac{V_B - V_{th1}}{mV_T}\right) + exp\left(\frac{V_B - V_{th4}}{mV_T}\right)\right] \tag{4}$$

where $|V_{thp}|$ represents the threshold voltage of the PMOS transistor. It can be observed from (4) that V_{out} is nonlinear with respect to the threshold voltages V_{th1} and V_{th4} of the NMOS tail current sources. V_{th1} and V_{th4} fluctuate with random doping variations, and they provide the uncertainty of V_{out}. The other threshold voltages of NMOS transistors also introduce the randomness into the PUF. Both of them serve as the entropy sources of the sub-threshold PUF.

Same as (4), the output voltage of the upper array can be expressed as (5) when the challenge c_i is applied.

$$V_{out1} = V_{DD} - |V_{thp}| - \ln\left(\frac{I_S^{mV_T}}{I_{S1}}\right) - mV_T \cdot \ln\left[\sum_{i=1}^{N} c_i \cdot \exp\left(\frac{V_B - V_{thi}}{mV_T}\right)\right] \tag{5}$$

The upper and lower array has the symmetric structure. The output voltage of the lower array is computed as (6) accordingly, where V_{thi} and $V_{thi}^{'}$ represent the i[th] entropy source in the upper and lower array, respectively.

$$V_{out2} = V_{DD} - |V_{thp}| - \ln\left(\frac{I_S^{mV_T}}{I_{S1}}\right) - mV_T \cdot \ln\left[\sum_{i=1}^{N} c_i \cdot \exp\left(\frac{V_B - V_{thi}^{'}}{mV_T}\right)\right] \tag{6}$$

Then the corresponding response of c_i is expressed as (7) and (8).

$$Res = sgn(V_{out1} - V_{out2}) \tag{7}$$

$$Res = sgn\left\{\begin{matrix} mV_T \cdot \ln\left[\sum_{i=1}^{N} c_i \cdot \exp\left(\frac{V_B - V_{thi}^{'}}{mV_T}\right)\right] \\ -\ln\left[\sum_{i=1}^{N} c_i \cdot \exp\left(\frac{V_B - V_{thi}}{mV_T}\right)\right] \end{matrix}\right\} \tag{8}$$

The function $sgn(x)$ is introduced by the comparator. This attack exploits the characteristics of the sign function to simplify the nonlinearity of $\ln\left[\sum_{i=1}^{N} c_i \cdot \exp\left(\frac{V_B - V_{thi}^{'}}{mV_T}\right)\right]$ function. The result of $sgn(x)$ function only depends on the sign of x. The term mV_T is a positive constant, which has no influence on the sign of $(V_{out1} - V_{out2})$. The terms $[\ln(\sum_{i=1}^{N} c_i \cdot \exp\left(\frac{V_B - V_{thi}^{'}}{mV_T}\right))-\ln(\sum_{i=1}^{N} c_i \cdot \exp\left(\frac{V_B - V_{thi}}{mV_T}\right))]$ and $(\sum_{i=1}^{N} c_i \cdot \exp\left(\frac{V_B - V_{thi}^{'}}{mV_T}\right) - \sum_{i=1}^{N} c_i \cdot \exp\left(\frac{V_B - V_{thi}}{mV_T}\right))$ have the same sign, since $ln(x)$ function is both monotonic and increasing. Thus, the equation (8) can be simplified to (9).

$$\begin{aligned} Res &= sgn\left\{\left[\sum_{i=1}^{N} c_i \cdot \exp\left(\frac{V_B - V_{thi}^{'}}{mV_T}\right)\right] - \left[\sum_{i=1}^{N} c_i \cdot \exp\left(\frac{V_B - V_{thi}}{mV_T}\right)\right]\right\} \\ &= sgn\left\{\sum_{i=1}^{N} c_i \cdot \left[\exp\left(\frac{V_B - V_{thi}^{'}}{mV_T}\right) - \exp\left(\frac{V_B - V_{thi}}{mV_T}\right)\right]\right\} \end{aligned} \tag{9}$$

Another technique is used to simplify the nonlinearity of $\exp(\frac{V_B - V_{thi}}{mV_T})$. As shown in (9), the entropy sources $V_{thi}^{'}$ and V_{thi} are always extracted in pairs, and they only provide a single fluctuation value. So an equivalent effective entropy source Δ_i is able to substitute $\left[\exp\left((V_B - V_{thi}^{'})/mV_T\right) - \exp((V_B - V_{thi})/mV_T)\right]$. Then (9) is simplified to (10).

$$Res = sgn\{c_1 \cdot \Delta_1 + c_2 \cdot \Delta_2 + \cdots\} \tag{10}$$

The relationship between the response and challenge is derived from (7) to (10). The original nonlinear part $(V_{out1} - V_{out2})$ in (8) is simplified to a linear part in (10). The non-differentiable function $sgn(x)$ is replaced by the $sigmoid(x)$ function as (11), in order to apply the gradient descent method.

$$Res = sigmoid\left\{\sum_{i=1}^{N} c_i \cdot \Delta_i\right\} \tag{11}$$

Then the linear mathematical model is realized in an ANN model as shown in Fig. 3. The ANN consists of the input layer with N input neurons and the output layer with one neuron. The input layer corresponds to the N-bit challenge vector, and the output layer computes the prediction of the response. The synapses between the two layers correspond to the effective entropy sources $\Delta_1 \sim \Delta_N$. When the ratio of weights on synapses are trained to be similar to the ratio of effective

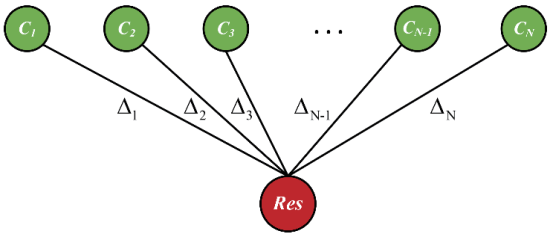

Fig. 3. The ANN of the single-stage sub-threshold voltage divider array PUF.

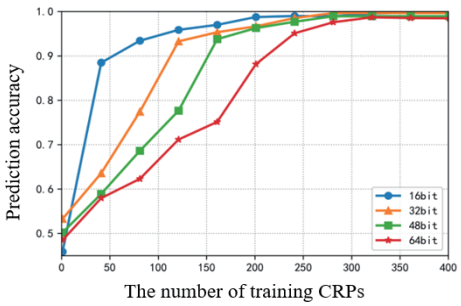

Fig. 4. Attack result of the single-stage sub-threshold PUF.

entropy sources, the ANN model is able to predict the unknow CRPs with high accuracy.

The attack test simulations are carried out on the single-stage sub-threshold PUF. Assumes that the attacker can access up to half of the total CRPs of the target PUF. The modeling attack tests consist of two phases: the training phase and the testing phase. The CRPs are collected from the target PUFs with different sizes. Part of these CRPs is selected randomly and serves as the training set. The remaining CRPs are used as the testing set. The proposed ANN model is trained by the training set. Then the trained ANN model is utilized to predict the CRPs in the testing set, and the prediction accuracy is recorded. The attacks are performed on the target PUFs with 16-bit, 32-bit, 48-bit and 64-bit arrays, and the prediction accuracies of the responses are shown in Fig. 4. It shows that the prediction accuracy of the proposed ANN model is above 97% with 350 training CRPs. While the prediction accuracies of the LR, SVM and MLP methods are reported to be 60% with 8000 CRPs in [8]. It proves that the sub-threshold PUF is not invincible to the modeling attacks.

B. The Attack Tests on the PUF with Multi-Stage Sub-Threshold Voltage Divider Array

To enhance the security of sub-threshold PUF, Abilash Venkatesh et al. proposed the PUF with multi-stage sub-threshold voltage divider array [8], as shown in Fig. 5. It is referred as the multi-stage sub-threshold PUF in this paper. The multi-stage sub-threshold PUF consists of several single-stage sub-threshold PUFs, and these single-stage PUFs are connected in series. These single-stage PUFs are used as the sub-PUFs in the multi-stage sub-threshold PUF. The outputs of a single-stage PUF determine the first input bits of its successor single-stage PUFs. The last single-stage PUF generates the final response. The multi-stage sub-threshold PUF utilizes the same PUF units as the single-stage sub-threshold PUF.

Although the multi-stage sub-threshold PUF can have more stages of sub-PUFs, [8] opted for the three-stage configuration for the following reasons. In an ideal scenario, the accuracy of

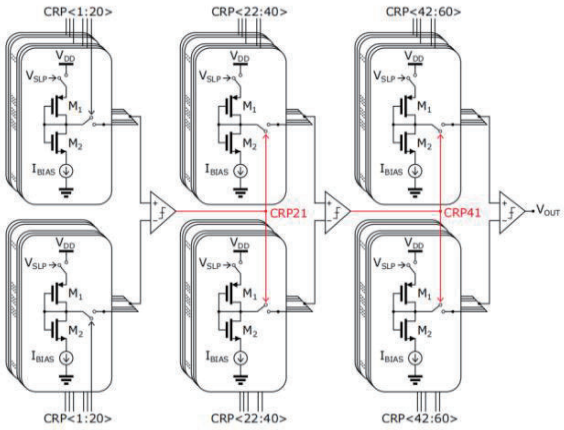

Fig. 5. A three-stage sub-threshold voltage array PUF [8].

modeling attack exhibits a significant decline as the number of cascading stages increases. Furthermore, to enhance the resistance against the modeling attack, a minimum of seven PUF units per stage is recommended [8]. However the multiple stages consume more area for hardware implementation. Additionally, for the PUF with more stages, it amplifies the total comparator noise, since each stage has its own comparator. It also deteriorates the reliability of PUF because more energy is consumed in the multi-stage sub-threshold PUF. The higher energy consumption results in more temperature fluctuations, voltage noise and current leakage, thereby the reliability of PUF becomes worse.

The mathematical model is established for the three-stage sub-threshold PUF in Fig. 5 as an example. The working principle of each sub-PUF is similar as that of the single-stage sub-threshold PUF. The only difference lies in the additional input bit from the precursor stage. A three-stage sub-threshold PUF with $2N$ PUF units can accept $(N-2)$-bit challenge. Two challenge bits are substituted by the outputs of the first and second sub-PUFs. The responses of the three sub-PUFs are computed as (12), (13) and (14), respectively.

$$Res1 = sigmoid\left[\sum_{i=1}^{p} c_i \cdot \Delta_i\right] \quad (12)$$

$$Res2 = sigmoid\left[Res1 \cdot \Delta_{p+1} + \sum_{i=p+2}^{q} c_i \cdot \Delta_i\right] \quad (13)$$

$$Res = sigmoid\left[Res2 \cdot \Delta_{q+1} + \sum_{i=q+2}^{N} c_i \cdot \Delta_i\right] \quad (14)$$

where p, $(q-p)$, and $(N-q)$ are the number of PUF units in the first, second and third sub-PUFs, respectively. The ANN model is designed for the attack test of the three-stage sub-threshold PUF accordingly, as shown in Fig. 6. The ANN model consists of four layers with p, $(q-p)$, $(N-q)$ and 1 neuron, respectively. Every two adjacent layers work as the ANN model in Fig. 3. The weights between each two adjacent layers represent the values of entropy sources of the corresponding sub-PUFs. The first layer has p neurons, which represent the p-bit input vector of the first sub-PUF. Same as the ANN model in Fig. 3, the neuron $Res1$ in the second layer computes the output of the first sub-PUF. Then $Res1$ serves as an input bit of the second sub-PUF. The output of the second sub-PUF $Res2$ is determined by $Res1$ and the other input bits in the second layer. Similarly, the output neuron Res computes the final response with $Res2$ and other input bits in the third layer.

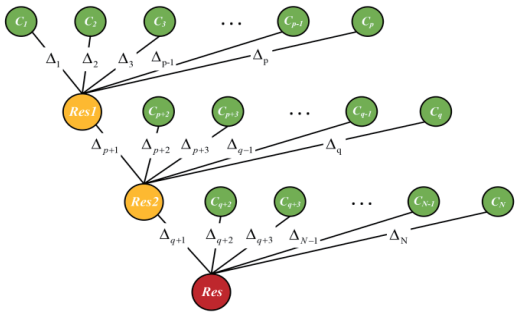

Fig. 6. The ANN model for the three-stage sub-threshold voltage divider array PUF.

The attack test simulations are performed on the multi-stage sub-threshold PUFs with array sizes of 20 bit×2 stage, 20 bit×3 stage, 20 bit×4 stage, 20 bit×5 stage, respectively. The values of entropy sources, i.e. the sub-threshold voltages, are randomly generated according to the normal distribution. 1000 pairs of CRPs are collected from each PUF. These CRPs are randomly divided, and are used for the training and the testing. The number of training CRPs increases 100 for each step, and the corresponding prediction accuracy is recorded.

Fig. 7 presents the relationship between the prediction accuracy of the testing set and the training CRP count. The required size of the training set to achieve the same prediction accuracy increases with the number of stages. The overall trend remained consistent with Fig. 7. A 'peak prediction accuracy' occurs at around 300 training CRPs, the growth rate of the prediction accuracy slows down significantly if more training CRPs are applied. The introduction of multi-stage configuration does not increase the training CRP counts required to reach the peak prediction accuracy. However, the maximum prediction accuracy decreases as the number of stages increases, when the training CRPs are 800. It shows that the extra stages enhance the PUF's resistance against attacks. This reduction in the maximum prediction accuracy is attributed to the accumulation of training errors in the multi-stage ANN model. With each additional sub-PUF, an extra layer of approximation is introduced in the ANN model, which results in a decrease in the prediction accuracy.

To deal with the accuracy decrease caused by the accumulation of errors, another simplified attack method is further proposed. It only considers the final stage of sub-PUF and neglects the influence of the precursor stages on the PUF response. This approach simplifies the multi-layer ANN model back to the two-layer ANN, and can attack the target PUF with a high prediction accuracy.

For the three-stage sub-threshold PUF, only the first input bit of the last sub-PUF is related to the precursor sub-PUFs. The simplified attack test ignores the first input bit, and computes the final response by the rest of input bits as (15).

$$Res = sigmoid\left[\sum_{i=q+2}^{N} c_i \cdot \Delta_i\right] \tag{15}$$

In such cases, the m-stage sub-threshold PUF with n-bit challenge is attacked as a $(n/m-1)$-bit single-stage sub-threshold PUF. The ANN model of the$(n/m-1)$-bit single-stage sub-threshold PUF can also be utilized in the simplified attack on the m-stage sub-threshold PUF. Accordingly, the attack test simulations are carried out, and the results are

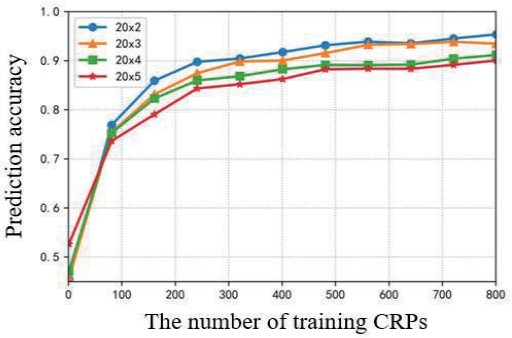

Fig. 7. The attack results of multi-stage sub-threshold PUF with the ANN model in Fig.6.

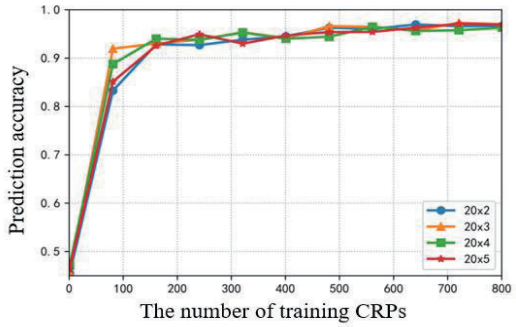

Fig. 8. The attack result of multi-stage sub-threshold PUF with the simplified attack method.

shown in Fig. 8. Fig. 7 and Fig. 8 show that, the simplified method achieves almost the same prediction accuracy when the number of stages of sub-PUF is 2 or 3, and the simplified method achieves even higher prediction accuracy when the number of stages is 4 or 5.

The simplified method is effective in the modeling attacks of the multi-stage sub-threshold PUF. For instance, the maximum prediction accuracy improves from around 90% to approximately 96% for a five-stage PUF. This improvement is attributed to the avoidance of error introduced by the approximation of multi-stage models when using only the single-stage model for the attack tests. Another advantage of the simplified method is that, the prediction accuracy reaches the peak value with less training CRPs, because the ANN model of the simplified method is simpler and easier to be trained. It can also be observed from Fig. 8 that, the peak prediction accuracy is the same for the multi-stage sub-threshold PUFs with the different number of stages. It shows that, the error does not increase with the number of the stages, although the first input bit induced error is ignored.

III. The Security Enhancement of the Multi-stage Sub-threshold PUF

The simulation results of the attack tests reveal the weak resistance of the multi-stage sub-threshold PUFs against modeling attacks. This weak resistance comes from the limited influence of each challenge bit on the final response [9]. And the limited influence results in two weaknesses for the security of multi-stage sub-threshold PUFs: the poor randomness of CRPs [10] and the minimal effect from the precursor sub-PUFs [11]. Firstly, the randomness of the multi-stage sub-threshold PUF is measured. 1000 PUF samples are randomly generated for the multi-stage sub-threshold PUF with array size of 32 bit×3 stage. The CRPs of these samples

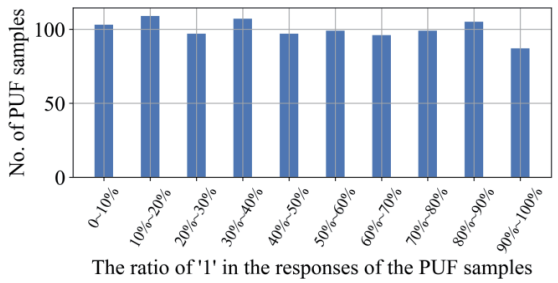

Fig. 9. The randomness of the single-stage sub-threshold PUF.

are obtained by simulation, and Fig. 9. records the ratio of '1' in the responses. It is observed from Fig. 9 that these PUF samples with different ratios of '1' do not follow a normal distribution, but they almost conform to a uniform distribution. About 60% PUF samples provide more than 70% or less than 30% '1' in the total responses. The poor randomness of multi-stage sub-threshold PUF significantly undermines the resistance against attack. Therefore the improvement on the randomness of the multi-stage sub-threshold PUF is a promising way to enhance its resilience against the modeling attack. Secondly, all the precursor sub-PUFs only control one bit of entropy source in the last sub-PUF. And the variation in this 1-bit entropy source has a minimal impact on the output response. A high prediction accuracy can still be achieved even when this 1-bit input is ignored.

According to the analysis above, the security of the multi-stage sub-threshold PUF can be enhanced from two perspectives: improving the randomness and enhancing the control of the challenge on the entropy source. An improved scheme for the sub-threshold PUF is proposed as shown in Fig. 10. In the original PUF, the PUF units in the upper array can only be connected to the upper input of the comparator. Inspired by the structure of Arbiter PUF, a connection network is introduced, which is configured by the challenges. A PUF unit is connected to the upper or lower input of the comparator through the network. During the working process, half of the parallel-connected PUF units are connected to the upper input of comparator, which generate one output voltage. And the remaining PUF units provide another output voltage to the lower input of the comparator. The two output voltages are compared to generate the final response.

The Res of the improved single-stage sub-threshold PUF in Fig. 10 is computed as (16).

$$Res = sgn\left\{\begin{array}{l}\varphi_1\left[\exp\left(\frac{V_B - V'_{th1}}{mV_T}\right) - \exp\left(\frac{V_B - V_{th1}}{mV_T}\right)\right] + \\ \varphi_2\left[\exp\left(\frac{V_B - V'_{th2}}{mV_T}\right) - \exp\left(\frac{V_B - V_{th2}}{mV_T}\right)\right] + \cdots\end{array}\right\} \quad (16)$$

$$= sgn\left(\sum_{i=1}^{n}\varphi_i\Delta_i\right)$$

where

$$\varphi_i = \prod_{j=i}^{n}(2c_j - 1)$$

$$\Delta_i = \exp\left(\frac{V_B - V'_{thi}}{mV_T}\right) - \exp\left(\frac{V_B - V_{thi}}{mV_T}\right)$$

The improved multi-stage sub-threshold PUF consists of several serially connected improved single-stage sub-threshold PUFs, as presented in Fig. 11. The output of the precursor sub-PUF serves as the last challenge bit, i.e. C_N in

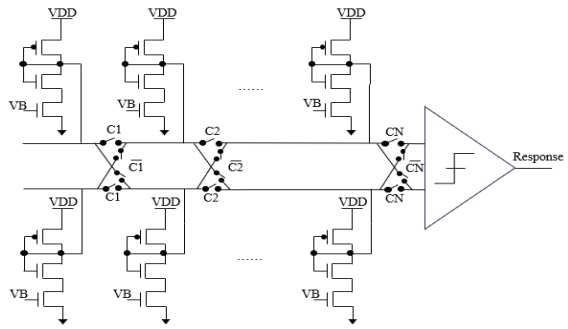

Fig. 10. The improved PUF entropy source control scheme for the single-stage sub-threshold PUF.

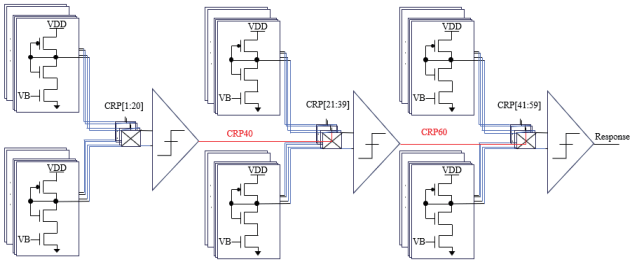

Fig. 11. The improved PUF entropy source control scheme of sub-threshold PUF with array size of 20 bit×3 stage.

Fig. 10. This structure can maximize the impact of the precursor sub-PUF, because C_N can influence most of the PUF units. The challenge bit is able to influence more PUF units, if the challenge bit is closer to the comparator. For example, as in Fig. 10, C_N can influence the connection of all the PUF units. But C_1 can only control the connection of the two PUF units on the far left of Fig. 10.

The mathematical model of the improved PUF scheme is expressed as (17), (18) and (19).

$$Res1 = sgn\left(\sum_{i=1}^{p}\varphi_i\Delta_i\right), \quad \varphi_i = \prod_{j=i}^{p}(2c_j - 1) \quad (17)$$

$$Res2 = sgn\left(Res1\Delta_q + \sum_{i=p+1}^{q-1}\varphi_i\Delta_i\right), \varphi_i = Res1 \times \prod_{j=i}^{q-1}(2c_j - 1) \quad (18)$$

$$Res3 = sgn\left(Res2\Delta_N + \sum_{i=q+1}^{N-1}\varphi_i\Delta_i\right), \varphi_i = Res2 \times \prod_{j=i}^{N-1}(2c_j - 1) \quad (19)$$

The attack test is carried out based on the ANN model in Fig. 6. Fig. 12 shows the simulation results of the attack tests of the improved PUF scheme. It shows that the anti-modeling-attack capability of the improved scheme is enhanced significantly. The prediction accuracy keeps below 55% even when the training CRP count are more than 10^4. In contrast, the CRPs of original PUF can be predicted with a high accuracy. The security improvement comes from the increased influence of the precursor sub-PUFs on the final response. The outputs of the precursor sub-PUFs affect the connection of all the PUF units in the last sub-PUF. As a result, the simplified method becomes ineffective. Besides that, the simulation results show that, the ANN model in Fig. 6 also cannot attack the improved PUF successfully. The errors accumulate in the feed-forward process of the multi-layer ANN. In the original scheme, these errors can be ignored due to their little influence on the final response. But the impact of these errors significantly increases in the improved scheme,

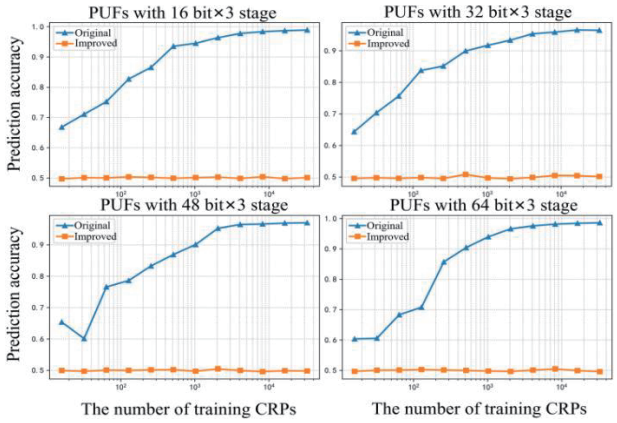

Fig. 12. The attack results of the improved PUF entropy source control scheme of the three-stage sub-threshold PUF.

because the outputs of the precursor sub-PUFs can control all the PUF units in the last sub-PUF. It amplifies the impact of errors, and makes the ANN more difficult to output a correct prediction. Furthermore, all the PUF units in the new scheme are connected to one of the two inputs of the comparator. It means that each entropy source jointly determines the response during the generation of each response bit. Whereas, the original scheme may utilize only several PUF units to generate the corresponding responses. Compared with the original scheme, the response of the new scheme is generated by more entropy sources. The security of PUF is enhanced by increasing the number of fluctuation parameters during the generation of each CRP.

2000 PUF samples are randomly generated for the 64-bit original and improved PUF schemes. 500 CRPs are collected from each PUF samples to compute the values of randomness and uniqueness for the PUF schemes. The threshold voltage of MOS transistors increases randomly after aging. Assumes that the values of the randomness increase following the normal distribution with a standard deviation of 5%. The CRPs of the PUF samples are collected to compute the reliability before and after aging.

The indicators of these PUF schemes are simulated in Table. I. The randomness of the improved schemes is closer to 0.5, and their standard variation values are smaller. It shows that the improved schemes have the better randomness than the original schemes for both the single-stage and the three-stage sub-threshold PUFs. The uniqueness of the original and the improved scheme is both near the ideal value (50%). And the variation of the improved schemes is smaller than that of the original schemes. The reliability of the improved scheme is only lower by approximately 0.007 compared to that of the original scheme.

IV. Conclusion

This work proposes a new test method for the subthreshold PUF. The dedicated ANN models are constructed for the single-stage and multi-stage sub-threshold voltage divider array PUFs. The non-linearity of the sub-threshold PUF is simplified during the design of the ANN model. The maximum prediction accuracy of 97% indicates that both the single-stage sub-threshold PUF and the multi-stage sub-threshold PUF are not invincible to the modeling attack. During the attack test simulations, the two weaknesses of the original sub-threshold PUF are identified, that is the poor PUF randomness and the poor controllability of the challenge to the entropy sources. Accordingly, this work proposes an improved PUF scheme to enhance the security. The attack results show that the prediction accuracy of the improved PUF is below 55%, which verifies the strong anti-modeling-attack capability of the improved PUF.

Table. I. The comparison of the PUF schemes (64-bit).

PUF Type	Randomness		Uniqueness		Reliability
	μ	σ	μ	σ	
Original Single-stage	0.483	0.138	0.503	0.167	0.972
Improved Single-stage	0.501	0.008	0.499	0.061	0.964
Original Three-stage	0.556	0.159	0.499	0.174	0.970
Improved Three-stage	0.499	0.009	0.500	0.062	0.963

Acknowledgment

This work was supported by National Natural Science Foundation of China under Grant No. 92373206, Shenzhen Science and Technology Innovation Committee u nder Grant No. JCYJ20220818100814033 and No.KQTD20 20082011310-5004.

References

[1] Cui, A., Zhou, W., Qu, G., & Li, H. (2018). "A New Scheme to Extract PUF Information by Scan Chain," in 2018 IEEE 27th Asian Test Symposium (ATS), Hefei, China, 2018, pp. 104-108.

[2] Song, X., Fan, G., & Rao, M. (2005). "Automatic CRP Mapping Using Nonparametric Machine Learning Approaches," IEEE Transactions on Geoscience and Remote Sensing, vol. 43, no. 4, pp. 888-897, April 2005.

[3] U. Rührmair and M. van Dijk, "PUFs in Security Protocols: Attack Models and Security Evaluations," 2013 IEEE Symposium on Security and Privacy, Berkeley, CA, USA, 2013, pp. 286-300, doi: 10.1109/SP.2013.27.

[4] Suh, G. E., & Devadas, S. (2007). "Physical Unclonable Functions for Device Authentication and Secret Key Generation," in Proceedings of the 44th Annual Design Automation Conference, 2007, pp. 9-14.

[5] Uddin, M., Majumder, M. B., and Rose, G. S. (2017). "Robustness Analysis of a Memristive Crossbar PUF Against Modeling Attacks," IEEE Transactions on Nanotechnology, vol. 16, no. 3, pp. 396-405.

[6] Sahoo, D. P., Mukhopadhyay, D., Chakraborty, R. S., et al. (2017). "A Multiplexer-Based Arbiter PUF Composition with Enhanced Reliability and Security," IEEE Transactions on Computers, vol. 67, no. 3, pp. 403-417.

[7] Venkatesh, A., Venkatasubramaniyan, A. B., Xi, X., and Sanyal, A. (2020). "0.3 pJ/Bit Machine Learning Resistant Strong PUF Using Subthreshold Voltage Divider Array," IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 67, no. 8, pp. 1394-1398, Aug. 2020, doi: 10.1109/TCSII.2019.2943121.

[8] Venkatesh, A., and Sanyal, A. (2019). "A Machine Learning Resistant Strong PUF using Subthreshold Voltage Divider Array in 65nm CMOS," in 2019 IEEE International Symposium on Circuits and Systems (ISCAS), IEEE, pp. 1-5.

[9] Kumar, A., Mishra, R. S., and Kashwan, K. R. (2016). "Challenge-Response Generation Using RO-PUF with Reduced Hardware," in 2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Jaipur, India, 2016, pp. 1305-1308.

[10] Xi, X., Zhuang, H., Sun, N., et al. (2017). "Strong Subthreshold Current Array PUF with 265 Challenge-Response Pairs Resilient to Machine Learning Attacks in 130nm CMOS," in 2017 Symposium on VLSI Circuits, IEEE, pp. C268-C269.

[11] Chen, Y., Cui, X., Ye, W., et al. (2021). "The Security Enhancement Techniques of the Double-Layer PUF Against the ANN-Based Modeling Attack," in 2021 IEEE International Test Conference (ITC), IEEE, pp. 63-72.