基于敏感度混淆机制的控制型物理不可克隆函数研究

徐金甫 吴 缙* 李军伟 曲彤洲 董永兴

(信息工程大学 郑州 450001)

摘 要：为了克服物理不可克隆函数(PUF)面对建模攻击的脆弱性，该文提出一种基于敏感度混淆机制的控制型PUF架构。根据PUF的布尔函数定义及Walsh谱理论，推导出各个激励位具有不同敏感度，分析并归纳了与混淆值位宽奇偶性有关的位置选取规则。利用该规则指导了多位宽混淆算法(MWCA)的设计，构建了具有高安全性的控制型PUF架构。将基础PUF结构作为控制型PUF的防护对象进行实验评估，发现基于敏感度混淆机制的控制型PUF所产生的响应具有较好的随机性。采用逻辑回归算法对不同PUF结构进行建模攻击，实验结果表明，相比基本ROPUF、仲裁器PUF以及基于随机混淆机制的OB-PUF，基于敏感度混淆机制的控制型PUF能够显著提高PUF的抗建模攻击能力。

关键词：信息安全；机器学习；布尔函数；敏感度

中图分类号：TP331 **文献标识码**：A **文章编号**：1009-5896(2019)07-1601-09

DOI: 10.11999/JEIT180775

Controlled Physical Unclonable Function Research Based on Sensitivity Confusion Mechanism

XU Jinfu WU Jin LI Junwei QU Tongzhou DONG Yongxing

(*The Information Engineering University*, *Zhengzhou* 450001, *China*)

Abstract: In order to overcome the vulnerability of Physical Unclonable Function (PUF) to modeling attacks, a controlled PUF architecture based on sensitivity confusion mechanism is proposed. According to the Boolean function definition of PUF and Walsh spectrum theory, it is derived that each excitation bit has different sensitivity, and the position selection rules related to the parity of the confound value bit width are analyzed and summarized. This rule guides the design of the Multi-bit Wide Confusion Algorithm (MWCA) and constructs a controlled PUF architecture with high security. The basic PUF structure is evaluated as a protective object of the controlled PUF. It is found that the response generated by the controlled PUF based on the sensitivity confusion mechanism has better randomness. Logistic regression algorithm is used to model different PUF attack. The experimental results show that compared with the basic ROPUF, the arbiter PUF and the OB-PUF based on the random confusion mechanism, the controlled PUF based on the sensitivity confusion mechanism can significantly improve the PUF resistance capabilities for modeling attack.

Key words: Information security; Machine learning; Boolean function; Sensitivity information

1 引言

物理不可克隆函数(Physical Unclonable Function, PUF)作为保障信息安全的一种新兴技术，具有不可克隆和不可预测等特性，其在轻量级安全协议、可信根构建和数字密钥存储等领域广泛应用。许多研究已经证明，基于机器学习算法的建模攻击能够成功破解PUF的激励响应行为[1]。

为了提高PUF在硬件防护应用中的安全性，国内外学者提出了许多安全增强机制。文献[2]提出了控制型PUF架构，旨在通过附加控制逻辑来增强PUF激励响应行为的隐蔽性，但仅提供了设计思路，没有给出具体实现来验证该架构的有效性。文献[3]在控制型PUF研究基础上，提出了基于哈希算法的逻辑可重构PUF模型，哈希算法用于绑定和隐藏PUF的激励响应行为。但敌手可以通过破解哈希算法来攻击内部的PUF模型，文献[4]证明了该PUF模型的不可预测性没有得到提高。文献[5]提出了一种新的可重构PUF模型，通过配置内部结构的状态信息来重构激励响应行为，使得学习算法难以对其进行预测。该模型以学习算法无法访问状态信息为前提假设，在实际应用中，为实现状态信息

收稿日期：2018-08-06；改回日期：2019-02-11；网络出版：2019-03-23
*通信作者： 吴缙 woshi57890@163.com

的隐秘性，需要承受额外的安全威胁。文献[6]针对轻量级认证协议需求设计了OB-PUF系统，采用随机混淆机制来防护针对仲裁器PUF(Arbiter Physical Unclonable Function, APUF)的建模攻击。随机混淆机制采用了混淆值位宽固定且插入位置随机选取的方法。PUF是依赖于器件物理特性且具有激励响应行为的概率函数[7]，文献[6]没有分析激励和响应的映射关系，在激励中随机地插入混淆值，无法达到最优的混淆效果。

综上分析，现有的PUF安全增强机制各有优势，但也有严重缺陷。为了有效增强PUF对建模攻击算法的防护能力，本文提出了一种具体的控制型PUF架构，其控制逻辑设计以敏感度混淆模型结论为理论依据。该模型详细分析了不同混淆值插入方案对敏感度的影响规律，并归纳出混淆位置选取规则，利用该规则来指导设计合理的混淆机制，作为基本PUF结构的防护机制，并通过实验评估其重要性。

2 敏感度混淆模型

控制型PUF通用架构如图1所示。本文中控制逻辑采取混淆激励响应间映射关系的方法，防护针对PUF的建模攻击。通过建立敏感度混淆模型，来分析混淆值位置选择对激励响应映射关系的影响规律。首先，根据布尔函数的Walsh谱理论[7]推导出PUF激励响应间的相关性。然后，基于敏感度概念分析PUF响应翻转概率期望的分布情况。最后，归纳出相应的混淆位置选取规则。

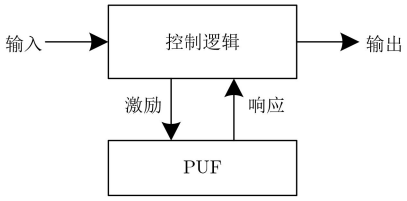


图 1 控制型PUF架构

2.1 布尔函数的Walsh谱理论

PUF的激励响应映射关系的数学模型可表示为2元域F_2上的布尔函数 $f_{\mathrm{PUF}}:\{0,1\}^n\to\{0,1\}$[7]。令 $V_n=\{c_1,c_2,\cdots,c_n\}$表示布尔变量(激励位)的集合，每个变量标记为 “1” 或 “0” 。任意布尔函数都能唯一表示成代数正规型的形式，则有

$$f(x)=\sum_{S\subseteq\{1,2,\cdots,n\}}a_S\prod_{i\in S}x_i \tag{1}$$

Walsh谱理论是分析布尔函数的输入与输出的相关性的重要工具[8]。在应用Walsh谱理论研究布尔函数时，通过编码 $\chi(0_{F_2}):=+1,\chi(1_{F_2}):=-1$ 将2元域扩展到复数域，沿用复数值的Walsh谱理论进行研究。

定义 1 设 $f_{\mathrm{PUF}}:Z_2^n\to Z_2$ 是布尔函数，$w\in Z_2^n$，则称

$$W_{(f_{\mathrm{PUF}})}(w)=\frac{1}{2^n}\sum_{x\in Z_2^n}(-1)^{f_{\mathrm{PUF}}(x)\oplus w\cdot x} \tag{2}$$

为 f_{PUF}在 w点的Walsh循环谱。

布尔函数 $f_{\mathrm{PUF}}(x)$转化为 ±1值函数 $(-1)^{f_{\mathrm{PUF}}(x)}$，复数值函数 $\tilde{f}(x)=(-1)^{f_{\mathrm{PUF}}(x)}$，故有

$$\begin{aligned}W_{(f_{\mathrm{PUF}})}(w)&=\frac{1}{2^n}\sum_{x\in Z_2^n}(-1)^{f_{\mathrm{PUF}}(x)\oplus w\cdot x}\\&=\frac{1}{2^n}\sum_{x\in Z_2^n}(-1)^{f_{\mathrm{PUF}}(x)}(-1)^{w\cdot x}\\&=\frac{1}{2^n}\sum_{x\in Z_2^n}\tilde{f}(x)(-1)^{w\cdot x}=W_{\tilde{f}}(w)\end{aligned} \tag{3}$$

据此就可由基于Walsh谱 $W_{\tilde{f}}(w)$的理论，得到布尔函数的Walsh循环谱理论。

定义 2 设 ξ,η 为2元随机变量，则称 $\rho(\xi,\eta)=2p(\xi=\eta)-1$ 为ξ与η的相关系数，其绝对值称为ξ与η的相关优势。

引理1[7] 设 $f_{\mathrm{PUF}}:Z_2^n\to Z_2$ 是布尔函数，$w\in Z_2^n$，则有

$$\begin{aligned}W_{(f_{\mathrm{PUF}})}(w)&=2p(f_{\mathrm{PUF}}(x)=w\cdot x)-1\\&=p(f_{\mathrm{PUF}}(x)=w\cdot x)\\&\quad-p(f_{\mathrm{PUF}}(x)\neq w\cdot x)\end{aligned} \tag{4}$$

其中，$p(f_{\mathrm{PUF}}(x)=w\cdot x)=\dfrac{1}{2^n}\#\{x\in Z_2^n:f_{\mathrm{PUF}}(x)=w\cdot x\}$。

定理1说明，布尔函数在w点的Walsh循环谱就是$f_{\mathrm{PUF}}(x)$与线性函数$w\cdot x$的相关系数，因而Walsh循环谱绝对值的大小，直接反映了$f_{\mathrm{PUF}}(x)$的输出与$f_{\mathrm{PUF}}(x)$输入的w线性组合$w\cdot x$之间的相关优势。

引理2[7] (能量守恒定理) 设$f_{\mathrm{PUF}}:Z_2^n\to Z_2$是布尔函数，则有

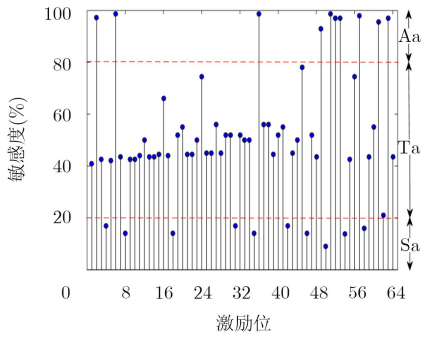
$$\sum_{w\in Z_2^n}\left[W_{(f_{\mathrm{PUF}})}(w)\right]^2=1 \tag{5}$$

引理2说明，一个布尔函数$f_{\mathrm{PUF}}(x)$不可能与输入的所有的线性组合$w\cdot x$的相关优势都是0。如果$f_{\mathrm{PUF}}(x)$在一个点的Walsh循环谱的绝对值过小，必然导致在其它点的Walsh循环谱的绝对值偏大。文献[8]证明了f_{PUF}至多为$n-2$阶相关免疫函数，即 $f_{\mathrm{PUF}}(c)$ 至多与$n-2$个变量$c_{i_1},c_{i_2},\cdots,c_{i_{n-2}}\in V_n$统计独立，对于$\forall w\in F(i_1i_2\cdots i_{n-2}),w\neq0$，有

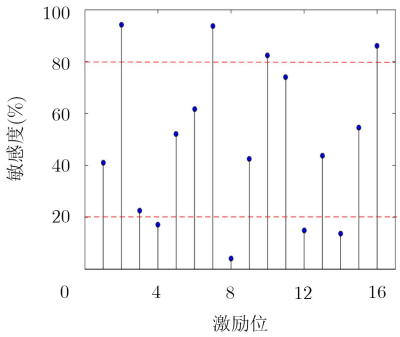
$W_{(f)}(w)=0$。结合能量守恒定理，PUF中不同激励位对生成的响应的影响不同。

2.2 敏感度混淆模型

根据2.1小节结论，从学习算法攻击的角度出发，PUF的不同激励位为预测响应值提供的熵不同。本文将激励位对生成响应的影响程度定义为敏感度(Sensitivity Information, SeI)，取值空间为[0,100%]。仅翻转该激励位比特值时，所产生的PUF响应翻转概率越大，则该激励位的敏感度越高。图2(a)为对激励位宽为64 bit的仲裁器PUF[9]进行敏感度测量的结果。图2(b)为具有512个环形振荡器的ROPUF[10](激励位宽为16 bit)的敏感度测量结果。



(a) 冲裁气PUF



(b) ROPUF

图 2 敏感度测量结果

考虑只选择了敏感度过低的激励位的极端情况，所插入混淆值将不会影响学习算法对PUF的预测结果。因此，基于敏感度设计合理的激励位选择方案对优化混淆策略具有重要作用。根据敏感度大小将激励位大致划分为3个区域：活跃区(Active area, Aa)，过渡区(Transition area, Ta)，寂静区(Silent area, Sa)，区域之间的敏感度大小为

$$\mathrm{SeI}.A > \mathrm{SeI}.T > \mathrm{SeI}.S \tag{6}$$

假设混淆值位宽N_{OB} bit, $N_{\mathrm{OB}} \le n$，所选择的激励位为$c_{i_1}, c_{i_2}, \cdots, c_{i_{N_{\mathrm{OB}}}}$。假设其它激励位保持不变，混淆值翻转而导致PUF响应的翻转为独立事件，所以插入$2k$ bit的混淆值可以看作$2k$个伯努利实验[11]。偶数个翻转事件发生将导致PUF响应不翻转，因此，PUF响应翻转的必要条件为奇数个翻转事件同时发生。从$2k$个伯努利实验中任意选择奇数个翻转事件，所有可能的组合数为$N = C_{N_{\mathrm{OB}}}^{2m-1}$, $m = \{1, 2, \cdots, \lfloor N_{\mathrm{OB}}/2 \rfloor\}$。PUF响应翻转概率的期望为

$$E_{\mathrm{Pr}} = \frac{1}{\lfloor N_{\mathrm{OB}}/2 \rfloor} \cdot \sum_{m=1}^{\lfloor N_{\mathrm{OB}}/2 \rfloor} \left(\frac{1}{N} \cdot \sum_{a=1}^{N} \left(\prod_{l=1}^{2m-1} \mathrm{SeI}\left(c_{j_l}\right) \cdot \prod_{i_d \ne j_l}^{N_{\mathrm{OB}}} \left(1 - \mathrm{SeI}\left(c_{i_d}\right)\right) \right) \right) \tag{7}$$

2.3 模型分析

混淆值位宽不宜过长，否则将导致生成随机混淆值的资源开销过大。以$N_{\mathrm{OB}} = 1, 2, 3, 4$为例，混淆值插入位置所有可能的选择情况如表1所示。

表 1 $N_{\mathrm{OB}} = 1, 2, 3, 4$时混淆值位置选择所有组合情况

N_{OB}	(Aa,Ta,Sa)
1	(1,0,0);(0,1,0);(0,0,1)
2	(1,1,0);(1,0,1);(0,1,1);(2,0,0);(0,2,0);(0,0,2)
3	(1,1,1);(2,1,0);(2,0,1);(1,2,0);(0,2,1);(1,0,2);(0,1,2); (3,0,0);(0,3,0);(0,0,3)
4	(2,1,1);(1,2,1);(1,1,2);(2,2,0);(2,0,2);(0,2,2);(3,1,0);(3,0,1); (1,3,0);(0,3,1);(1,0,3);(0,1,3);(4,0,0);(0,4,0);(0,0,4)

表1中3元组中的数字代表选择对应区域的混淆值位数n_{ob}, $n_{\mathrm{ob}}=0$表示不在该区域选择，$n_{\mathrm{ob}}=1$表示单区选择，$n_{\mathrm{ob}} \ge 2$表示同区选择，两个或两个以上区域中$n_{\mathrm{ob}} \ne 0$表示异区选择。

(1) $N_{\mathrm{OB}} = 1$时只存在单区选择。根据式(7)可得出3种组合对应的E_{Pr}分别为$\mathrm{SeI}.A\left(c_i\right)$, $\mathrm{SeI}.T\left(c_i\right)$和$\mathrm{SeI}.S\left(c_i\right)$。结合式(6)可知

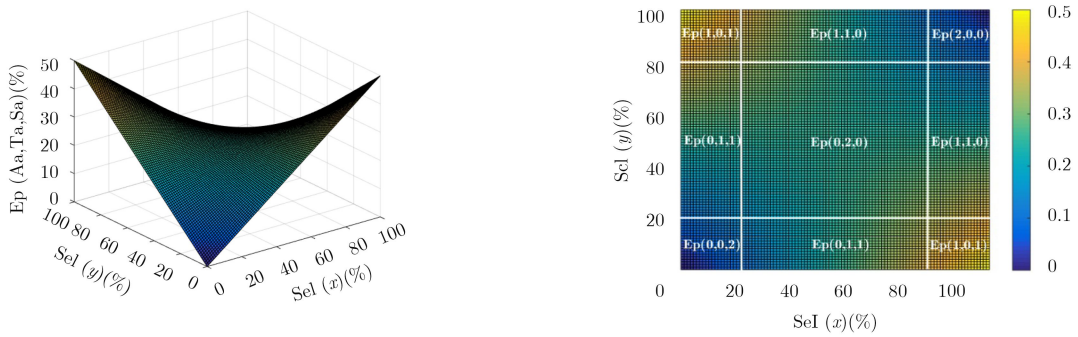
$$\mathrm{SeI}.A\left(c_i\right) > \mathrm{SeI}.T\left(c_i\right) > \mathrm{SeI}.S\left(c_i\right) \tag{8}$$

(2) $N_{\mathrm{OB}} = 2$时存在异区选择和同区选择。根据式(7)计算$N_{\mathrm{OB}} = 2$时的PUF响应翻转概率的期望为

$$E_{\mathrm{Pr}}\left(\mathrm{Aa}, \mathrm{Ta}, \mathrm{Sa}\right) = \frac{\mathrm{SeI}\left(c_{i_1}\right) \cdot \left(1 - \mathrm{SeI}\left(c_{i_2}\right)\right) + \left(1 - \mathrm{SeI}\left(c_{i_1}\right)\right) \cdot \mathrm{SeI}\left(c_{i_2}\right)}{2} \tag{9}$$

图3为根据式(9)所画的概率期望$E_{\mathrm{Pr}}\left(\mathrm{Aa}, \mathrm{Ta}, \mathrm{Sa}\right)$的3维空间及平面分布图。

(3) $N_{\mathrm{OB}} = 3$时还存在异区和同区选择同时出现的情况(如$(2,1,0)$)。根据式(7)计算该PUF响应翻转概率的期望为



(a) 3维空间图 (b) 平面分布图

图 3 $N_{\mathrm{OB}}=2$时概率期望$E_{\mathrm{Pr}}(\mathrm{Aa},\mathrm{Ta},\mathrm{Sa})$的分布图

$$\begin{aligned}E_{\mathrm{Pr}}(\mathrm{Aa},\mathrm{Ta},\mathrm{Sa})&=\frac{1}{2}\cdot\left[\frac{1}{\mathrm{C}_3^1}\cdot\sum_{a=1}^{C_3^1}\left(\mathrm{SeI}(c_{i_a})\cdot\prod_{b\neq a}(1-\mathrm{SeI}(c_{i_b}))\right)+\frac{1}{\mathrm{C}_3^3}\cdot\mathrm{SeI}(c_{i_1})\cdot\mathrm{SeI}(c_{i_2})\cdot\mathrm{SeI}(c_{i_3})\right]\\&=\frac{1}{6}\cdot(\mathrm{SeI}(c_{i_1})+\mathrm{SeI}(c_{i_2})+\mathrm{SeI}(c_{i_3}))+\mathrm{SeI}(c_{i_1})\cdot\mathrm{SeI}(c_{i_2})\cdot\mathrm{SeI}(c_{i_3})\\&\quad-\frac{1}{3}\cdot(\mathrm{SeI}(c_{i_1})\cdot\mathrm{SeI}(c_{i_2})+\mathrm{SeI}(c_{i_2})\cdot\mathrm{SeI}(c_{i_3})+\mathrm{SeI}(c_{i_1})\cdot\mathrm{SeI}(c_{i_3}))\end{aligned}\tag{10}$$

根据式(10)的对偶性，令任意一个$\mathrm{SeI}(c_{i_a})$取Aa, Ta和Sa区域中心值(90%, 50%和10%)，分别画出概率期望$E_{\mathrm{Pr}}(\mathrm{Aa},\mathrm{Ta},\mathrm{Sa})$的空间及平面分布，如图4所示。

(4)$N_{\mathrm{OB}}=4$时的选择情况与(3)相比，还存在两个区域都出现同区选择的情况(如$(2,2,0)$)。根据式(7)计算此时PUF响应翻转概率的期望为

$$\begin{aligned}E_{\mathrm{Pr}}(\mathrm{Aa},\mathrm{Ta},\mathrm{Sa})&=\frac{1}{2}\cdot\left[\frac{1}{\mathrm{C}_3^1}\cdot\sum_{a=1}^{C_4^1}\left(\mathrm{SeI}(c_{i_a})\cdot\prod_{b\neq a}(1-\mathrm{SeI}(c_{i_b}))\right)+\frac{1}{\mathrm{C}_4^3}\cdot\sum_{l=1}^{C_4^3}\left(\mathrm{SeI}(c_{i_l})\cdot\prod_{d\neq l}(1-\mathrm{SeI}(c_{i_d}))\right)\right]\\&=\frac{1}{8}\cdot(1-\mathrm{SeI}(c_{i_1})-\mathrm{SeI}(c_{i_2})+2\cdot\mathrm{SeI}(c_{i_1})\cdot\mathrm{SeI}(c_{i_2}))\\&\quad\cdot(\mathrm{SeI}(c_{i_3})+\mathrm{SeI}(c_{i_4})-2\cdot\mathrm{SeI}(c_{i_3})\cdot\mathrm{SeI}(c_{i_4}))\\&\quad+\frac{1}{8}\cdot(1-\mathrm{SeI}(c_{i_3})-\mathrm{SeI}(c_{i_4})+2\cdot\mathrm{SeI}(c_{i_3})\cdot\mathrm{SeI}(c_{i_4}))\\&\quad\cdot(\mathrm{SeI}(c_{i_1})+\mathrm{SeI}(c_{i_2})-2\cdot\mathrm{SeI}(c_{i_1})\cdot\mathrm{SeI}(c_{i_2}))\end{aligned}\tag{11}$$

令$g(x,y)=\mathrm{SeI}(c_{i_x})+\mathrm{SeI}(c_{i_y})-2\cdot\mathrm{SeI}(c_{i_x})\cdot\mathrm{SeI}(c_{i_y})$, $(x,y)\in\{(1,2),(3,4)\}$且$0\le g(x,y)\le 1$代入式(11)得

$$\begin{aligned}E_{\mathrm{Pr}}(\mathrm{Aa},\mathrm{Ta},\mathrm{Sa})&=\frac{1}{8}\cdot(g(1,2)+g(3,4)\\&\quad-2\cdot g(1,2)\cdot g(3,4))\end{aligned}\tag{12}$$

根据式(12)的对偶性以及$g(x,y)$函数“鞍”形图(见图3)的中心对称性可知

$$\begin{cases}g(1,2)\to 0\\g(3,4)\to 0\end{cases}\text{或}\begin{cases}g(1,2)\to 1\\g(3,4)\to 1\end{cases}\Rightarrow E_{\mathrm{Pr}}(\mathrm{Aa},\mathrm{Ta},\mathrm{Sa})\to 0\tag{13}$$

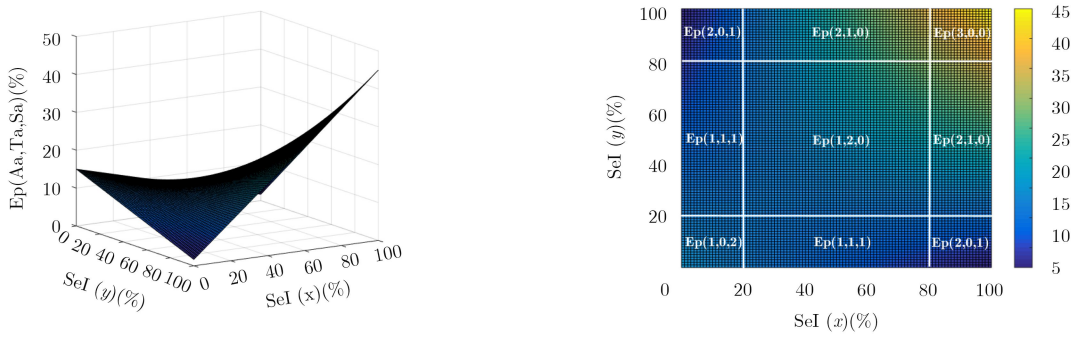
$$\begin{cases}g(1,2)\to 1\\g(3,4)\to 0\end{cases}\text{或}\begin{cases}g(1,2)\to 0\\g(3,4)\to 1\end{cases}\Rightarrow E_{\mathrm{Pr}}(\mathrm{Aa},\mathrm{Ta},\mathrm{Sa})\to 1\tag{14}$$

式(13)和式(14)中“$\to$”表示趋近于右侧数值，解式(13)得$(\mathrm{Aa},\mathrm{Ta},\mathrm{Sa})=(4,0,0)$, $(0,0,4)$或$(2,0,2)$，解式(14)得$(\mathrm{Aa},\mathrm{Ta},\mathrm{Sa})=(3,0,1)$或$(1,0,3)$。

$N_{\mathrm{OB}}>4$时情况的分析方法与$N_{\mathrm{OB}}=3$和$N_{\mathrm{OB}}=4$相似，这里不再赘述。在不考虑剩余激励位的平均敏感度时，PUF响应翻转概率的期望值越大，则混淆值插入方案对建模攻击的混淆程度越高。综合上述分析，可归纳出基于混淆值位宽奇偶性的位置选取规则。

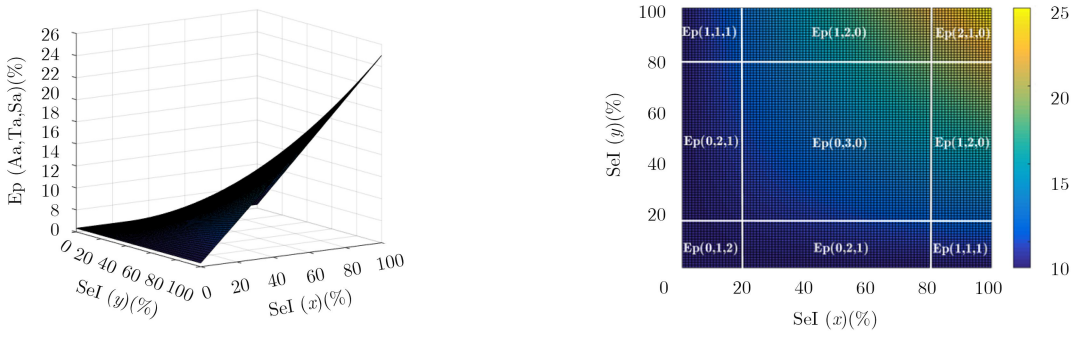
偶数规则Rule_E：混淆值位宽为偶数时，异区选择比同区选择的混淆程度更高，所选激励位中同时包含Aa区(高敏感度)和Ta区(低敏感度)时混淆程度最高，但需要保证两个区域的混淆值数量都为奇数。

奇数规则Rule_O：混淆值位宽为奇数时，选择Aa区的混淆值数量越多混淆程度越高，即尽可



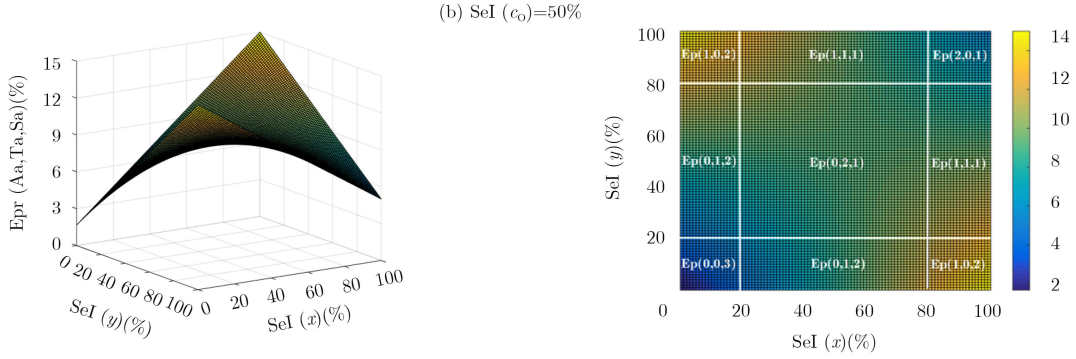
(a$_1$) 3维空间图　　(a$_2$) 平面分布图

(a) SeI (c_O)=90%



(b$_1$) 3维空间图　　(b$_2$) 平面分布图

(b) SeI (c_O)=50%



(c$_1$) 3维空间图　　(c$_2$) 平面分布图

(c) SeI (c_O)=10%

图 4 $N_{OB}=3$时概率期望$E_{Pr}(\mathrm{Aa},\mathrm{Ta},\mathrm{Sa})$的分布图

能多的选择敏感度高的激励位作为混淆值的插入位置，该结论对于异区选择和同区选择都适用。

3　基于敏感度混淆机制的控制型PUF架构设计

所归纳的位置选取规则能够增强对激励响应映射关系的混淆程度，依赖于该规则的防护机制称为敏感度混淆机制。敏感度混淆机制采用多位宽缺失策略，即混淆值位宽可在一定整数区间内选择，根据Rule _E和Rule_O选择相应的插入位置。假设混淆值位宽为$x=\{1,2,\cdots,N_{\max}\}$，随机混淆机制和敏感度混淆机制能够产生的最大激励响应对(Challenge-Response Pairs, CRPs)数量分别为$2^{n-N_{\max}}$和$\sum_{i=1}^{N_{\max}}2^{n-i}$，故敏感度混淆机制能够产生更大的CRPs空间。为了便于区分，将插入混淆值前的激励称为混淆激励C_{OB}，将插入之后的激励称为完整激励C_f。在多位宽缺失策略中，C_{OB}和C_f之间不是一一映射，且C_f集合的元素都有多个原象与之对应，则基于敏感度混淆机制的控制型PUF的CRPs空间为非均匀分布，故针对该PUF的学习算法所建立的模型还可能过拟合[12]。

图5为基于敏感度混淆机制的控制型PUF架构。控制逻辑采用敏感度混淆机制来隐藏激励响应的映射关系，提高学习算法成功破解PUF的计算复杂度和预测错误率等参数，从而增加交付最终模型

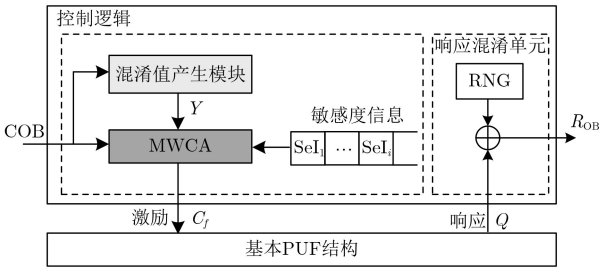


图 5 基于敏感度混淆机制的控制型PUF架构

的成本[13]。本文设计的控制逻辑主要包括混淆值产生模块、MWCA算法模块和响应混淆单元。根据多位宽缺失机制，混淆激励$C_{\rm OB}$的位宽可以取不同值，从而增大CRPs空间。混淆值产生模块能够生成多种位宽的随机混淆值Y。MSCA算法模块根据输入$C_{\rm OB}$和混淆值Y，调用敏感度信息，输出完整激励C_f。响应混淆单元通过随机数对PUF响应Q进行异或掩码，产生混淆响应$R_{\rm OB}$。

3.1 混淆值产生模块

由于PUF在响应生成过程中具有鲁棒性，完整激励C_f需要能够重复生成，混淆值产生模块同样应该具有鲁棒性。最优解决方法为基于APUF设计混淆值产生模块，如图6所示。混淆激励$c_{\rm OB}$=$\{c_1, c_2, \cdots, c_t\}$将传输到$N-t$个并行的APUF，以产生$(N-t)$ bit的混淆值。时序脉冲发生器(Timing Pulse Generator, TPG)用于产生APUF所需的时序脉冲。

本文在传统设计的APUF结构上进行了改进，在延迟开关链的末端增加了$(N-t-1)$个仲裁器单元$\rm Arb_i$，其输出连接到同一个数据选择器上，根据所需的混淆值位宽配置数据选择器的索引值Adr，其位宽取$\lfloor \lg(N-t) \rfloor$ bit，相应仲裁器单元的输出即为混淆值。例如，完整激励位宽为16 bit，混淆激励最小位宽为12 bit，混淆值最大位宽为4 bit，需要并行4个APUF，每个延迟开关链末端连接3个仲裁器单元，索引值位宽为2 bit。根据混淆值位宽从小到大变化，索引值可取$(00, 01, 10, 11)$，并译码为使能信号EN = $(1000, 1100, 1110, 1111)$，使能信号为高有效。当混淆激励位宽取值变化时，基于APUF的混淆值产生模块能够稳定生成相应长度的随机混淆值，满足多位宽缺失机制和PUF鲁棒性的需求。

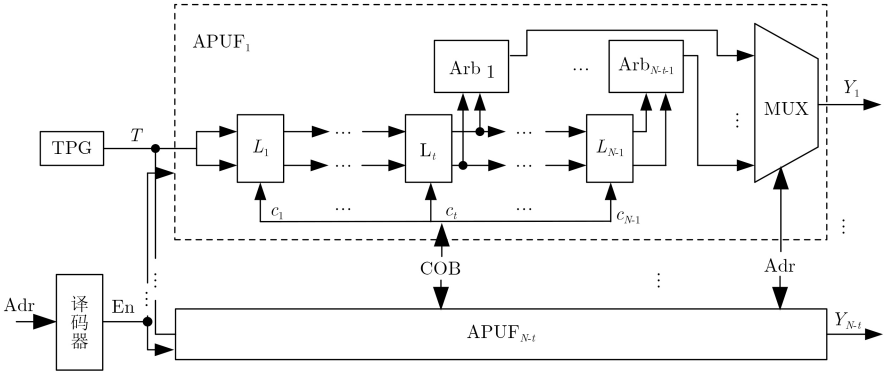


图 6 基于APUF的混淆值产生模块

3.2 多位宽混淆算法MWCA

本文设计了多位宽混淆算法(Multi-bit Wide Confusion Algorithm, MWCA)来产生完整激励C_f，具体步骤如表2所示。该算法的主要思想是：给每个混淆激励$c_{\rm OB}^k$绑定标识值a_k(初始值为0)，每次输入混淆激励$c_{\rm OB}^k$时，若标识值a_k等于1，则判断该混淆激励为重复输入，调用相同的位置选择策略P_k来产生完整激励C_k；否则，根据$c_{\rm OB}^k$位宽奇偶性调用规则Rule_E或Rule_O，生成相应的位置选择策略P_k, a_k加1；每轮运算结束时对a_k和P_k进行存储。

4 实验评估

为了验证本文的设计，将基础PUF结构(如仲裁器PUF[9], ROPUF[10])作为控制逻辑的保护对象，在Xilinx Spartan-6 FPGA上实现了基于敏感度混淆机制的控制型PUF。考虑FPGA的资源开销问题，将敏感度混淆机制中的混淆值最大位宽$N_{\max}$设置为4 bit。下面对控制型PUF的唯一性、可靠性和随机性等性能指标进行评估。

4.1 可靠性和唯一性评估

唯一性通过测量多个PUF设备的片间汉明距离P_s来评估，可靠性通过重复测量相同激励作用下PUF生成响应之间的片内汉明距离P_d来评估。理想情况下，PUF的唯一性和可靠性分别为50%和100%。下面以ROPUF为例来评估控制型PUF的功能完整性。控制型PUF响应的汉明距离分布如图7所示，其唯一性和可靠性分别为51.24%和98.44%。

将本文设计的控制型PUF与基本ROPUF[10],

表 2　MWCA的具体算法

输入：混淆激励集合$C_{\mathrm{OB}}=\left\{c_{\mathrm{OB}}^{k}|k\in\left\{1,2,\cdots,\sum_{i=1}^{t}2^{N-i}\right\}\right\}$；

混淆值集合$Y=\left\{y_k|k\in\left\{1,2,\cdots,\sum_{i=1}^{t}2^{N-i}\right\}\right\}$

过程：

(1) for $\forall k\in\left\{1,2,\cdots,\sum_{i=1}^{t}2^{N-i}\right\}$ do

(2) 输入混淆激励集合$c_{\mathrm{OB}}^{k}=\{c_1,c_2,\cdots,c_t\}$、混淆值$y_k$和标识值$a_k$;

(3) if $a_k=1$ then

(4) 调用位置选择策略P_k;产生完整激励C_k=$P_k(c_{\mathrm{OB}}^{k},y_k)$;return

(5) else

(6) 计算$r=(N-t)\%2$;

(7) end if

(8) if $r=0$ then

(9) 调用偶数规则Rule_E,生成位置选择策略P_k, $a_k=1$;

(10) else

(11) 调用奇数规则Rule_O,生成位置选择策略P_k, $a_k=1$;

(12) end if

(13) 存储标识值a_k和策略P_k,用于之后激励混淆步骤的判断依据;

(14) 按照策略将混淆值插入混淆激励c_{OB}^{k},产生完整激励C_k=$P_k(c_{\mathrm{OB}}^{k},y_k)$;return

(15) end for

输出：完整激励集合$C_f=\{C_k|k\in\{1,2,\cdots,\sum_{i=1}^{t}2^{N-i}\}\}$。

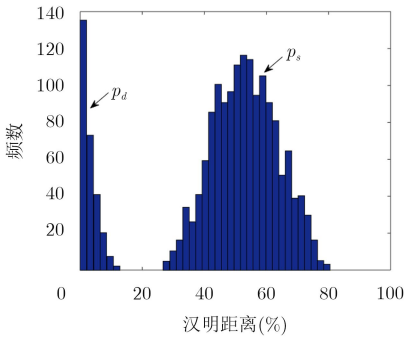


图 7 控制型PUF响应的汉明距离频率分布图

LR-PUF[5], OB-PUF[6]的平均汉明距离进行对比，如表3所示。控制型PUF的P_s值与理想值50%间的距离略大于OB-PUF，略小于基本ROPUF和LR-PUF；控制型PUF的P_d值与理想值0之间的距离略小于基本ROPUF和LR-PUF，而略大于OB-PUF。对比结果表明，本文设计的控制型PUF在唯一性和可靠性上与现有PUF结构基本保持相同水平。

4.2　NIST随机数测试

为了评估随机性指标，采用了NIST (National Institute of Standards and Technology) 随机数测试程序[14]对PUF产生的响应进行测试，测试结果如表4所示，当P_value值高于0.0001时，表示测试通过。测试结果表明，基本ROPUF仅通过了块内频数、游程和序列检验测试项目，本文的控制型PUF通过了表中所列的全部随机数测试项目。由此可知，基于敏感度混淆机制的控制型PUF相比基本ROPUF设计具有更好的随机性，在实际应用中能够获得更高的安全性。

表 3　可靠性和唯一性指标对比结果

PUF结构	基本ROPUF[10]	LR-PUF[5]	OB-PUF[6]	本文
P_s(%)	49.97	≈50	47.31	51.24
P_d(%)	1.59	2	0.86	1.56

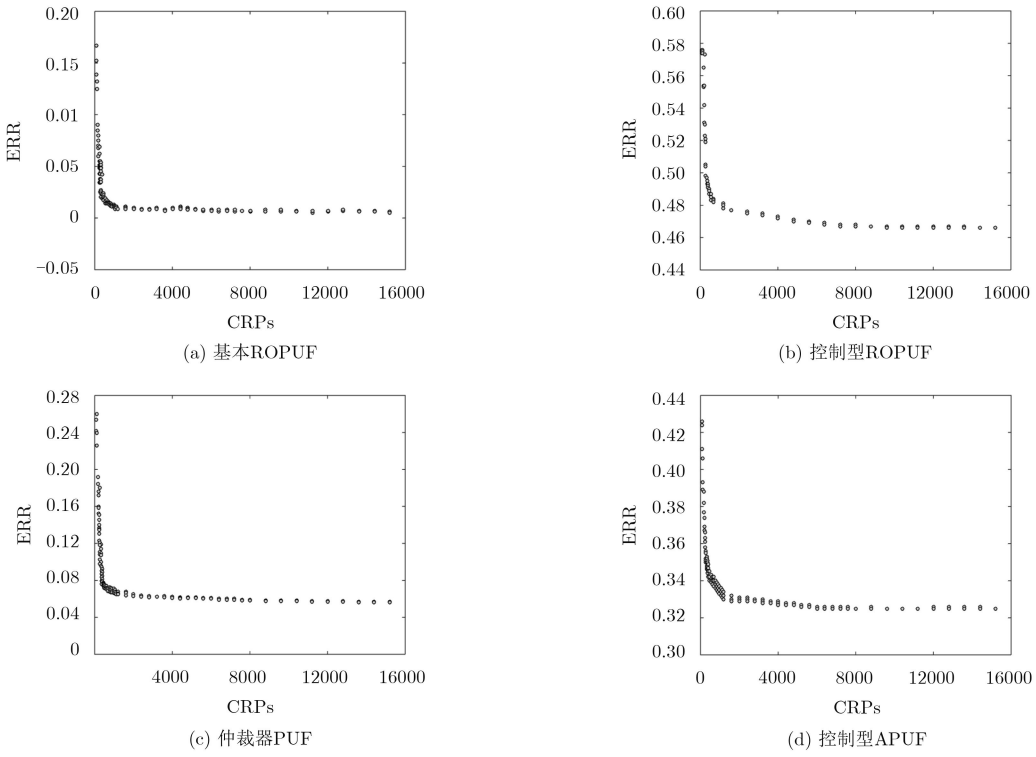
表 4　NIST随机数测试结果

测试项目	基本ROPUF[10]		本文	
	P_value	结果	P_value	结果
频率检验	0.000003	失败	0.829896	通过
块内频数检验	0.050764	通过	0.580358	通过
向前累加和检验	0.000000	失败	0.502594	通过
向后累加和检验	0.000000	失败	0.347179	通过
游程检验	0.302788	通过	0.329404	通过
块内最长游程检验	0.000062	失败	0.024892	通过
近似熵检验	0.000001	失败	0.693147	通过
向前序列检验	0.070160	通过	0.024892	通过
向后序列检验	0.192277	通过	0.756264	通过

4.3　安全性评估

为了验证安全性增强效果，本实验采用逻辑回归算法[15]来进行建模攻击。因为建模攻击结果与混淆机制和训练集数量有关，为了更加清楚的查看实验结果，分别对基本ROPUF、仲裁器PUF以及相应的基于敏感度混淆机制的控制型PUF(控制型ROPUF和控制型APUF)进行攻击，每次选取不同数量的训练集。为了保证结果的准确性，采用了10折交叉验证法。最终结果如图8所示。其中横坐标是训练集的数量(CRPs)，纵坐标是对测试数据的预测误差率(ERR)。

从图8中可以清楚地看出，随着训练数据的增多，预测误差率逐渐降低，当训练数据达到一定值时，继续增大其数量，预测误差率将稳定不变。不同PUF结构被成功攻击所需的训练集数量、攻击时间以及预测成功率如表5所示。OB-PUF采用的是随机混淆机制，OB-PUF(i)表示在激励中随机插入i比特混淆值的OB-PUF。由表5可知，随着i值的增大，攻击OB-PUF所需要的训练数据和攻击时间会增加，而预测成功率会降低。相比随机混淆机制，



(a) 基本ROPUF (b) 控制型ROPUF

(c) 仲裁器PUF (d) 控制型APUF

图 8 不同PUF结构的建模攻击结果

表 5 不同PUF结构的攻击结果对比

PUF类型	训练集数量	攻击时间(s)	预测成功率(%)
OB-PUF(1)	4400	26	87.6
OB-PUF(2)	5200	30	84.4
OB-PUF(3)	6000	62	73.1
OB-PUF(4)	8000	128	69.8
基本ROPUF	5200	8	99.5
控制型ROPUF	7200	262	53.4
仲裁器PUF	4800	15	93.9
控制型APUF	6500	187	67.2

采用敏感度混淆机制对预测成功率的降低效果明显。相比基本ROPUF和仲裁器PUF，逻辑回归算法对控制型PUF的预测成功率分别降低了46.1%和26.7%。因此，敏感度混淆机制能够显著提高PUF的抗建模攻击能力，使其更安全的应用于硬件防护领域。

5 结束语

针对现有的PUF结构抗建模攻击能力弱的问题，本文从激励响应之间的相关性出发，研究了在激励位中插入混淆值对响应产生的影响，利用布尔函数的Walsh谱理论得出不同激励位具有不同敏感度的结论，进一步归纳出混淆位置选择策略的奇偶规则，并指导了基于敏感度混淆机制的控制型PUF架构的设计。

实验评估了控制型PUF的唯一性和可靠性，分别为51.24%和98.44%，与现有PUF结构基本保持相同水平。通过NIST随机数测试程序对控制型PUF和基本ROPUF进行测试对比，结果表明控制型PUF具有更高的随机性。采用逻辑回归算法对不同PUF结构进行建模攻击，结果表明，基于敏感度混淆机制的控制型PUF架构能够显著提高PUF的抗建模攻击能力。

参 考 文 献

[1] RÜHRMAIR U, SÖLTER J, SEHNKE F, *et al*. PUF modeling attacks on simulated and silicon data[J]. *IEEE Transactions on Information Forensics and Security*, 2013, 8(11): 1876–1891. doi: 10.1109/TIFS.2013.2279798.

[2] GASSEND B, VAN DIJK M, CLARKE D, *et al*. Controlled physical random functions[J]. *ACM Transactions on Information and System Security*, 2008, 10(4): 23–25.

[3] KATZENBEISSER S, KOÇABAS Ü, VAN DER LEEST V, *et al*. Recyclable PUFs: Logically Reconfigurable PUFs[M]. Berlin, Germamy: Springer, 2011: 374–389.

[4] LAO Yingjie and PARHI K K. Reconfigurable architectures for silicon physical unclonable functions[C]. Proceedings of 2011 IEEE International Conference on Electro/Information Technology, Mankato, USA, 2011: 1–7.

[5] MAJZOOBI M, KOUSHANFAR F, and POTKONJAK M. Techniques for design and implementation of secure reconfigurable PUFs[J]. *ACM Transactions on*

Reconfigurable Technology and Systems, 2009, 2(1): 5.

[6] GAO Yansong, AL-SARAWI S F, ABBOTT D, *et al*. Modeling attack resilient reconfigurable latent obfuscation technique for PUF based lightweight authentication[J]. arXiv:1706.06232, 2017.

[7] 许道云, 韦立, 王晓峰. 布尔函数的学习与性质测试[J]. 武汉大学学报: 理学版, 2012, 58(2): 125–134.
XU Daoyun, WEI Li, and WANG Xiaofeng. Learning and testing of properties for Boolean functions[J]. *Journal of Wuhan University*: *Natural Science Edition*, 2012, 58(2): 125–134.

[8] GANJI F, TAJIK S, FÄßLER F, *et al*. Strong machine learning attack against PUFs with no mathematical model[C]. Proceedings of the 18th International Conference on Cryptographic Hardware and Embedded Systems, Santa, USA, 2016: 391–411.

[9] ZALIVAKA S S, PUCHKOV A V, KLYBIK V P, *et al*. Multi-valued arbiters for quality enhancement of PUF responses on FPGA implementation[C]. Proceedings of 201621st Asia and South Pacific Design Automation Conference, Macau, China, 2016: 533–538.

[10] GASSEND B, CLARKE D, VAN DIJK M, *et al*. Silicon physical random functions[C]. Proceedings of the 9th ACM Conference on Computer and Communications Security, USA, 2002: 148–160.

[11] LAO Yingjie and PARHI K K. Statistical analysis of MUX-based physical unclonable functions[J]. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2014, 33(5): 649–662. doi: 10.1109/TCAD.2013.2296525.

[12] 庞子涵, 周强, 高文超, 等. FPGA物理不可克隆函数及其实现技术[J]. 计算机辅助设计与图形学学报, 2017, 29(9): 1590–1603. doi: 10.3969/j.issn.1003-9775.2017.09.002.
PANG Zihan, ZHOU Qiang, GAO Wenchao, *et al*. Hardware implementation of physical unclonable function on FPGAs[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2017, 29(9): 1590–1603. doi: 10.3969/j.issn.1003-9775.2017.09.002.

[13] DENG R, WENG Jian, REN Kui, *et al*. Security and Privacy in Communication Networks[M]. Cham: Springer, 2016: 675–693.

[14] KODÝTEK F and LÓRENCZ R. Proposal and properties of ring oscillator-based PUF on FPGA[J]. *Journal of Circuits*, *Systems and Computers*, 2016, 25(3): 1640016. doi: 10.1142/S0218126616400168.

[15] KODÝTEK F, LÓRENCZ R, and BUČEK J. Improved ring oscillator PUF on FPGA and its properties[J]. *Microprocessors and Microsystems*, 2016, 47: 55–63. doi: 10.1016/j.micpro.2016.02.005.

徐金甫：男，1965年生，教授，硕士生导师，研究方向为专用集成电路设计技术.
吴　缙：男，1994年生，硕士生，研究方向为专用集成电路设计技术.
李军伟：男，1988年生，讲师，研究方向为安全芯片设计.
曲彤洲：男，1994年生，硕士生，研究方向为可重构计算与信息安全.
董永兴：男，1994年生，硕士生，研究方向为专用集成电路设计技术.