基于状态映射的 AES 算法硬件混淆设计

张跃军[①] 潘 钊[①] 汪鹏君*[①②] 丁代鲁[①] 李 刚[①]
[①](宁波大学电路与系统研究所 宁波 315211)
[②](温州大学物理与电子信息工程学院 温州 325035)

摘 要：代码混淆利用系统自身逻辑来保护内部重要信息和关键算法，常用于软件代码的安全防护，确保开发者和用户的利益。如何在硬件电路上实现混淆、保护硬件 IP 核的知识产权，也是亟待解决的问题。该文通过对硬件混淆和 AES 算法的研究，提出一种基于状态映射的 AES 算法硬件混淆方案。该方案首先利用冗余和黑洞两种状态相结合的状态映射方式，实现有限状态机的混淆；然后，采用比特翻转的方法，实现组合逻辑电路的混淆；最后，在 SMIC 65 nm CMOS 工艺下设计基于状态映射的 AES 算法硬件混淆电路，并采用 Toggle、数据相关性和代码覆盖率等评价硬件混淆的效率和有效性。实验结果表明，基于状态映射的 AES 算法硬件混淆电路面积和功耗分别增加 9%和 16%，代码覆盖率达到 93%以上。

关键词：状态映射；代码混淆；AES 算法；逆向工程；IP 核安全
中图分类号： TP331 **文献标识码**： A **文章编号**：1009-5896(2018)03-0750-08
DOI: 10.11999/JEIT170556

Design of Hardware Obfuscation AES Based on State Deflection Strategy

ZHANG Yuejun[①] PAN Zhao[①] WANG Pengjun[①②] DING Dailu[①] LI Gang[①]
[①](*Institute of Circuits and Systems*, *Ningbo University*, *Ningbo* 315211, *China*)
[②](*College of Physics and Electronic Information Engineering*, *Wenzhou University*, *Wenzhou* 325035, *China*)

Abstract: Obfuscation is used to safeguard lawful rights and interests of developers and users in software security, by protecting critical information and algorithms with the system logic relation. Also, how to achieve obfuscation method to protect the hardware IP core is becoming an urgent problem. In this paper, a hardware obfuscation scheme based on deflection strategy is proposed by studying the obfuscation method and the AES algorithm. The deflection strategy with redundancy and black hole states are used to realize the Finite State Machine (FSM) obfuscation, and the bit flip method is used to realize the combinational logic obfuscation. Finally, the proposed hardware obfuscation AES algorithm is designed in SMIC 65 nm CMOS process. The parameters of toggle, data correlation and code coverage are selected to evaluate the efficiency and effectiveness of hardware confusion. Experimental results show that the area and power consumption of the hardware obfuscation AES algorithm is increased by 9% and 16% respectively, and the code coverage rate is over 93%.

Key words: State deflection; Hardware obfuscation; AES algorithm; Reverse engineering; IP security

1 引言

随着超大规模集成电路和信息技术的发展，软/硬件系统已经获得广泛应用，在极大地方便人们生活的同时也为开发者带来合法的经济收益。但是知识产权盗用事件时有发生，如珠海矽微电子科技有限公司侵犯商业秘密案等[1]。如何有效地提高 IP 核的知识产权保护，已经成为安全机构和 IP 核开发商的关注重点。混淆的概念最早在软件代码混淆领域被提出，在软件保护、数字水印等领域有着实际的应用[2]。代码混淆借助程序本身的逻辑来保护内部的重要信息和关键算法，确保系统的开发者和用户的利益不受侵害。代码混淆已经成为当前密码学领域的研究热点。在密码学领域，Barak 等人[3]引入混淆概念并提出虚拟黑盒混淆的形式化定义及安全性要求，对密码函数的安全通用混淆研究具有非常重要

收稿日期：2017-06-09；改回日期：2017-11-15；网络出版：2017-12-12
*通信作者：汪鹏君 wangpengjun@nbu.edu.cn
基金项目：浙江省自然科学基金(LY18F040002)，国家自然科学基金(61404076, 61474068)，浙江省公益项目(2015C31010, 2016C31078)，宁波市自然科学基金(2014A610148, 2015A610107)，王宽诚幸福基金
Foundation Items: The Zhejiang Provincial Natural Science Foundation (LY18F040002), The National Natural Science Foundation of China (61404076, 61474068), The S&T Plan of Zhejiang Provincial Science and Technology Department (2015C 31010, 2016C31078), The Ningbo Natural Science Foundation (2014A610148, 2015A610107), The K. C. Wong Magna Fund in Ningbo University, China

的理论意义。代码混淆研究的不断进步，与随机预言机、全同态加密、零知识证明等密码原语的相互结合，对具体密码函数的安全混淆在云计算、代理计算等领域具有实际的应用价值。文献[3]已经证明在标准定义下不存在通用的安全混淆，因此后续的程序混淆方面的研究工作主要集中在 3 个方面：对具体函数类的混淆实现、混淆的新模型研究以及混淆与其它密码模型的关系研究及应用[4]。但是这些研究主要针对软件系统的安全问题，如使用 C, C++ 和 Java 等编写的程序代码[5,6]，很少涉及硬件 IP 核的混淆保护。

IP 核指通过设计验证且具有特定功能的电路模块，根据存在形式不同可分为软核、固核和硬核。随着 SoC 的集成度越来越大，产品开发周期越来越长。市场驱动以 IP 核为基础进行开发来缩短设计周期、提高产品竞争力，相应地出现专门从事 IP 核设计、开发和营销的公司，大量可复用 IP 核获得有效推广。同时不法分子采用非法传播、反向设计与过度制造等手段谋取利益，侵犯 IP 核开发者和授权用户的权益。现有 IP 核保护方法主要包括《合同法》、《专利法》、《集成电路布图设计保护条例》等知识产权部门法与相关条例，主要从法律和道德上约束；技术上，License 也无法从根本上解决 IP 核盗用的问题。硬件混淆技术可增强 IP 核的安全性，为有效解决 IP 核知识产权提供保障。Alkabani 等人[7]提出远程激活的方式管理数字产权，将集成电路自身特性应用于 FSM 的转换条件，实现远程 IC 控制。文献[8]在电路网表级提出 HARPOON 混淆技术，该方法在面积和功耗开销小于 5%的前提下，可以同时实现混淆和验证。文献[9]提出基于密钥混淆控制和数据流低开销的 RTL 硬件知识产权保护技术，将 RTL 变换成状态控制和数据流图，实现有限状态机的硬件混淆。Zhang 等人[10]提出 PUF-FSM 的硬件混淆，有效保护 FPGA 器件的 IP 核，实现 Pay-Per-Device 的强制付费许可机制。Koushanfar 等人[11]提出采用 IC metering 的方法实现知识产权拥有者控制每个芯片工作状况，避免芯片尾货流入市场。但是上述的硬件混淆方法，只对进入 FSM 有效状态前进行控制，当 FSM 有效状态正常工作后，安全防护机制将失效，且存在寄存器强写攻击和硬件木马攻击等威胁。鉴此，通过对已有安全混淆技术和 AES 算法的研究，提出一种基于状态映射的 AES 硬件混淆算法。该算法采用冗余状态和黑洞状态相结合的方式模糊有效状态；然后，利用状态映射方法实现时序逻辑电路混淆，采用按位取反的方式实现组合逻辑电路混淆；其次，结合代码覆盖率、Toggle 和数据相关性等评估混淆算法的有效性；最后，在 SMIC 65 nm CMOS 工艺下，实现基于状态映射的 AES 硬件混淆算法。

2　攻击模型和状态映射方法

2.1　攻击模型

IC供应链中的攻击模型如图1所示。攻击者的权限是由他在供应链中的地位所决定，由于所掌握的资源不同，导致实施的攻击模式也不同。例如，供应商或设计人员参与IP核的设计过程，可实施IP盗用、硬件木马等方式的攻击；系统集成人员、晶圆代工厂以及芯片测试公司可以访问整体的芯片，可实施逆向工程、过度生产和旁道分析等攻击[12,13]。

现有的电路安全混淆方法通过修改FSM实现

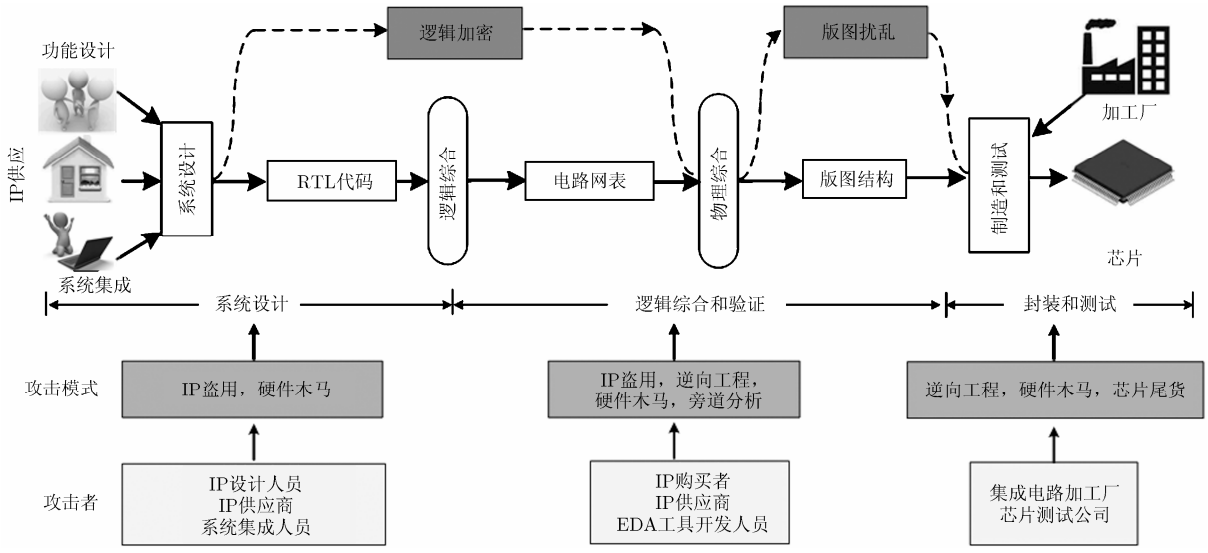
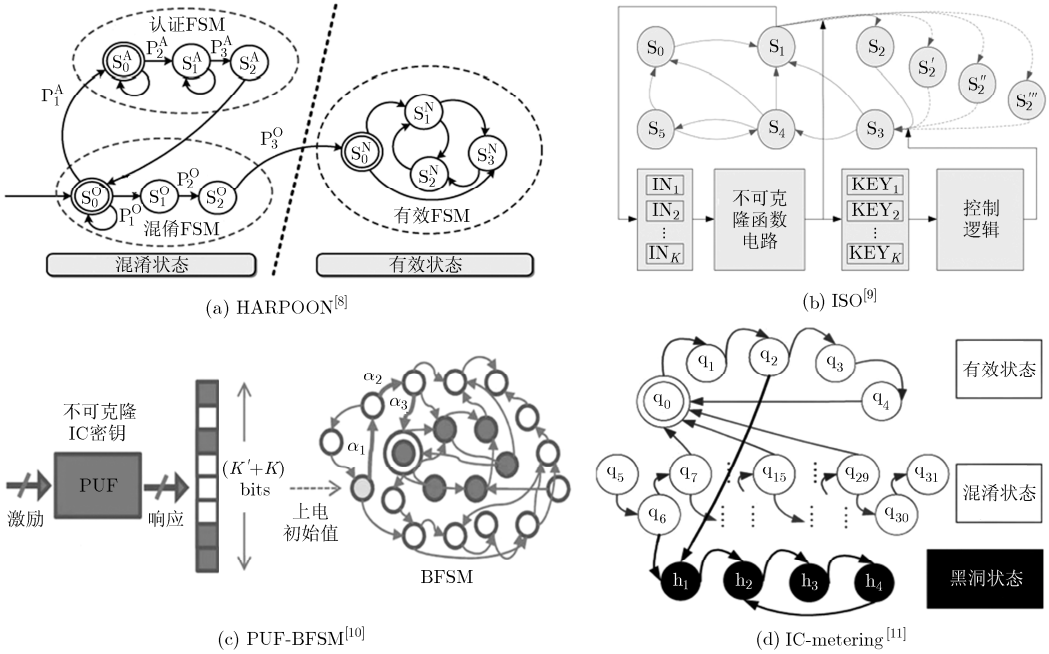


图 1　IC供应链中的攻击模型

IP 核保护，主要方法如图 2 所示[8−11]。上述方法存在以下局限性：(1)安全混淆方法主要针对 SoC 系统整体保护，但是内部 IP 核存在被反向攻击的可能，尤其是借助如 TetraMAX, Formality 以及 ESL 等工具软件。(2)原有防御方法的前提是攻击者只能通过穷举法来查找所有的密钥序列，但是假如攻击者可以接触 IP 软核(即攻击者来自公司内部或者 IP 核设计团队)，就可以改写寄存器的存储状态达到盗用 IP 核的目的。(3)现有的安全混淆 FSM 设计，当电路进入正常工作模式之后，就没有额外的保护措施。

2.2 状态映射方法

针对上述 IP 核存在的威胁，本节提出一种状态映射方法提高安全性。在混淆模糊处理 FSM 状态的基础上，状态映射方法通过增加混淆状态 Sx 和黑洞状态 Bx，实现保护 FSM 有效状态的目的，如图 3 所示。其中，KS 为状态控制密钥，混淆状态和有效状态之间可以互相转换，黑洞状态和有效状态之间不能转换。一旦系统进入黑洞状态，FSM 将不会返回到有效状态。当输入错误密钥 KS 时，FSM 有效状态将跳变到黑洞状态 Bx。为了保护 FSM 所有的有效状态，提出在错误密钥下具有状态映射功能的黑洞状态动态分配方案。为了进一步提高安全性，我们采用黑洞状态的群集方法，所有黑洞状态 Bx 都不相同，且可以跳转到其它黑洞状态。图 3 给出两种状态映射方法的流程图。网表级 IP 核在状态映射方法的保护下，攻击者就无法确认有效状态和混



(a) HARPOON[8] (b) ISO[9]

(c) PUF-BFSM[10] (d) IC-metering[11]

图 2 电路级混淆方法

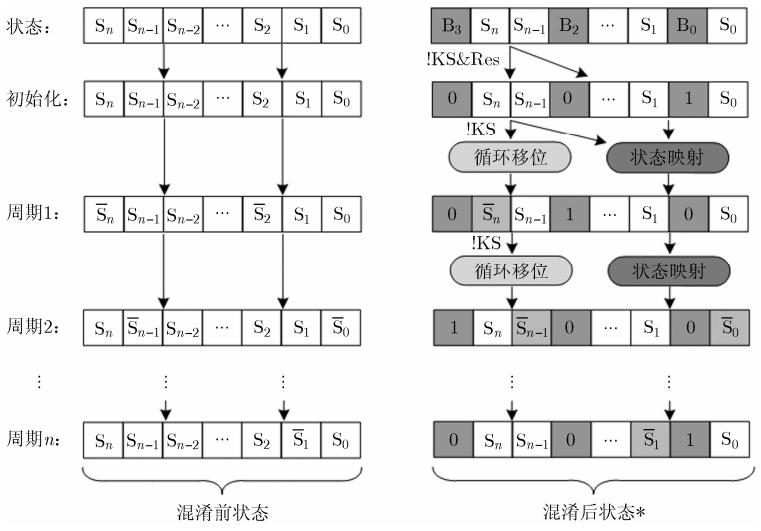


图 3 状态变化流程图

淆状态。为了实现输入密钥KS错误依然保持寄存器翻转的功能，提出状态翻转方法实现输出信号的状态变化，电路方案如图4所示。当Bx为1时候，输出按位翻转，当Bx为0时，输出保持不变。

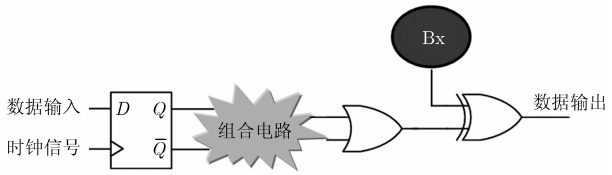


图 4 状态翻转的硬件混淆模型

3　基于状态映射的安全混淆 AES 算法设计

AES 算法已经成为当前广泛应用的加密标准之一[14,15]，根据密钥长度的不同分为 128 位、192 位和 256 位，通常需要完成 12 轮迭代运算，每一轮运算包括以下 4 个步骤：轮密钥加，字节变化，行变化和列混淆。本文将混淆状态嵌入到 AES 算法的有限状态机，并在组合逻辑电路中增加条件控制的按位取反功能，实现 AES 密码算法的安全混淆。基于状态映射的 AES 硬件混淆算法的有限状态机如图 5 所示。安全混淆 AES 算法需要在原有算法的基础上增加 6 个步骤：第 1 步，指定混淆状态和混淆密钥，并定义状态翻转和状态映射；第 2 步，在端口列表中增加位宽的信号；第 3 步，进行密钥序列的判断，并插入状态映射算法；第 4 步，按照循环移位算法切换新添加的状态触发器的内容；第 5 步，如果状态翻转不全为零，则按位翻转有效状态；第 6 步，如果当前的状态属于黑洞状态，则按位翻转输出数据。上述的操作均在 RTL 代码上完成，接下来将详细介绍 AES 算法硬件混淆的加密过程和解密过程。

3.1 安全混淆 AES 加密算法

安全混淆AES加密算法电路结构如图6所示，输入明文和密钥均为128位的D_{in}和Key，输出密文为128位的D_{out}。具体如下：步骤1，初始化第1轮状态和密钥的第1轮Key，判断状态密钥KS，错误进入黑洞状态，正确进入步骤2。步骤2, N_{r-1} 轮中间运算。每一轮的运算过程如下：(1)行变化，对状态的行进行移位操作，第0行不移位，第1行循环左移1位，第2行循环左移2位，第3行循环左移3位；(2)字节变化，通过查找表S-box实现，状态中每个字节根据它的值作为地址进行查表，从而得到一个新的字节作为输出；(3)列混淆，对状态的每列独立地进行操作，每列的每个字节被映射为一个新值，此值由该列中的4个字节通过函数变换得到；(4)轮密钥加，状态的当前行与第 N_{r-1} 轮密钥按位作异或操作；(5)判断状态密钥KS，错误时D_{out}输出取反，正确时输出D_{out}，进行状态的下一行计算，N_{r-1} 加1同时跳转到行变化，直到 N_{r-1} 为12，执行下一步运算。最后一次轮运算，先将状态前3行复制到状态最后3行，然后将状态进行字节变化，将状态的当前行与密钥的最后一轮密钥做异或操作。当KS正确时输出D_{out}，完成整体操作；当KS错误时，将D_{out}按位取反后输出。

3.2 安全混淆 AES 解密算法

安全混淆AES解密过程与加密时相反，从最后一轮开始运算。具体如下：步骤1，初始化最后一轮状态和最后一轮密钥进行异或操作，当KS错误时进入黑洞状态，当KS正确时进入步骤2。步骤2，N_{r-1}

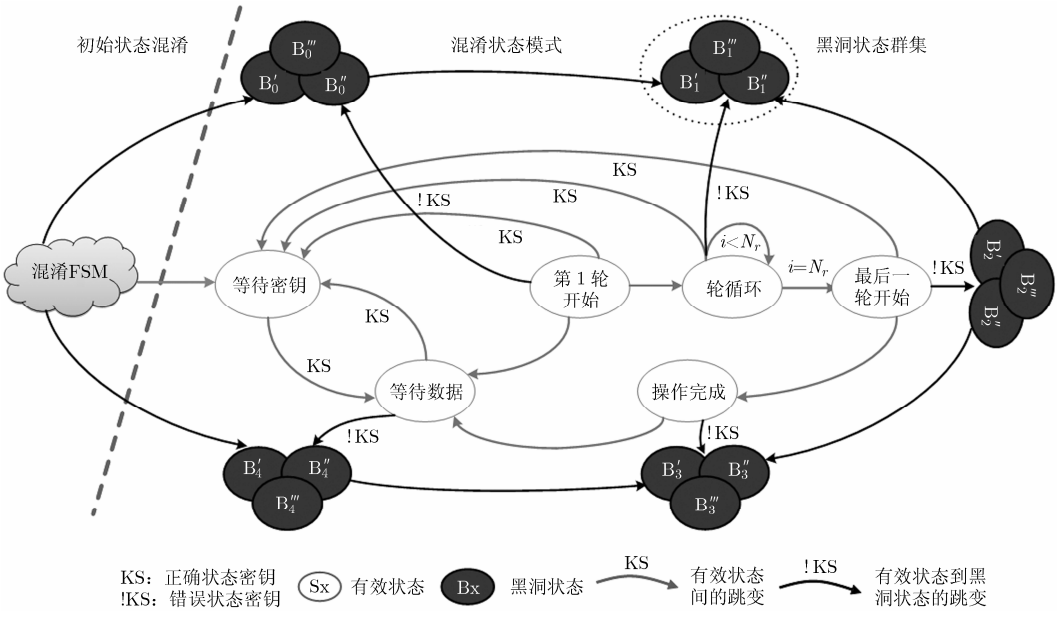


图 5 基于状态映射混淆的AES状态机模型

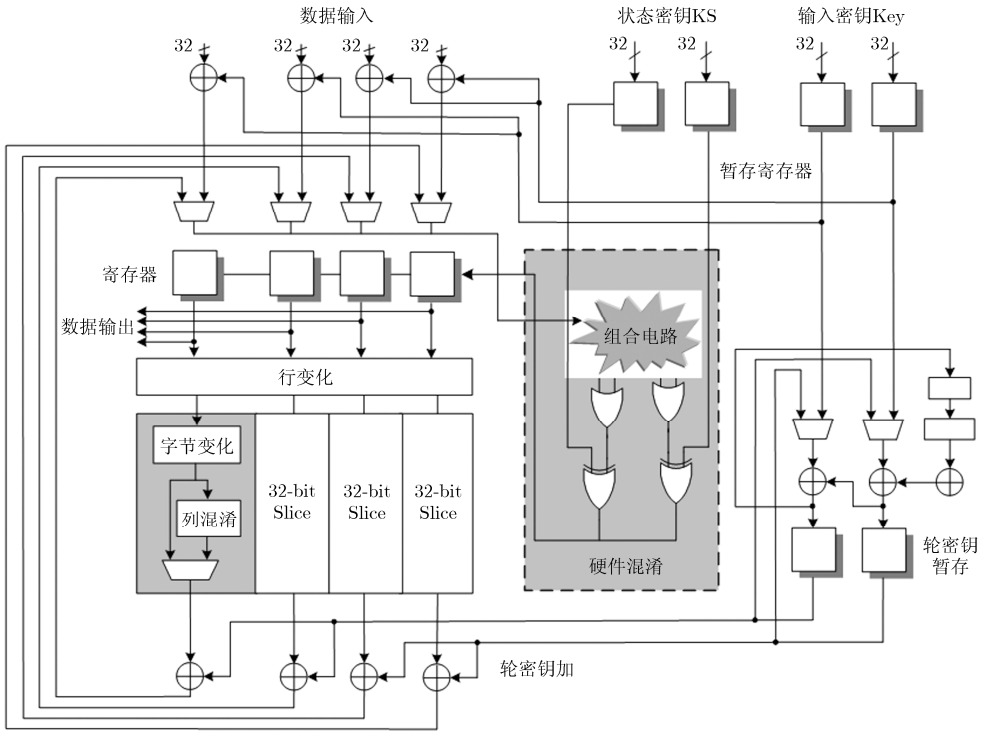


图 6 AES算法安全混淆加密电路结构图

轮中间运算，每一轮的运算过程如下：(1)行变化，对状态的行进行移位操作，第0行不移位，第1行循环左移1位，第2行循环左移2位，第3行循环左移3位；(2)字节变化，通过一个查找表S-box实现，状态中每个字节根据它的值作为地址进行查表，从而得到一个新的字节作为输出；(3)状态的当前行和第n轮密钥进行异或操作；(4)列混淆，对状态的每列独立地进行操作，每列的每个字节被映射为一个新值，此值由该列中的4个字节通过函数变换得到；状态密钥KS错误时，D_{out}输出取反，进行状态的下一行计算。n加1同时跳转到行变化，直到n为12；步骤3，最后一次轮运算，首先将状态的最后3行复制到状态的前3行，然后将状态进行字节变化，最后状态的当前行和第1轮密钥作异或操作。当KS正确时输出D_{out}，完成整体操作；当KS错误时，将D_{out}按位取反后输出。AES算法安全混淆解密电路结构如图7所示。

4 实验结果与分析

在SMIC 65 nm CMOS工艺下设计基于状态映射的AES硬件混淆算法，涉及NClaunch, DC, TetraMAX和Matlab等工具软件。在电路分析中，攻击者可以利用状态是否翻转来区分有效状态和混淆状态。即在正确密钥KS下内部寄存器翻转，错误密钥KS下内部寄存器不翻转，则可判定该寄存器为混淆电路。在此，状态翻转率Toggle可以用来作为混淆指标。在混淆后AES算法网表级电路中，借助TetraMAX软件产生激励信号，然后利用NClaunch仿真随机记录内部128位寄存器的数据，根据记录的中间数据产生如图8所示的AES加密和解密Toggle统计结果。从图8可以看出，加密过程中，错误密钥的翻转率约为正确密钥的56%；解密过程中，两者的翻转率几乎一致。实验结果表明，无论在错误和正确密钥下，混淆后AES算法都存在寄存器翻转，可以防御寄存器翻转率的分析攻击。

自相关函数用来描述不同状态之间的相关性，相关性越少，被攻击的概率越低。在NClaunch工具软件的支持下，通过对混淆后AES算法RTL代码仿真，记录内部寄存器数据，结合自相关函数分析获得如图9所示的数据相关性。从图中可以看出，加密过程数据相关性逐渐趋近0，置信区间为89%；解密过程数据相关性为0，且置信区间72%。

在SMIC 65 nm CMOS工艺下，在TT环境下，工作频率约束为100~900 MHz，采用DC综合，可获得基于状态映射的AES算法硬件混淆电路的面积、速度和功耗，如图10所示。其中，DEC为AES解密算法，ENC为AES加密算法，DECO为混淆后的AES解密算法，ENCO为混淆后的AES加密算法。从图10(a)中可以看出，电路面积受工作频率影响较小。从图10(b)中可以看出，电路延时随着约束工作频率的增加呈线性降低。从图10(c)中可以看出，电

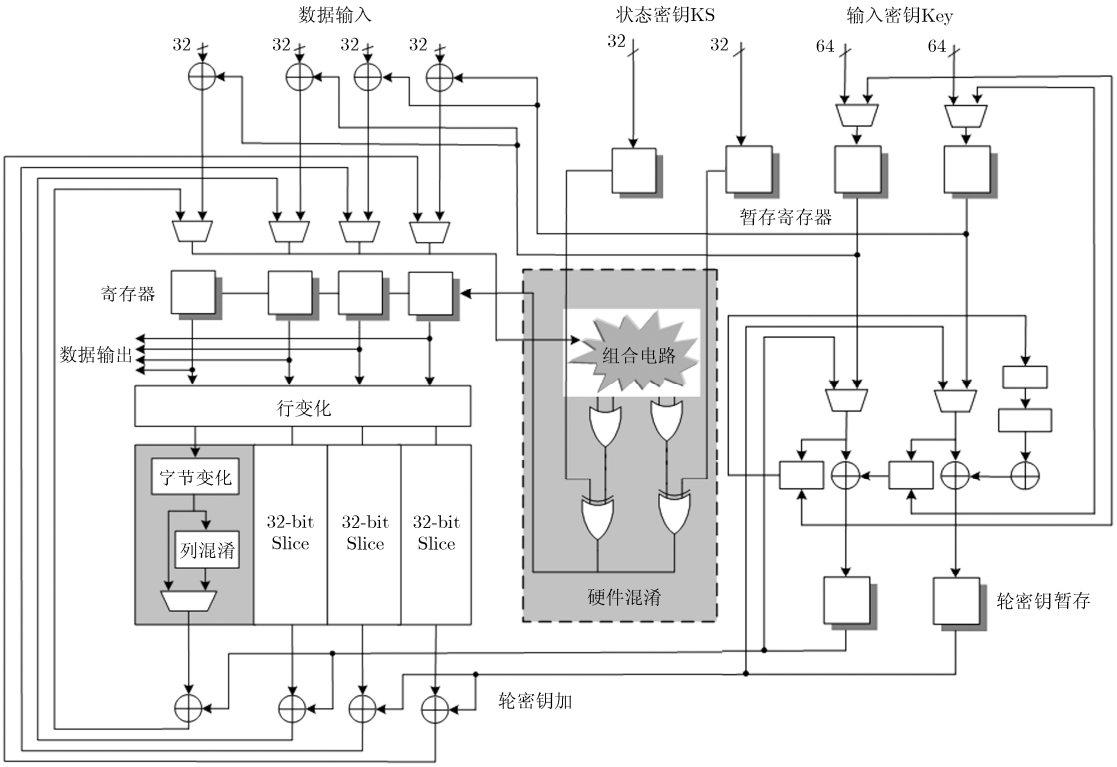


图 7　AES算法安全混淆解密电路结构图

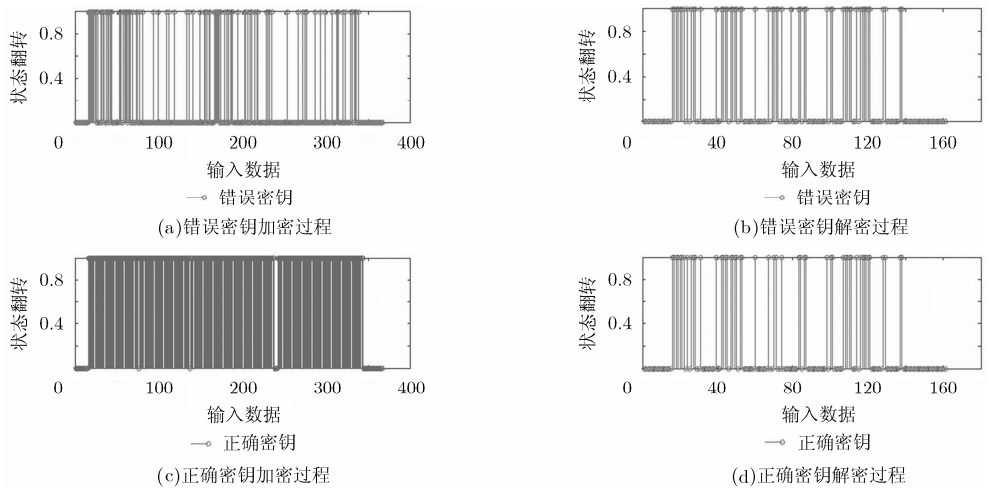


图 8　Toggle仿真结果

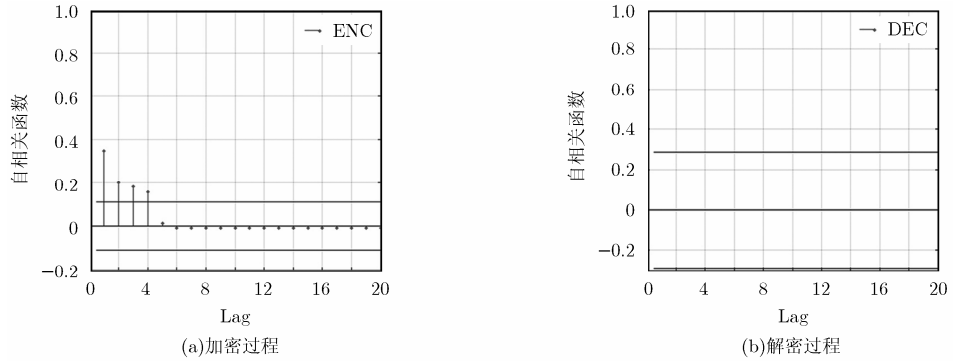


图 9 数据相关性仿真结果

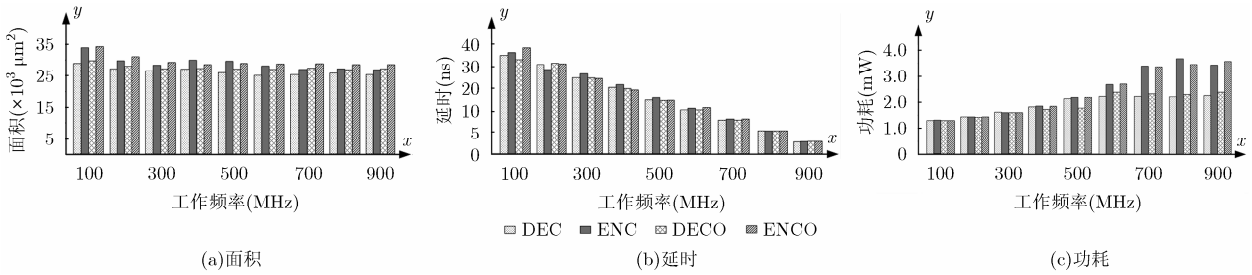


图10 不同频率下的性能比较

路功耗随着工作频率的增加呈线性递增。工作频率为 800 MHz 时混淆电路的功耗增加最大。与相关文献的比较结果如表 1 所示，基于状态映射的 AES 硬件混淆算法与传统 AES 算法比较，面积约增加 9%、功耗约增加 16%；基于状态映射的 AES 硬件混淆算法的非线性特性大大提高，代码覆盖率达到 93%以上。

5 结论

本文设计了基于状态映射的 AES 硬件混淆算法。该算法利用状态映射方法控制 AES 算法的有限状态机，实现时序逻辑电路的混淆，结合状态翻转的方法实现组合逻辑电路的混淆，并采用代码覆盖率、Toggle 和数据相关性等评估混淆算法的有效性。该方法可以阻止攻击者进入 FSM 的有效状态，达到防御逆向攻击和寄存器强写攻击的目的。最后在 SMIC 65 nm CMOS 工艺下，实现基于状态映射的 AES 硬件混淆算法。实验结果表明，在面积和功耗分别增加 9%和 16%的前提下，代码覆盖率达到 93%。

表 1 与相关文献的比较结果

算法	加密类型	面积(μm²)	功耗(mW)	代码覆盖(%)	是否混淆
VLSI 2005[14]	AES-ENC	23526	0.6305	95.8	否
	AES-DEC	21140	0.6648	97.2	否
DATA 2014[15]	AES	38482	1.0167	95.8	否
本文算法	AES-ENC	25983	0.7558	95.0	是
	AES-DEC	23611	0.7958	93.0	是

参 考 文 献

[1] 2015 年度检察机关保护知识产权十大典型案例[OL]. http://news.xinhuanet.com/legal/2016-05/05/c_128959767.htm, 2016.

[2] 杨宇波. 代码混淆模型研究[D]. [博士论文], 北京邮电大学, 2015.
YANG Yubo. Research on code obfuscation model[D]. [Ph.D. dissertation], Beijing University of Posts and Telecommunications, 2015.

[3] BARAK B, GOLDREICH O, IMPAGLIAZZO R, *et al*. On the Impossibility of obfuscating programs[J]. *Lecture Notes in Computer Science*, 2001, 2139: 1–18. doi: 10.1007/3-540-44647-8-1.

[4] 林水明, 吴伟民, 陶桂华, 等. 基于主成分分析的代码混淆有效性综合评估模型[J]. 计算机应用研究, 2016, 33(9): 2819–2822, 2840. doi: 10.3969/j.issn.1001-3695.2016.09.059.
LIN Shuiming, WU Weimin, TAO Guihua, *et al*. PCA-based code obfuscation effective comprehensive assessment model [J]. *Application Research of Computers*, 2016, 33(9): 2819–2822, 2840.. doi: 10.3969/j.issn.1001-3695.2016.09.059.

[5] 赵玉洁, 汤战勇, 王妮, 等. 代码混淆算法有效性评估[J]. 软件学报, 2012, 23(3): 700–711. doi: 10.3724/SP.J.1001.2012.03994.
ZHAO Yujie, TANG Zhanyong, WANG Ni, *et al*. Evaluation of code obfuscating transformation[J]. *Journal of Software*, 2012, 23(3): 700–711. doi: 10.3724/SP.J.1001.2012.03994.

[6] 谢鑫, 刘粉林, 芦斌, 等. 基于多层次属性加权的代码混淆有效性量化评估[J]. 计算机科学, 2015, 42(3): 167–173. doi: 10.11896/j.ssn.1002-137X.2015.3.035.
XIE Xin, LIU Fenlin, LU Bin, *et al*. Quantitative evaluation for effectiveness of code obfuscation based on multi-level weight attributes[J]. *Computer Science*, 2015, 42(3): 167–173. doi: 10.11896/j.ssn.1002-137X.2015.3.035.

[7] ALKABANI Y, KOUSHANFAR F, and POTKONJAK M. Remote activation of ICs for piracy prevention and digital right management[C]. 2007 IEEE/ACM International Conference on Computer-Aided Design, San Jose, CA, 2007:

674–677. doi: 10.1109/ICCAD.2007.4397343.

[8] CHAKRABORTY R S and BHUNIA S. HARPOON: An obfuscation-based SoC design methodology for hardware protection[J]. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2009, 28(10): 1493–1502. doi: 10.1109/TCAD.2009.2028166.

[9] CHAKRABORTY R S and BHUNIA S. RTL hardware IP protection using key-based control and data flow obfuscation [C]. 23rd International Conference on VLSI Design, Bangalore, 2010: 405–410. doi: 10.1109/VLSI.Design.2010.54.

[10] ZHANG J, LIN Y, LÜ Y, *et al*. A PUF-FSM binding scheme for FPGA IP protection and pay-per-device licensing[J]. *IEEE Transactions on Information Forensics and Security*, 2015, 10(6): 1137–1150. doi: 10.1109/TIFS.2015.2400413.

[11] KOUSHANFAR F. Provably secure active IC metering techniques for piracy avoidance and digital rights management[J]. *IEEE Transactions on Information Forensics and Security*, 2012, 7(1): 51–63. doi: 10.1109/TIFS.2011.2163307.

[12] CHANG Chiphong and POTKONJAK M. Secure System Design and Trustable Computing[M]. Switzerland: Springer International Publishing, 2016: 269–299.

[13] ZHANG J. A practical logic obfuscation technique for hardware security[J]. *IEEE Transactions on Very Large Scale Integration* (*VLSI*) *Systems*, 2016, 24(3): 1193–1197. doi: 10.1109/TVLSI.2015.2437996.

[14] FISCHER V, DRUTAROVSKY M, CHODOWIEC P, *et al*. InvMixColumn decomposition and multilevel resource sharing in AES implementations[J]. *IEEE Transactions on Very Large Scale Integration* (*VLSI*) *Systems*, 2005, 13(8): 989–992. doi: 10.1109/TVLSI.2005.853606.

[15] WANG Y, YU H, SYLVESTER D, *et al*. Energy efficient in-memory AES encryption based on nonvolatile domain-wall nanowire[C]. Design, Automation & Test in Europe Conference & Exhibition (DATE), Dresden, 2014: 1–4. doi: 10.7873/DATE.2014.196.

张跃军： 男，1982 年生，副教授，研究方向为低功耗、高信息密度集成电路理论和设计、安全芯片理论和设计.

潘　钊： 男，1994 年生，硕士生，研究方向为安全芯片理论和设计、低功耗集成电路理论和设计.

汪鹏君： 男，1966 年生，教授，研究方向为低功耗、高信息密度集成电路理论和设计、安全芯片理论和设计、多媒体技术及相关理论.

丁代鲁： 男，1991 年生，硕士生，研究方向为高信息密度集成电路理论和设计、纳米级老化传感器设计.

李　刚： 男，1988 年生，博士生，研究方向为低功耗、高信息密度集成电路理论和设计、安全芯片理论和设计.