计算机研究与发展
Journal of Computer Research and Development
DOI:10.7544/issn1000-1239.2018.20180427
55(10):2269-2277,2018

基于独立分量技术的类 GIFT 算法 S 盒逆向分析

马向亮[1,2,3] 李 冰[3] 习 伟[4] 陈 华[1] 陈财森[5]
[1](中国科学院软件研究所可信计算与信息保障实验室 北京 100190)
[2](中国科学院大学 北京 100049)
[3](国家信息技术安全研究中心 北京 100084)
[4](南方电网科学研究院 广州 510663)
[5](陆军装甲兵学院演训中心 北京 100072)
(maxiangliang@163.com)

Reverse-Analysis of S-Box for GIFT-Like Algorithms Based on Independent Component Analysis Technology

Ma Xiangliang[1,2,3], Li Bing[3], Xi Wei[4], Chen Hua[1], and Chen Caisen[5]
[1](*Trusted Computing and Information Assurance Laboratory*, *Institute of Software*, *Chinese Academy of Sciences*, *Beijing* 100190)
[2](*University of Chinese Academy of Sciences*, *Beijing* 100049)
[3](*National Research Center for Information Technology Security*, *Beijing* 100084)
[4](*Southern Power Grid Science Research Institute*, *Guangzhou* 510663)
[5](*Department of Training Center*, *Academy of Army Armored Force*, *Beijing* 100072)

Abstract In the security evaluation of practical crypto system or module, the reverse analysis of unknown cryptographic algorithm is an important evaluation content. At present, the reverse analysis methods of cryptographic algorithms mainly contain mathematical analysis and physical bypass analysis, among which the latter is more popular due to its low cost and high universality. The side-channel analysis technology based on independent component analysis technology recovers the intermediate state value directly, bypassing the limitation of the "guess-and-determine" attack idea in the traditional side-channel analysis. This paper studies the safety of GIFT-like algorithm under the reverse analysis. We successfully recovered the content of the S-box by independent component analysis (ICA) technology, taking advantage of the characteristics of GIFT algorithm structure and taking the inputs of P permutation as observation conditions. The result of this paper is one of the reverse analysis results of the GIFT-like algorithm, and the method is also of reference significance to the reverse analysis of other unknown algorithms.

Key words independent component analysis (ICA); reverse analysis; S box; GIFT algorithm; power analysis

摘 要 在实际密码系统或模块的安全性评估中,对未知密码算法的逆向分析是一项重要的评估内容.目前关于密码算法的逆向分析方式主要分为数学分析和物理旁路分析 2 种,后者因其代价低、通用性高

收稿日期:2018-06-12;**修回日期**:2018-08-03
基金项目:国家重点研发计划(2018YFB0904901,2016YFF0204005,2016YFF0204003,2016YFF0204002);国家自然科学基金项目(61402528);"十三五"装备预研领域基金项目(6140002020115);CCF-启明星辰"鸿雁"科研计划(2017003)
This work was supported by the National Key Research and Development Program of China (2018YFB0904901, 2016YFF0204005, 2016YFF0204003, 2016YFF0204002), the National Natural Science Foundation of China (61402528), the "Thirteen-Fifth" Equipment Pre-research Foundation Program (61400002020115), and the CCF-Venus "Hongyan" Research Program (201703).

等优点更为流行.基于独立分量技术的侧信道分析技术绕过传统侧信道分析中的“先猜测后确定”的攻击思路限制,直接恢复中间状态值.研究了类 GIFT 算法在逆向分析下的安全性,利用 GIFT 算法结构的特点,将 P 置换输入作为独立分量攻击观测条件,利用独立分量技术成功恢复出了 S 盒内容.该结果是最早关于类 GIFT 算法的逆向分析结果之一,其方法对于其他未知算法的逆向分析也具有参考意义.

关键词 独立分量分析;逆向分析;S 盒;GIFT 算法;能量分析

中图法分类号 TP309.1

尽管目前密码学的一个基本安全假设为,密码算法的安全性依赖于密钥及密钥的长度,而不是算法本身的保密性.然而在很多时候,政府或国防军队为了加强安全级别,或企业为了商业应用,对一些密码算法、部件进行保密或修改.在此情况下,对带有未知算法实现的密码系统或模块进行安全性评估至关重要,其中逆向分析是进行安全性评估的前提.逆向分析将直接影响到密码算法实现测评的力度.

目前,已存在的逆向分析方法主要有数学逆向分析[1]、基于故障注入技术的逆向分析[2]和基于侧信道技术的逆向分析[3]等.和数学分析方法相比,基于侧信道和故障注入等旁路技术的逆向分析具有代价低、通用性强等优势.另外,芯片解剖逆向[4]也是一种逆向分析方式,通过还原电路进行逻辑判断、分析,然而由于周期长、代价高、可行性低等原因,该方法不适宜现在工艺发达、高度集成的加密电路芯片.

1) 数学逆向分析

在欧洲密码年会(EUROCRYPT2001)上,文献[1]提出了一种结构分析方法,目的是在减少轮上进行 SPN 机构的逆向分析.考虑到分析的高复杂度,为了建立未知操作输入和输出之间的关系,利用该方法通过线性和非线性操作发现多重集的特征.在 2011 年文献[5]使用这种方法实现了类 PRESENT 算法的逆向 S 盒分析.

2) 基于故障注入技术的逆向分析

基于差分故障技术的逆向分析模型[6]是 1997 年由 Biham 等人最初提出的,即错误发生位置可以控制,但错误的结果未知.他们提出在一个封闭的防篡改设备(如智能卡)中可以逆向恢复未知 S 盒结构.2014 年 Clavier 等人基于无效故障模型对类 AES 算法进行了逆向分析[7].在该模型下,即使敌手在不知道密钥,对未保护的或带有常用布尔掩码和乱序防护措施的算法软实现,依然能够完全逆向恢复所有未知参数.

3) 基于侧信道逆向分析

① 传统侧信道逆向分析

2008 年文献[8]提出基于硬件设计的未知 Feistel 结构的逆向分析方案,使用选择明文攻击的侧信道攻击可以逆向分析恢复.2010 年文献[9]提出 2 种攻击方法,一种是攻击线性反馈寄存器结构算法的未知线性部件;另一种是恢复未知的非线性函数如 S 盒.2011 年文献[10]使用故障注入同时对 2 个常用的 DES 和 AES 算法逆向恢复未知 S 盒.2016 年文献[11]提出了基于线性回归攻击的新型侧信道逆向分析方法,与之前的侧信道逆向分析相比,该方法使用更少的先验知识,并能给出更多的密码实例.

② 基于独立分量攻击(ICA)的逆向分析

2017 年文献[12]提出了一种新型的侧信道分析技术,即基于独立分量技术[13](independent component analysis, ICA)的侧信道分析方法.盲源分离问题是指在原始信号被混合输出后,如何从一组被观测的混合信号中分离出没有被观测到的原始信号.独立分量技术可将侧信道泄露转化为有效的独立分量观测,并直接从泄露的曲线中恢复出中间轮的某个中间状态.与传统的侧信道分析策略不同,该方法没有采取“先猜测后确定”,而是采用了直接恢复的分析策略.该方法可用于未知密码算法的逆向分析中.相比于已有方法,该方法对算法类型和实现条件的限制较少,攻击复杂度较低,适用性好.文献[12]同时给出了一个针对 DES 实现的攻击实例,结果显示该方法恢复出 S 盒的等价形式仅需几百条能量迹.

综合以上相关工作,侧信道在逆向分析中具有很大的优势,在一定程度上降低了逆向分析的复杂度,可行性较强.进一步,独立分量技术在侧信道逆向分析中的应用使得侧信道逆向分析攻击的适用性更强、复杂度更低.鉴于独立分量技术的特点,在侧信道逆向分析中,特别适用于密码算法具有比特打乱或比特置换部件.

GIFT 算法[14]是在密码硬件与嵌入式系统会议(CHES 2017)上提出的一种轻量级分组密码算法.自该算法提出后,已出现一些针对它的分析结果,但尚没有针对它抵抗逆向分析的安全性评估.

本文提出了一种基于独立分量攻击技术的类 GIFT 算法 S 盒逆向分析方法. 我们针对 GIFT 算法中 P 置换的比特置换特点,构造了有效的独立分量观测值,并从得到的功耗曲线中直接恢复出了第一轮 P 置换输出值. 因此在明文已知的情况下,S 盒的内容可以恢复出来. 另外我们也进行了 GIFT 算法的逆向分析实验,实验结果验证了本方法是可行的. 最后,我们针对独立分量技术的攻击特点给出了相关防御建议.

1 预备知识

1.1 ICA 介绍

1.1.1 ICA 定义

ICA 是信号处理中解决盲源分离[12](blind source separation, BSS)问题的一种常用技术. 利用该技术实现了鸡尾酒会假设的问题,即在酒会上多人在一起讲话,房间的多个麦克风将大家的声音混合在一起经信号转换后进行记录,如何由记录的混合信号还原回原始的各个独立的信号.

ICA 模型假设观察到的混合信号变量 $\boldsymbol{Y}=(y_1,y_2,\cdots,y_m)^{\mathrm{T}}$,这些变量是由 n 个独立未知源信号 $\boldsymbol{S}=(s_1,s_2,\cdots,s_n)$线性混合而成. 因此混合信号 y_j 和源信号 $\boldsymbol{S}$ 之间存在关系 $y_j=a_{j,1}s_1+a_{j,2}s_2+\cdots+a_{j,n}s_n$,其中 $a_{i,j}$是未知的混合系数. 混合信号也可以使用向量表示为 $\boldsymbol{Y}=\boldsymbol{AS}$,$\boldsymbol{A}$ 为未知混合系数矩阵. 在实际信号处理过程中,通常会考虑噪声的干扰,ICA 模型修正为 $\boldsymbol{Y}=\boldsymbol{AS}+\boldsymbol{N}$,$\boldsymbol{N}$ 为未知的不可预测噪声,符合正态高斯分布. ICA 的作用是仅仅通过观察 $\boldsymbol{Y}$ 来恢复源信号 $\boldsymbol{S}$ 和混合系数矩阵 $\boldsymbol{A}$.

1.1.2 ICA 约束与不确定性

由 ICA 定义可知,在混合系数矩阵 $\boldsymbol{A}$ 未知情况下,由 $\boldsymbol{Y}$ 直接恢复 $\boldsymbol{S}$ 是非常有挑战性的. 使用 ICA 方法解决这个难题,需要满足以下 3 个条件:

1) 观测信号相互独立;

2) 观测信号必须非高斯分布;

3) 观察个数大于等于源信号个数 ($m\geqslant n$).

由于源信号 $\boldsymbol{S}$ 和混合系数矩阵 $\boldsymbol{A}$ 未知,基于 ICA 模型恢复必然带来了一些不确定结果,不确定性主要有以下 2 个方面:

1) 无法确定源信号的次序. 这个是显而易见的,选择一个可逆矩阵 $\boldsymbol{P}$,由 $\boldsymbol{Y}=\boldsymbol{AS}$,则 $\boldsymbol{Y}=\boldsymbol{AP}^{-1}\boldsymbol{PS}$,若 $\boldsymbol{A}$ 和 $\boldsymbol{S}$ 是解,则 $\boldsymbol{AP}^{-1}$和 $\boldsymbol{PS}$ 也是解.

2) 无法确定源信号的维度. 任何一个源信号都可以通过乘一个标量然后在矩阵 $\boldsymbol{A}$ 中对应的列中除以该标量抵消掉.

1.1.3 ICA 算法

ICA 自从提出开始已有大量的研究成果,不同的处理算法相继提出. 按步骤分类为以下 3 种:

1) 批处理算法. 成对数据旋转法(Jacobi)、特征矩阵的联合近似对角法(JADE)等;

2) 自适应算法. 串行矩阵更新及其自适应算法、扩展的 Infomax 法等;

3) 探查性投影追踪法. 梯度算法、固定点算法(Fast ICA)等.

以上算法中 Fast ICA 类最常用,本文使用该算法进行实验.

1.2 GIFT 算法介绍

GIFT 算法是 Banik 等人在 CHES2017 年会议上提出的 SPN 结构轻量级算法,是 PRESENT 算法的升级版,修正了 PRESENT 算法已知的缺点. 该算法有 2 个版本,密钥都是128 b,明文分组有区别分别为 64 b 和 128 b. 本文实验使用的是 64 b 版本,以 64 b 为例介绍,但是 64 b 和 128 b 攻击效果是一样的.

GIFT 算法一共 28 轮运算,每轮运算由 3 个步骤组成:S 盒替换、比特 P 置换和与密钥异或运算.

1.2.1 初始化

将加密的 64 位明文 $b_{n-1}b_{n-2}\cdots b_0$ 作为一个密码状态 $\boldsymbol{S}$,其中 b_0 是最低有效位. 也可以将 $\boldsymbol{S}$ 表示为 4 位半字节排列,即 $\boldsymbol{S}=w_{s-1}\parallel w_{s-2}\parallel\cdots\parallel w_0$,其中$s=16$. 密钥 128 位可以表示为 16 位的字排列,即 $K=k_7\parallel k_6\parallel\cdots\parallel k_0$.

1.2.2 S 盒替换

S 盒是 GIFT 算法中唯一的非线性部件,将状态 S 中的每半个字节经查表替换为相应的值. $w_i\leftarrow GS(w_i)$,$\forall\, i\in\{0,\cdots,s-1\}$. S 盒 GS 如表 1 所示,表 1 中的数据使用十六进制.

1.2.3 比特置换

比特置换用来打乱原来的位顺序,将第 i 位的状态置换为第 $P(i)$ 位,即 $b_{P(i)}\leftarrow b_i$, $\forall\, i\in\{0,\cdots,n-1\}$. P 置换表如表 2 所示.

1.2.4 密钥异或

32 位的轮密钥被分为 2 个 16 位字排列,即 $RK=U\parallel V=u_{s-1}\cdots u_0\parallel v_{s-1}\cdots v_0$. U 和 V 分别和密码状态 b_{4i+1} 和 b_{4i} 进行异或运算. 即 $b_{4i+1}\leftarrow b_{4i+1}\oplus u_i$,$b_{4i}\leftarrow b_{4i}\oplus v_i$, $\forall\, i\in\{0,\cdots,15\}$.

Table 1 GIFT S Box GS

表 1 GIFT 算法的 S 盒 GS

x	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
$GS(x)$	1	a	4	c	6	f	3	9	2	d	b	7	5	0	8	e

Table 2 GIFT Bit Permutation

表 2 GIFT 算法 P 置换表

i	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$P_{64}(i)$	0	17	34	51	48	1	18	35	32	49	2	19	16	33	50	3
i	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
$P_{64}(i)$	4	21	38	55	52	5	22	39	36	53	6	23	20	37	54	7
i	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47
$P_{64}(i)$	8	25	42	59	56	9	26	43	40	57	10	27	24	41	58	11
i	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63
$P_{64}(i)$	12	29	46	63	60	13	30	47	44	61	14	31	28	45	62	15

与轮密钥异或运算完的密码状态的第 63,23,19,15,11,7 和 3 位分别与 1 和 6 位轮常数进行异或运算. 轮常数为 $C=c_5c_4c_3c_2c_1c_0$,即

$$b_{n-1}\leftarrow b_{n-1}\oplus 1, b_{23}\leftarrow b_{23}\oplus c_5,$$
$$b_{19}\leftarrow b_{19}\oplus c_4, b_{15}\leftarrow b_{15}\oplus c_3,$$
$$b_{11}\leftarrow b_{11}\oplus c_2, b_7\leftarrow b_7\oplus c_1, b_3\leftarrow b_3\oplus c_0.$$

1.2.5 密钥编排和轮常数

轮密钥 $RK32$ 位由 2 个 16 位字排列组合而成,即 $RK=U\parallel V. U\leftarrow k_1, V\leftarrow k_0$. 密钥每轮更新一次,$k_7\parallel k_6\parallel\cdots\parallel k_0\leftarrow k_1\ggg 2\parallel k_0\ggg 12\parallel\cdots\parallel k_3\parallel k_2$. 其中,$\ggg i$ 指的是 16 位字循环右移 i 位. 轮常数使用 6 位线性反馈寄存器进行更新,6 位初始值为 0,需要先更新才能使用,更新函数为 $(c_5,c_4,c_3,c_2,c_1,c_0)\leftarrow(c_4,c_3,c_2,c_1,c_0,c_5\oplus c_4\oplus 1)$. 每轮常数如表 3 所示:

Table 3 Rounds constants

表 3 轮常数

Rounds	Constants
1~16	01,03,07,0F,1F,3E,3D,3B,37,2F,1E,3C,39,33,27,0E
17~32	1D,3A,35,2B,16,2C,18,30,21,02,05,0B,17,2E,1C,38
33~48	31,23,06,0D,1B,36,2D,1A,34,29,12,24,08,11,22,04

2 ICA 逆向分析

2.1 ICA 逆向分析模型及算法

在本节中,我们详细介绍如何将 ICA 用于侧信道逆向恢复. 依赖数据侧信道泄露模型以及数据之间的独立与非高斯性,通过深入研究算法的特定结构,构造符合 ICA 攻击条件,将侧信道中需要恢复的中间值数据问题转换为 ICA 问题.

侧信道常用的数据依赖映射模型为汉明重量模型. 对于中间值数据 n 位的能量泄露值可以表示为

$$L(\boldsymbol{x})=a_1x_1+a_2x_2+\cdots+a_nx_n, a_i\in\{0,1\},$$

其中,x_i 表示中间值 $\boldsymbol{x}$ 的二进制表示. 能量泄露值的汉明重量模型满足 ICA 线性混合条件. 在侧信道环境中,x_i 是比特信号,取值 0 或者 1,服从伯努利分布,天然满足非高斯和独立性. 在此基础上,构造符合 ICA 观测成为 ICA 在侧信道逆向分析的关键,由于该问题需要结合特定算法结构进行分析解决,针对本文的 GIFT 算法观测将在 3 节介绍.

通过结合算法结构和侧信道相关理论,使得 ICA 算法在侧信道环境中适用于逆向分析,通过观察 T 个中间状态 $\boldsymbol{x}=(x(1),x(2),\cdots,x(T))$. 由比特置换可知,每个比特之间互相独立,可表示为 $\boldsymbol{x}_i=(x_i(1),x_i(2),\cdots,x_i(T))$. 相对应的观测为 $\boldsymbol{Y}=(y_1,y_2,\cdots,y_m)^{\mathrm{T}}$,而每比特观测为 $\boldsymbol{y}_i=(y_i(1),y_i(2),\cdots,y_i(T))$. 如算法 1 所示.

算法 1. 用于侧信道逆向的 ICA 算法[15].

输入:T 条能量迹;

输出:二元信号 $\hat{\boldsymbol{x}}$.

① 收集 m 个观测 $\boldsymbol{Y}=(y_1,y_2,\cdots,y_m)^{\mathrm{T}}$.

② 利用 Fast ICA 算法,得到原观测的 n 个实值独立分量 $(\boldsymbol{IC}_1,\boldsymbol{IC}_2,\cdots,\boldsymbol{IC}_n)$.

③ 找到 IC 最接近的二元源信号 $\tilde{\boldsymbol{x}}=(\tilde{x}_1,\tilde{x}_2,\cdots,\tilde{x}_n)$.

④ 以源信号 $\bar{x}$ 为初始猜测,估计初始参数 $\theta^{(0)}$,进入 EM 算法迭代

while $\Delta L > threshold$ do

E-Step:用当前的参数 $\theta^{(j)}$,计算对数似然期望函数 $L(\theta^{(j)})$;

M-Step:计算最大化对数似然期望函数 L 的参数($\theta^{(j+1)}$);

end while

⑤ 找出当前最大化对数似然期望函数 L 的源信号 $\hat{x}$.

⑥ return 输出找到的二元信号 $\hat{x}$.

电子噪声在能量曲线采集中不可避免,噪声会改变能量迹中的能量值,从而对逆向分析带来了难度.一个简单直接的办法就是改善采集硬件条件尽可能降低电子噪声和转换噪声,另外就是通过多次采集求均值、滤波技术、奇异谱分析技术等预处理技术,减少噪声的干扰.S 盒是分组密码算法的关键非线性部件,它是逆向分析研究的主要分析对象.在针对 S 盒的逆向分析中,利用算法的特殊结构,ICA 可以从能量迹上的能量值直接恢复 S 盒的输出,本文提出基于 ICA 的逆向分析模型如图 1 所示:

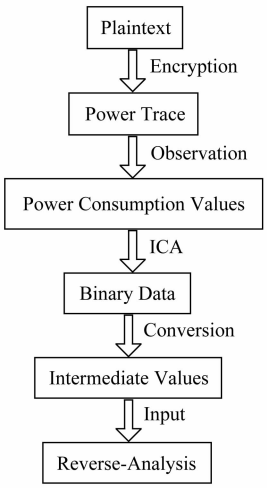

Fig. 1 ICA model for SCARE

图 1 ICA 侧信道逆向分析模型

由图 1 该模型下逆向分析 S 盒的步骤如下:

1) 选择随机明文,经加密算法 M 次加密运算,得到 M 条能量迹,也就得到一组 M 长的源信号观测,而且相互之间是独立的;

2) 选取 S 盒输出的能量迹上的相应值观测作为能量值.能量值依据示波器的采用率和分辨率,而且和采样时间长短有关.需要从能量迹上识别出所需的点;

3) 使用 Fast ICA 算法对能量值进行攻击,恢复出 S 盒输出二进制序列;

4) 将二进制转化为 16 进制数据,作为算法 S 盒输出;

5) 由于 S 盒输入已知,由输入和输出逆向分析出 S 盒结果.

2.2 ICA 逆向分析成功率

由于 ICA 的不确定性,返回的信号是位置换且反转信号得出的结果应该也是正确的,由此选择 $\boldsymbol{X}_{\hat{k}}$ 和 $\boldsymbol{X}_r$ 中最接近的一个.在下面公式中,$\| v \|_1$ 代表 v 的曼哈顿距离,而 $\boldsymbol{X}_{k,i}$ 代表的是 $\boldsymbol{X}_k$ 中第 i 位.

$$D(\boldsymbol{X}_r, \boldsymbol{X}_c) = \min_B \left(\sum_{i=1}^{n} dist(\boldsymbol{X}_{k,i}, \boldsymbol{X}_{r,B(i)}) \right),$$

其中,

$$dist(\boldsymbol{X}_{k,i}, \boldsymbol{X}_{r,B(i)}) =$$

$$\min(\| \boldsymbol{X}_{k,i} - \boldsymbol{X}_{r,B(i)} \|_1, \| \boldsymbol{X}_{k,i} - \overline{\boldsymbol{X}_{r,B(i)}} \|_1).$$

由于 ICA 方法直接恢复密码算法的某个中间值,可以使用正确信号($\boldsymbol{X}_c$)与 ICA 计算恢复信号($\boldsymbol{X}_r$)之间的最小距离来判断恢复结果的有效性.随着能量迹的增多,$D(\boldsymbol{X}_r, \boldsymbol{X}_c)$将变大,因此 ICA 逆向分析成功率可定义为

$$Succ(\boldsymbol{X}_r) = 1 - \frac{D(\boldsymbol{X}_r, \boldsymbol{X}_c)}{nM}.$$

由 $D(\boldsymbol{X}_r, \boldsymbol{X}_c)$定义可知,ICA 逆向恢复成功率最小为 0.5;若成功率比较大,远远大于 0.5,则说明 ICA 成功恢复出 S 盒中间值;若成功率和 0.5 差不多,则说明 ICA 计算返回的中间值不可信.

3 GIFT 算法逆向分析

本节将详细介绍 GIFT 算法利用 ICA 技术进行逆向分析.GIFT 算法主要由 3 部分组成,S 盒置换、P 置换和与轮密钥、常数异或.因此在 GIFT 算法 ICA 观测中,选择比特置换作为观测源,满足 ICA 观测需求,如图 2 所示.

在这里我们将分析第 1 轮 S 盒输出(它同时也是第 1 轮 P 置换的输入)的能量泄露,由于 P 置换是逐比特进行的,比特之间相互独立,这种方法特别适用于构造 ICA 观测.$\boldsymbol{x}_i = (x_i(1), x_i(2), \cdots, x_i(T))$.相对应的观测为 $\boldsymbol{Y} = (y_1, y_2, \cdots, y_m)^{\mathrm{T}}$,而每比特观测为 $y_i = (y_i(1), y_i(2), \cdots, y_i(T))$.通过调用 ICA 算法,可以直接恢复出 S 盒的输出值.

随机选择 2000 组明文数据,使用 GIFT 算法对该数据进行加密,得到 2000 条能量迹.由上面观测

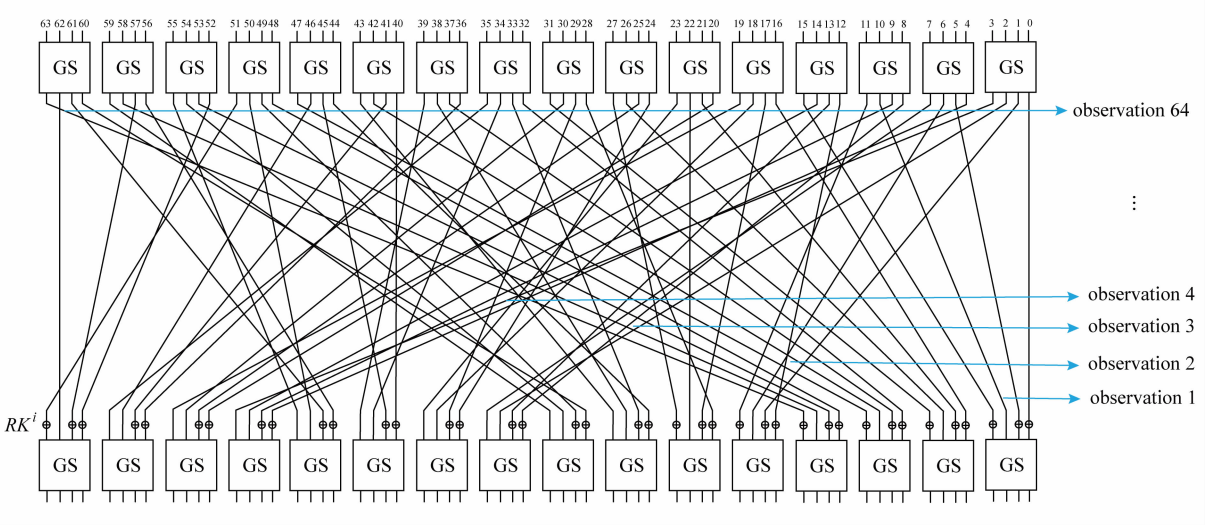

Fig. 2 Observations GIFT

图 2 GIFT 算法观测图

图 2 和先验知识可知,第 1 个 S 盒输出比特置换位分别为 0,17,34 和 51,这 4 个比特之间相互独立满足 ICA 观测要求,因此能量迹上存在 4 个点产生的能量值作为观测值. 以这种观测方式,得到 2 000 组中间值观测,其他 S 盒执行类似观测.

将 2 000 组观测值作为 ICA 逆向分析算法输入数据进行计算. 算法计算完成返回一个 2 000 行 4 列二进制数据 0 或 1 矩阵. 将矩阵中每行 4 列二进制数据按照相应权重计算为一个 16 进制数据,由此得到一个 2 000 行向量,每个值对应为 S 盒输出数据. 由于 2 000 个明文已知,对应的 S 盒输出数据已由 ICA 算法计算获得,从而完成 GIFT 算法 S 盒逆向恢复.

本实验使用的能量迹来自侧信道通用分析平台 Riscure 公司的 INSPECTOR 软件仿真,考虑到噪声在真实环境中肯定存在,而且会对能量迹及分析结果产生一定的影响,本文使用多组基于正态高斯分布的不同噪声,噪声大小有的甚至已远远超过真实环境配置,已有公开成果分析得出真实环境的噪声符合正态高斯分布[15]. 噪声对侧信道分析产生的影响不是由噪声大小决定的,而是由信噪比决定的,即 $SNR=\frac{Var(signal)}{Var(noise)}$. 本文由于仿真,并且可以控制信噪比的大小,采用信噪比分别为 0.01,0.015 6,0.04,0.25,1,4 和没有噪声共 7 组不同信噪比仿真产生能量迹 2 000 条. 能量迹横轴代表采集的点数,纵轴代表的能量值(mv)也就是 GIFT 算法开始执行时随时间变化的能量值. 本文列出 3 组不同信噪比的能量迹,信噪比为 0.01 的能量迹仿真如图 3 所示,信噪比为 0.04 的能量迹仿真如图 4 所示. 信噪比为 1 的能量迹仿真如图 5 所示.

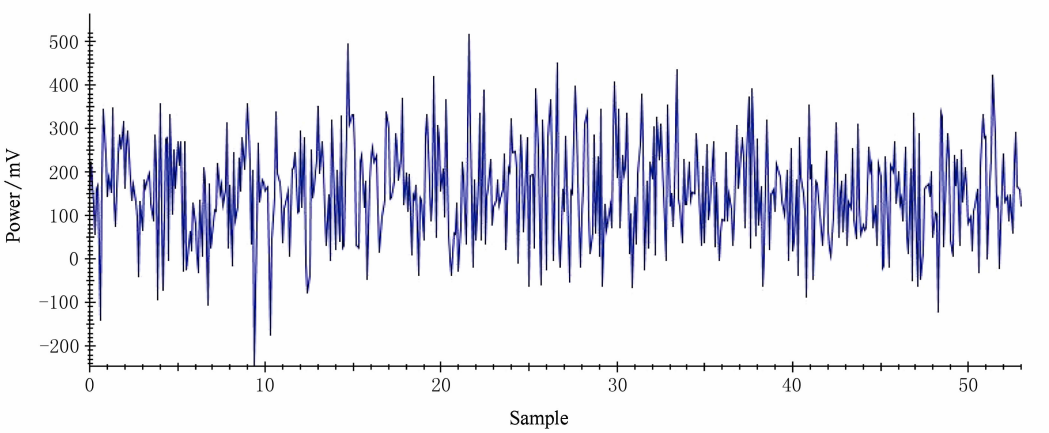

Fig. 3 GIFT trace with $SNR=0.01$

图 3 信噪比为 0.01 的 GIFT 能量迹

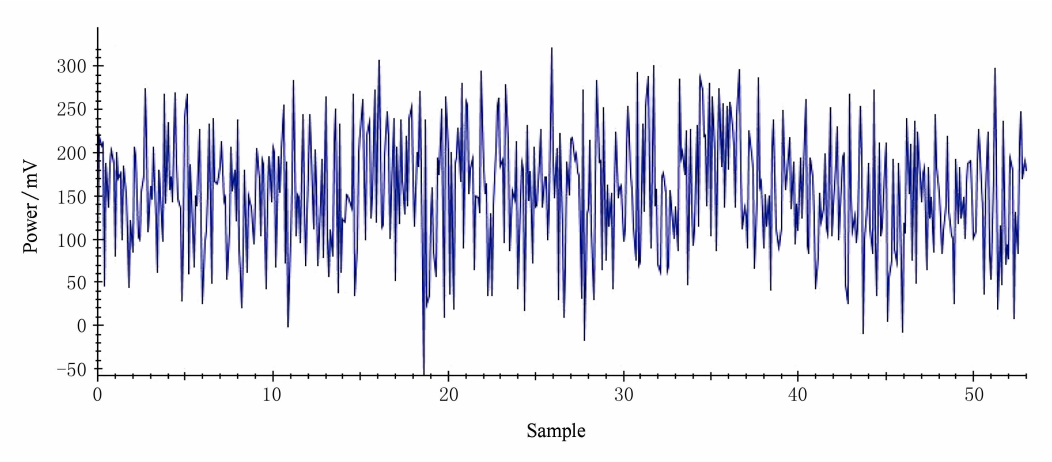

Fig. 4 GIFT trace with $SNR=0.04$

图 4 信噪比为 0.04 的 GIFT 能量迹

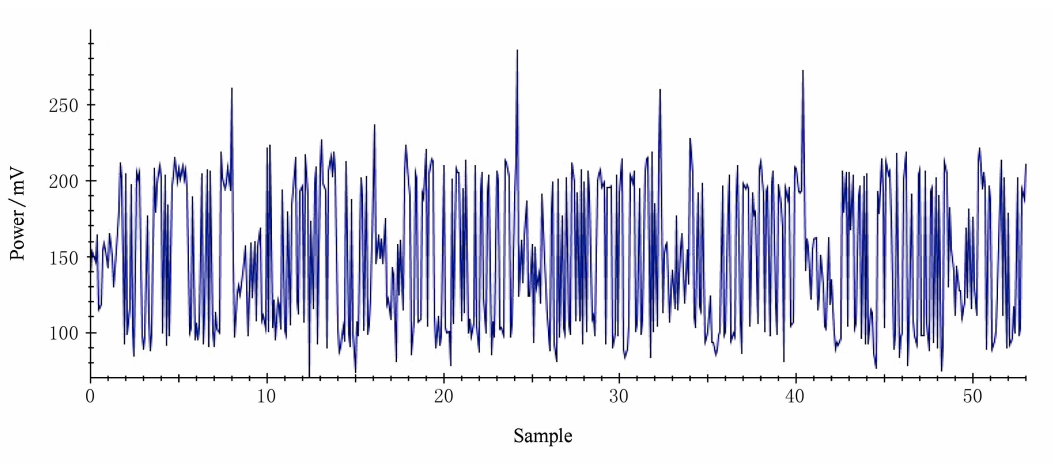

Fig. 5 GIFT trace with $SNR=1$

图 5 信噪比为 1 的 GIFT 能量迹

由图 3 和图 5 能量迹可以直观得出,噪声对侧信道分析的影响也可以从能量迹上直观的看到,信噪比小的能量迹波动很大. 信噪比为 1 的能量迹波动明显小于信噪比为 0.01 的能量迹波动. 利用比特置换 P 为观测源,通过选取能量迹上 4 个点能量值作为 ICA 观测值,通过 ICA 算法对上述仿真的 7 组不同信噪比能量迹进行逆向分析的成功率如表 4 所示.

图 6 给出了 GIFT 算法首轮第一个 S 盒 ICA 逆向恢复结果. 实验中,没有噪声的观测值恢复很快,无疑正确的恢复出了结果;信噪比在接近 1 以及

Table 4 Success Rates Based on ICA for Recovering S Box with Different *SNR*

表 4 不同信噪比 S 盒逆向恢复成功率

SNR	Success Rates
0.01	0.6275
0.0156	0.6613
0.04	0.7575
0.25	0.9950
1	1
4	1
∞	1

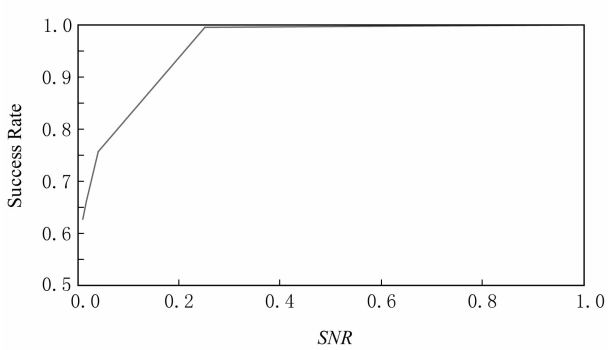

Fig. 6 Success rates based on ICA for recovering GIFT S box

图 6 ICA 逆向分析 GIFT 算法 S 盒成功率

大于 1 时的成功率都为 1;可以看出信噪比为 0.04 的逆向成功率为 0.757 5,效果也比较理想;但在信噪比为 0.01 时的成功率较低,仅有 0.627 5.

实验中,统一使用的能量迹 2 000 条,成功率为 1 的逆向分析实际上用不了 2000 条. 成功率小于 1 的随着能量迹条数的增多,成功率会逐渐增大,但是信噪比较低的,成功率不会有大的改善. 另外,本实验没有使用任何降噪技术,如果能量迹进行降噪预处理后再进行观测,那么效果会得到改善,可以预计的是攻击条数会有相应减少,信噪比太低的,成功率仍然不会有很大提高,这是由于有用信息已被噪声淹没掉.

4 结 论

对带有未知算法实现的密码系统或模块进行逆向分析具有重要意义,侧信道方法具有一定的优势,在一定程度上降低了逆向分析的复杂度,而且可行性强. 与传统的侧信道逆向分析相比,本文提出的一种基于独立分量攻击技术的类 GIFT 算法 S 盒逆向分析方法,不需要进行猜测和确定. 我们针对该算法中 P 置换的比特置换特点,构造了有效的独立分量观测值,并从仿真得到的能量迹中直接恢复出了第一轮 P 置换输出值,从而成功进行逆向分析.

本文提出的基于 ICA 的逆向分析方法和模型,基于 P 置换的比特置换特点. 对于某些算法结构中,并不存在比特置换特点如国密 SM4 算法,相互之间不独立,因此无法进行有效观测. 如何构建符合 ICA 攻击的独立观测是一个难点,可以尝试选择特定明文或使明文与一个常数异或,使得该算法在第二轮加密时达到某些观测值之间相互独立.

如果算法工程采用软实现方式,针对本文提出的基于独立分量技术的逆向分析,使用大噪声干扰是一种方案,该方法在一定程度上增加了该攻击难度,如果噪声足够大,则信号将淹没,信噪比很低,无法恢复出有用信号. 对于硬实现,暂时没有找到合适的线性模型,观测值之间相互不独立,分离信号有一定的难度,是目前该方法的局限性,因此硬实现是一种比较理想的防护方案. 最后统筹考虑安全因素和性能代价,选择没有比特置换特点的算法,使得攻击者无法构建符合 ICA 攻击的独立观测. 尽管该方法不够完美,有一定的局限性,但本方法为密码分析者和测评机构提供了一种新的思路和评估手段,对于其他未知算法的逆向分析也具有参考意义.

参 考 文 献

[1] Biryukov A, Shamir A. Structural cryptanalysis of sasas [C] //Proc of the Int Conf on the Theory and Applications of Cryptographic Techniques. Berlin: Springer, 2001: 395-405

[2] Ming Ting, Qiu Zhenlong, Peng Hongbo, et al. Toward reverse engineering on secret s-boxes in block ciphers [J]. Science China (Information Sciences), 2014, 57(3): 1-18

[3] Daudigny R, Ledig H, Muller F, et al. SCARE of the des [C] //Proc of the Int Conf on Applied Cryptography and Network Security. Berlin: Springer, 2005: 393-406

[4] Torrance R, James D. The state-of-the-art in IC reverse engineering [C] //Proc of the Workshop on Cryptographic Hardware and Embedded Systems. Berlin: Springer, 2009: 363-381

[5] Borghoff J, Knudsen L R, Leander G, et al. Cryptanalysis of present-like ciphers with secret s-boxes. [G] //LNCS6733. Berlin: Springer, 2011: 270-289

[6] Biham E, Shamir A. Differential fault analysis of secret key cryptosystems [C] //Proc of the Int Cryptology Conf. Berlin: Springer, 1997: 513-525

[7] Clavier C, Isorez Q, Marion D, et al. Complete reverse-engineering of aes-like block ciphers by SCARE and FIRE attacks [J]. Cryptography & Communications, 2015, 7(1): 121-162

[8] Réal D, Dubois V, Guilloux A M, et al. SCARE of an Unknown Hardware Feistel Implementation [C] //Proc of the Int Conf on Smart Card Research and Advanced Applications. Berlin: Springer, 2008: 218-227

[9] Guilley S, Micolod J, Micolod J. Defeating any secret cryptography with SCARE attacks [C] //Proc of the Int Conf on Progress in Cryptology: Cryptology and Information Security in Latin America. Berlin: Springer, 2010: 273-293

[10] Pedro M S, Soos M, Guilley S. FIRE: Fault injection for reverse engineering [C] //Proc of the Int Conf on Information Security Theory and Practice: Security and Privacy of Mobile Devices in Wireless Communication. Berlin: Springer, 2011: 280-293

[11] Gao Si, Chen Hua, Wu Wenling, et al. Linear regression attack with F-test: A new SCARE technique for secret block ciphers [C] //Cryptology and Network Security. Berlin: Springer, 2016: 3-18

[12] Gao Si, Chen Hua, Wu Wenling, et al. My traces learn what you did in the dark: Recovering secret signals without key guesses [C] //Topics in Cryptology-CT-RSA 2017. Berlin: Springer, 2017: 363-378

[13] Hyvärinen A. Fast and robust fixed-point algorithms for independent component analysis [J]. IEEE Trans on Neural Networks, 1999, 10(3): 626-634

[14] Banik S, Pandey SK, Peyrin T, et al. GIFT: A small present towards reaching the limit of lightweight encryption

[C] //Proc of the Workshop on Cryptographic Hardware and Embedded Systems. Berlin: Springer, 2017: 321-345

[15] Mangard S, Thoswald E, Popp T. Power Analysis Attack [M]. Translated by Feng Dengguo, Zhou Yongbin, Liu Jiye, et al. Beijing: Science China Press, 2010 (in Chinese) (Mangard S, Thoswald E, Popp T. 能量分析攻击[M]. 冯登国, 周永彬, 刘继业, 等译. 北京: 科学出版社, 2010)

Ma Xiangliang, born in 1986. PhD candidate. His main research interests include information security and side channel attack.

Li Bing, born in 1962. Senior engineer. His main research interests include cyberspace security and cryptographic algorithm application.

Xi Wei, born in 1981. Master, senior engineer. His main research interests include smart grid and electric power chip development.

Chen Hua, born in 1976. PhD. Professor, PhD supervisor in the Institute of Software, Chinese Academy of Sciences. Her research interests include side channel cryptanalysis, automatic cryptanalysis and randomness test.

Chen Caisen, born in 1983. PhD, lecturer. His main research interests include side channel attack and cache timing attack.