doi:10.3969/j.issn.1000-436x.2013.07.015

基于汉明重的 LED 代数旁路攻击研究

冀可可[1]，王韬[1]，郭世泽[2]，赵新杰[1]，刘会英[1]
(1. 军械工程学院 信息工程系，河北 石家庄 050003；2. 北方电子设备研究所，北京 100083)

摘 要：对 CHES 2011 会议提出的 LED 轻型分组密码抗代数旁路攻击能力进行了评估。给出了密码算法代数旁路攻击模型及 LED 密码代数方程表示方法；利用示波器采集微控制器 ATMEGA324P 上的 LED 实现功耗泄露，选取功耗特征较为明显的部分泄露点，基于 Pearson 相关系数方法推断加密中间状态汉明重；分别基于可满足性问题、伪布尔优化问题、线性编程问题给出了 LED 密码和汉明重泄露的 3 种代数方程表示方法；使用 CryptoMinisat 和 SCIP 2 种解析器对建立的代数方程求解恢复密钥，在已知明文、未知明密文、容错等场景下进行了大量的攻击实验。结果表明，LED 易遭受代数旁路攻击，一条功耗曲线的 1 轮汉明重泄露分析即可恢复 64 bit 完整密钥。
关键词：汉明重；LED；代数旁路攻击；可满足性；伪布尔优化；线性编程
中图分类号：TP393.08 **文献标识码：**A **文章编号：**1000-436X(2013)07-0134-09

Research of Hamming weight-based algebraic side-channel attack on LED

JI Ke-ke[1], WANG Tao[1], GUO Shi-ze[2], ZHAO Xin-jie[1], LIU Hui-ying[1]
(1. Department of Information Engineering, Ordnance Engineering College, Shijiazhuang 050003, China;
2. The Institute of North Electronic Equipment, Beijing 100083, China)

Abstract: The security of LED against the algebraic side-channel attack (ASCA)was evaluated, which is a lightweight block cipher proposed in CHES 2011. Firstly, the attack model of ASCA was analyzed, and then the design and algebraic representations of LED were described. Secondly, the power leakages of LED on ATMEGA324P microcontroller were measured by a digital oscilloscope; some leakage points with obvious power patterns were chosen as the targeted points and used to deduce the Hamming weight via computing the Pearson correlation factor; satisfiability-based, Pseu-do-Boolean optimization-based, linear programming-based methods were used to representing Hamming weights with algebraic equations. Finally, the CryptoMinisat and the SCIP solver were applied to solve for the key and many attacks are conducted under different scenarios. Experiment results demonstrate that LED is vulnerable to ASCA, full 64 bit master key can be derived via analyzing the HW leakages of the first round in LED.
Key words: Hamming weight; LED; ASCA; satisfiability; pseudo-Boolean optimization; linear programming

1 引言

密码实现需依附具体物理设备平台，其运行过程中会泄漏出时间[1]、功耗[2]、电磁[3]等旁路信息，通过采集并分析这些信息进行密钥破解的方法称为旁路攻击。自从 Kocher 于 1996 年首次提出旁路攻击[1]的概念以来，攻击研究取得了丰硕的成果，攻击对象趋于多样化，分析方法趋于复杂化。当前，旁路分析方法研究有 2 个热点，一是寻找新的旁路区分器，如互信息区分器[4]、KS 检验区分器[5]等；二是结合数学分析方法，如碰撞旁路分析[6]、Cube 旁路分析[7]、代数旁路分析[8~11]。本文主要研究代数旁路分析 (ASCA, algebraic side-channel analysis) 方法。

代数旁路分析思想由 Renauld 等[8,9]在 2009 年提出，攻击将代数和旁路攻击有机结合起来，利用代数

收稿日期：2012-01-09；修回日期：2012-04-16
基金项目：国家自然科学基金资助项目(61173191)
Foundation Item: The National Natural Science Foundation of China (61173191)

方法构建密码算法的布尔方程组，利用旁路攻击手段获取加密中间状态值并转化为额外代数方程组，与密码算法方程组联立以加速密钥求解。

Renauld 等基于可满足性(SAT, satisfiability)[13]问题使用 zChaff 解析器对 PRESENT、AES 密码进行了分析；随后 CHES 2010 会议上 Oren 等[10]基于伪随机优化问题(PBOPT, pseudo-Boolean optimization)使用 SCIP 解析器对 Keeloq 分组密码进行了容错代数旁路分析，并将 AES 容错代数旁路攻击作为一个公开问题；COSADE 2012 会议上，赵新杰等[11]提出了一种基于可满足性问题的多推断代数旁路攻击方法(MDASCA, multiple deductions-based ASCA)。结果表明 MDASCA 可用于开展容错 ASCA 攻击，利用 CryptoMiniSAT 解析器，首次基于汉明重泄露模型对 ATMEGA324P 控制器下的 AES 实现进行了容错代数旁路攻击；同时，MDASCA 可用于挖掘新泄露模型，首次指出 MDASCA 可用于挖掘访问驱动和踪迹驱动 2 种 Cache 泄露模型，优化 Cache 攻击，对 AES 在不同场景下进行了攻击实验。

代数旁路攻击[8~11]是数学分析方法与旁路攻击发展的必然结果，克服了传统代数攻击方程组求解复杂度较高的缺陷，弥补了旁路攻击中旁路信息利用率低、分析轮数少、样本量大等不足，在未知明密文、掩码防护等场景下攻击使用一条功耗曲线仍能成功实施，极大地提高了旁路攻击适用性和效率。加强代数旁路攻击研究对于改进算法设计、加强算法防护、提高算法安全性具有重要意义。

随着信息技术和电子元器件的发展，密码设备发展呈现出轻型化的趋势，如何在 RFID 标签等轻型设备上实现分组密码算法已成为近年来密码研究的新热点。轻型分组密码要求在追求快速的执行效率的同时又能够提供足够的安全保障，在安全性、价格及执行效率之间寻求平衡。LED[12]是在 CHES 2011 上提出的轻型分组密码，采用了类 AES 的 SPN 结构，抗差分、线性、代数攻击能力极强，但抗代数旁路攻击能力在设计时并未考虑，对其进行代数旁路攻击对其安全性评估及防护有重要意义。本文对 LED 代数旁路攻击进行了研究，主要贡献如下。

1) 给出了 LED 算法的代数方程表示方法。

LED 代数攻击中，非线性 S 盒部件和列混淆部件对应的代数方程组构建是其中的重点和难点问题。本文基于布尔积理论，给出了 LED 的 S 盒代数表示方法，并首次给出 LED 列混淆函数的代数表示方法，在此基础上构建了 LED 算法的全轮方程组。

2) 首次实现了基于汉明重泄露模型的 LED 密码代数旁路攻击。

通过示波器采集微控制器上 LED 算法执行过程中的功耗泄露，基于 Pearson 相关系数方法推断出加密中间状态的汉明重值并转化为代数方程组，然后通过方程组求解恢复密钥。结果表明：LED 易遭代数旁路攻击威胁，已知明文条件下 1 轮汉明重泄露分析、未知明密文条件下 5 轮汉明重泄露分析即可恢复全部初始密钥。

3) 首次对不同方程组求解策略的代数旁路攻击进行了研究。

当前，代数旁路攻击中利用的方程组求解策略有 2 种，一是基于可满足性问题 SAT[13]求解，典型解析器有 zChaff[14]、MiniSAT[15]、CryptoMiniSAT[16]；二是基于混合整数编程（MIP, mixed integer programming）[17]问题，如伪随机优化 PBOPT 问题求解，典型解析器有 SCIP[18,19]、Gurobi[20]。

前人代数旁路攻击研究中[8,11]，解析器大都只能输出一个唯一的满足解或者最优解。本文研究表明：基于 SCIP 解析器可以输出多个可能解集合，且正确解总在该集合中。笔者还发现 MIP 中的线性编程（LP, linear programming）问题也可用于代数旁路攻击。本文分别选取 CryptoMiniSAT（用于 SAT 问题求解）、SCIP（用于 PBOPT、LP 问题求解），对 LED 代数旁路攻击中建立的方程组进行求解，并对二者的执行效率进行了对比分析。

结果表明：在汉明重推断全部正确时，CryptoMiniSAT 解析器效率较高，攻击所需时间较小。汉明重推断部分错误时，SCIP 解析器能充分利用错误汉明重信息，输出多重解，攻击复杂度较低。

2　LED 代数旁路攻击基础

2.1　LED 算法设计

LED 算法分组长度为 64 bit，支持 64/128 bit 的密钥长度，加密轮数为 32 轮。本文主要针对 64 bit 长度密钥的 LED 密码算法。算法第一轮之前进行一次轮密钥加，以后每 4 轮进行一次轮密钥加操作。算法的 1 步包括 4 轮，每一轮包括轮常量加、S 盒、行移位和列混淆 4 个操作。LED 算法流程如图 1 所示。为提高加密速度及减小硬件实现规模，LED 算

法没有密钥扩展算法，轮密钥即为初始密钥。实现只需 966 个门电路，是同类分组密码中最少的，适于硬件实现但仍保留了合理的软件实现能力。算法状态采用 GF(2^4)上的 4×4 矩阵，每个元素 4 bit。

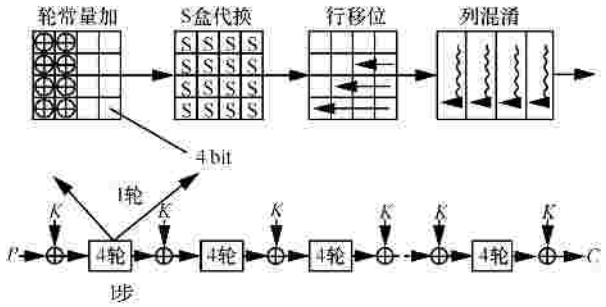

图 1 LED 算法流程

算法主要步骤如下。

1) 轮常量加 *AC*：6 bit 的轮常量参数(rc_5, rc_4, rc_3, rc_2, rc_1, rc_0)的初始值为 0，在每一轮使用前依次左移 1 bit，新的 rc_0 用 $rc_5 \oplus rc_4 \oplus 1$ 更新。与状态矩阵按位异或的轮常量矩阵如下：

$$\begin{bmatrix} 0\ (rc_5 \| rc_4 \| rc_3)\ 0\ 0 \\ 1\ (rc_2 \| rc_1 \| rc_0)\ 0\ 0 \\ 2\ (rc_5 \| rc_4 \| rc_3)\ 0\ 0 \\ 3\ (rc_2 \| rc_1 \| rc_0)\ 0\ 0 \end{bmatrix}$$

2) S 盒代换 *SB*：算法采用了 16 个 4 进 4 出的 S 盒，S 盒沿用 PRESENT 密码 S 盒，如表 1 所示。

表 1 LED S 盒输入输出

x	*S*[*x*]	*x*	*S*[*x*]
0	C	8	3
1	5	9	E
2	6	A	F
3	B	B	8
4	9	C	4
5	0	D	7
6	A	E	1
7	D	F	2

3) 行移位 *SR*：状态矩阵的第 *i* 行向左移 *i* bit，*i*=0,1,2,3。

4) 列混淆 *MC*：状态矩阵的每一列由混淆矩阵和该列向量相乘所得的新的向量替换更新。

$$\begin{vmatrix} 4 & 1 & 2 & 2 \\ 8 & 6 & 5 & 6 \\ B & E & A & 9 \\ 2 & 2 & F & B \end{vmatrix}$$

2.2 代数旁路攻击

代数旁路攻击可分为密码算法方程组表示、旁路泄露采集及利用、代数方程组求解 3 个步骤。

Step1 密码算法方程组表示

将密码算法表示为关于明文 *P*，中间变量 *X*，密文 *C*，密钥 *K* 的代数方程组。在 LED 代数方程组构造中，最为关键的部分是如何构造非线性部件 S 盒及复杂线性变换列混淆的代数方程组。本文选取文献[21]中布尔积方法来构建 LED 的 S 盒方程组。

Step2 旁路泄露采集及利用

给定一个密码算法的代数方程组，密钥恢复等价于代数方程组求解。由于直接进行密码代数方程组求解是一个 N-P 难题，故需要利用旁路泄漏信息加速方程组求解。实际攻击中，根据攻击者的测量能力和密码实现平台差异，攻击者选取一个密码实现旁路泄漏模型 *M*，然后通过采集旁路泄露信息 *L*，根据泄漏模型 *M* 将 *L* 转化为加密中间状态 *D*，最后将 *D* 使用额外的代数方程组表示出来。由于 LED 实现基于 4 bit，本文中 LED 代数旁路攻击基于 4 bit 汉明重泄漏模型实现。

Step3 代数方程组求解

将旁路泄露信息带入方程组之后，需要对其求解得到密钥。目前，典型的代数方程组求解主要包括基于 Grobner 基方法、线性化方法、SAT 问题、MIP 问题等方法。本文首先基于 SAT 问题进行代数方程组求解，选取 CryptoMinisat 解析器作为求解工具；然后基于 MIP 问题进行代数方程组求解，将代数方程组分别转化为基于 PBOPT 和 LP 的 MIP 问题，选取 SCIP 解析器作为求解工具。

3 LED 算法代数方程组构建

3.1 整体方程组设计

为建立 LED 算法加密代数方程组，结合算法结构，本文引入了 9 种类型的 64 bit 中间状态变量，对应标记及含义如表 2 所示。

表 2 新增变量及说明

变元	说明	变元	说明
P	明文	K	密钥
X_1	AK 输出	X_2	AC 输出
X_3	SB 输出	X_4	SR 输出
X_5	MC 输出	C	密文
RC	轮常量		

引入变元之后，结合 LED 具体算法设计，可建立其对应的布尔方程组，如图 2 所示。

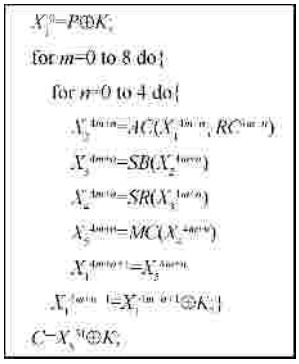

图 2　LED 代数方程组

3.2　S 盒非线性部件方程组设计

LED 算法中非线性 S 盒部件和列混淆部件对应的代数方程组构建是重点和难点问题。令 S 盒输入为(x_0, x_1, x_2, x_3)，输出为(y_0, y_1, y_2, y_3)，本文基于布尔积理论、参考文献[21]中的方法，将 LED 的 S 盒用 4 个等式表示，在 x_i 和 y_i（$0 \le i<4$）之外又引入 14 个变量。LED 的 S 盒对应代数方程组为

$$\begin{cases} 1+x_0+x_2+x_3+x_1x_2+x_0x_1x_3+x_1x_2x_3+x_0x_2x_3+y_0=0 \\ 1+x_0+x_1+x_0x_2+x_0x_3+x_2x_3+x_0x_1x_3+x_0x_2x_3+y_1=0 \\ x_0+x_2+x_0x_1+x_0x_2+x_0x_1x_3+x_0x_2x_3+x_1x_2x_3+y_2=0 \\ x_0+x_1+x_3+x_1x_2+y_3=0 \end{cases} \tag{1}$$

3.3　列混淆线性部件方程组设计

列混淆输出矩阵由混淆矩阵和输入状态矩阵相乘得到，根据混淆矩阵中的 11 个不同元素值，将 4 bit 输入(x_0, x_1, x_2, x_3)转化为对应 4 bit 输出(y_0, y_1, y_2, y_3)，不同矩阵元素对应输出如表 3 所示。

表 3　　LED 列混淆变换

元素	y_0	y_1	y_2	y_3
1	x_0	x_1	x_2	x_3
2	x_1	x_2	x_0+x_3	x_0
4	x_2	x_0+x_3	x_0+x_1	x_1
5	x_0+x_2	$x_0+x_1+x_3$	$x_0+x_1+x_2$	x_1+x_3
6	x_1+x_2	$x_0+x_2+x_3$	x_1+x_3	x_0+x_1
8	x_0+x_3	x_0+x_1	x_1+x_2	x_2
9	x_3	x_0	x_1	x_2+x_3
A	$x_0+x_1+x_3$	$x_0+x_1+x_2$	$x_0+x_1+x_2+x_3$	x_0+x_2
B	x_1+x_3	x_0+x_2	$x_0+x_1+x_3$	$x_0+x_2+x_3$
E	$x_0+x_1+x_2+x_3$	$x_1+x_2+x_3$	x_2+x_3	$x_0+x_1+x_2$
F	$x_1+x_2+x_3$	x_2+x_3	x_3	$x_0+x_1+x_2+x_3$

此外，考虑轮常量加、行移位等部件，每轮 LED 加密可引入 512 个变量，816 个 ANF 等式。最后加入轮密钥加操作后，全部 32 轮 LED 算法共引入 17 089 个变量，26 689 个 ANF 等式。

4　LED 运行泄露汉明重推断

4.1　LED 运行泄露汉明重选取

选取功耗泄露较为明显的算法操作对保证汉明重推断的准确率十分重要。实验中选取了轮密钥加、S 盒代换、列混淆变换输出 3 类功耗泄露，并进行汉明重推断，攻击利用的汉明重泄露点选取位置如图 3 所示。

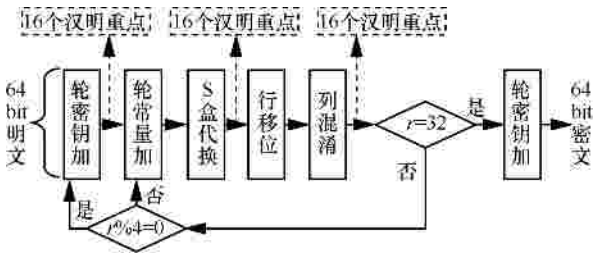

图 3　攻击利用汉明重泄露点选取

由图 3 可知，对于 LED 算法，每轮选取了 32 个汉明重泄露点，每 4 轮选取了 144 个汉明重泄露点，全部 32 轮共有 8×(16+4×(16×2))+16=1 168 个汉明重信息采集点。在实际攻击中，一般仅需采集部分轮的汉明重信息即可。

图 4 给出的是 LED 第一轮 S 盒代换对应的功耗曲线，可以看出存在 16 个较为明显的泄露点，分别对应 S 盒代换的 16 个 4 bit 的汉明重泄露。

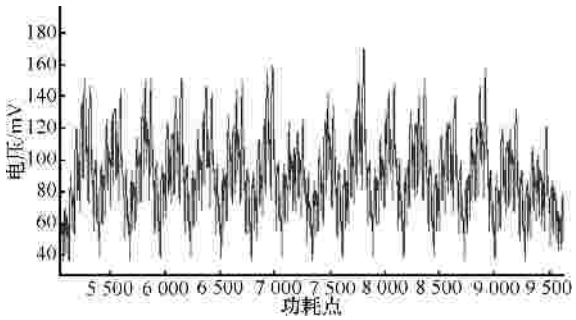

图 4　LED 第一轮 S 盒代换功耗泄露曲线

4.2　基于 Pearson 相关系数的汉明重推断

对于微控制器等平台上的密码实现，在处理某个中间状态，如 1 byte，其功耗大小常与处理状态的汉明重成正比，这也是很多经典的功耗分析攻击的基础[2,23]。由于 LED 是以 nibble（4 bit）为基本运

算单位的加密算法，因此其实现过程中会泄露出每 4 bit 的汉明重。为便于表述，后面提到的 LED 加密中字节均指 4 bit 的中间状态。

为推断出攻击的 LED 加密过程中各个字节的汉明重值，常需首先建立各个汉明重值对应的功耗轨迹模板，然后利用模板匹配方式进行汉明重推断。常见的汉明重推断方法有基于模板分析方法[22]、Pearson 相关系数方法[23]。本文采用后者方法。

假设汉明重功耗模板向量为 U，待匹配功耗向量为 V，分别由 m 个功耗点组成。则 U 和 V 之间的匹配度可通过 Pearson 相关系数进行计算

$$r<U,V>=\frac{\sum_{0}^{m}(U_x-\overline{U})(V_x-\overline{V})}{\sqrt{\sum_{0}^{m}(U_x-\overline{U})^2}\sqrt{\sum_{0}^{m}(V_x-\overline{V})^2}} \tag{2}$$

计算 V 和 5 个可能汉明重对应的功耗模板向量的匹配度后，匹配度最高的值作为汉明重推断值。为提高汉明重推断正确率，攻击中可通过对同一个选定明文加密进行多次采集，计算功耗曲线均值并同模板曲线进行匹配分析。

实际实验表明，每个样本加密 1 次对应汉明重推断成功率为 25%；重复加密 2 次的功耗均值曲线对应的汉明重推断成功率为 90%；重复加密 3 次的功耗均值曲线对应的汉明重推断成功率为 99.9%。

5 基于 3 种求解策略的代数方程表示

5.1 基于 SAT 问题的代数方程表示

基于 SAT 问题求解方程组的基本思想包括线性化方程组、将线性化方程组转化为 CNF 2 个步骤。其中最重要的部分是方程组的线性化问题，通过降幂来完成。非线性多元布尔方程组为

$$x_0\oplus x_1\oplus x_2\oplus x_3\oplus x_4\oplus x_1x_2x_3=0 \tag{3}$$

对于其中的高次单项式 $x_1x_2x_3$，为了达到降幂的目的，需引进一个未知变量 q，使得 $q=x_1x_2x_3$，降幂方法为

$$x_1x_2x_3=q\Rightarrow\begin{cases}x_1\vee\overline{q}=1\\ x_2\vee\overline{q}=1\\ x_3\vee\overline{q}=1\\ \overline{x_1}\vee\overline{x_2}\vee\overline{x_3}\vee q=1\end{cases} \tag{4}$$

得到线性方程组之后还需将其转化为长度一定的 CNF 子句，方便解析器求解。式(5)给出了切割方法，即引入中间变量 s，将方程组划分为最长为 4 的子句。

$$\begin{aligned}&x_0\oplus x_1\oplus x_2\oplus x_3\oplus x_4\oplus q=0\\ \Rightarrow&\begin{cases}x_0\oplus x_1\oplus x_2\oplus s=0\\ s\oplus x_3\oplus x_4\oplus q=0\end{cases}\end{aligned} \tag{5}$$

切割之后，将其表示为 CNF 子句。以式(5)中切割后的第 1 个式子为例，表示为 CNF 子句后为

$$\begin{aligned}&(\overline{x_0}\vee x_1\vee x_2\vee s)\wedge(x_0\vee\overline{x_1}\vee x_2\vee s)\wedge\\ &(x_0\vee x_1\vee\overline{x_2}\vee s)\wedge(x_0\vee x_1\vee x_2\vee\overline{s})\wedge\\ &(\overline{x_0}\vee\overline{x_1}\vee\overline{x_2}\vee s)\wedge(\overline{x_0}\vee\overline{x_1}\vee x_2\vee\overline{s})\wedge\\ &(\overline{x_0}\vee x_1\vee\overline{x_2}\vee\overline{s})\wedge(x_0\vee\overline{x_1}\vee\overline{x_2}\vee\overline{s})\end{aligned} \tag{6}$$

根据 4.2 节中的方法得到加密各个字节的汉明重后，攻击者需要将其转化为加密中间状态代数方程组进行表示。对于一个 4 bit 的值 $X=(x_0, x_1, x_2, x_3)$，假设其对应汉明重表示为 $H(X)$，则 $0\le H(X)\le 4$。易见，$H(X)$可用一个 3 bit 值来表示。定义 $Y=(y_0, y_1, y_2)$来表示 $H(X)$。x_0 和 y_0 分别表示 X 和 Y 的最高有效位，则 Y 可以用关于 X 的布尔函数表示为

$$\begin{cases}y_0=\prod_{i=0}^{3}x_i\\ y_1=\sum_{t=0}^{5}a_tx_ix_j\ (0\le i<j\le 3)\\ y_2=\sum_{i=0}^{3}x_i\end{cases} \tag{7}$$

由于 CryptoMinisat 可以自动将布尔方程组转化为 CNF 求解，所以线性化方程组之后不需要再进行切割及转化为 CNF 操作。实际攻击中首先将布尔方程组通过变量与索引值的代换，生成符合 CryptoMinisat 输入格式的文本文件，然后以命令行方式调用 crypt.exe 程序进行求解。

5.2 基于 PBOPT 问题的代数方程表示

代数旁路攻击可以被描述为一个最优化问题：给定密码算法、设备运行时的功耗信息和实际的功耗模型，则求解密钥的过程就是寻找一组密钥，该密钥运行时的功耗最接近于攻击者采集到的泄露功耗。PBOPT 的求解策略为设定一个基于布尔变量的目标函数和一组约束不等式，求满足约束不等式并使目标函数值最小的解。

伪布尔语言的表达能力很强，可以用简单的式子表达出复杂的约束条件，相对于 SAT 要简单。异或操作可以表示为

$$x_1 \oplus x_2 \oplus x_3 = 0 \Rightarrow -x_1 + \overline{x_1} x_2 + x_1 \overline{x_2} = 0 \quad (8)$$

对于高次单项式 $x_1x_2x_3$，为达到降幂的目的，引进未知变量 $q = x_1x_2x_3$ 后，可表示为

$$q - x_1x_2x_3 = 0 \quad (9)$$

例如对于一个 4 bit 的值 $X=(x_0, x_1, x_2, x_3)$，其汉明重 $H(X)$可直接使用一个整数 Y 表示

$$Y - x_0 - x_1 - x_2 - x_3 = 0 \quad (10)$$

在实际攻击中，由于噪声的存在，攻击者得到的汉明重推断常有一定的偏差，对于微控制器平台来说，汉明重推断偏差约为±1[10]。由于基于 SAT 的代数方程求解是基于严格的布尔问题进行的，任何比特的错误都将导致求解失败。而基于 PBOPT 问题的方程求解是基于整数编程策略的，如果汉明重有±1 的偏差，可在表示每个汉明重泄露时引入 2 个中间变量 t_0、t_1，式(10)可转化为

$$Y - x_0 - x_1 - x_2 - x_3 - t_0 + t_1 = 0 \quad (11)$$

此时，即使攻击者而得到的汉明重 Y 值是错误的，只要误差在±1 范围内，基于 PBOPT 问题进行方程求解仍可恢复密钥。

5.3　基于 LP 问题的代数方程表示

线性编程问题是最优化问题的一种，求满足约束条件并使目标函数最小或最大的解，和 PBOPT 问题相比不同的是，LP 问题的约束条件必须为线性约束。其汉明重表示方法和式(9)相同。降幂操作需要引入变量，对于单项式 x_1x_2，降幂方法如式(12)所示。式(13)给出了异或操作的表示方法。

$$x_1x_2 = q \Rightarrow \begin{cases} q - x_1 \le 0 \\ q - x_2 \le 0 \\ x_1 + x_2 - q \le 0 \end{cases} \quad (12)$$

$$x_1 \oplus x_2 \oplus x_3 = 0 \Rightarrow \begin{cases} x_1 + x_2 + x_3 \le 2 \\ -x_1 + x_2 + x_3 \le 2 \\ x_1 - x_2 + x_3 \le 2 \\ x_1 + x_2 - x_3 \le 2 \\ -x_1 - x_2 - x_3 \le 0 \\ x_1 - x_2 - x_3 \le 0 \\ -x_1 + x_2 - x_3 \le 0 \\ -x_1 - x_2 + x_3 \le 0 \end{cases} \quad (13)$$

由于 LP 问题也是基于整数编程求解策略，和 PBOPT 问题类似，也能在汉明重有错误的情况下进行方程组求解。和 PBOPT 问题不同，LP 问题可自然地将汉明重的偏差范围用不等式来表示，而无需引入新的中间变量。对于误差在±1 的汉明重偏差，式(10)基于 LP 问题可转化为式(14)。

$$-1 \le Y - x_0 - x_1 - x_2 - x_3 \le 1 \quad (14)$$

同 SAT 相比，基于 MIP 问题的 PBOPT 和 LP 求解还具有支持多重解输出的优点。特别是在采集汉明重值较少或者错误率较大情况下，基于 PBOPT 和 LP 问题求解常常会得到多组解，其中包含正确解；此时基于 SAT 问题的求解在搜寻到任意一个可满足解后即退出，而该解在很大概率上是错误解。

6　LED 代数旁路攻击实验

6.1　实验环境

实验中，以 8 bit AVR 微控制器 ATMEGA324P 为攻击对象，系统晶振为 20 MHz。为测量 LED 加密在某一时刻的能量功耗，在微控制器和稳压电源 GND 端之间串联了一个阻值为 18.2 O 的电阻，如图 5 所示。加密过程中适时提供触发信号以便示波器采集电阻两端电压，并通过 USB 数据线将采集到的功耗轨迹传到 PC 机。实验中，电压设置为 5 V，微控制器工作频率为 8 MHz，示波器采样频率为 100 MS/s。表 4 给出了攻击中所用设备的性能参数。

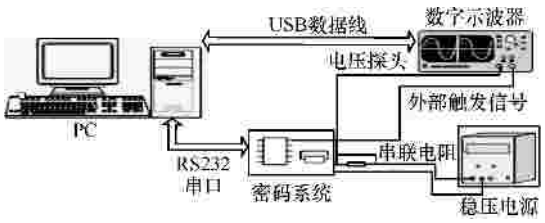

图 5　功耗采集实验环境

由表 4 可知，实验中使用了 2 种方程组求解器。其中，CryptoMinisat 解析器是基于 MiniSAT 内核开发的、专门适用于密码分析的 SAT 解析器，具有求解速度快，可靠性高的优点。SCIP 解析器基于整数编程和约束编程问题来解决优化问题，常将问题通过算法分解为子问题后线性求解每个问题，最后将结果组合起来。其中的线性求解部件是一个独立的 LP 解析器 SoPlex。为规范攻击成功率，当求解时间超过 3 600 s 时，视为攻击失败。

表 4 ASCA 设备的性能参数

名称	性能参数
PC 机	Athlon64 3000+ CPU、1.81 GHz、1 GB 内存、Windows XP 操作系统。
微控制器	ATMETGA324P 单片机：1 KB EEPROM, 2 KB SRAM，32 KB Flash，20 MHz 时钟频率。
数字示波器	MSO6012A：最大采样速率 2 GSa/s，最小电压分辨率 0.312 5 mV。
解析器	CryptoMinisat：版本号 2.9.0； SCIP 版本号 2.0.1，其中，SoPlex 版本号为 1.5.0。

6.2 基于 SAT 的密钥恢复

1) 基础攻击实验

攻击实验考虑了已知明文和未知明密文两种攻击场景，结果如表 5 所示。已知明文时，利用初始轮密钥加、第一轮 S 盒代换和列混淆变换共 48 个汉明重泄露即可求解出 LED 完整密钥；未知明密文条件下 5 轮汉明重泄露可恢复全部密钥。

表 5 不同攻击场景下实验结果

攻击场景	泄露轮数
已知明文	1
未知明密文	5

下面给出已知明文时的一个攻击实例。攻击中，首先产生随机明文 P=0x33E330861ECFC43F，密钥 K=0x395457051 C9B42A8，根据第 3 节方法建立 LED 代数方程组；然后利用示波器采集 LED 加密功耗，根据 4.2 节方法推断出初始轮密钥加、第一轮 S 盒代换和列混淆泄露的汉明重（如表 6 所示）；利用 5.1 节方法将汉明重表示为代数方程；使用 CryptoMinisat 进行密钥求解，结果如表 7 所示。

表 6 实例中汉明重泄露值

加密部件	泄露的汉明重值
初始轮密钥加	0,2,3,3,2,3,1,2,0,1,2,1,1,2,2,3
第一轮 S 盒代换	2,4,1,3,3,2,2,3,2,2,0,2,1,3,3,3
第一轮列混淆	2,1,1,4,3,2,2,2,1,2,2,1,2,1,1,2

表 7 中编号 2 至 65 依次表示 64 bit 的初始密钥比特值，编号为正表示对应密钥位为 1，否则为 0。根据表 7，求解出的密钥为 0011100101010100010101110000010100011100100110110100001010101000，转化为 16 进制即 0x395457051 C9B42A8，同真实密钥一致，攻击成功。

表 7 实例中 CryptoMinisat 密钥求解结果

编号	K	编号	K	编号	K	编号	K
- 2	0	- 18	0	- 34	0	- 50	0
- 3	0	19	1	- 35	0	51	1
4	1	- 20	0	- 36	0	- 52	0
5	1	21	1	37	1	- 53	0
6	1	- 22	0	38	1	- 54	0
- 7	0	23	1	39	1	- 55	0
- 8	0	24	1	- 40	0	56	1
9	1	25	1	- 41	0	- 57	0
- 10	0	- 26	0	42	1	58	1
11	1	- 27	0	- 43	0	- 59	0
- 12	0	- 28	0	- 44	0	60	1
13	1	- 29	0	45	1	- 61	0
- 14	0	- 30	0	46	1	62	1
15	1	31	1	- 47	0	- 63	0
- 16	0	- 32	0	48	1	- 64	0
- 17	0	33	1	49	1	- 65	0

2) 考虑汉明重推断错误的攻击

实际攻击中，受测试计量仪器精度、被测平台噪声等因素的影响，旁路信息的采集会存在一定误差，导致汉明重推断错误。由于代数方程组求解对输入要求十分严格，1 bit 的输入错误即可导致解析器无解。为此，本文通过对应用 4.2 节方法计算出的 5 个可能汉明重匹配度进行分析，如果最大匹配度低于某阈值（该值的选取决定于攻击平台，本文实验中选取为 0.9），即可视为汉明重推断错误，然后将其丢弃，不再转化为方程组用到攻击中。

应用上述策略，攻击利用的汉明重泄露点随机分布在加密轮中，由于未充分利用每轮的所有汉明重点，此时实验成功所需攻击轮数有一定增加。实验中，定义推断失败的汉明重数量和攻击利用的总的汉明重数量之比为错误率，表 8 给出了不同错误率下攻击所需轮数、成功率和平均求解时间。

表 8 离散条件下错误率与所需汉明重轮数

错误率/%	轮数	成功率/%	时间/s
40	2	90	148
50	3	100	1.2
60	3	100	18
70	4	90	150
80	6	50	250

由表 8 易见，在汉明重推断错误率高达 80%时，通过采集 1 条 LED 加密功耗曲线 6 轮的汉明重泄

露仍可快速恢复 64 bit 完整密钥。

6.3　基于 PBOPT 和 LP 的密钥恢复

1) 基础攻击实验

在汉明重推断正确的理想条件下，已知明文时利用 1 轮汉明重泄露可求解出 LED 密钥；未知明密文时 PBOPT 5 轮、LP 6 轮汉明重泄露可恢复全部密钥。结果如表 9 所示。

表 9　不同攻击场景下实验结果

攻击场景	错误率	泄露轮数	求解时间/s
PBOPT	已知明文	1	2.5
	未知明密文	5	379
LP	已知明文	1	2.8
	未知明密文	6	53

在已知和未知明密文条件下，和 CryptoMinisat 相比，SCIP 的 2 种求解策略耗时都较长，这是因为 PBOPT 和 LP 都属于优化问题，在汉明重信息足够时求解过程较可满足性问题复杂，从而耗时较长。

2) 考虑汉明重推断错误的攻击

当汉明重推断错误时，采取和 SAT 问题不同的策略，当汉明重值的推断错误在±1 范围内时可以继续代入方程组求解，即可以容忍部分错误存在，这对于存在一定推断错误概率的实际攻击有很现实的意义。在 1 轮汉明重泄露时，错误率、多重解的个数及平均求解时间如表 10 所示。

表 10　错误率与多重解的个数

攻击场景	错误率	个数	时间/s
PBOPT	10%	2.8	4.7
	20%	7.4	9.2
	30%	86.2	21.6
	40%	319	34
	50%	3 575	99.7
LP	10%	2.2	11.7
	20%	8.8	34.7
	30%	161.6	88
	40%	1 748	211.1
	50%	6 618.7	552.1

在汉明重推断存在错误的情况下，由于 CryptoMinisat 解析器基于可满足性策略进行方程组求解，若解不唯一则只给出满足要求的一个解，求解成功率较低，所以攻击中需要有足够的汉明重信息加入方程组，以便使唯一的可满足解就是正确解。随着错误率增加，基于 SAT 的攻击所需汉明重轮数也相应增加。此时 SCIP 解析器表现出了一定优势，基于 PBOPT 策略时仅需 1 轮汉明重信息就可在最多 50%的错误率下给出 3 575 个多重解，其中包含正确解，即将密钥空间从 2^{64} 降低到了 $2^{11.8}$；基于 LP 策略时仅需 1 轮汉明重信息可在最多 50%的错误率下将密钥空间从 2^{64} 降低到 $2^{12.7}$。随着错误率的增大，基于 PBOPT 问题的求解方法比基于 LP 问题的求解方法效率更高，多重解的个数更少。

6.4　攻击结果分析

通过实验结果可知，CryptoMinisat 解析器求解速度较高，可攻击轮数较多，适合可以获取多轮旁路泄露信息的情况；SCIP 解析器对错误汉明重的利用率较高，并能输出多重解最大化降低密钥空间，但对多轮代数方程组的求解速度较慢，适合仅有少量泄露信息并对时间要求较低的情况。在实际代数旁路攻击中可根据具体需求选用合适的解析器。

作为轻型分组密码，LED 需要在加密速度和安全性之间寻求平衡，而基于微控制器实现的 LED 由于串行运行且运行噪声较低，功耗泄露曲线较为明显，推断错误率较低，更易遭受代数旁路攻击。根据本文实验，未采取保护措施的 LED 实现仅需一个样本就能被破解，在实际使用中并不安全。这对密码防护提出了更高要求，使用者应该根据情况通过加入随机时延、随机掩码等方式来提高其安全性。

7　结束语

本文提出一种针对 LED 密码的代数旁路攻击方法，针对 8 bit ATMEGA324P 微控制器上的 LED 实现，通过功耗旁路采集获取加密中间状态汉明重，将其转化为额外方程组后利用多种代数方程组求解策略进行密钥恢复，开展了大量的攻击实验。结果表明：LED 易遭受代数旁路攻击；攻击所需样本量较小，在已知明文和未知明密文场景下 1 次 LED 加密部分轮汉明重泄露即可恢复全部初始密钥；此外，对于汉明重推断部分错误的情况也能成功实施攻击。鉴于代数旁路攻击对密钥安全性的极大威胁，给出攻击的防御措施是我们未来的研究方向。

参考文献：

[1] KOCHER P C. Timing attacks on implementations of diffie-hellman, RSA, DSS, and other systems [A]. CRYPTO 1996[C]. Berlin: Springer, 1996.104-113.

[2] KOCHER P C, JAFFE J, JUN B. Differential power analysis[A]. CRYPTO 1999[C]. Berlin: Springer, 1999. 388-397.

[3] QUISQUATER J J, SAMYDED D. Electromagnetic analysis (EMA): measures and countermeasures for smart cards[A]. e-Smart 2001[C]. Berlin: Springer, 2001.200-210.

[4] BATINA L, GIERLICHS B, PROUFF E. Mutual information analysis: a comprehensive study[J]. Journal of Cryptology, 2011, (24):269-291.
[5] WHITNALL C, OSWALD E, MATHER L. An exploration of the Kolmogorov-Smirnov test as competitor to mutual information analysis[EB/OL]. http://eprint.iacr.org/2011/380.pdf.
[6] BOGDANOV A. Improved side-channel collision attacks on AES[A]. SAC 2007, LNCS 4876[C]. Berlin:Springer, 2007.84-95.
[7] DINUR I, SHAMIR A. Side channel cube attacks on block ciphers[EB/OL]. http://eprint.iacr.org/2009/127.
[8] RENAULD M, STANDAERT F -X. Algebraic side-channel attacks[A]. INSCRYPT 2009[C]. Berlin:Springer,2009.393-410.
[9] RENAULD M, STANDAERT F X, VEYRAT C N. Algebraic side-channel attacks on the AES:Why time also matters in DPA[A]. CHES 2009[C]. Berlin:Springer, 2009.97-111.
[10] OREN Y, KIRSCHBAUM M, PPOPP T. Algebraic side-channel analysis in the presence of errors[A]. CHES 2010[C]. Berlin:Springer, 2010.428-442.
[11] ZHAO X J, ZHANG F, GUO S Z. MDASCA:an enhanced algebraic side-channel attack for error tolerance and new leakage model exploitation[A]. Proceedings of COSADE 2012[C]. 2012.
[12] GUO J, PEYRIN T, POSCHMANN A. The LED bock cipher[A]. CHES 2011[C]. Berlin:Springer, 2011.326-341.
[13] NICOLAS T. C, GREGORY V. B. Algebraic cryptanalysis of the data encryption standard[A]. 11-th IMA Conference[C]. Cirencester, UK, 2007. 152-169
[14] FU Z, MARHAJAN Y, MALIK S. zChaff SAT solver[EB/OL]. http://www.princeton.edu/ chaff/.
[15] E´EN N, SÖRENSSON N. An open-source SAT solver package[EB/OL]. http://www.cs.chalmers.se/Cs/Research/FormalMethods/Mi iSat/.
[16] SOOS M, NOHL K, CASTELLUCCIA C. Extrending SAT solvers to cryptographic problems[A]. SAT 2009[C]. Berlin:Springer. 2009.244-257.
[17] TOBIAS A. Constraint Integer Programming[D]. TU Berlin, 2007.
[18] BERTHOLD T, HEINZ S, PFETSCH M E. SCIP- solving constraint integer programs[A]. SAT 2009[C]. Berlin:Springer, 2009.
[19] BERTHOLD T, HEINZ S, PFETSCH M E. Nonlinear pseudo-boolean optimization:Relaxation or propagation? [A]. SAT 2009[C]. Berlin:Springer, 2009.441–446.
[20] CARLA P. G, ASHISH S, Handbook of Satisfiability[M]. IOS Press, 2009.633-654.
[21] KNUDSEN L R, MIOLANE C V. Counting equations in algebraic attacks on block ciphers[J]. International Journal of Information Security, 2010, 9(2):127-135.
[22] CHARI S, RAO J R, ROHATAJ P. Template attacks[A]. CHES 2002[C]. Berlin:Springer, 2002.13-28.
[23] BRIER E, CLAVIER C, OLIVIER F. Correlation power analysis with a leakage model[A]. CHES 2004[C]. Berlin:Springer, 2004. 16-29.

作者简介：

冀可可（1988-），女，河南上蔡人，军械工程学院硕士生，主要研究方向为分组密码代数旁路分析。

王韬（1964-），男，河北石家庄人，博士，军械工程学院教授、博士生导师，主要研究方向为信息安全、密码学。

郭世泽（1969-），男，河北石家庄人，博士，主要研究方向为信息安全、密码学。

赵新杰（1986-），男，河南开封人，军械工程学院博士生，主要研究方向为分组密码旁路分析和故障分析。

刘会英（1984-），男，湖北黄石人，军械工程学院博士生，主要研究方向为图像加密和密码旁路分析。

（上接第 133 页）

[19] SHEHORY O, KRAUS S. A kernel-oriented model for coalition formation in general environments: implementation and res lts[A]. Proceedings of the 3th National Conference on Artificial ntelligence (AAAI-96)[C]. Anaheim, CA, 1996. 715-723.
[20] Java agent development framework (JADE)[EB/OL]. http://jade.tilab.com.

张霞（1965-），女，辽宁沈阳人，博士，东北大学教授，主要研究方向为软件工程、数据库、面向服务智能应用等。

作者简介：

邓寒冰（1984-），男，辽宁沈阳人，东北大学博士生，主要研究方向为软件 agent **与本体技术。**

刘积仁（1955-），男，辽宁沈阳人，博士，东北大学教授、博士生导师，主要研究方向为软件工程、数据库、计算机图形学、医学影像计算、网络安全与管理和嵌入式技术与系统等。