doi:10.3969/j.issn.1000-436x.2014.03.018

基于功耗预处理优化的 LED 密码模板攻击研究

王小娟[1]，郭世泽[2]，赵新杰[2,3]，宋梅[1]，张帆[4]

(1. 北京邮电大学 电子工程学院，北京 100876；2. 北方电子设备研究所，北京 100083；
3. 军械工程学院 信息工程系，河北 石家庄 050003；4. 康涅狄格大学 计算机科学与工程系，康涅狄格州 斯托斯 06269)

摘 要：对 CHES 2011 会议提出的轻量级分组密码 LED 抗功耗模板攻击能力进行了评估，从功耗曲线预处理优化的角度对模板攻击提出了改进：利用功耗曲线频域上的相位相关性计算偏移量，消除了模板构建过程中的数据干扰；利用明文片段对功耗曲线聚类划分的特征差异，提出了一种基于类间距离的特征提取方法，可实现不同泄露点的功耗数据自动切割；利用均值和噪声信息评估模板区分度，提出了一种基于聚类有效度的动态选点策略，提高了旁路信息利用率。实验结果表明：数据对齐和切割提高了匹配度的区分效果，降低了模板构建和攻击所需功耗曲线数量；聚类有效度选点策略与现有策略相比，攻击数据复杂度低，2 条功耗曲线即可使成功概率收敛于 1。

关键词：功耗预处理；数据对齐；数据切割；有效点选取；模板攻击；LED

中图分类号：TP393.08 **文献标识码**：A **文章编号**：1000-436X(2014)03-0157-11

Research of power preprocessing optimization-based template attack on LED

WANG Xiao-juan[1], GUO Shi-ze[2], ZHAO Xin-jie[2,3], SONG Mei[1] , ZHANG Fan[4]

(1. School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China;
2. Institute of North Electronic Equipment, Beijing 100083, China;
3. Department of Information Engineering, Ordnance Engineering College, Shijiazhuang 050003, China;
4. Department of Computer Science and Engineering, University of Connecticut, Storrs 06269, USA)

Abstract: The security of LED, a lightweight block cipher proposed in CHES 2011, was evaluated by the template attack (TA). Several improvements of TA from the perspective of the preprocessing optimization was proposed. Firstly, the noise offset was calculated by using the phase-only correlation factor in the frequency view of the power trace to eliminate the data interference in the template building phase. Secondly, a novel character extracting method was proposed based on calculating the cross-cluster offset of different clusters classified by the plaintexts to cut the different leakage points from the power traces automatically. Thirdly, a dynamic effective power points choosing strategy was proposed by utilizing the mean value and the noises of the of power traces to evaluate the differences between different templates and improve the utilization of side channel information. Experiment results demonstrate that the proposed techniques of data alignment and automatically data cutting enlarge the differences of templates and reduce the number of the required power trace in both the template building and attacking phase. The proposed effective power points choosing strategy reduces the data complexity of the attack and only two power traces are required to launch the attack with the success rate of 100%.

Key words: power preprocessing; data alignment; data cutting; interesting points selection; template attack; LED

1 引言

在现实世界中，密码算法总要依赖于一个物理载体来实现，如 PC、智能卡或者嵌入式处理器。密码在物理载体上运行时会产生时间[1]、功耗[2]、电磁[3]等旁路信息泄漏，这些泄露同密码运算中的数据或操作存在相关性，可用于密钥恢复。基于旁路泄露分析的密码攻击称之为旁路攻击。自从

收稿日期：2012-11-10；修回日期：2013-09-11
基金项目：国家自然科学基金资助项目(61173191, 61272491, 61309021)
Foundation Item: The National Natural Science Foundation of China (61173191, 61272491, 61309021)

Kocher 等[1]首次提出通过分析 RSA 实现过程中的时间信息泄露可恢复密钥的思想以来，旁路攻击越来越受到工业界和学术界的重视。

在旁路攻击中，功耗分析由于攻击效果的强大和攻击实施的便捷而引起广泛关注。Kocher 等[2]提出了经典的简单功耗分析(SPA, simple power analysis)和差分功耗分析（DPA, differential power analysis）方法；Brier 等[4]基于相关性分析方法，提出了一种 DPA 的变种：相关功耗分析（CPA, correlation power analysis）。在 SPA 中，攻击者直观分析功耗曲线波形特征，通过单条功耗曲线推断密钥；在 DPA 中，攻击者根据多条功耗曲线的特征，利用统计学的方法来预测中间值，通过中间值的一位或者多位将功耗曲线分为 2 个集合，然后检测 2 个均值曲线的差分轨迹是否出现尖峰来推断预测密钥是否正确。SPA 的前提十分严格，要求分析者对芯片内部的运算实现细节非常了解，并且运算的相关信息在功耗信号轨迹上能表现明显特征；DPA 仅仅利用中间值的一位或者几位，运算过程中其他位信息都当作噪声来对待，旁路信息利用率较低，攻击样本量较大。随着对功耗分析方法的改进，Chari 等[5]提出了模板攻击(TA, template attack)。TA 分为 2 个阶段：模板构建阶段，攻击者需获取同目标密码芯片相同的密码设备（模板设备），并能对模板设备使用不同密钥进行加密，建立不同密钥对应的功耗模板；模板匹配阶段，攻击者采集至少一条目标密码芯片的功耗轨迹，通过和预搭建模板进行匹配，利用计算二者的相似度来进行密钥推断。Rechberger 等[6]首次对 TA 的实际问题进行了阐述。Medwed 等[7]给出了 TA 的 ECDSA 实现。

TA 所需样本量较少，主要原因如下：1) 构建的模板中除了均值信息外，还将噪声因素进行考虑，功耗泄露利用率较高；2) 攻击利用最大似然估计作为相似度量，该方法与 CPA 中的相关性分析相比适用范围更广（相关性分析只能刻画不同功耗曲线间的线性依赖[8]）。如何降低模板构建和目标密码芯片攻击所需功耗曲线数量是 TA 研究的热门方向。相关研究主要从 2 个方面开展：一是消除功耗曲线的噪声和干扰，如 Gebotys 等[9]使用傅里叶变换消除噪声，Nakajima 等[10]在功耗曲线频域特征进行攻击；另一方面是通过有效点选取方法，减少模板构建和匹配过程中所用功耗点的数量[5]，降低功耗分析复杂度。在有效点选取中，一种策略是直接从功耗曲线中选取有效点，如 Rechberger 等[6]使用的累积均值差方法，Gierlichs 等[11]提出的 T 检验方法；另外一种策略是通过对功耗曲线点进行重新映射抽取而来，如 Archambeau 等[12]提出的主成分分析法[13] (PCA, principal component analysis)，Hastie 等[14]提出的 Fisher 线性判别分析(LDA, linear discriminant analysis)。

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 标签等轻量级设备上实现密码算法[15,16]已成为近年来密码研究的新热点。LED[17]是在 CHES 2011 上提出的轻量级分组密码，算法的设计充分借鉴了 AES 密码设计思想，抗差分、线性、代数攻击能力较强。未来，LED 算法有望被广泛应用到 RFID 标签、无线传感器设备等资源受限环境中，面临着功耗旁路攻击的现实威胁。

在 LED 旁路攻击研究方面，李玮等[18]、Jeong 等[19]、Jovanovic 等[20]使用基于差分故障分析方法对 LED 进行了密钥恢复，最好的结果为基于半字节故障模型 1 次故障注入恢复完整密钥[20]；Kreuzer 等[21]、赵新杰等[22]基于代数故障分析方法对 LED 进行了密钥恢复，结果表明基于字节故障模型 1 次故障注入分析仍可恢复完整密钥[22]，代数故障分析可用于评估故障攻击后的 LED 密钥搜索空间；冀可可等对基于汉明重模型[23]和碰撞模型[24]的 LED 密码代数功耗攻击进行了研究，在较低的数据复杂度下成功恢复 LED 密钥。在 LED 功耗模板攻击方面，并未看到相关工作。

本文主要基于功耗模板攻击，对功耗预处理优化方法进行了研究，并以 LED 密码为例进行了验证实验，主要研究贡献如下。

1) 提出了一种基于相位相关性计算偏移量的功耗点对齐方法，可以消除模板中的数据干扰。

由于触发机制、高级程序语言运行时间不定、插入随机空操作等防护措施和噪声因素影响，攻击者采集的功耗曲线很难在时域上绝对对齐。不同时间点上功耗噪声的交错混淆，已成为时域信号模板分析的一个致命缺陷。本文提出了一种新的功耗点对齐方法，首先利用傅里叶变换将功耗曲线从时域映射到频域空间，然后根据相位相关性计算 POC 函数，将函数的尖峰值作为相似性的度量，相似性最大时对应重组功耗曲线即为对齐后的曲线。

2) 提出了一种基于聚类中心距离的功耗特征

提取方法，可实现不同泄露点功耗曲线自动切割。

功耗攻击主要基于“分而治之”思想，通过某轮扩展密钥的不同片段（计算单元）在密码运算过程中的功耗曲线，分析得到每个扩展密钥片段值，在此基础上通过拼接恢复密钥。在模板攻击中，如何提取不同密钥片段对应的功耗曲线，对于提高攻击实用性十分重要。本文提出了一种新的功耗曲线切割方法，利用不同明文片段值对某段对应功耗曲线进行聚类划分，通过计算聚类中心距离辨识不同密钥片段对应功耗曲线的分割点，从而实现数据切割的自动化。

3) 提出了一种基于聚类有效度的动态选点策略，降低了模板攻击成功所需样本量。

有效点选取策略是降低模板攻击计算复杂度、提高模板区分度、减少攻击数据复杂度的关键。本文提出了一种新的选点策略，利用聚类有效度来评估模板之间区分度，通过规定候选集的最小距离来实现有效点的动态选择。LED 攻击实验表明，与现有选点策略相比，提出的动态选定策略提高了旁路信息利用率，降低了攻击所用的样本量。

2　LED 模板攻击基础

2.1　LED 算法设计

LED 算法[17]分组长度为 64 bit，支持 64/128 bit 的密钥长度，加密轮数为 32 轮。LED 算法流程如图 1 所示。首先进行轮密钥加操作，以后每 4 轮进行一次轮密钥加，其中，1 轮包括轮常量加、*S* 盒代换、行移位和列混淆 4 个操作。LED 算法将轮密钥作为初始密钥，没有密钥扩展算法，从而提高了加密速度，减小了硬件实现规模。LED 算法实现只需 966 个门电路，是同类分组密码中最少的，适于硬件实现并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。

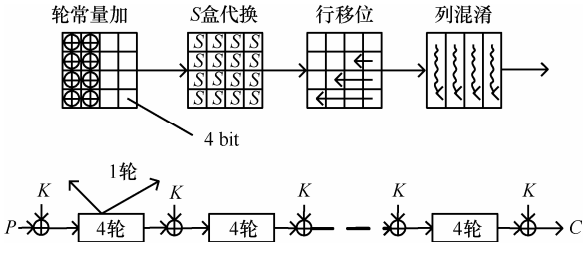


图 1　LED 算法流程

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*]
0	C
1	5
2	6
3	B
4	9
5	0
6	A
7	D
8	3
9	E
A	F
B	8
C	4
D	7
E	1
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　微控制器实现功耗泄露

微控制器中寄存器的基本结构单元是 CMOS 反相器，根据微控制器芯片功耗泄露机理，微控制器指令的功耗泄露特征与处理数据具有相关性，即功耗曲线与($m \oplus k$)的值相关，其中，m 为明文，k 为密钥，k_i 为第 i 个泄露点对应的子密钥。功耗攻击主要利用密码执行中间状态或操作和功耗泄漏之间的相关性进行密钥分析。图 2 给出了 LED 算法第一轮操作中的轮密钥加所对应的功耗曲线，以及重复加密 k_2 和 k_{15} 得到功耗曲线的均值曲线。可以看出，功耗曲线存在 16 个较为明显的泄露点；

当 $m_i \oplus k_i$ 取不同值时，功耗曲线具有差异性，而这种差异性是构建模板的前提。

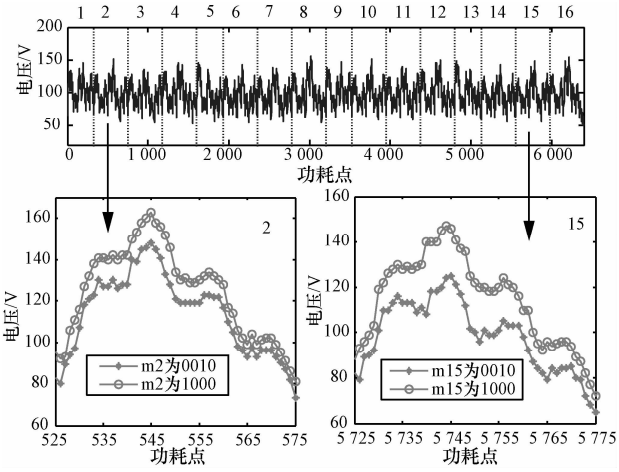


图 2 LED 中轮密钥加所对应的功耗泄露曲线

2.3 标准模板攻击

模板攻击利用旁路信息与被操作数据的相关性推断密钥，分为 2 步：第 1 步，假设攻击者能够掌控一台同目标密码服务器一样的模板密码服务器，并能对其使用已知密钥进行加密操作，根据设备运行过程中泄露的旁路信息构建模板；第 2 步，攻击者采集待攻击设备的一组旁路信息，与事先构建的模板进行匹配，相似度最高为猜测密钥。模板攻击的方法有很多种，下面介绍由 Rechberger 等[6]提出的一种比较经典的模板攻击。

2.3.1 模板构建

为对所有可能的某轮扩展密钥的第 m 个片段 k_m 构建模板$(\mu(k_m), \boldsymbol{\Sigma}(k_m))$。

1) 使用已知某轮扩展密钥片段 k_m 为随机明文样本执行加密操作，得到功耗曲线 $T=\{t_{ij}(k_m)\}$，其中，$t_{ij}(k_m)$表示第 i 个样本的第 j 个功耗点，设模板曲线的数目为 N_c。

2) 均值模板 $\mu(k_m)$的第 j 个元素的值计算如下

$$\mu_j(k_m)=\frac{1}{N_c}\sum_{i=1}^{N_c} t_{ij}(k_m) \tag{1}$$

3) 噪声模板 $\boldsymbol{\Sigma}(k_m)$是协方差矩阵，第 u 行 v 列元素计算如下

$$\boldsymbol{\Sigma}_{uv}(k_m)=\frac{1}{N_c-1}\sum_{i=1}^{N_c}(t_{iu}(k_m)-\mu_u(k_m))(t_{iv}-\mu_v(k_m)) \tag{2}$$

2.3.2 模板匹配

对于待攻击设备，模板匹配步骤如下。

1) 采集功耗曲线 $T=\{t_{ij}\}$，设攻击曲线数目为 N_a；

2) 假设某轮扩展密钥的片段为 k_m 的前提下，计算功耗曲线 t_i 和模板$(\mu(k_m), \Sigma(k_m))$的匹配概率：

$$p(t_i \mid k_m)=(2\pi)^{-\frac{N_p}{2}}\left|\boldsymbol{\Sigma}(k_m)\right|^{-\frac{1}{2}} \mathrm{e}^{-\frac{1}{2}(t_i-\mu(k_m))^T \boldsymbol{\Sigma}_{k_m}^{-1}(t_i-\mu(k_m))} \tag{3}$$

其中，N_p 为功耗点数目；

3) 根据贝叶斯原理，计算功耗曲线 T 的条件下密钥为 k_m 的概率为

$$p(k_m \mid T)=\frac{(\prod_{i=1}^{N_a} p(t_i \mid k_m))\cdot p(k_m)}{\sum_{k'=1}^{N_{k_m}}((\prod_{i=1}^{N_a} p(t_i \mid k_m'))\cdot p(k_m'))} \tag{4}$$

其中，N_{k_m} 为候选密钥数目；

4) 选定匹配概率 $p(k_m|T)$最高的为猜测密钥。

3 问题提出

本节从功耗预处理优化的角度(如图 3 所示)分析模板攻击中存在的问题。

3.1 对齐问题

模板攻击的核心思想是匹配，这要求每条功耗曲线在时间线上是完全对齐的。然而由于防护措施和噪声等因素影响，导致密码算法执行同一个运算对应的功耗点泄漏在时域上很难做到绝对对齐。在模板攻击过程中，可将功耗曲线看作多维变量，功耗点数据在时域上的不对齐特性将导致属于不同运算的功耗点在时域上产生交错，进而降低了模板准确性。如何克服功耗曲线在时域上的不对齐问题是改进模板攻击的一个关键问题。

3.2 切割问题

功耗攻击通过拼接每个扩展密钥片段值恢复密钥，这就需要将不同密钥片段使用时对应的功耗曲线提取出来。相同的运算对应的功耗曲线在波形上也是相似的，仅凭观测波形很难实现精确切割。如何提取相同运算在功耗曲线上的特征，实现数据切割自动化也是真实攻击中需解决的问题。

3.3 选点问题

对于给定的均值、噪声模板，攻击者需要选定功耗曲线中特征明显的点作为模板攻击的有效点。有效点需充分反映不同模板之间的差异性，否则将会削弱功耗曲线的统计特征，降低匹配的精度。累积均值差[6]通过计算均值模板的差值和来选定有效点，但是没有考虑噪声的影响；T 检验方法[11]和

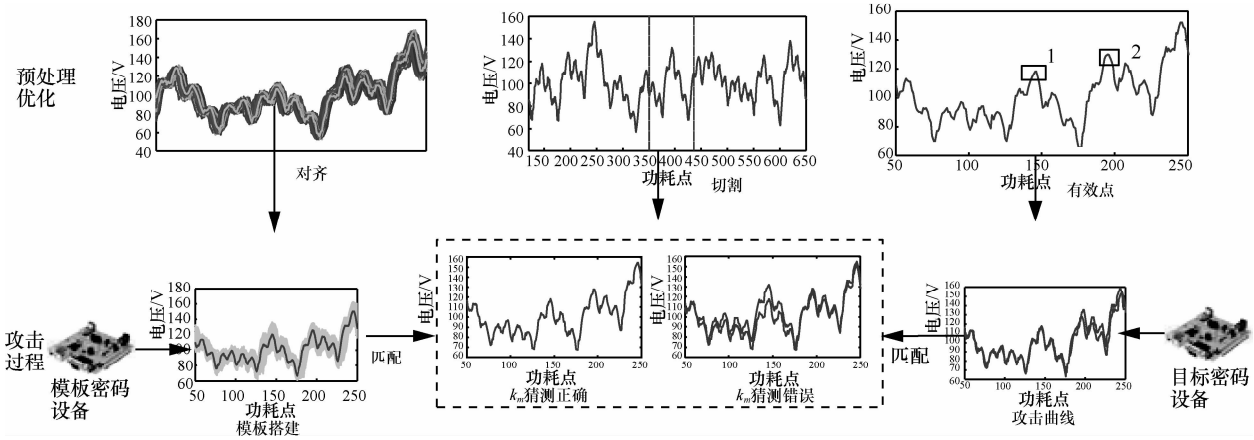


图 3　问题提出

PCA[12]综合考虑了均值和噪声，但对噪声的利用仅限于样本的方差；PCA[12]和 LDA[14]通过对所有功耗点进行降维，但是所有点的重构削弱了特征明显功耗点的区分效果。如何利用功耗曲线信息合理选择有效点，增强匹配效果是改进模板攻击的一个重要方向。

4　模板攻击改进

4.1　数据对齐

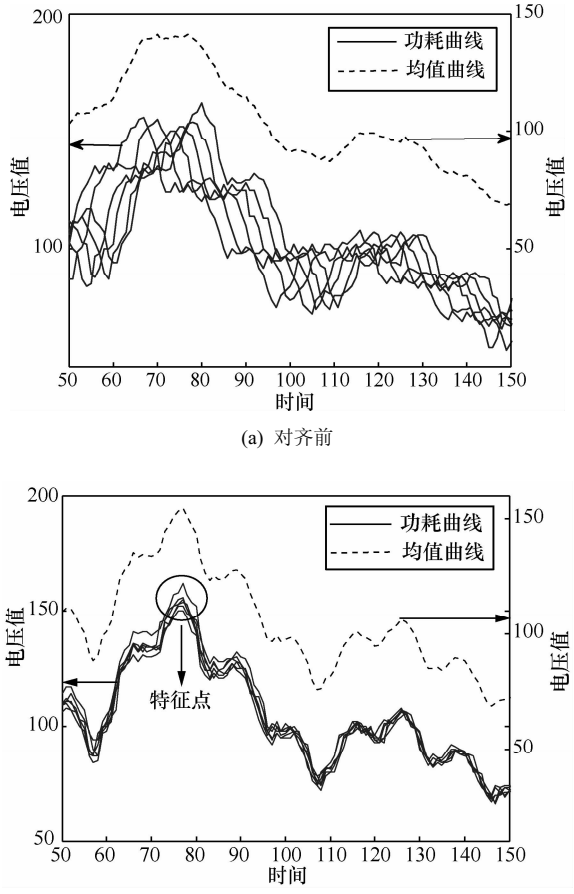
功耗曲线的位移误差通常都不大，但是会造成信息损失。图 4 给出了随机选择的 6 条功耗曲线，采用双坐标显示（实线对应左纵轴，虚线对应右纵轴），可以看出，在时间轴上存在偏移量。对齐后的均值曲线(图 4(b))相比对齐前(图 4(a))存在明显特征点，而明显的特征点是模板构建的重要前提。所以在实施模板攻击之前，需要对数据进行预处理，实现数据对齐，保证均值、协方差等统计量具有实际意义，提高模板的准确性。

本节利用相位相关法(POC, phase-only correlation)[25]来计算功耗曲线的偏移量，POC 是图像匹配中的常用方法，通过波峰位置来确定偏移量，对噪声的容忍度较高。POC 的理论基础是傅里叶变换，令 $f(n)$和 $g(n)$为长度相同的功耗曲线，其中，$n\in[-M, M]$。对 $f(n)$和 $g(n)$进行离散傅里叶变换(DFT, discrete Fourier transforms)

$$F(k)=\sum_{n=-M}^{M} f(n)W_N^{kn}=A_F(k)\mathrm{e}^{\mathrm{j}\theta_F(k)} \tag{5}$$

$$G(k)=\sum_{n=-M}^{M} g(n)W_N^{kn}=A_G(k)\mathrm{e}^{\mathrm{j}\theta_G(k)} \tag{6}$$

其中，$W_N=\mathrm{e}^{-\mathrm{j}(2\pi/N)}$，$A_F(k)$和 $A_G(k)$表示振幅，$\mathrm{e}^{\mathrm{j}\theta_F(k)}$ 和 $\mathrm{e}^{\mathrm{j}\theta_G(k)}$ 表示相位。POC 计算方法如下



(a) 对齐前

(b) 对齐后

图 4　功耗曲线对齐前后对比

$$POC(f,g)=\frac{1}{N}\sum_{-M}^{M}\frac{F(k)\overline{G(k)}}{\left|F(k)\overline{G(k)}\right|}W_N^{-kn} \tag{7}$$

其中，$\overline{G(k)}$ 是 $G(k)$的共轭，POC 实际上就是 $F(k)$和 $G(k)$相位余弦的反离散傅里叶变换(IDFT, inverse discrete Fourier transform)。如果 $f(n)$和 $g(n)$相似，那

么 POC 函数存在一个明显的尖峰。尖峰的值可以作为相似性的度量。下面给出数据对齐的具体方法。

1) 选定功耗曲线 t_1 作为基准曲线，滑动窗口设为$[-l, l]$。

2) 对于第 i 次功耗曲线 t_i，设偏移量 $\sigma\in[-l,l]$，选定长度为 L 的向量

$\boldsymbol{X}=\{ t_1(l), t_1(l+1),\cdots, t_1 (L+l-1)\}$

$\boldsymbol{Y}=\{ t_i(l+\sigma), t_i(l+\sigma+1),\cdots, t_i(L+\sigma+l-1)\}$

3) 计算 $POC(X, Y)$，得到不同偏移量 σ 对应的尖峰的值 $peak(\sigma)$。

4) 计算 t_i 的平移量 $move(t_i)=\max(peak(\sigma))$。

5) 根据每条功耗曲线的偏移量$move(t_i)$调整起始点，采集长度为 L 的功耗曲线作为对齐后的数据样本。

4.2 数据切割

设 k_i 为某轮扩展密钥的片段 i，m_i 为对应的第 i 个明文片段值，$t(k_i)$为 k_i 对应的功耗曲线。攻击者首先从采集的功耗曲线中提取 $t(k_i)$，然后利用模板攻击思想推断出 k_i，最后拼接 k_i 恢复密钥 k。观测轮密钥加的功耗曲线(如图 2 所示)，可以看出，密钥片段 k_i 操作的相似性，导致 k_i 所对应功耗曲线在波形上也是相似的。所以仅凭观测波形很难找到 k_i 的分割点。

根据功耗曲线与$(m\oplus k)$的依赖关系，当 k_i 相同的情况下，$t(k_i)$只和 m_i 相关。也就是说，依据 m_i 对 $t(k_i)$进行聚类，是有明显聚类效果的。如图 5 所示，依据 m_2 对 $t(k_2)$的划分，其中每条曲线代表同一 m_2 所对应的功耗曲线的均值。可以看出有明显的聚类效果(如图 5(a)所示)；而依据 m_2 对 $t(k_3)$的划分则没有这种聚类效果(如图 5(b)所示)。根据这个特性可以计算不同密钥片段对应功耗曲线的分割点，实现数据切割的自动化。

设 $\mu(k_i)$为 k_i 加密随机明文得到的功耗曲线均值，$\mu(k_i)|m_i$ 为 k_i 加密明文 m_i 得到的功耗曲线均值。将 $\mu(k_i)$作为聚类中心，$\mu(k_i)|m_i$ 作为子类中心，计算类间距离

$$d(m_i)=\sum_{m_i} d^2(\mu(k_i),\mu(k_i)\,|\,m_i) \tag{8}$$

$d(m_i)$越大，聚类效果越明显。下面基于 $d(m_i)$给出数据切割的方法。

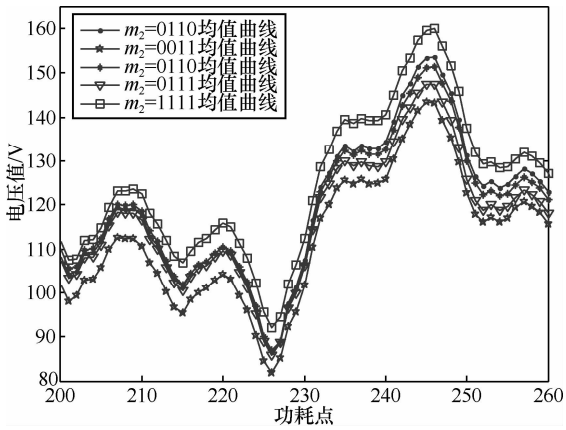
1) 从初始点开始，设 i=0，$x(i)$=1。

2) 以 $x(i)$作为起始点，对待切割功耗曲线进行数据对齐。

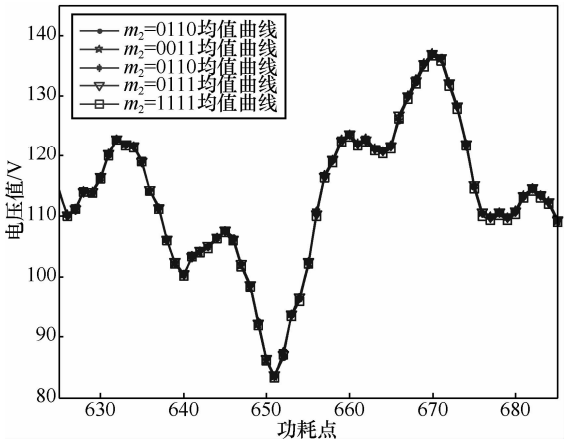
3) 计算依据 m_i 和 m_{i+1} 划分的类间距离的差值，选定 $d(m_i)-d(m_{i+1})$=0 作为数据的切割点，记为 $x(i+1)$。

4) 令 $i=i+1$，返回步骤 2)。

5) 记录切割出的 $t(k_i)$，起点为 $x(i)$，终点为 $x(i+1)$。



(a) m_2 对 $t(k_2)$的划分



(b) m_2 对 $t(k_3)$的划分

图 5 功耗曲线根据 m 的聚类划分

4.3 选点策略

为节省时间开销，需要选择一部分特征明显的功耗点即有效点来进行模板攻击。有效点选取应该满足以下条件[6]。

1) 选定的有效点所对应的模板之间应具有区分度，这样才能保证给定的功耗曲线仅能匹配一个模板即正确模板。

2) 有效点之间的距离不能太近，距离比较近的点包含的信息是类似的，造成资源的浪费。

如图 6 所示，深色曲线为均值曲线，浅色曲线为功耗曲线，从给出的 2 个模板的情况可以看出，特征点 3 的模板曲线之间没有区分，不能作为有效点；特征点 4 虽然模板曲线之间具有区分度，但是

对于给定的功耗曲线样本，2 个模板之间有重叠，导致在匹配时 2 个模板相互干扰；特征点 1 对应的模板之间没有重叠，区分度较好，适合作有效点；特征点 2 虽然也满足分度的条件，但是和特征点 1 包含信息具有重复性，所以不能同时选择二者作为模板攻击的有效点。

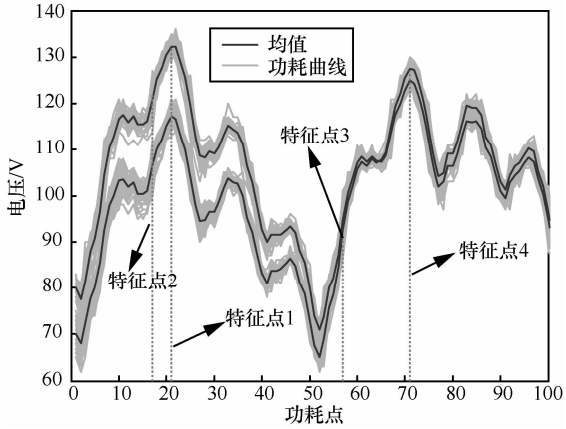


图 6　不同特征点选取

条件 1)可以通过评估聚类有效度来实现。将功耗曲线的均值 $\mu(k_i)$作为数据集的中心，将根据明文 m_i 划分得到的均值 $\mu(k_i)|m_i$ 作为子类的中心。对于给定划分，子类中心与数据集中心越分散，子类中的数据与子类中心越紧密聚类效果越明显，本文采用 Calinski-Harabasz（CH）指标来度量聚类划分的效果

$$CH=\frac{\frac{1}{(\sum_{m_i} n(m_i))-1}\sum_{m_i} n(m_i)d^2(\mu(k_i),\mu(k_i)\,|\,m_i)}{\frac{1}{L-(\sum_{m_i} n(m_i))}\sum_{m_i}\sum_{j\in C(m_i)} d^2(t_j,\mu(k_i)\,|\,m_i)} \tag{9}$$

其中，$n(m_i)$是明文 m_i 所对应的功耗曲线的数目。*CH* 越大说明聚类效果越好，聚类有效度不仅考虑了类内的差距还考虑了类间的差异。

条件 2)可以在 *CH* 计算的基础上，进行动态选点，使得有效点的候选集之间的距离大于下限值 *d*。

选点方法如下：首先计算所有点的 *CH* 值，并对其进行排序；按照从大到小的顺序对有效点进行判定，如果该点与候选集之间的距离大于 *d*，加入候选集，否则判定下一个。

5　LED 模板攻击实验

本节针对 64 bit 长度密钥的 LED 密码算法进行模板攻击实验，攻击对象为 8 bit AVR 微控制器 ATMEGA324P。

首先在微控制器和稳压电源 GND 端之间串联一个阻值为 18.2 Ω 的电阻；然后根据 LED 加密过程中提供的触发信号，利用示波器采集电阻两端电压；最后将采集到的功耗轨迹传到 PC 机。其中，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 2 给出了攻击实验中所用设备的性能参数。

表 2　　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

表 3 给出了实验分析中所涉及的符号。

表 3　　符号汇总

符号	意义	符号	意义
N_p	有效点个数	*S*	选点策略
N_c	模板曲线条数	*CH*	聚类有效度
N_a	攻击曲线条数	T-test	T 检验
N	实验次数	Mean	累积均值差
p_{succ}	攻击成功的概率	PCA	主成分分析法
l	滑动窗口长度	*d*	选点距离
p	功耗曲线片段数		

5.1　标准模板攻击实验

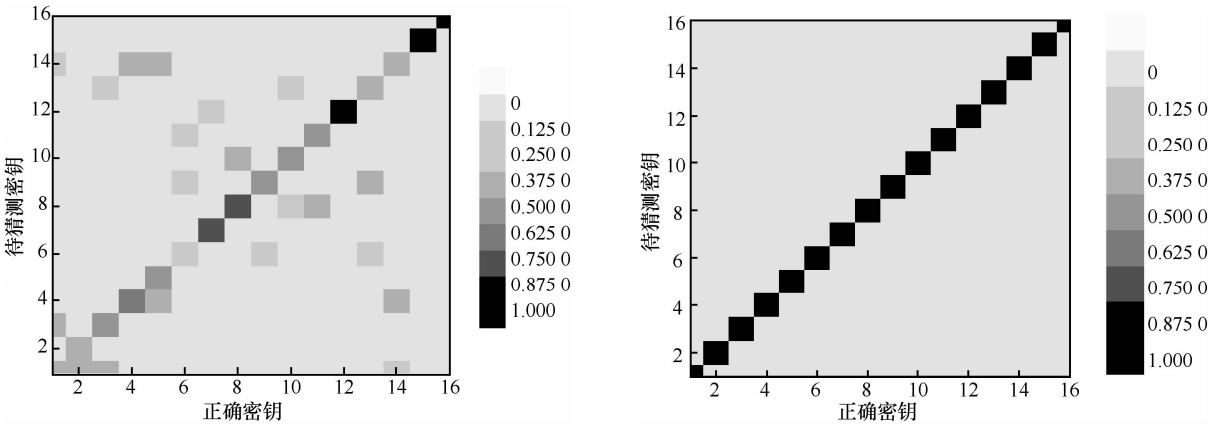
根据 2.3 节提出的标准模板攻击方法，对 LED 密码进行攻击，设定攻击参数为 p=2，N_p=30，N_c=200，N_a=10，S=Mean。攻击成功的结果如表 4 所示，表中每一行是原始密钥与 16 个匹配密钥计算得到的相似度，概率最高的为猜测密钥。可以看出，虽然正确密钥的匹配概率大于错误密钥，但是区分效果不是很明显。

5.2　数据对齐效果

为了分析功耗曲线的对齐效果，设定攻击参数为 p=2，N_p=30，N_c=200，N_a=10，S=Mean，l=70。从攻击结果可以看出，对齐后匹配概率(图 7(b))与对齐前(图 7(a))相比，区分度明显加强，并且正确密钥的匹配概率接近于 1。

表 4 功耗曲线标准模板攻击结果

原始密钥	匹配密钥															
	0000	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
0000	**0.31**	0.12	0.25	0.07	0.01	0.01	0	0	0.03	0	0	0	0.04	0.14	0	0
0001	0.32	**0.37**	0.12	0.01	0	0.03	0	0	0.06	0	0	0	0.02	0.08	0	0
0010	0.28	0.03	**0.39**	0.02	0	0.02	0	0	0.07	0	0	0	0.15	0.04	0	0
0011	0.02	0.01	0.02	**0.55**	0.11	0	0	0	0	0	0	0	0	0.3	0	0
0100	0	0	0	0.25	**0.48**	0	0	0	0	0	0	0	0	0.26	0	0
0101	0.03	0.01	0.09	0	0	**0.25**	0.04	0.04	0.2	0.01	0.22	0	0.12	0	0	0
0110	0	0	0	0	0	0.02	**0.75**	0	0.01	0	0	0.22	0	0	0	0
0111	0	0	0	0	0	0	0	**0.7**	0	0.27	0.03	0	0	0	0	0
1000	0.11	0.07	0.03	0	0	0.23	0	0	**0.49**	0	0	0	0.05	0.01	0	0
1001	0	0	0.02	0	0	0.02	0	0.14	0.03	**0.47**	0.08	0	0.24	0	0	0
1010	0	0	0.02	0	0	0.06	0.01	0.32	0.02	0.04	**0.48**	0	0.05	0	0	0
1011	0	0	0	0	0	0	0.11	0	0	0	0	**0.88**	0	0	0	0
1100	0.05	0.01	0.08	0	0	0.22	0	0	0.29	0.02	0.01	0	**0.29**	0.02	0	0
1101	0.17	0.08	0.07	0.32	0.01	0	0	0	0	0	0	0	0.01	**0.33**	0	0
1110	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**	0
1111	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**



(a) 对齐前 (b) 对齐后

图 7 对齐前后匹配概率对比

攻击曲线的条数 N_a 是衡量攻击效果的重要指标，攻击模型越强大，攻击成功所用 N_a 越少。设定 p=2，N_p=30，N_c=100，N=100，S=Mean，l=70，N_a 与 p_{succ} 之间的关系如图 8 所示。可以看出对于相同 p_{succ}，对齐后所用 N_a 明显少于对齐前，数据对齐后仅需要 6 条功耗曲线，p_{succ} 就收敛于 1，与对齐前的 66 条相比，降低了攻击所用样本量。

模板曲线条数 N_c 代表了模板的顽健性，模板信息提取越精确，攻击成功所用 N_c 越少。设定 p=2，N_p=30，N_a=10，N=100，S=Mean，l=70，N_c 与 p_{succ} 之间的关系如图 9 所示。可以看出：1)对于相同 p_{succ}，对齐后所用 N_c 明显少于对齐前；2) 随着 N_c 的增加，对齐后 p_{succ} 的收敛速度明显快于对齐前。数据对齐后相同时间点上的功耗对应的是相同操作，均值、噪声能够反映功耗曲线的统计特征，用较少的模板曲线就能提取足够的信息。

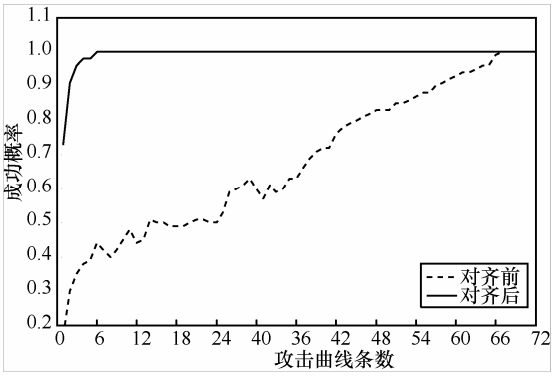


图 8　对齐前后 N_a 与 p_{succ} 关系

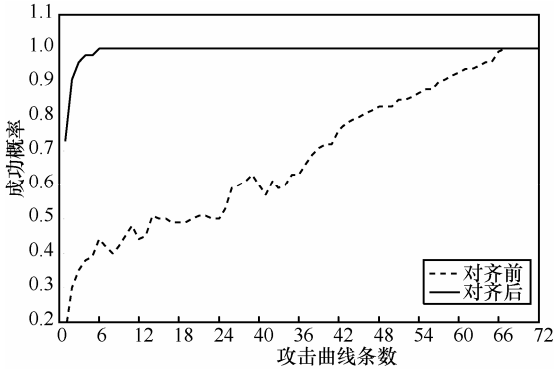


图 9　对齐前后 N_c 与 p_{succ} 关系

5.3　数据切割效果

图 10 给出了功耗曲线 $t(k_2)$和 $t(k_3)$的切割效果，可以看出，类间距离的差值存在明显的分割点。

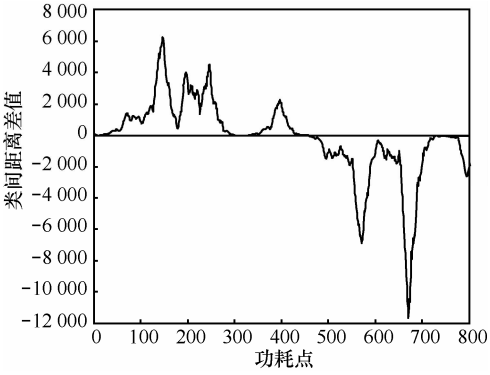


图 10　功耗曲线 $t(k_2)$和 $t(k_3)$的切割效果

基于类间距离将功耗曲线切割成 16 个片段(如图 11 所示)，其中，第 1、5、9 和 13 片段长度为 325，其余 12 组长度均为 425。长度不一致的原因是由于加密代码循环所致。

对 $t(k_2)$切除 100 个点之后和 $t(k_1)$趋势相似(如图 12 所示)，也就是说 16 个片段的操作从本质上来说是类似的，即具有内在联系。

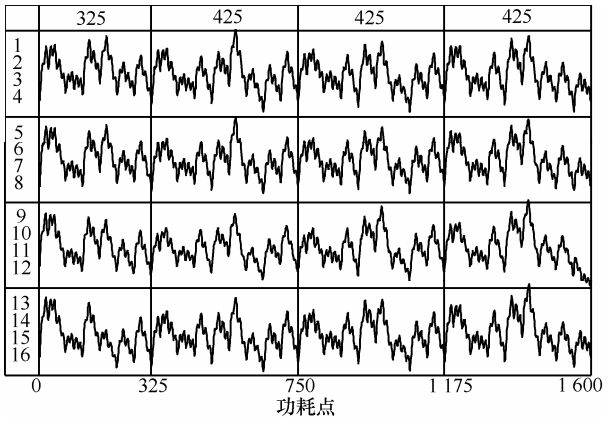


图 11　数据切割效果

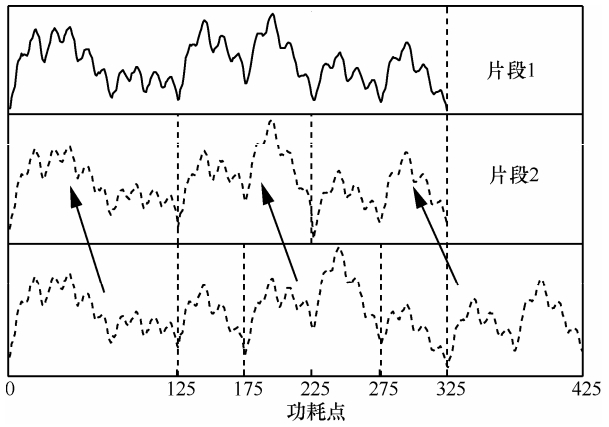


图 12　功耗曲线 $t(k_1)$和 $t(k_2)$的内在关系

5.4　有效点选取效果

为了分析 CH 选点策略效果，同 Mean[6]、T-test[11]、PCA[13]方法做对比。设定攻击参数为 p=2，N_c=100，N_a=2，N_p(CH,T-test,Mean)=30，N=100，d=3，N_p(PCA)=10，不同选点策略下的攻击结果如图 13 所示。可以看出，除正确密钥 k_2=0010 的情况，*CH* 的 p_{succ} 略低于 Mean 的外，CH 的 p_{succ} 最高，并且相对稳定。

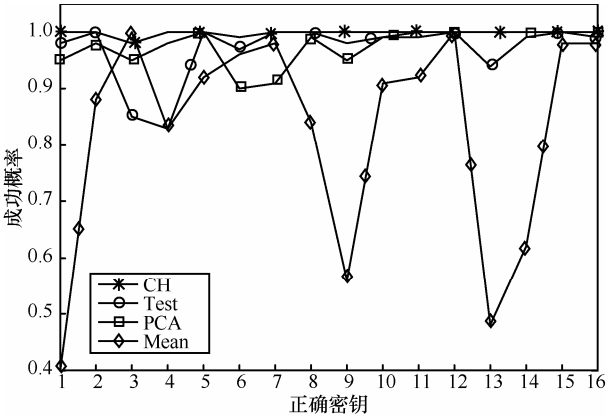


图 13　不同选点策略攻击结果

设定攻击参数为 p=2，N_p(CH,T-test,Mean)=30，N_c=100，N_p(PCA)=10，N=100，d=3，N_a 与 p_{succ} 之间的关系如图 14 所示。可以看出，CH 利用均值和噪声信息，信息利用率最高，效果最好；Test 和 PCA 利用均值和方差信息，效果次之；Mean 只利用了均值信息，效果最差。实验表明，利用 CH 选点策略，仅需 2 条攻击曲线即可令 p_{succ} 收敛于 1。

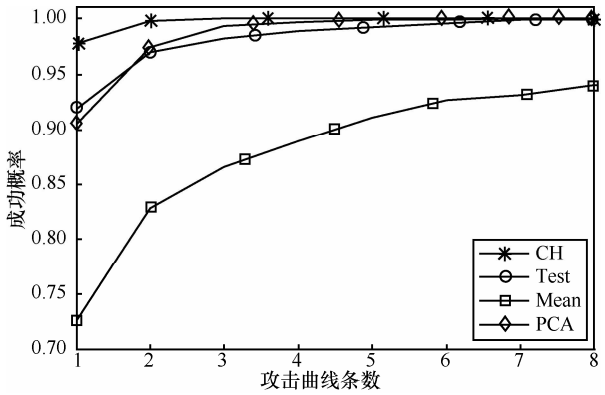


图 14 不同选点策略 N_a 与 p_{succ} 关系

设定攻击参数为 p=2，N_p(CH,T-test, Mean)=30，N_a=2，N_p(PCA)=10，N=100，d=3，N_c 与 p_{succ} 之间的关系如图 15 所示。可以看出，当 N_c<6 时，受到模板曲线数目的限制，噪声信息不能有效反映模板特征；当 N_c >6 时，攻击效果与选点策略的信息利用率正相关，CH 显示出优越性，20 条模板曲线即可令 p_{succ} 收敛于 1。

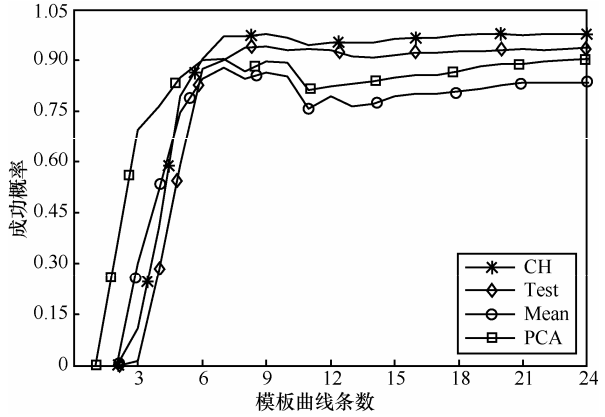


图 15 不同选点策略 N_c 与 p_{succ} 关系

6 结束语

本文从功耗预处理的角度，对数据对齐、数据切割、数据选点算法分别进行了改进，提高了模板攻击实用性。针对 LED 密码的验证实验表明：提出的数据对齐和切割算法增强了模板的攻击效果，提出的选点策略仅需 2 条曲线即可令攻击成功概率收敛于 1。

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, SAMYDE D. Electromagnetic analysis (EMA): measures and countermeasures for smart cards[A]. E-Smart 2001[C]. Berlin: Springer, 2001.200-210.

[4] BRIER E, CLAVIER C, OLIVIER F. Correlation power analysis with a leakage model[A]. CHES 2004[C]. Berlin: Springer, 2004.16-29.

[5] CHARI S, RAO J R, ROHATGI P. Template attacks[A]. CHES 2002[C]. Berlin: Springer, 2002.13-28.

[6] RECHBERGER C, OSWALD E. Practical template attacks[A]. WISA 2004[C]. Berlin: Springer, 2004. 440-456.

[7] MEDWED M, OSWALD E. Template attacks on ECDSA[A]. WISA 2008[C]. Berlin: Springer, 2008. 14-27.

[8] OLSSON U. Maximum likelihood estimation of the polychoric correlation coefficient[J]. Psychometrika, 1979, 44(4):443-460.

[9] WADDLE J, WAGNER D. Towards efficient second-order power analysis[A]. CHES 2004[C]. Berlin: Springer, 2004.1-15.

[10] GEBOTYS C H, HO S, TIU C C. EM analysis of rijndael and ECC on a wireless java-based PDA[A]. CHES 2005[C]. Berlin: Springer, 2005. 250-264.

[11] GIERLICHS B, LEMKE-RUST K, PAAR C. Templates vs stochastic methods[A]. CHES 2006[C]. Berlin: Springer, 2006.15-29.

[12] ARCHAMBEAU C, PEETERS E, STANDAERT F X, *et al*. Template attacks in principal subspaces[A]. CHES 2006[C]. Berlin: Springer, 2006.1-14.

[13] VIDAL R, MA Y, SASTR Y. Generalized principal component analysis (GPCA)[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 2005,(27): 1945-1959.

[14] ZEITOUNI K. A survey of spatial data mining methods databases and statistics point of views[A]. IRMA[C]. Alaska, USA, 2000.487-491.

[15] BOGDANOV A, KNUDSEN L R, LEANDER G, *et al.* PRESENT: an ultra-lightweight block cipher[A]. CHES 2007[C]. Berlin: Springer, 2007.450-466.

[16] HONG D, SUNG J, HONG S. HIGHT: a new block cipher suitable for low-resource device[A]. CHES 2006[C]. Berlin: Springer, 2006.46-59.

[17] GUO J, PEYRIN T, POSCHMANN A, *et al*. The LED block cipher[A]. CHES 2011[C]. Berlin: Springer, 2011.326-341.

[18] 李玮, 谷大武, 赵辰等. 物联网环境下LED轻量级密码算法的安全性分析[J]. 计算机学报, 2012, 35(3): 434-445.
LI W, GU D W, ZHAO C, *et al*. Security analysis of the LED lightweight cipher in the internet of things[J]. Chinese Journal of Computers, 2012, 35(3): 434-445.

[19] JEONG K, CHANGHOON L. Differential fault analysis on block

cipher LED-64[A]. DFIS 2012[C]. Berlin: Springer, 2012.747-755.

[20] JOVANOVIC P, KREUZER M, POLIAN I. A fault attack on the LED block cipher[A]. Proceedings of COSADE 2012[C]. Berlin: Springer, 2012.120-134.

[21] KREUZER M. Algebraic fault attacks webinar[EB/OL]. http://web.stevens.edu/algebraic/Files/SCPQ/SCPQ-2012-04-19-talk-kreuzer.pdf.

[22] ZHAO X J, GUO S Z, ZHANG F, *et al*. Improving and evaluating differential fault analysis on LED with algebraic techniques[A]. FDTC2013[C]. Santa Barbara, USA, 2013.41-51.

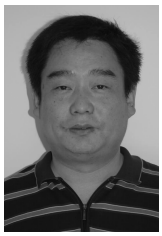
[23] 冀可可, 王韬, 郭世泽. 基于汉明重的 LED 代数旁路攻击研究[J]. 通信学报, 2010, 31(12): 82-89.
JIN K K, WANG T, GUO S Z. Research of Hamming weight-based algebraic side-channel attack on LED[J]. Journal on Communications, 2010, 31(12): 82-89.

[24] 冀可可, 王韬, 赵新杰等. 基于碰撞模型的 LED 代数旁路攻击[J]. 计算机应用研究, 2013, 30(1): 270-272.
JIN K K, WANG T, ZHAO X J. Collision model-based algebraic side-channel attack on LED[J]. Application Research of Computers, 2013, 30(1): 270-272.

[25] ZHANG L, ZHANG D. Finger-knuckle-print verification based on band-limited phase-only correlation[A]. Proceedings of the 13th International Conference on Computer Analysis of Images and Patterns[C]. Berlin: Springer, 2009.141-148.

作者简介：

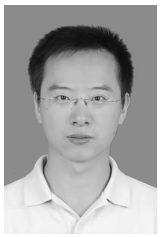
王小娟（1985-），女，河北保定人，北京邮电大学硕士生，主要研究方向为分组密码功耗、电磁旁路分析。



郭世泽（1969-），男，河北石家庄人，博士，北方电子设备研究所研究员、博士生导师，主要研究方向为信息安全和密码学。

赵新杰（1986-），男，河南开封人，博士，北方电子设备研究所工程师，主要研究方向为分组密码旁路分析、故障分析和组合分析。

宋梅（1960-），女，天津人，北京邮电大学教授、博士生导师，主要研究方向为下一代网络与服务、移动互联网、通信系统与集成电路。



张帆（1978-），男，浙江杭州人，康涅狄格大学博士生，主要研究方向为密码旁路分析、计算机体系结构和无线传感器网络安全。