计 算 机 研 究 与 发 展
Journal of Computer Research and Development
https://crad.ict.ac.cn
62(10): 2441−2454, 2025

SM4 密钥扩展算法的单能迹攻击

吴 震[1,2] 赵 洋[1] 王 敏[2]

[1]（电子科技大学信息与软件工程学院 成都 610054）

[2]（成都信息工程大学网络空间安全学院（芯谷产业学院） 成都 610225）

（wzhen@cuit.edu.cn）

Single-Trace Attack on SM4 Key Expansion Algorithm

Wu Zhen[1,2], Zhao Yang[1], and Wang Min[2]

[1]（*School of Information and Software Engineering*, *University of Electronic Science and Technology of China*, *Chengdu* 610054）

[2]（*School of Cybersecurity* (*Xin Gu Industrial College*), *Chengdu University of Information Technology*, *Chengdu* 610225）

Abstract In the context of large-scale applications of distributed Internet of things (IoT), cryptographic technologies, as the underlying infrastructure for information security, are facing severe challenges from side-channel attacks (SCAs)—a physical-layer security threat. The SM4 block cipher algorithm, as China's independently developed commercial cryptographic standard, has been deeply integrated into distributed IoT security protocols. However, its implementation-level vulnerabilities to side-channel attacks remain unresolved. Research on side-channel attacks targeting the SM4 key expansion algorithm is scarce, with existing methods primarily relying on multi-trace statistical characteristics, while single-trace attack approaches are under-explored. This paper proposes a single-trace side-channel attack method integrating Bayesian networks with side-channel attack modeling. By constructing a probabilistic graphical model and combining it with the belief propagation algorithm, the method efficiently infers round subkeys from a single power trace, ultimately recovering the master key. Simulation and practical experiments demonstrate the effectiveness of this approach, achieving an 85.74% success rate for master key recovery under ideal conditions. Even when substantial Gaussian white noise is added to practical power traces (resulting in a signal-to-noise ratio of only 10 dB), the success rate remains at 70%. Compared with traditional methods, this approach exhibits significant advantages in success rate, required power traces, and attack duration, offering new insights and technical tools for side-channel attack research on cryptographic devices in distributed IoT systems, as well as theoretical foundations for related defensive designs.

Key words SM4 key expansion algorithm; single-trace attack; Bayesian network; belief propagation algorithm; side-channel attack (SCA); distributed Internet of things

摘 要 在分布式物联网的大规模应用背景下，各实体设备中密码技术作为信息安全的底层支撑架构，正面临着侧信道攻击（SCA）这一物理层安全威胁的严峻挑战. SM4 分组密码算法作为我国自主研制的商用密码算法标准，已深度集成于分布式物联网安全协议中，但其实现层面的侧信道脆弱性问题亟待解决. 针对 SM4 密钥扩展算法的侧信道攻击研究存在空白，现有攻击方法多依赖多能迹统计特性，而单能迹攻击研究匮乏. 研究提出一种基于贝叶斯网络结合建模侧信道攻击的单能迹侧信道攻击方法，针对单条能量轨迹，通过构建概率图模型，结合置信传播算法，实现对轮子密钥的高效推测，进而恢复主密钥. 仿真实

收稿日期：2025-06-04；**修回日期**：2025-07-13

基金项目：国家社会科学基金项目（23BSH061）

This work was supported by the National Social Science Fundation of China (23BSH061).

通信作者：王敏（wmcuit@cuit.edu.cn）

验与实测实验表明该攻击方法有效,在理想实测环境下主密钥恢复成功率达 85.74%,即使在实测能迹中添加大量高斯白噪声,使得信噪比仅为 10 dB 的条件下,成功率仍可达 70%.与传统方法相比,所提方法在成功率、所需能量轨迹数量和攻击时间等方面优势显著,为分布式物联网系统含密设备的侧信道攻击研究提供了新的思路与技术手段,也为相关防护设计提供了理论依据和参考.

关键词 SM4 密钥扩展算法;单能迹攻击;贝叶斯网络;置信传播算法;侧信道攻击;分布式物联网

中图法分类号 TP309.1

DOI: 10.7544/issn1000-1239.202550382 **CSTR:** 32373.14.issn1000-1239.202550382

在分布式物联网的大规模应用背景下,密码技术作为信息安全的底层支撑架构,通过建立基于密码学的设备身份认证机制、端到端加密传输协议以及细粒度访问控制策略,构建了多维协同的安全防护体系.值得关注的是,侧信道攻击(side-channel attack, SCA)作为物理层安全的新型威胁向量,已对物联网终端设备构成严重安全挑战.研究表明,攻击者可利用物理接触或远程渗透方式获取设备控制权,通过高精度测量设备在密码运算过程中泄露的能量轨迹、电磁辐射、时间延迟等物理侧信息,运用统计分析与机器学习方法实现密钥的渐进式恢复.此类攻击手段已成功渗透智能家居(如生物识别门锁、环境监测终端)和工业物联网(如 PLC 控制器、DCS 系统传感器)等关键场景,密钥泄露将直接导致数据完整性破坏、系统逻辑劫持等最严重的级联式安全风险.作为我国自主研制的商用密码算法标准,SM4 分组密码算法[1]凭借其 ASIC 友好型架构和可证明安全性,已在分布式物联网安全协议中实现深度集成,但其实现层面的侧信道脆弱性问题仍需解决.

侧信道分析攻击自 Kocher 提出计时攻击[2]和差分能量攻击(differential power attack, DPA)[3]以来,已演变为密码分析领域的重要分支.攻击者通过分析密码设备在运行过程中产生的物理侧信道信息,结合数理统计技术[4]、建模技术[5]或机器学习技术[6]等,可有效恢复密钥或敏感信息.目前侧信道攻击技术已呈现出多样化发展趋势,包括时序攻击、能量分析攻击、缓存攻击和故障攻击等多种类型,且随着攻击技术不断演进,防御难度日益加大.

当前针对 SM4 算法的侧信道攻击研究呈现以下特征:

1)非建模类攻击的主导地位.多数研究集中于 DPA/CPA 对 SM4 算法的攻击,通过选择中间值泄露函数(如汉明重量、汉明距离)构建区分器.此类方法虽在实验环境下取得成功,但在实际场景中易受噪声干扰,也需精确寻求攻击的兴趣点.

2)建模类攻击的技术探索.部分研究尝试利用模板攻击或机器学习模型(如随机森林、卷积神经网络)构建从中间值到物理泄漏的映射关系[7].然而,此类方法面临中间变量选择困难、模型泛化能力不足等挑战,尤其在单能迹攻击场景下性能显著下降.

3)单能迹攻击的研究空白.现有方法多依赖多轨迹统计特性,对单条能量轨迹的攻击研究匮乏.单能迹攻击可突破样本数量限制,但需解决噪声抑制、信息提取效率等难题.

需着重指出的是,现有针对 SM4 算法实现的侧信道攻击[8-12]主要聚焦于加解密算法的核心运算模块.尽管这些研究成果有效揭示了 SM4 加密流程在物理实现层面的安全漏洞,却未充分重视密码系统另一核心组件——密钥扩展算法的侧信道攻击安全性评估.密钥扩展算法作为 SM4 算法的核心模块之一,其安全性直接关系到主密钥的保护.若攻击者能通过侧信道泄露正确恢复轮密钥生成过程中的敏感信息,进而可通过数学推导恢复出主密钥,这将严重威胁整个密码系统的安全性.因此,针对 SM4 密钥扩展算法的侧信道攻击研究不仅具有理论价值,更对密码设备的实际防护设计具有指导意义.

然而相较于已获得广泛关注的 SM4 加解密算法,针对 SM4 密钥扩展算法的侧信道攻击研究目前尚缺乏系统性探索.究其原因,是由于密钥扩展算法在运行过程中无法直接获取其输入输出参数,使得传统基于 SM4 加解密算法实现的侧信道攻击方法难以被直接套用.这种技术壁垒使得建立针对密钥扩展算法的安全评估体系,不仅需要突破现有攻击范式的局限性,更需从泄露机理建模层面开展研究.

利用单能迹攻击 SM4 密钥扩展算法面临 2 个技术难点.

1)传统攻击方法的局限性.现有针对 SM4 密钥扩展算法的侧信道攻击主要存在两大技术瓶颈:非建模类攻击(如 DPA/CPA)依赖多能迹统计特性,但密钥扩展算法缺乏已知明文/密文输入,无法构建假

设检验框架；而建模类攻击（如模板攻击）在单能迹条件下，由于中间值恢复错误率的累积效应，导致主密钥恢复成功率急剧下降.

2）密钥扩展算法的特殊性挑战. 相较于加解密模块，密钥扩展算法在运行过程中无法直接获取输入输出参数，且其运算流程具有强关联性和状态连续性特征. 这导致传统基于独立中间值分析的攻击范式难以适用，需要突破现有方法论构建新型攻击模型.

本研究针对 SM4 密钥扩展算法的实现，构造了基于能量特征建模与贝叶斯推理网络相结合的单能迹分析模型，通过建立泄露信号与初始密钥状态之间的概率依赖关系，实现了仅用 1 条能量迹就完成对主密钥的恢复. 本研究提出了一种基于数学化建模的含密设备侧信道攻击分析方案，为分布式物联网系统安全架构提供技术保障.

1 预备知识

1.1 SM4 密码算法

SM4 算法作为中国商用密码体系中的重要组成部分，采用了分组密码的加密原理，其轮函数与密钥扩展机制是保障数据安全的关键. 轮函数是 SM4 算法中实现数据混淆的核心操作，它通过一系列复杂的数学变换，确保了数据在加密过程中的不可预测性. 密钥扩展机制则负责将用户提供的主密钥转换为一系列轮密钥，这些轮密钥在加密过程中依次被使用，为数据提供逐轮加密的保护.

SM4 加密算法如图 1 所示.

设算法的输入明文为 128 b，记为(X_0,X_1,X_2,X_3)，其中$X_i(i=0,1,2,3)$为 32 b；输出密文为 128 b，记为(Y_0,Y_1,Y_2,Y_3)，其中$Y_i(i=0,1,2,3)$为 32 b；轮子密钥记为$(rk_0,rk_1,\cdots,rk_{31})$，其中$rk_i,(i=0,1,\cdots,31)$为 32 b. SM4 密码算法的加密运算如下式迭代：

$$X_{i+4}=F(X_i,X_{i+1},X_{i+2},X_{i+3},rk_i)=X_i\oplus T(X_{i+1}\oplus X_{i+2}\oplus X_{i+3}\oplus rk_i).$$

最终输出的密文为：$(Y_0,Y_1,Y_2,Y_3)=(X_{35},X_{34},X_{33},X_{32})$，

其中$T(\cdot)$为合成置换，由非线性变换$\tau(\cdot)$和线性变换$L(\cdot)$复合而成，即$T(\cdot)=L(\tau(\cdot))$.

非线性变换$\tau(\cdot)$是 SM4 中唯一的非线性运算，由 4 个相同 S 盒并行构成，表示为$Sbox(\cdot)$. 设τ变换的输入为$A=a_0\|a_1\|a_2\|a_3,a_i(i=0,1,2,3)$为 8 b，输出为$B=b_0\|b_1\|b_2\|b_3,b_i(i=0,1,2,3)$为 8 b，则$\tau$变换为：

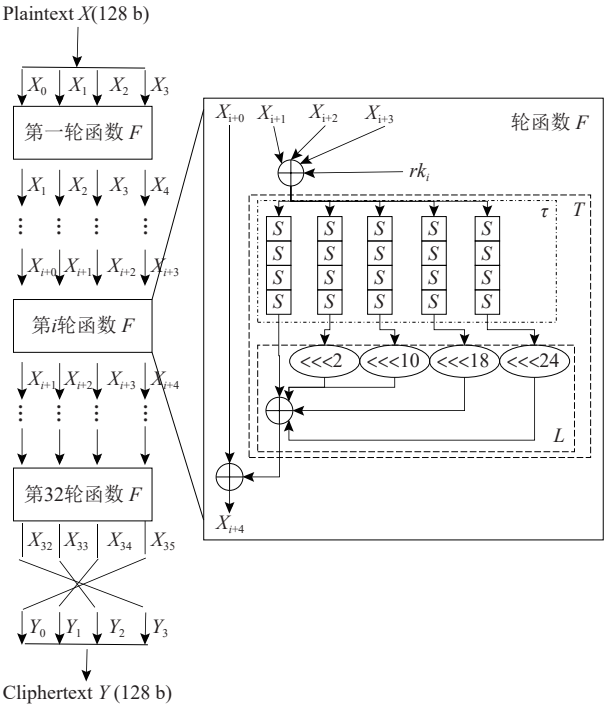


Fig. 1 Schematic diagram of SM4 encryption operation process

图 1 SM4 加密运算过程示意图

$$(b_0,b_1,b_2,b_3)=\tau(A)=(Sbox(a_0),Sbox(a_1),Sbox(a_2),Sbox(a_3)),$$

其中τ变换的输出B是线性变换$L(\cdot)$的输入，其输出为C，则线性变换$L(\cdot)$可表示为：

$$C=L(B)=B\oplus(B\lll 2)\oplus(B\lll 1)\oplus(B\lll 18)\oplus(B\lll 24).$$

1.2 针对对称密码的侧信道攻击

侧信道攻击（side-channel attack）是一种基于密码设备运行过程中泄露的物理侧信道信息（包括能量、电磁辐射、时间特征等）进行密钥或敏感信息提取的密码分析方法.

侧信道信息泄露与测量可建立以下形式化模型，模型间关系如图 2 所示.

1）对称密码算法的基本模型

在密码学理论中，对称密码算法的形式化表达可严格定义为一个五元组：(*E*, *D*, *K*, *X*, *C*).

其中：

加密函数 *E*，通常表示为 *E*: $K\times X\to C$，接受密钥和明文作为输入，输出密文.

解密函数 *D*，通常表示为 *D*: $K\times C\to X$，接受密钥和密文作为输入，输出明文.

密钥空间 *K*，是所有可能的密钥组成的集合.

明文空间 *X*，是所有可能的明文组成的集合.

密文空间 *C*，是所有可能的密文组成的集合.

该五元组需要满足一些基本条件才能被作为一

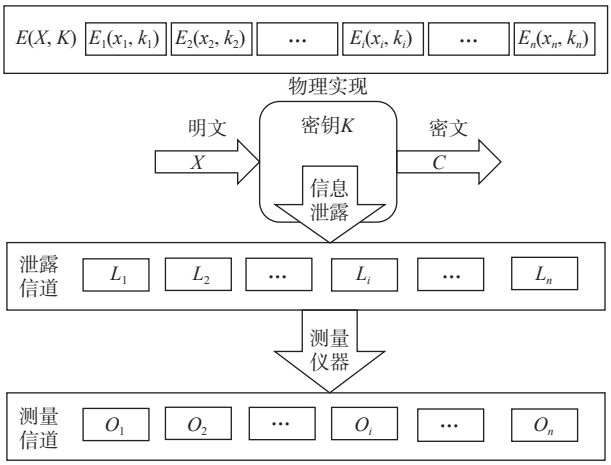


Fig. 2 Schematic diagram of the relationship between side-channel leakage and measurement model

图 2 侧信道泄露与测量模型关系示意图

个对称密码算法使用,而经典的密码学分析就是基于这些基本条件来进行算法攻击研究.

2)中间值计算的形式化模型

加解密函数实现过程并非简单的直接映射,而是经过一系列非线性迭代.

以加密函数 E 为例,E 的算法实现需要完成一系列的前后相关的中间值运算 $E_1,E_2,\cdots,E_n$. 对于每个中间值计算 E_i,在有 1)的形式化定义条件下,可形式化为一个四元组

$$(E_i,X_i,K_i,C_i),\ i=1,2,\cdots,n,$$

其中:

中间值变换函数 E_i,表示为 $E_i:X_i\times K_i\to C_i$,接受密钥分片和可预计算值作为输入,输出中间值.

密钥分片空间 K_i,由密钥分片所有可能的值组成的集合. 密钥分片是由密钥扩展算法以及密钥使用方法决定.

可预计算值空间 X_i,是所有可能的可预计算值组成的集合. 可预计算值就是通过已知明文 X 和已知密钥分片集合 $\{K_1,K_2,\cdots,K_{i-1}\}$ 的某个子集计算出的值.

中间值空间 C_i,是所有可能的中间值组成的集合.

3)侧信道信息泄露的形式化模型

对于存在中间值侧信道信息泄露的算法实现,在有 1)和 2)的形式化定义条件下,中间值的侧信道信息泄露可形式化表达为一个三元组:

$$(H_j,C_j,L_j),\ j=1,2,\cdots,m(m\leqslant n),$$

其中:

泄露函数 H_i,通常表示为 $H_j:C_j\to L_j$,接受存在泄露中间值作为输入,输出泄露向量.

存在泄露中间值空间 C_j,是所有可能的存在信息泄露中间值的集合. 为算法实现中间值空间 C_i 的一个子空间.

泄露向量空间 L_j,是所有可能的泄露向量值组成的集合.

4)侧信道信息泄露测量的形式化模型

对于中间值的侧信道信息泄露向量的测量(又称采集),在有 1)~3)的形式化定义条件下,可形式化表达为一个三元组:

$$(M_k,O_k,L_k),\ k=1,2,\cdots,r(r\leqslant m),$$

其中:

测量函数 M_k,通常表示为 $M_k:L_k\to O_k$,接受被测量泄露向量作为输入,输出观测向量.

被测量泄露向量空间 L_k,是所有可能的被测量泄露向量值组成的集合. 该空间为泄露向量空间 L_j 的一个子空间.

观测向量空间 O_k,是所有可能的测量向量值组成的集合.(在能量分析攻击中,观测向量即为能量轨迹,简称能迹.)

与上文所述加密函数 E 相对应,亦可对解密函数 D 所涉及的侧信道信息泄露及测量展开形式化定义. 鉴于二者在形式化定义层面具有高度同构性,为避免重复,此处不再展开详述.

根据以上的形式化定义可知,侧信道攻击就是利用收集到的若干观测向量的集合 O_k,来预测各个密钥分片 K_i,再据此推导出原始密钥 K 的攻击方法.

依据攻击模型构建方式,现有侧信道分析技术可分为非建模类分析和建模类分析两大类[13].

在非建模类侧信道分析攻击中,攻击者需基于目标密码算法的结构特性,选取特定运算环节中具有信息区分度的中间变量 C_i,如 SM4 算法第 1 轮 S 盒输出值,并构建相应的能量泄露假设模型,如泄露函数为 $L_j=H(C_j)=HW(Sbox(X_j\oplus K_j))$ 的数学表征,其中 $HW(\cdot)$ 表示汉明重量函数. 鉴于密码系统通常采用分片调度策略(如 SM4 算法 S 盒 8 位轮子密钥调用机制),攻击者在已知明文 X 条件下可对密钥分片 K_j 实施穷举遍历. 通过建立理论泄露模型 $L=H(x,k)$ 与实测物理观测向量 $\boldsymbol{O}$ 的关联映射,攻击者可采用统计推断方法(如皮尔逊相关系数或互信息量分析)对遍历密钥空间生成的候选泄露向量进行假设检验. 差分能量攻击(DPA)作为经典非建模方法,其核心在于利用密钥猜测导致的能量差异进行统计显著性检验,而相关能量攻击(correlation power attack, CPA)则通过优化相关系数计算显著提升攻击效率. 实验

研究表明，泄露模型与观测向量失配将导致假设检验统计效力的显著衰减.

建模类分析攻击方法则通过构建参数化的物理泄漏模型$O=G_x(k)$，再求解 $k=G_x^{-1}(O)$来实现密钥恢复，其技术演进可分为传统模板攻击[5]和机器学习驱动攻击[6]2 个阶段. 攻击者在模型训练阶段需利用已知明文-密钥对$\{(x, k)\}$及其对应侧信道轨迹集合 O，通过传统统计建模或基于机器学习算法（如随机森林、深度神经网络等）构建从中间值到物理泄露的非线性参数化映射模型$O=G_x(k)$. 在建模类分析攻击框架下，攻击者需为每个候选密钥分片建立多维高斯分布模型，再通过计算被攻击轨迹与各模板之间的似然概率实现密钥恢复.

1.3 概率图模型

概率图模型（probabilistic graphical model, PGM）是一类通过图形化表示来描述变量之间概率关系的数学模型，是概率论、图论和统计学相结合的产物. 它在机器学习、人工智能、统计学、信息论、计算机视觉、自然语言处理等诸多领域都有广泛的应用，是现代数据分析和知识表示的重要工具之一.

1）基本概念

概率图模型利用图的结构来表示变量之间的概率依赖关系，图中的节点代表随机变量，边或弧表示变量之间的概率关联. 根据图的类型，概率图模型主要分为 2 类：有向无环图（directed acyclic graph, DAG）表示的贝叶斯网络（Bayesian networks）和无向图表示的马尔可夫随机场（Markov random fields, MRFs），本文使用贝叶斯网络的精确推理来实现单能迹侧信道攻击.

贝叶斯网络是一种有向无环图模型，它通过有向边表示变量之间的因果关系或概率依赖关系. 每个节点代表一个随机变量，有向边从父节点指向子节点，表示父节点对子节点的概率影响. 贝叶斯网络的核心是条件概率分布，即每个变量在其父节点给定的条件下，遵循某种概率分布.

贝叶斯网络由有向无环图 G 和条件概率表 P 组成，记作 $B=(G, P)$，其中：

①有向无环图. 图 $G=(V, E)$，其中 V 是有限非空节点集合，表示随机变量集合，记作 $V=\{X_1, X_2, \cdots, X_n\}$. E 是有向边集合，表示变量之间的直接依赖关系. 图中不存在有向环.

②条件概率表（conditional probability table, CPT）. 对于每个节点 $X_i \in V$，都有一个条件概率分布 $P(X_i|Parents(X_i))$，其中 $Parents(X_i)$表示节点 X_i在图 G 中的所有父节点的集合.

③联合概率分布. 贝叶斯网络定义了变量的联合概率分布，对于所有变量 $X_1, X_2, \cdots, X_n$，其联合概率分布可表示为：

$$P(X_1, X_2, \cdots, X_n)=\prod_{i=1}^{n} P(X_i|Parents(X_i)).$$

贝叶斯网络的精确推理是在给定证据变量的情况下，计算其他变量的后验概率.

2）置信传播算法

概率图模型提供了高效的推理算法，如变量消除法（variable elimination, VE）、置信传播（belief propagation, BP）算法等，能够在给定部分变量的观测值的情况下，快速推断出其他变量的概率分布. 这种推理能力使得概率图模型在实际应用中能够有效地解决诸如预测、诊断、决策等问题.

BP 算法是一种在概率图模型中用于计算变量边缘概率分布的高效算法. 它属于消息传递算法的一种，通过在图模型的节点之间传递概率信息来实现对变量概率分布的推断. BP 算法在贝叶斯网络和马尔可夫随机场等概率图模型中都有广泛的应用，是现代概率图模型推理算法中的重要组成部分.

BP 算法的基本原理是将复杂的联合概率分布分解为局部的边缘概率分布和条件概率分布的乘积，然后通过迭代的方式在图模型的节点之间传递这些局部概率信息，最终达到收敛并得到每个变量的边缘概率分布.

在 BP 算法中，每个节点（变量）会向其相邻的节点发送消息，这些消息包含了关于变量概率分布的信息. 对于有向图（如贝叶斯网络），消息的传递方向沿着边的方向进行；对于无向图（如马尔可夫随机场），消息的传递是双向的. 消息的计算通常基于变量的条件概率分布和相邻节点发送过来的消息.

BP 算法的步骤概括为 4 个阶段：

1）初始化. 为每个节点的初始消息设定一个初始值，通常为均匀分布或根据先验知识设定.

2）消息传递. 在图模型中，每个节点根据其条件概率分布和接收到的消息，计算并发送新的消息给其相邻的节点.

3）边缘概率计算. 每个节点根据其条件概率分布和所有接收到的消息，计算其边缘概率分布.

4）收敛判断. 检查所有节点的边缘概率分布是否收敛. 如果满足收敛条件（如边缘概率分布的变化小于某个阈值），则算法结束；否则，返回消息传递阶段继续迭代.

BP 算法在概率图模型的推理和学习中有着广泛的应用, 在本次研究中, 是利用其在概率图模型中进行变量推断的能力, 即在给定部分中间变量的观测值的情况下, 使用 BP 算法可以推断出其他变量的边缘概率分布, 从而实现对轮子密钥值的预测和估计.

2 SM4 密钥扩展算法的单能迹攻击

2.1 SM4 密钥扩展算法

SM4 分组密码算法的加密密钥长度为 128 b, 表示为 $MK=(MK_0,MK_1,MK_2,MK_3)$, 其中 $MK_i(i=0,1,2,3)$ 为 32 b. 轮子密钥表示为 $(rk_0,rk_1,\cdots,rk_{31})$, 其中 $rk_i(i=0,1,\cdots,31)$ 为 32 b.

固定参数 $FK=(FK_1,FK_2,FK_3,FK_4)$ 以及 $CK=(CK_0,CK_1,\cdots,CK_{31})$ 用于密钥扩展算法, 其中 $FK_i(i=0,1,2,3)$, $CK_i(i=0,1,\cdots,31)$ 均为 32 b. 中间变量表示为 $K=(K_0,K_1,\cdots,K_{39})$, 其中 $K_i(i=0,1,\cdots,39)$ 为 32 b.

SM4 密钥扩展算法 $Exp(\cdot)$ 可用下式描述:

$$Exp(MK)=(rk_0,rk_1,\cdots,rk_{31}).$$

算法输入为主密钥 MK, 算法输出为 32 个轮子密钥 $(rk_0,rk_1,\cdots,rk_{31})$.

轮子密钥生成算法数据流图如图 3 所示.

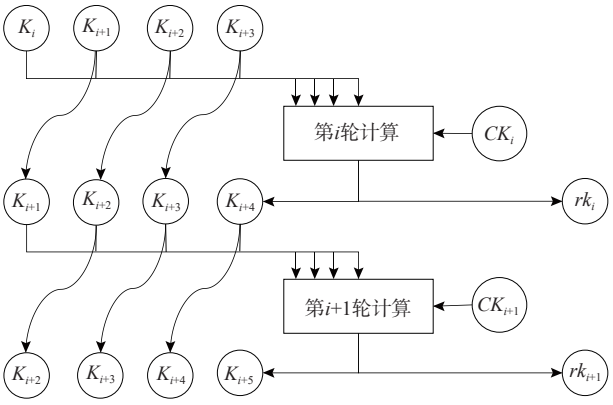


Fig. 3 Schematic diagram of data flow of SM4 key expansion algorithm

图 3 SM4 密钥扩展算法数据流示意图

步骤如下:

用主密钥 MK 初始化中间变量 (K_0,K_1,K_2,K_3):

$$K_0=MK_0\oplus FK_0,$$
$$K_1=MK_1\oplus FK_1,$$
$$K_2=MK_2\oplus FK_2,$$
$$K_3=MK_3\oplus FK_3.$$

用前 4 轮中间变量迭代产生当前轮的轮子密钥:

$$rk_i=k_{i+4}=k_i\oplus T'(k_{i+1}\oplus k_{i+2}\oplus k_{i+3}\oplus ck_i),\ i=0,1,\cdots,31,$$

其中 $T'(\cdot)$ 是将加密算法中合成置换 $T(\cdot)$ 的线性变换 L 替换为 L':

$$L'(B)=B\oplus(B\lll 13)\oplus(B\lll 23).$$

SM4 密钥扩展算法的第 i 轮运算数据流如图 4 所示, 其对应的算法实现数据流如图 5 所示.

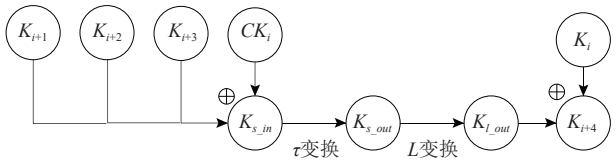


Fig. 4 Schematic diagram of the i-th round operation data flow of SM4 key expansion algorithm

图 4 SM4 密钥扩展算法第 i 轮运算数据流示意图

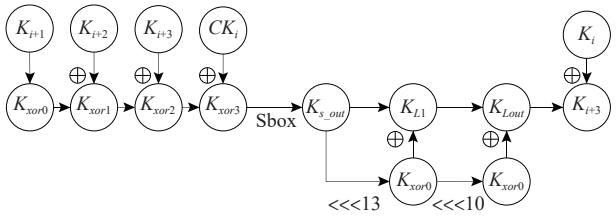


Fig. 5 Schematic diagram of the i-th round algorithm data flow of SM4 key expansion algorithm

图 5 SM4 密钥扩展算法第 i 轮算法数据流示意图

2.2 SM4 主密钥恢复方法

利用连续 4 轮的轮子密钥 $(rk_i,rk_{i+1},rk_{i+2},rk_{i+3})$, 其中 $i=0,1,\cdots,28$ 可恢复出其对应的主密钥, 主密钥恢复算法 $Rec(\cdot)$ 可表示为:

$$MK=(MK_0,MK_1,MK_2,MK_3)=$$
$$Rec(rk_i,rk_{i+1},rk_{i+2},rk_{i+3}),\ i=0,1,\cdots,28.$$

推导如下:

根据 SM4 算法密钥扩展规则:

$$rk_i=k_{i+4}=k_i\oplus T'(k_{i+1}\oplus k_{i+2}\oplus k_{i+3}\oplus ck_i).$$

可建立如下方程:

$$k_i=rk_i\oplus T'(k_{i+1}\oplus k_{i+2}\oplus k_{i+3}\oplus ck_i),$$

以及:

$$k_i=rk_i\oplus T'(rk_{i-3}\oplus rk_{i-2}\oplus rk_{i-1}\oplus ck_i).$$

通过递归应用上述方程, 可依次求解 k_3,k_2,k_1,k_0 的数值, 最终主密钥恢复公式表述为:

$$MK_0=FK_0\oplus k_0,$$
$$MK_1=FK_1\oplus k_1,$$
$$MK_2=FK_2\oplus k_2,$$
$$MK_3=FK_3\oplus k_3.$$

由上述推导可知, 若能使用某攻击方法恢复出正确连续 4 轮的轮子密钥, 则必可恢复出正确的原始主密钥.

2.3 SM4 密钥扩展算法的单能迹攻击

2.3.1 问题的提出

由 2.2 节描述可知，通过对 SM4 密钥扩展算法的实现实施侧信道分析攻击，完全有可能恢复原始主密钥. 但由于密钥扩展算法自身特性，在未知密钥情况下无法使用 DPA、CPA 等非建模类侧信道分析攻击方法实施攻击，而建模类侧信道攻击在单能迹条件下，攻击成功率也极低.

在传统非建模类侧信道攻击中，为能进行统计分析，均要求多个已知不同的明文 X(或密文)和相同未知密钥 K 为输入的条件下，测量得到对应的密码算法以实现运行时信息泄露观测值(在能量分析攻击场景下即为多条能迹)，才能得以实施该类侧信道攻击. 而通过 2.1 节中对 SM4 密钥扩展算法的分析可知，算法输入为被攻击的敏感值密钥 K，输出亦为被攻击的敏感值轮子密钥序列，不存在已知明文或密文，故而无法建立相应的假设检验框架，不能使用传统非建模类侧信道攻击对 SM4 密钥扩展算法的实现实施攻击.

在传统的建模类侧信道攻击中，首先要求能对被攻击对象进行建模. 对于 SM4 密钥扩展算法的实现来讲，该建模阶段是可实施的. 原因在于被测设备在密码算法运算调用时，至少存在以下 2 种模式：一种为外部密钥模式，它使用外部输入的明文 X、密钥 K_{ext}来进行密码运算，返回相应密文 C；另一种为内部密钥模式，它使用外部输入的明文 X，与设备内部的秘密密钥 K_{int}结合来进行密码运算，返回相应密文 C. 而在算法运行阶段，这 2 种模式都先调用密钥扩展算法得到轮子密钥序列，再用轮子密钥序列进行加解密运算. 因此攻击者可利用外部密钥模式对密钥扩展算法进行建模，再利用内部密钥模式攻击密钥扩展算法，对秘密密钥 K_{int}进行恢复.

然而传统的建模类侧信道攻击，在单个测量向量下的成功率极低，一般在实施过程中均需要提高测量次数，获取多个测量向量以提高其攻击成功率.

在能量分析攻击场景下，主密钥恢复成功率低下的原因在于传统建模类分析攻击框架中，攻击者需为每个中间值与能量轨迹中多个样本点之间建立多维高斯分布模型，再通过计算被攻击轨迹与各模板之间的似然概率以及中间值与主密钥间关系来实现密钥恢复. 因此主密钥恢复的成功率依赖于各轮子密钥的正确恢复，而各轮子密钥恢复正确率又依赖于各中间值恢复的成功率，只要若干中间值中任意一个恢复错误，则主密钥恢复必然失败. 由于模板的概率特性，中间值恢复时存在一定的错误概率，导致恢复错误.

在针对 SM4 密码算法实现的能量分析攻击中，令被测量的 n 个中间值为 $C=\{C_1,C_2,\cdots,C_n\}$，对应的观测值为 $O=\{O_1,O_2,\cdots,O_n\}$，常规建模类攻击的方法是分别按极大似然概率求出这 n 个中间值的估值，$\hat{C}_j = ArgmaxP(C_j|O_j), i=1,2,\cdots,n$，再用这些估值计算出对应轮子密钥 rk_i. 若令 $Err(\cdot)$ 为错误率，$Acc(\cdot)$ 为正确率，$Succ(\cdot)$ 为成功率，$Fail(\cdot)$ 为失败率，又令某轮子密钥 rk_i正确恢复时，需依赖 n 个中间值的正确恢复，其集合表示为 $C=\{C_1,C_2,\cdots,C_n\}$，则有：

$$Succ(MK)=Acc(MK)=\prod_{i=1}^{4}Acc(rk_i)=\prod_{i=1}^{4}\left(\prod_{j=1}^{n}Acc(C_j)\right)=(1-Err(C_j))^{4n},$$

$$Fail(MK)=1-(1-Err(C_j))^{4n}.$$

SM4 算法中某轮子密钥 rk_i的恢复，根据密钥切片建模需求，至少依赖 4 个中间值. 若每个中间值错误率为 10%，根据以上推导可知主密钥恢复的成功率将小于 18.53% 和失败率大于 81.47%. 而在实际攻击中，由于噪声的存在，单能迹下单个中间值恢复失败率远超 10%，这意味着主密钥的正确恢复几乎是不可能的. 因此传统建模攻击类在实际攻击中，需利用多能迹和去噪技术等技术手段，降低中间值恢复的失败率，从而达到提升密钥恢复成功率的目标.

2.3.2 结合概率图推理的单能迹建模攻击

本文提出一种使用建模侧信道技术与概率图推理相结合的方法来对 SM4 密钥扩展算法实施单轨迹建模类攻击. 在建模阶段，通过已知输入的密码运算推导，将得到的算法中间值与对应测量到的能迹建立多维联合分布概率的模型，同时根据算法实现的数据流图，用概率图模型对各中间值及轮子密钥进行建模. 在攻击阶段，用未知密钥所对应的单能迹通过已建立的概率图模型和多维联合分布概率，求出主密钥候选集合，最后通过加解密检验筛选出正确主密钥 MK.

2.3.2.1 原理

本文提出的单能迹攻击方法借鉴低密度奇偶校验(LDPC)码的迭代纠错机制[14]，将单轨迹侧信道攻击的密钥恢复问题，转化为基于物理泄露信息观测值的编码序列最优解码过程. 该方法的技术框架继承了 LDPC 解码的核心思想，通过构建动态概率图模型实施贝叶斯统计推断，在单条能量轨迹多观测点数

据集合的条件下,将原本孤立的中间变量最大似然估计问题,重构为全局联合概率最大化的参数估计问题.进一步依托中间变量间的密码算法约束关系,构建多维参数协同优化机制,在迭代过程中实现各节点置信度的动态传播与修正,进而有效纠正传统方法对中间值的误判问题,使得攻击成功率获得显著提升.相关纠错算法的完备数学推导可参见文献[14].

在密码算法实现中,各中间值$C=\{C_1,C_2,\cdots,C_n\}$通过各种密码运算存在着依赖关系,有效利用各中间值之间关系,寻求各中间值联合最大化而非各中间值局部独立最大,成为提高密钥恢复成功率的一个突破点.

将针对SM4密钥扩展算法的单轨迹建模攻击表示成:

$$\widehat{MK}=Rec(\widehat{rk}_i,\widehat{rk}_{i+1},\widehat{rk}_{i+2},\widehat{rk}_{i+3}),\quad i=0,1,\cdots,28,$$

其中:

$$\widehat{rk}_i=\arg\max_{rk_i}P(rk_i|C).$$

令由中间值求解轮子密钥的函数为$Drk(\cdot)$,则有

$$\widehat{rk}_i=Drk\left(\arg\max_{C}P(C|O)\right).$$

由上述推导可知,对轮子密钥的估值,是在全体的观测值O下,各中间值C最大联合似然率所对应的解,可利用概率图模型来完成数学表达的求解.将密码运算中的各中间值作为图的节点、各中间值之间的依赖关系作为边,则可利用贝叶斯推理算法,求解出在单能迹下全局最优的轮子密钥估值.

贝叶斯网络的精确推理是在给定证据变量的情况下,计算其他变量的后验概率.

1)证据变量.假设观测到证据变量集合$E=\{e_1,e_2,\cdots,e_m\}$,其中每个e_j是一个特定变量X_{ij}的特定取值.

2)查询变量.设查询变量为$Q\subseteq V$,是攻击者感兴趣的未观测变量.

3)后验概率计算.目标是计算在给定证据E的情况下,查询变量Q的后验概率分布$P(Q|E)$.

后验概率$P(Q|E)$可以通过联合概率分布和边缘化(求和)其他未观测变量来计算:

$$P(Q|E)=\frac{P(Q,E)}{P(E)},$$

其中:

$P(Q,E)$是包含查询变量Q和证据变量E的联合概率:

$$P(Q,E)=\sum_{X\notin Q\cup E}P(X_1,X_2,\cdots,X_n),$$

$P(E)$是证据变量的边缘概率,用于归一化:

$$P(E)=\sum_{X\notin E}P(X_1,X_2,\cdots,X_n).$$

在实际计算中,可使用BP算法来高效地计算这些概率,以避免直接计算所有变量的联合概率而带来的巨大计算开销.

由图4和图5可看出,密码算法实现的数据流图实质是一个有向无环图,再给出图中各数据的条件概率表,则可获得算法实现对应的一个贝叶斯网络描述.将建模期间得到的被观测中间值C作为贝叶斯网络的证据变量E,而轮子密钥rk_i作为贝叶斯网络的查询变量Q,用建模期间得到的被观测中间值C的概率分布作为证据变量E的初始分布,其余变量的初始分布为均匀分布,边上的一步转移矩阵用确定系统描述,可得到完整的条件概率表,至此就可使用BP算法推导出在当前的被观察中间值C条件下,轮子密钥rk_i的边缘概率$P(rk_i|C)$,进而用最大似然估计对rk_i进行估值,达到对主密钥MK恢复成功率最大化的目的.

为进一步提高主密钥MK恢复成功率,在计算复杂度可容忍的情况下,提供一个主密钥候选集合M,通过遍历验证M中所有的候选主密钥来确定正确的主密钥值MK.

在SM4密钥扩展算法中,M集合中的候选值数量$|M|$由连续4轮轮子密钥候选值的组合决定,若某轮子密钥rk_i的候选值数量为d,则有$|M|=d^4$和$d=\log_4|M|$,据此可根据主密钥候选值数量$|M|$,决定轮子密钥的候选值d.

2.3.2.2 形式化描述

针对SM4密钥扩展算法的单能迹攻击形式化描述如下:

1)符号定义

MK表示SM4算法的主密钥.

rk_i表示SM4算法第i轮轮子密钥,i=1, 2, …, 32.

k_{ij}表示轮子密钥的密钥分片.

X表示SM4密钥扩展算法实现过程中的中间变量.

E表示攻击者采集到的单条能量迹.

O表示从能量迹E中提取的与中间值相关的物理信息观测量(例如能量变化特征).

N表示能量迹采集过程中的噪声.

$H(x)$表示变量x的泄露函数.

$P(A|B)$表示在事件B发生的条件下,事件A发生的概率.

2)能量泄露观测模型

假设物理信息观测量O与中间变量X存在信息泄露被观测到的关系，并受到噪声N的影响，可建立如下能量泄露观测模型$G(\cdot)$：

$$O = G(X) = \alpha \cdot H(X) + N,$$

其中：

$H(X)$为中间值X的能量泄露函数；α为线性模型的系数，表示能量泄露对观测量的影响程度；N假设为高斯白噪声，即$N \sim N(\mu, \sigma^2)$.

3）贝叶斯网络建模

将 SM4 密钥扩展算法的实现数据流图建模为贝叶斯网络$B = (G, P)$，其中：

$G = (V, E)$为有向无环图，表示变量之间的概率依赖关系. V为网络中所有节点的集合，包括：观测节点O表示物理信息观测量；隐变量节点X_i表示 SM4 密钥扩展算法实现过程中的中间变量，例如轮子密钥rk_i、密钥分片k_{ij}以及算法执行过程中的中间状态C_i. E为有向边的集合，表示变量之间的依赖关系. 例如，从X_{i-1}指向X_i的边表示X_i的值依赖于X_{i-1}的值. 在侧信道攻击中，有向边只存在于隐变量节点X_i之间以及隐变量节点X_i到观测节点之间.

CPT 为网络中所有节点的条件概率分布（CPD）的集合，即$P(X_i|Par(X_i))$，其中$Par(X_i)$表示节点X_i的父节点集合. 由于密码运算是确定系统，隐变量节点X_i之间的 CPD 值非 0 即 1. 而隐变量节点X_i到对应观测节点O_i之间的 CPD 则由能量泄露观测模型G决定.

2.3.2.3　单能迹攻击过程

单能迹攻击旨在通过侧信道能量泄露观测值O恢复 SM4 算法主密钥MK，采用基于贝叶斯网络建模的侧信道攻击方法，其攻击流程如图 6 所示，可分为建模和攻击 2 个阶段. 建模阶段构建算法实现数据流图的贝叶斯网络及多维联合分布模型，攻击阶段则利用实测单能迹，通过贝叶斯推理和置信传播算法迭代更新变量节点概率，最终推测出轮子密钥并恢复主密钥.

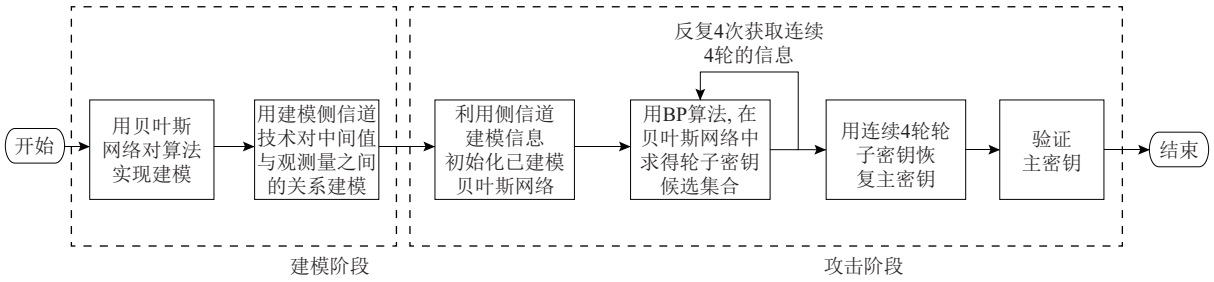


Fig. 6　Flow chart of single-trace attack

图 6　单能迹攻击流程图

在建模阶段，其步骤如下：

1）变量与网络定义

设$X = \{X_1, X_2, \cdots, X_n\}$为 SM4 密钥扩展算法实现过程中的中间变量集合，包括轮子密钥rk_i、密钥分片k_{ij}以及算法执行过程中的中间状态C.

通过 SM4 密钥扩展算法第i轮实现数据流图以及中间值信息泄露情况，构建贝叶斯网络B=(G, P).

2）能量泄露建模

假设能量泄露与中间变量X的汉明重量$H(X)$存在线性关系，并受噪声N影响，建立能量泄露模型$O = \alpha \cdot H(X) + N$.

3）建模能量迹集合采集与预处理

攻击者准备n个已知主密钥MK_i，$i = 1, 2, \cdots, n$. 以所有已知主密钥MK_i为输入，用外密钥模式在被攻击设备上执行 SM4 密钥扩展算法，采集被攻击设备在执行 SM4 密钥扩展算法时的能量迹E_i，形成能量迹集合E.

对能量迹集合E中的各个能迹进行预处理，例如去噪、对齐、特征提取等，得到物理信息观测量集合O.

4）参数估计与模型确定

根据已知主密钥MK的输入，推算出中间值X. 将X和对应的观测值O，通过优化算法（如最大似然估计）估计能量泄露模型$H(X)$的系统参数α和σ^2等值，确定条件概率$P(O|X)$的分布.

根据 SM4 密钥扩展算法的实现细节，确定贝叶斯网络中各节点间的转换矩阵，即条件概率分布$P(X_i|Par(X_i))$，从而构建多维联合分布模型，完整描述所有隐变量节点与观测节点间的概率依赖关系.

在攻击阶段，攻击者的目标是利用单条能迹E，恢复出 SM4 密钥扩展算法的主密钥MK. 由于算法的轮子密钥rk_i是由主密钥MK通过一个可逆变换运算得到，因此攻击者可以通过恢复部分或全部轮子密钥rk_i，进而推导出主密钥MK. 该阶段关键步骤如下：

1）攻击能量迹采集与预处理. 攻击者采集被攻击设备在执行 SM4 密钥扩展算法时的单条能量迹E.

对能量迹 E 进行预处理, 例如去噪、对齐、特征提取等, 得到物理信息观测量 O.

2)证据设置. 将提取的物理信息观测量 $O=o$ 作为贝叶斯网络 B 的证据变量, 即设置观测节点 O 的值为 o(o 为从能量迹中提取的具体观测值).

3)概率计算. 利用贝叶斯定理和贝叶斯网络 B 的条件概率分布 P, 计算在给定证据 $O=o$ 的条件下, 各个隐变量节点(特别是密钥分片节点 K_{ij})的后验概率分布 $P(X_i|O=o)$(特别是 $P(K_{ij}|O=o)$).

4)置信传播. 为了高效计算后验概率分布, 可以使用 BP 算法在贝叶斯网络中进行概率推理. BP 算法通过迭代的方式在图模型的节点之间传递概率信息(消息), 最终得到每个变量的边缘概率分布.

5)密钥推测. 根据计算得到的密钥分片节点 $P(K_{ij}|O=o)$ 推导出轮子密钥节点 rk_i 的后验概率分布 $P(rk_i|O=o)$, 按后验概率降序将前 d 个 rk_i 值作为对第 i 轮轮子密钥 rk_i 的候选集合 $\widehat{rk}_i$.

针对同一能量迹对应的连续 4 轮加密过程, 重复上述过程, 可依次提取轮次 i~i+3 的轮子密钥候选集合 $\{\widehat{rk}_i,\widehat{rk}_{i+1},\widehat{rk}_{i+2},\widehat{rk}_{i+3}\}$.

6)主密钥恢复与候选验证. 根据推测出的轮子密钥候选集合 $\{\widehat{rk}_i,\widehat{rk}_{i+1},\widehat{rk}_{i+2},\widehat{rk}_{i+3}\}$, 利用 SM4 密钥扩展算法的逆过程, 可计算得到主密钥 $\widehat{MK}$ 候选集合 M.

借助加解密验证实验, 将所有候选密钥输入 SM4 加解密算法验证器逐一进行正确性校验, 最终从主密钥候选集合 M 中筛选出与实测结果一致的唯一正确主密钥 MK.

3 实验与分析

3.1 实验设置

1)实验软硬件环境

本研究攻击靶标为智能卡芯片, 芯片中 SM4 算法为软实现, 能迹采集、分析及攻击依托 Inspector SCA 侧信道分析系统搭建, 在配备高精度示波器及可控温湿度的电磁屏蔽环境中, 成功捕获了该芯片运行时的能量轨迹信号.

2)实验建模所使用的能量泄露观测模型

本研究针对存储器存取的能量泄露物理机制, 采用基于变量 x 汉明重量的线性建模方法, 构建物理信息观测量 O 与算法实现变量 x 及噪声的数学关联模型, 其表达式可表述为: $O=a\times HW(x)+b+N(0,\sigma^2)$. 其中: $HW(x)$ 表征变量 x 的汉明重量, $N(0,\sigma^2)$ 表示均值为零、标准差为 σ 的高斯白噪声随机分量, a 为模型线性权重系数, b 为系统固有偏置量. 该模型通过线性组合方式有效整合了算法变量的汉明重量特征与系统的噪声干扰.

根据以上模型可知, 观测量 O 可表示为一个正态分布的概率模型 $N(\mu,\sigma)$, 其均值 $\mu=a\times HW(x)+b$.

3.2 攻击有效性验证实验

为了系统评估所提出的单能迹攻击方法在不同实验条件下对 SM4 分组密码算法主密钥恢复的有效性, 设计并实施了以下有效性验证实验. 有效性验证实验分为仿真实验和实测实验.

1)仿真实验

仿真实验所使用隐变量数据, 是在 Python 中复现智能卡安全芯片中的 SM4 密钥扩展算法中实现的, 数据包括主密钥 MK、轮子密钥、密钥分片、运算中间值等. 另一方面侧信道能量观测值数据模拟是使用 3.1 节所述的能量泄露观测模型.

通过设置不同系统参数 a、b 和 σ^2, 可计算得到信噪比(signal to noise, SNR), 借此可分析能量泄露观测模型的噪声对单轨迹攻击成功率的影响.

作为对比, 实验还实现了 DPA、CPA 以及常规模板攻击, 得到对应成功率结果. 实验结果如表 1 所示. 由于针对 SM4 密钥扩展算法, DPA、CPA 是无法实施的, 因此表中的值为 N/A. 对于常规模板攻击, 即使在完全无噪声影响下(即 $SNR=+\infty$), 由于泄露模型使用汉明重量模型而非数值模型, 除非常特殊的中间值(如全 1、全 0)外, 中间值恢复平均错误率大于 30% 以上, 导致常规模板攻击主密钥恢复成功率不到 1%. 在本次仿真实验环境下, 针对随机生成 10 000 组主密钥对应的仿真单能迹实施常规模板攻击测试, 结果表明所有攻击尝试均未能成功完整恢复主密钥信息.

2)实测实验

实测实验是仿真实验基础上, 使用在实验设置中的真实硬件平台所获取的能量轨迹进行的实验.

Table 1 Success Rates of the Four Types of Side-Channel Attack in the Simulation Experiments Under Different *SNR*s

表 1 仿真实验中 4 种侧信道攻击在不同 *SNR* 下的成功率

SNR/dB	单轨迹攻击/%	DPA/%	CPA/%	常规模板攻击/%
+∞	100	N/A	N/A	0
30	100	N/A	N/A	0
20	100	N/A	N/A	0
10	100	N/A	N/A	0

注: 数据值 N/A 为 Not Applicable (不适用) 的缩写, 表示该攻击无法完成; 而 0% 表示该攻击可以完成, 但密钥恢复成功率为 0%.

在建模阶段，实验采用选择密钥的方式，通过对有泄露的关键中间值的可能取值进行独立遍历，分别推导出相应主密钥 *MK* 的取值，再在相同硬件平台上输入这些主密钥，采集对应的能量轨迹来构造中间值的模板. 由于攻击对象是软实现，使用相同寄存器的不同中间值，其泄露函数是相同的，因此仅需分析算法实现中中间值与寄存器使用之间的关系，通过对少量寄存器建模即可完成本阶段的目标.

在攻击阶段，实验采用随机生成的方式构造了 10 000 组不同的密钥. 每组数据均独立且均匀分布于 SM4 算法可能的密钥空间. 随后，在相同硬件平台上对每组轮子密钥生成过程采集对应的能量消耗轨迹，形成用于后续攻击分析的单能迹样本集.

针对每一条采集到的能量轨迹，应用本文所提出的单能迹攻击方法进行处理. 具体而言，攻击流程包括预处理、特征提取、假设构建、边缘概率计算、密钥候选排序以及验证等关键步骤. 最终根据攻击输出是否准确恢复出原始主密钥作为判断攻击是否成功的标准.

实验结果如表 2 所示. 在实测 *SNR* 未知的情况下，全部 10 000 组实验样本中，共有 8 574 次单能迹攻击成功恢复出正确的 SM4 主密钥，攻击成功率达到 85.74%. 这一结果表明，所提出的单能迹攻击方法在大多数情况下能够有效识别并还原出加密过程中的主密钥信息，显示出较高的攻击能力.

Table 2　Success Rates of the Four Types of Side-Channel Attacks in the Actual Measurement Experiment

表 2　4 种侧信道攻击在实测实验中的成功率

SNR	单轨迹攻击/%	DPA 攻击/%	CPA 攻击/%	常规模板攻击/%
未知	86%	N/A	N/A	0%

注：数据值 N/A 为 Not Applicable（不适用）的缩写，表示该攻击无法完成；而 0% 表示该攻击可以完成，但密钥恢复成功率为 0%.

剩余 1 426 次未能成功恢复主密钥的实验中，进一步分析显示其失败原因主要包括 3 个方面：1）能量信号噪声干扰较大，影响特征提取准确性；2）实测能量信息泄露中存在非线性相关，与实验建模所使用线性概率模型存在一定差异，导致预测模型精度下降；3）某些特定密钥组合下，能量泄露特征差异不明显，造成统计分析失效.

3.3　鲁棒性验证试验

为评估本文所提出的单能迹攻击方法在实际侧信道攻击场景中对噪声干扰的适应能力与稳定性，设计并实施了鲁棒性验证实验.

在上述有效性验证的实测实验基础上，采用高斯白噪声作为干扰源，在能量轨迹处理阶段对原始信号进行人为加噪处理. 具体而言，分别设置 *SNR* 为 10 dB、20 dB 和 30 dB 三个典型水平，以代表从低到高的噪声干扰强度. 根据设定的 *SNR* 值，分别将对应的高斯白噪声与建模阶段采集的建模能迹集和攻击阶段采集的攻击能迹集中的原始能迹中进行加性叠加，形成有高斯白噪声注入的建模能迹集合攻击能迹集.

攻击实施与有效性验证相同，通过建模能迹集构造中间值模板，再用攻击能迹集中的单能迹实施攻击，最终根据攻击输出是否准确恢复出原始主密钥作为判断攻击是否成功的标准.

噪声注入环节采用标准的高斯分布模型，通过 Python 编程实现对原始轨迹的加噪处理. 攻击程序使用基于 Python 开发环境实现，包含信号预处理、特征提取、假设构建、相关性分析与密钥猜测等模块. 所有攻击均在统一软硬件平台上执行，避免因环境差异造成额外误差.

攻击实验的数据汇总如表 3 所示，随着噪声的削弱，*SNR* 增加，攻击成功率呈现明显的上升趋势，如图 7 所示. 同时实验结果说明，即使在较高噪声干扰下如 *SNR*=10 dB，噪声能量与原信号能量几乎相等，本文所提出的攻击方法仍能维持较高的主密钥恢复成功率，显示出良好的抗干扰能力.

进一步对未能成功恢复主密钥的攻击样本进行归类分析，发现其失败原因主要包括 3 个方面：1）在低 SNR 条件下，部分关键点的能量泄露被噪声淹没，导致特征提取失败；2）噪声能量增加导致中间值恢复错误率上升，整体纠错性能下降；3）部分失败样本还表现出一定的“边缘性”特征，即攻击虽然未能完全恢复轮子密钥，但能够准确识别出大部分字节，提示后续可通过结合其他辅助信息（如错误容忍机制或多轨迹融合策略）来进一步提升整体攻击成功率.

Table 3　Success Rates of the Four Types of Side-Channel Attack in the Real Experiments Under Different *SNR*s

表 3　实际实验中 4 种侧信道攻击在不同 *SNR* 下的成功率

SNR/dB	单能迹攻击/%	DPA/%	CPA/%	常规模板攻击/%
+∞	86	N/A	N/A	0
30	85	N/A	N/A	0
20	80	N/A	N/A	0
10	70	N/A	N/A	0

注：数据值 N/A 为 Not Applicable（不适用）的缩写，表示该攻击无法完成；而 0% 表示该攻击可以完成，但密钥恢复成功率为 0%.

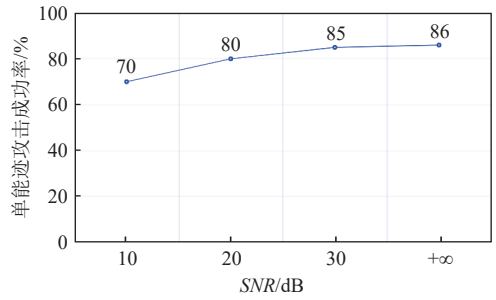


Fig. 7 The relationship between *SNR*s after adding Gaussian white noise and the success rates of single-trace attack

图 7 实测信号添加高斯白噪声后的信噪比 *SNR* 与单能迹攻击成功率的关系

本次鲁棒性验证实验的结果有力支持了本文所提出单能迹攻击方法在实际侧信道攻击中的可行性与有效性.

3.4 实验结论

通过上述验证实验, 评估了针对 SM4 密钥扩展算法的单能迹攻击方法的有效性及鲁棒性. 实验结果表明, 本文提出的单能迹攻击方法在不同条件下能够有效恢复主密钥, 具有较高的成功率和效率, 同时在不同噪声水平下表现出良好的鲁棒性. 此外, 与传统侧信道攻击方法相比, 单能迹攻击方法在攻击成功率、所需能量轨迹数量和攻击时间等方面均表现出显著优势. 这些结果表明, 单能迹攻击方法是一种高效、鲁棒且实用的侧信道攻击手段, 能够对基于 SM4 算法的密码系统构成潜在威胁.

4 结论与展望

4.1 研究成果总结

本文针对 SM4 分组密码算法的密钥扩展过程, 提出了一种基于贝叶斯网络结合建模侧信道攻击的单能迹侧信道攻击方法. 该方法通过构建概率图模型, 结合置信传播算法, 实现了对轮子密钥的高效推测, 进而恢复主密钥. 实验结果表明, 该方法在无噪声仿真环境下达到 100% 的成功率, 在实测环境下能够以 85.74% 的成功率恢复 SM4 主密钥, 且在不同 *SNR* 条件下均表现出良好的鲁棒性. 即使在 *SNR* 仅为 10 dB 的极端低信噪比条件下, 攻击成功率仍可达 70%. 此外, 该方法在多种硬件平台上均能有效实施, 具有较高的普适性. 与传统侧信道攻击方法相比, 单能迹攻击方法在攻击成功率、所需能量轨迹数量和攻击时间等方面均展现出显著优势, 为侧信道攻击领域提供了新的研究方向.

为抵御单能迹攻击, 建议 SM4 密钥扩展算法的实现采取多项防护措施以增强其安全性.

1)增强硬件的防护. 通过增加硬件混淆措施, 如引入随机噪声、时钟抖动等, 以提高攻击者获取有效能量轨迹的难度.

2)增强指令序列的防护. 引入随机冗余计算和并行乱序执行等策略. 通过增加随机冗余计算, 算法执行中可引入额外的随机性, 使攻击者难以预测算法行为, 从而提升算法的安全性. 同时, 并行乱序执行可打乱算法执行顺序, 使攻击者难以确定算法当前执行状态, 进一步提升攻击难度.

3)实施随机掩码防护. 随机掩码防护是一种有效的防御策略, 其核心在于算法执行过程中引入随机性的数据. 具体实施时, 可将密钥或其部分与随机生成的掩码进行运算, 以隐藏真实密钥信息. 即便攻击者能够获取能量轨迹, 也难以准确推断出真实密钥值. 由于掩码是随机生成的, 每次运算可能不同, 攻击者面临更大的不确定性, 从而增加攻击难度和复杂性. 在实施随机掩码防护时, 还需确保掩码生成和管理的安全性, 防止掩码泄露, 以保障整体防护体系的有效性.

这些防护措施的综合应用, 将显著增强 SM4 密钥扩展算法在面对单能迹侧信道攻击时的安全鲁棒性.

4.2 研究的局限性与未来工作方向

尽管本研究提出的单能迹攻击方法在实验中取得了较为理想的结果, 但其仍存在一定的局限性, 未来的研究可以从 6 个方面进一步优化和拓展:

1)提升噪声鲁棒性. 当前方法在低 *SNR* 条件下仍存在一定的失败率, 这表明噪声对攻击效果的影响仍不可忽视. 未来的研究可以进一步优化能量轨迹的预处理和特征提取算法, 以增强对噪声的鲁棒性. 例如, 可以探索更先进的去噪技术(如小波阈值去噪、主成分分析等)或结合深度学习方法自动提取更鲁棒的特征, 从而提高攻击在复杂电磁环境下的成功率.

2)优化攻击效率. 尽管单能迹攻击方法已经显著减少了所需能量轨迹的数量, 但进一步优化攻击算法的计算效率仍有空间. 例如, 可以探索更高效的贝叶斯网络结构或改进置信传播算法的迭代策略, 以缩短攻击时间. 此外, 结合硬件加速技术(如 FPGA 或 GPU 加速)也可能是一个有效的优化方向.

3)扩展至其他算法. 本研究主要针对 SM4 算法展开, 未来可以将单能迹攻击方法推广到其他分组密码算法, 如 AES、DES 等, 探索其在不同算法结构

下的适用性和优化策略. 这不仅有助于验证该方法的通用性, 还能为密码学领域的侧信道攻击研究提供更广泛的参考.

4)硬件防护对策研究. 为了应对单能迹攻击的威胁, 未来可以研究针对侧信道攻击的硬件防护技术, 如掩码技术、随机化技术等, 以提高密码设备的安全性. 通过构建更复杂的物理防护机制, 可以有效降低侧信道信息的泄露, 从而提高密码系统的抗攻击能力.

5)跨平台攻击模型的构建. 当前的攻击方法主要针对特定硬件平台进行了优化, 未来可以探索构建跨平台的攻击模型, 提高攻击方法的通用性和适应性. 通过领域自适应算法或迁移学习技术, 可以进一步验证该方法在不同芯片工艺和架构下的有效性, 为跨设备侧信道攻击研究提供新的思路.

6)针对分布式物联网的侧信道攻击防护策略研究. 鉴于本研究可对分布式物联网环境下的设备产生实质性的安全威胁, 未来研究将着重探索适用于分布式物联网设备的轻量化侧信道攻击防护策略. 针对物联网设备资源受限的特点, 研究如何在保证设备性能的同时, 有效降低侧信道信息泄露风险.

综上所述, 本研究提出的单能迹攻击方法为侧信道攻击领域提供了新的视角和解决方案. 未来的研究将进一步优化该方法, 并探索其在更广泛场景中的应用和防护对策, 以推动侧信道攻击与防御技术的进一步发展.

作者贡献声明: 吴震提出了攻击的技术方案, 给出有效性形式化证明, 完成论文撰写; 赵洋设计了实验方案, 完善了鲁棒性验证方法, 完成稿件的部分修订; 王敏实现了攻击平台, 完成了基于该平台上的攻击实测, 完成论文的整体修订.

参 考 文 献

[1] State Administration for Market Regulation, China National Standardization Administration, GB/T 32905−2016 Information Security Technology SM3 Cryptographic Hash Algorithm[S]. Beijing: China Standards Press, 2017
(国家市场监督管理总局, 中国国家标准化管理委员会, GB/T 32907−2016 信息安全技术 SM4 分组密码算法[S]. 北京: 中国标准出版社, 2017)

[2] Kocher P. Timing attacks on implementations of Diffie-Hellman, RSA, DSS, and other systems[C]//Advances in Cryptology—CRYPTO'96: 16th Annual Int Cryptology Conf. Berlin: Springer, 1996: 104−113

[3] Kocher P, Jaffe J, Jun B. Differential power analysis[C]//Proc of Annual Int Cryptology Conf. Berlin: Springer, 1999: 338−397

[4] Gilbert G, Benjamin, Josh J, et al. A testing methodology for side-channel resistance validation[J]. NIST Non-invasive Attack Testing Workshop, 2011(7): 115−136

[5] Chari S, Rao J R, Rohatgi P. Template attacks[C]//Proc of Int Workshop on Cryptographic Hardware and Embedded Systems. Berlin: Springer, 2002: 13−28

[6] Hospodar G, Gierlichs B, De M E, et al. Machine learning in side-channel analysis: A first study[J]. Journal of Cryptographic Engineering, 2011, 1(4): 293−302

[7] Jin Xin, Feng Junhao, Huang Boyang. Side channel attack on SM4 algorithm with deep learning-based analysis[C]//Proc of 2022 IEEE Int Conf on Advances in Electrical Engineering and Computer Applications (AEECA). Piscataway, NJ: IEEE, 2022: 749−752

[8] Kang Xiaoyun, Huang Kaitian, Lan Tian, et al. Template attacks on SM4 cipher algorithm[J]. Journal of Chengdu University of Information Technology, 2021, 36(5): 499−502 (in Chinese)
(匡晓云, 黄开天, 兰天, 等. 针对 SM4 密码算法的模板攻击[J]. 成都信息工程大学学报, 2021, 36(5): 499−502)

[9] Du Zhibo, Wu Zhen, Wang Min, et al. Multi-point joint power analysis attack against SM4[J]. Journal of Computer Research and Development, 2016, 53(10): 2224−2229 (in Chinese)
(杜之波, 吴震, 王敏, 等. 针对 SM4 密码算法的多点联合能量分析攻击[J]. 计算机研究与发展, 2016, 53(10): 2224−2229)

[10] Zhang Qi, Wu Zhen, Wang Min, et al. Research on correlation power attack on Sbox-input of SM4 algorithm[J]. Computer Applications and Software, 2015, 32(12): 290−293, 333 (in Chinese)
(张琪, 吴震, 王敏, 等. SM4 算法 S 盒输入的相关性能量攻击的研究[J]. 计算机应用与软件, 2015, 32(12): 290−293, 333)

[11] Du Zhi Bo, Wu Zhen, Wang Min, et al. Improved chosen-plaintext power analysis attack against SM4 at the round-output[J]. Journal on Communications, 2015, 36(10): 85−91 (in Chinese)
(杜之波, 吴震, 王敏, 等. 针对 SM4 轮输出的改进型选择明文功耗分析攻击[J]. 通信学报, 2015, 36(10): 85−91)

[12] Wang Min, Rao Jintao, Wu Zhen, et al. Power analysis attack against SM4 in frequency domain[J]. Information Network Security, 2015(8): 14−19 (in Chinese)
(王敏, 饶金涛, 吴震, 等. SM4 密码算法的频域能量分析攻击[J]. 信息网络安全, 2015(8): 14−19)

[13] Gu Dawu, Zhang Chi, Lu Xiangjun. Progress of some comments on the research of side-channel attack for cryptosystems[J]. Journal of Xidian University, 2021, 48(1): 14−21, 49 (in Chinese)
(谷大武, 张驰, 陆相君. 密码系统的侧信道分析: 进展与问题[J]. 西安电子科技大学学报, 2021, 48(1): 14−21, 49)

[14] Gallager R. Low-density parity-check codes[J]. IRE Transactions on Information Theory, 1962, 8(1): 21−28



Wu Zhen, born in 1975. PhD candidate. His main research interests include cybersecurity and security of cryptographic implementations.

吴　震, 1975 年生. 博士研究生. 主要研究方向为网络空间安全、密码实现安全.



Zhao Yang, born in 1973. PhD, associate professor. His main research interests include privacy computing and network cyberspace security.

赵 洋, 1973 年生. 博士, 副教授. 主要研究方向为隐私计算、网络空间安全.



Wang Min, born in 1977. PhD, associate professor. Her main research interests include cybersecurity and side-channel attacks.

王 敏, 1977 年生. 博士. 副教授. 主要研究方向为网络空间安全、侧信道攻击.