

分组密码算法的仿射线性密码分析

曹文芹^{1,2,3}, 张文涛^{1,2}

¹中国科学院信息工程研究所信息安全国家重点实验室 北京 中国 100093

²中国科学院大学网络空间安全学院 北京 中国 100049

³山东理工大学数学与统计学院 淄博 中国 255000

摘要 线性密码分析是分组密码算法重要的统计分析方法之一。自从 20 世纪 90 年代初 Matsui 提出线性密码分析以来, 很多学者已经对线性密码分析及其推广的各种线性分析进行了大量的研究, 辉煌的成果不计其数, 其中多维线性密码分析就是一种极其重要的研究成果。线性密码分析利用正确密钥和错误密钥下线性逼近的相关系数的平方服从具有不同期望值的 Gamma 分布来恢复密钥。多维线性密码分析使用了线性子空间中所有的非零线性逼近, 通过组合多个 Gamma 分布来扩大正确密钥和错误密钥下容量的差异, 从而提高了密钥恢复攻击的效率。而线性子空间中的线性逼近对容量的贡献大小不一样, 贡献较小的线性逼近对提高密钥恢复攻击效率起的作用较小。为此, Nyberg 提出了仿射线性密码分析, 它是多维线性密码分析的一个新的变体。仿射线性密码分析舍弃了多维线性子空间中对容量贡献较小或者没有贡献的一半线性逼近, 仅从保留的仿射子空间提取信息, 构造了更有效的卡方检验统计量对分组密码进行攻击。为了进一步提高攻击的效率, Nyberg 猜想舍弃仿射统计量中得分较小的项, 利用剩余项仍可以构造服从卡方分布的统计检验统计量。本文利用统计手段证明了 Nyberg 猜想是正确的, 并给出了该猜想的一个应用方法。最后, 利用 PRESENT 和 Serpent 算法验证了 Nyberg 模型的有效性。对 26 轮 PRESENT, 利用 Nyberg 猜想进行了数据复杂度为 $2^{61.5}$, 时间复杂度为 $2^{68.37}$ 的密钥恢复攻击; 对 27 轮 PRESENT, 可以利用 Nyberg 猜想进行数据复杂度为 $2^{63.25}$, 时间复杂度为 $2^{68.37}$ 的密钥恢复攻击。另外, 分析了 4 轮 Serpent 算法仿射线性密码分析的数据复杂度。

关键词 多维线性分析; 仿射线性分析; PRESENT

中图法分类号 TP309 DOI 号 10.19363/J.cnki.cn10-1380/tn.2026.05.17

Affine Linear Cryptanalysis of Block Cipher

CAO Wenqin^{1,2,3}, ZHANG Wentao^{1,2}

¹ State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100093, China

² School of Cyber Security, University of Chinese Academy of Sciences, Beijing 100049, China

³ School of Mathematics and Statistics, Shandong University of Technology, Zibo 255000, China

Abstract Linear cryptanalysis is one of the important statistical cryptanalytic methods for block ciphers. Since the early 1990s, Matsui proposed linear cryptanalysis, and many scholars have researched the properties of linear cryptanalysis and proposed all kinds of promotion methods based on linear cryptanalysis. They have obtained numerous brilliant achievements. Among them, multidimensional linear cryptanalysis is a kind of extremely important method. Under the right and wrong keys, the squared correlation of a linear approximation has Gamma distributions with different means. Based on this property, linear cryptanalysis can perform recovery key attacks. Multidimensional linear cryptanalysis uses all non-zero linear approximations in linear subspace and combines multiple Gamma distributions to amplify the difference in capacity between right and wrong keys. Therefore, multidimensional linear cryptanalysis improves the efficiency of the key recovery attacks. In linear subspace, the contributions of the linear approximations to the capacity are different, and the one with a smaller contribution plays a smaller role in improving the efficiency of the attacks. In this light, Nyberg proposed affine linear cryptanalysis, which is a new variant of multidimensional linear cryptanalysis. Affine linear cryptanalysis discards a whole half-space of linear approximations, which contribute little or nothing to the capacity, and only extracts information from the reserved affine subspace to construct more effective test statistics to attack block ciphers. In order to further improve the efficiency of the attacks, Nyberg conjectured that discarding the terms with low scores of affine statistic, and the sum of the remaining terms is also a statistic that follows chi-square distribution. This paper proves that Nyberg conjecture is correct using statistical techniques, and gives an application method of this conjecture. Finally, we verify the validity of the new model using PRESENT and Serpent algorithms. We perform 26 rounds attack on PRESENT with about $2^{61.5}$ data complexity and $2^{68.37}$ time complexity. For 27 rounds PRESENT, we perform key recovery attack with about $2^{63.25}$ data complexity and $2^{68.37}$ time complexity. In addition, we analyze the data complexity of affine linear cryptanalysis on 4-round Serpent.

通讯作者: 张文涛, 博士, 研究员, Email: zhangwentao@iie.ac.cn。

本课题得到国家自然科学基金(No. 61379138)资助。

收稿日期: 2021-03-11; 修改日期: 2021-04-27; 定稿日期: 2023-08-29

Key words multidimensional linear cryptanalysis; affine linear cryptanalysis; PRESENT

1 引言

在 1993 年的欧密年会上, Matsui^[1-2]提出了一个重要的密码分析方法——线性密码分析。随后, 人们对其进行了大量的研究, 扩展出了许多线性分析方法的变体, 例如多重线性密码分析^[3-4]、多维线性密码分析^[5-8]、零相关线性密码分析^[9-10]等。现在, 线性密码分析已经成为分析分组密码安全性的重要工具之一, 设计新的分组密码的一个基本要求就是它能够抵抗线性攻击。线性密码分析作为一类重要的统计攻击方法, 通过分析攻击过程中有关统计量的分布, 借助假设检验思想计算攻击的成功概率, 进而推导出数据复杂度的计算表达式, 这是线性密码分析的一个重要的研究问题。在分组密码线性密码分析过程中, 关键是建立一个有效的统计区分器, 用该区分器来判断观测数据是来自某指定的分组密码, 还是来自一个理想的分组密码。统计区分器通常是一个检验统计量——有关明密文数据和密钥的随机变量。文献[11]提出攻击优势的概念, 利用检验统计量在正确密钥和错误密钥下的不同分布, 推导出简单线性密码分析的数据复杂度与成功概率和优势之间的关系式。

把分组密码的简单线性分析推广到使用多个线性逼近情况, 文献[3]提出多重线性密码分析。它同时使用多条线性逼近, 大大改进了攻击的效率。因为正确密钥和错误密钥下相关系数的平方服从具有不同期望值的伽马分布, 所以使用多重线性密码分析的想法是通过组合多个伽马分布来扩大正确密钥和错误密钥下统计量期望值的差异, 从而更高效地区分正确与错误密钥。与简单线性逼近相比较, 在相同的成功概率和优势下, 减小了数据复杂度。随后, 为了避免处理线性逼近之间不独立问题, 文献[6]提出了多维线性密码分析。多维线性密码分析利用某个线性子空间上除了零之外的所有线性逼近进行密钥恢复攻击。决定攻击效率的一个重要指标是容量(capacity), 它度量多维线性逼近的分布与均匀分布之间的偏离程度。容量值越大, 越有利于区分正确与错误密钥。文献[7]证明了攻击的数据复杂度与容量近似呈反比例关系, 明确地给出了多维线性攻击的数据复杂度与成功概率、优势之间的关系。

Nyberg^[12]基于已有的多维线性密码分析提出线性密码分析的一个新的变体——仿射线性密码分析。它舍弃线性子空间中对容量贡献较小或没有贡

献的一半线性逼近, 仅从保留的仿射子空间提取信息进行统计推断——使用仿射子空间的线性逼近的相关系数构造新的卡方检验统计量。仿射检验统计量减小了卡方统计量的自由度, 从而减小了统计量的方差。在相同的成功概率和优势下, 达到减小数据复杂度的目的。仿射线性分析方法舍弃线性空间中相关系数绝对值较小的线性逼近, 使得该方法比多维线性密码分析更有效。

为了进一步减小检验统计量的自由度, Nyberg猜想^[12]舍弃仿射统计量中贡献较小的项, 利用剩余项构造的统计量仍服从卡方分布。

本文首先证明了 Nyberg 猜想^[12]是正确的, 接着给出了应用该猜想的一个方法。最后, 利用 PRESENT^[13]和 Serpent^[14]算法验证了仿射线性分析方法的有效性。对 26 轮 PRESENT, 可以利用 Nyberg 猜想进行数据复杂度为 $2^{61.5}$, 时间复杂度为 $2^{68.37}$ 的密钥恢复攻击; 对 27 轮 PRESENT, 可以利用 Nyberg 猜想进行数据复杂度为 $2^{62.2}$, 时间复杂度为 $2^{77.27}$ 的密钥恢复攻击。另外, 对文献[6]中的 4 轮 Serpent 的 7 维线性逼近进行分析发现, Nyberg 猜想的攻击效率与 LLR 方法相当。攻击结果与以前主要的线性攻击结果比较见表 1 和表 2。

表 1 PRESENT 线性攻击结果比较

Table 1 Comparison of linear attack on PRESENT

#Round	Time	Data	Ref.
26	$2^{68.37}$	$2^{61.9}$	本文仿射线性分析
		$2^{61.5}$	本文 Nyberg 猜想
27	$2^{68.37}$	$2^{63.68}$	本文仿射线性分析
		$2^{63.25}$	本文 Nyberg 猜想
27	$2^{77.27}$	$2^{62.54}$	本文仿射线性分析
		$2^{62.2}$	本文 Nyberg 猜想
26	$2^{75.3}$	$2^{61.9}$	[21]多元线性分析
27	$2^{77.27}$	$2^{63.83}$	
25	2^{65}	$2^{62.4}$	[17]多维线性分析
26	2^{72}	2^{64}	

表 2 4 轮 Serpent 的 7 维线性分析的数据复杂度

Table 2 The data complexity of 7-dimensional linear cryptanalysis on 4-round of Serpent

方法	Data	方法	Data
χ^2 -method ^[7]	$2^{28.4}$	本文	$2^{27.96}$
LLR-method ^[7]	$2^{27.5}$	Nyberg 猜想	$2^{27.57}$

本文组织如下, 第 2 节介绍简单、多维和仿射线

性密码分析; 第3节介绍Nyberg猜想, 给出证明, 并阐述Nyberg猜想的一个应用方法; 第4节利用仿射线性分析对26、27轮PRESENT算法进行了密钥恢复攻击, 分析了4轮Serpent算法仿射线性分析的数据复杂度; 第5节进行总结。

2 简单、多维和仿射线性密码分析

2.1 简单线性密码分析

在线性密码分析中, 一个变换 $F: F_2^n \rightarrow F_2^m$ 的线性逼近 $(u, v) \in F_2^n \times F_2^m$ 是一个布尔函数:

$$x \mapsto u \cdot x + v \cdot F(x), \quad u \in F_2^n, v \in F_2^m$$

u 称为输入掩码, v 称为输出掩码。线性逼近 (u, v) 的偏定义为

$$\varepsilon_{(u,v)} = P(u \cdot x + v \cdot F(x) = 0) - 1/2,$$

相关系数定义为 $c_{(u,v)} = 2\varepsilon_{(u,v)}$, $c_{(u,v)} \in [-1, 1]$, 线性概率(LP)定义为相关系数的平方, 记作 $LP_{(u,v)} = c_{(u,v)}^2$ 。

假设随机抽取 N 对明密文 x , 数据样本记作 D 。线性逼近 (u, v) 的经验相关系数为

$$\hat{c}_{(u,v)}(D) = \frac{2}{N} \cdot (\#\{x \in D \mid u \cdot x + v \cdot F(x) = 0\} - 1),$$

简单线性密码分析就是根据 $|\hat{c}_{(u,v)}(D)|$ 的大小进行密钥恢复攻击。

给定迭代分组密码 $E(P, K): F_2^n \times F_2^k \rightarrow F_2^n$, 分组长度为 n 比特, 轮密钥长度为 k 比特。我们令 $(\tau_i, \tau_{i+1}), i = 0, 1, \dots, r-1$ 是迭代分组密码的一系列轮线性逼近, 这个序列称为线性迹, 记作 $\tau = (\tau_0, \dots, \tau_r)$ 。当各轮线性逼近相互独立时, 由堆积引理知线性迹的相关系数为

$$c(\tau) = \prod_{i=0}^{r-1} c(\tau_i, \tau_{i+1})$$

迭代分组密码的一条线性壳是指一对掩码 (u, v) , 线性壳 (u, v) 的相关系数为

$$c_{(u,v)} = \sum_{\tau, \tau_0=u, \tau_r=v} c(\tau)$$

线性密码分析中的另一个重要参数是平均线性概率(ELP), 定义如下。

定义 1(文献[15]). 考虑一个密钥交替分组密码, 它的线性逼近的平均线性概率(ELP)是线性逼近在随机密钥下线性概率的平均值。

定理 1(文献[15]). 考虑一个密钥交替分组密码, 线性逼近 (u, v) 的平均线性概率(ELP)是所有输入掩码为 u , 输出掩码为 v 的线性迹的线性概率的和, 记

作 $ELP(u, v)$,

$$ELP(u, v) = \sum_{\tau, \tau_0=u, \tau_r=v} LP_{\tau}$$

对于密钥交替迭代分组密码, 文献[16]提出了一种估计 ELP 的方法——信噪分解方法。考虑线性逼近 $(u, v), \tau = (\tau_0, \dots, \tau_r)$ 是满足 $\tau_0 = u, \dots, \tau_r = v$ 的 r 轮线性迹。

其相关系数可以表示为

$$c_{(u,v)}(K) = \sum_{\tau} (-1)^{\tau \cdot K} \rho_{\tau},$$

ρ_{τ} 表示线性迹 τ 的相关系数的绝对大小。

根据定理 1, 密钥交替密码算法的线性逼近 (u, v) 的平均线性概率为

$$ELP(u, v) = \sum_{\tau} \rho_{\tau}^2$$

假设 S 表示线性逼近 (u, v) 的显著线性迹(signal)集合, 相应的总相关系数表示为

$$Q(u, v)(K) = \sum_{\tau \in S} (-1)^{\tau \cdot K} \rho_{\tau}$$

令 $R(K)$ 表示线性逼近 (u, v) 剩余的非显著线性迹(noise)对相关系数的贡献

$$R(K) = c_{(u,v)}(K) - Q(u, v)(K) = \sum_{\tau \notin S} (-1)^{\tau \cdot K} \rho_{\tau}$$

由文献[15]可知, 随机置换的线性逼近的相关系数近似服从均值为 0, 方差为 2^{-n} 的正态分布。假设 $R(K)$ 与随机噪声一样, 那么有 $R(K) \sim N(0, 2^{-n})$ 。假设密码采用随机密钥, 根据该分布计算得

$$ELP(u, v) = \text{Exp}(c_{(u,v)}^2(K)) = \sum_{\tau \in S} \rho_{\tau}^2 + 2^{-n}$$

这就给出了 ELP 的一个估计。文献[16]用PRESENT算法验证了上述估计方法比文献[17]忽略“噪声”的估计 $\sum_{\tau \in S} \rho_{\tau}^2$ 要精确得多。

线性密码分析是一类统计密码攻击方法, 这类统计攻击的主要前提是利用密码算法的某个统计性质, 通过观测到的明密文数据建立检验统计量, 记作 $T(D)$, 借助假设检验的思想从所有可能候选密钥中判断出正确密钥。在线性攻击中, 通常把检验统计量看作样本数据的随机变量。对所有错误的候选密钥, 假设检验统计量是同分布的; 类似地, 对正确的候选密钥, 检验统计量也是同分布的^[18]。当候选密钥为正确密钥时, 样本数据 D 来自密码算法, 服从非均匀分布, 此时检验统计量记作 $T_R(D)$ 。当候选密钥为错误密钥时, 样本数据 D 来自随机置换, 服从均匀分布, 此时检验统计量记作 $T_W(D)$ 。不失一般性我们假设 $E(T_W(D)) \leq E(T_R(D))$, 检验的临界值记为 θ ,

判断的规则为: 如果 $T(D) \geq \Theta$, 则认为 $T(D)$ 服从 $T_R(D)$ 的分布; 如果 $T(D) < \Theta$, 则认为 $T(D)$ 服从 $T_W(D)$ 的分布。

文献[11]首次定义了攻击优势的概念, 攻击的 a 比特优势是指攻击的筛选过程中保留了 2^{-a} 比例的候选密钥。攻击者关注的是正确密钥是否在保留下来的密钥中。令 P_s 表示正确密钥被保留下来的概率, 那么 P_s 就是攻击的成功概率。利用假设检验的思想得到:

$$2^{-a} = P(T_W(D) \geq \Theta), P_s = P(T_R(D) \geq \Theta)$$

根据 $T_R(D)$ 和 $T_W(D)$ 分布, 可以确定攻击的数据复杂度、优势和成功概率之间的关系表达式。

2.2 多维线性密码分析

令 $F: F_2^n \rightarrow F_2^n$ 是一个分组长度为 n 比特的分组密码算法。简单的线性密码分析只考虑了一个显著的线性逼近, 但是对许多的分组密码而言, 它通常具有许多相关系数不可忽略的线性逼近。多维线性密码分析就是考虑了某线性子空间上的所有非零的线性逼近, 利用多个线性逼近提取信息, 构造检验统计量。它显著提高了线性攻击的效率, 并且有一个重要优点——不必考虑线性逼近之间是否独立的问题。

假设 $\lambda_i(x) = u_i \cdot x + v_i \cdot F(x)$, $i = 1, \dots, t$, 是线性无关的 t 个线性逼近, 称为线性逼近子空间的基。记 $\lambda(x) = (\lambda_1(x), \dots, \lambda_t(x))$, 该线性基构造的线性子空间记作 U , U 是 $F_2^n \times F_2^n$ 的某个子空间。任意掩码对 $(u, v) \in U$ 都是由 $\lambda(x)$ 生成的线性逼近。多维线性密码分析考虑以线性基构造的子空间 U 中所有非零线性逼近 $(u, v) \in U, (u, v) \neq 0$ 。为了进行多维线性攻击, 我们需要计算 t 维线性逼近 $\lambda(x)$ 的多维概率分布:

$$p_\xi = 2^{-n} \cdot \#\{x \mid \lambda(x) = \xi\}, \xi \in F_2^t$$

如果 F 是一个随机置换, 那么多维线性逼近的概率分布 $p = (p_1, \dots, p_{2^t-1})^T$ 服从均匀分布。因此, 多维线性攻击的基本原理就是分析其概率分布 p 与均匀分布的差异程度, 差异越大, 越容易区分密码算法与随机置换。这样就引入了多维线性攻击中的一个重要参数, 其定义如下。

定义 2. 给定两个离散随机分布 $p = (p_1, \dots, p_{2^t-1})^T$,

$q = (q_1, \dots, q_{2^t-1})^T$, 它们之间的容量(capacity)定义为

$$C(p, q) = \sum_{\xi=0}^{2^t-1} \frac{(p_\xi - q_\xi)^2}{q_\xi}$$

多维线性逼近的分布的容量(capacity)定义为

$$C_\lambda(K) = 2^t \cdot \sum_{\xi \in F_2^t} (p_\xi - 2^{-t})^2,$$

它度量了多维线性逼近的概率分布 p 与离散均匀分布之间的差异程度。容量的取值越大, 越容易区分密码算法与随机置换。文献[6]中推论 2 给出了容量的另一种表达式:

$$C_\lambda(K) = \sum_{(u,v) \in U, (u,v) \neq 0} c_{(u,v)}^2$$

根据平均线性概率的定义, 平均容量等于线性子空间上所有非零线性逼近的平均线性概率的和, 记作 C_λ , 即

$$C_\lambda = \text{Exp}(C_\lambda(K)) = \sum_{\substack{(u,v) \in U, \\ (u,v) \neq 0}} \text{ELP}_{(u,v)}$$

下面我们以最后一轮的密钥恢复攻击说明多维线性密钥恢复攻击的过程(见图 1)。给定一个 r 轮迭代分组密码: $E_K(\cdot): F_2^n \times F_2^k \rightarrow F_2^n$, 分组长度为 n 比特。明文 $x \in \{0, 1\}^n$, 扩展密钥 $K = (k_1, \dots, k_r)$, k_i 为密码算法的第 i 轮的轮密钥, $i = 1, \dots, r$, 密文 $y \in \{0, 1\}^n$, 满足: $y = E_K(x)$ 。假设密码算法的第 r 轮的轮函数为 G , 前 $r-1$ 轮的轮函数为 E' , $K' = (k_1, \dots, k_{r-1})$, k'_r, y' 分别表示攻击所涉及的第 r 轮的轮密钥和密文, 有 $y' = G_{k'_r} \circ E_{K'}(x)$ 。攻击者要从所有的候选密钥 κ 中确定正确密钥 k'_r 。如果候选密钥 κ 为正确密钥, $\kappa = k'_r$, 那么有, $G_{k'_r}^{-1}(y') = E_{K'}(x)$ 。

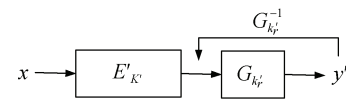


图 1 最后一轮密钥恢复攻击

Figure 1 The last round key-recovery attack

假设随机抽取 N 对明密文 x , 样本记作 D , 由样本数据 D 计算 $\lambda(x)$ 的经验概率 $\hat{p}_\xi(D)$:

$$\hat{p}_\xi(D) = N^{-1} \cdot \#\{x \in D \mid \lambda(x) = \xi\}$$

$\hat{c}_{(u,v)}(D)$ 表示线性逼近 $u \cdot x + v \cdot G_k^{-1}(E_{K'}(x))$ 的经验相关系数。多维线性密钥恢复攻击的检验统计量是多维线性逼近的经验容量, 记作 $T(D)$,

$$\begin{aligned} T(D) &= N \cdot \sum_{(u,v) \in U, (u,v) \neq 0} \hat{c}_{(u,v)}^2(D) \\ &= N 2^t \cdot \sum_{\xi \in F_2^t} (\hat{p}_\xi(D) - 2^{-t})^2. \end{aligned}$$

对所有可能的猜测密钥, 计算统计量 $T(D)$ 的值,

密钥恢复攻击的判断准则为: 若 $T(D) \geq \tau$, τ 为临界值, 则保留猜测密钥; 若 $T(D) < \tau$, 则舍弃猜测密钥。

根据假设检验的思想进行密钥恢复攻击需要知道统计量 $T(D)$ 分别在正确与错误密钥猜测下的概率分布。由文献[7]的结论可知, $T(D)$ 服从卡方分布:

$$T(D) \sim \begin{cases} \chi_{2^t-1}^2, & \text{样本 } D \text{ 来自均匀分布} \\ \chi_{2^t-1}^2(NC_\lambda), & \text{样本 } D \text{ 来自非均匀分布} \end{cases}$$

文献[7]利用 χ^2 方法确定了多维线性分析的数据复杂度表达式。给定成功概率 P_s 和攻击的 a 比特优势, 数据复杂度表达式为

$$N_{\chi^2} = \frac{2\sqrt{a(2^t-1)} + 4\Phi^{-2}(2P_s-1)}{C_\lambda} \quad (1)$$

这里 Φ 为标准正态分布的分布函数。

2.3 仿射线性密码分析

多维线性密码分析考虑了某线性子空间上的所有非零线性逼近, 它的一个主要优点是不需要对涉及的线性逼近做独立性假设。但从另一方面看, 线性子空间中包含许多弱的线性逼近, 从而降低了分析的效率。仿射线性分析考虑舍弃一半对攻击用途较弱的线性逼近, 即舍弃对容量贡献较小的一半线性逼近, 构造新的检验统计量, 提高线性分析方法的效率。

我们用一个典型的例子展示仿射线性密码分析的有效性。假设多维线性逼近的线性子空间为 U , $(u, v) \in U$ 。给定输入掩码 $u, u \neq 0$, 输出掩码 $v \in Q$, Q 为某线性子空间。包含形如 (u, v) 的掩码对的最小线性子空间为 $\{0, 1\} \times Q$ 。在该线性子空间中, 形如 $(0, v)$ 的线性逼近的相关系数为零, 它们对多维线性逼近的容量没有贡献, 但增大了线性子空间。显然,

$$\{0, u\} \times Q = (\{u\} \times Q) \cup (\{0\} \times Q),$$

仿射线性分析方法舍弃了 $\{0\} \times Q$ 中对线性攻击无用的线性逼近, 利用仿射子空间 $\{u\} \times Q$ 中的线性逼近构造仿射统计量进行攻击。

把 t 维线性逼近 λ 划分成两个部分 λ_1, λ_2 , 满足 $\lambda_1: F_2^n \rightarrow F_2, \lambda_2: F_2^n \rightarrow F_2^{t-1}, \lambda(x) = (\lambda_1(x), \lambda_2(x))$ 。 λ_1 是一维线性逼近, λ_2 是 $t-1$ 维线性逼近, 构成的线性子空间记作 V 。为了提高攻击的效率, 应该选择合适的线性逼近 λ_2 , 使得 λ_2 的容量 C_{λ_2} 尽可能小, 舍弃 V 中所有的线性逼近构造仿射线性分析的检验统计量。

下面我们先利用线性逼近 λ 的概率分布确定 λ_2 的概率分布, 再用这两个概率分布分别构造多维线性攻击统计量, 最后用这两个统计量的差构造仿射统计量。

令 $s = t-1$, 记 λ_2 的可能取值为 $\eta, \eta \in F_2^s, \lambda$ 的可能取值 $\xi = (b, \eta), b = 0, 1$ 。线性逼近 λ_2 的概率分布可以由 λ 的概率分布 p_ξ 确定:

$$\begin{aligned} P(\lambda_2(x) = \eta) &= P(\lambda(x) = (1, \eta)) + P(\lambda(x) = (0, \eta)) \\ &= p_{(1, \eta)} + p_{(0, \eta)}. \end{aligned}$$

可见, 由 λ 的概率分布可以确定 λ_2 的概率分布。随机抽取 N 对明密文 x , 样本记作 D 。由样本数据 D 可以计算经验概率 $\hat{p}_{(b, \eta)}(D), b = 0, 1$ 和经验相关系数 $\hat{c}_{(u, v)}^2(D), (u, v) \in U$ 。

由 λ 确定的 t 维线性攻击统计量为

$$\begin{aligned} T_\lambda(D) &= N2^{s+1} \cdot \sum_{b=0,1; \eta \in F_2^s} (\hat{p}_{(b, \eta)}(D) - 2^{-s-1})^2 \\ &= N \cdot \sum_{\substack{(u, v) \in U \\ (u, v) \neq 0}} \hat{c}_{(u, v)}^2(D). \end{aligned}$$

由 λ_2 确定的 $t-1$ 维线性攻击统计量为

$$\begin{aligned} T_{\lambda_2}(D) &= N2^s \cdot \sum_{\eta \in F_2^s} (\hat{p}_{(1, \eta)}(D) + \hat{p}_{(0, \eta)}(D) - 2^{-s})^2 \\ &= N \cdot \sum_{\substack{(u, v) \in V \\ (u, v) \neq 0}} \hat{c}_{(u, v)}^2(D). \end{aligned}$$

定义仿射检验统计量^[6] $T_{\text{aff}}(D)$:

$$T_{\text{aff}}(D) = T_\lambda(D) - T_{\lambda_2}(D)$$

引理 1(文献[12]), 如上定义, 我们有

$$T_{\text{aff}}(D) = N2^s \sum_{\eta \in F_2^s} (\hat{p}_{(1, \eta)}(D) - \hat{p}_{(0, \eta)}(D))^2$$

由引理 1 可见, 对于随机抽取的 x , 按照 λ_2 的取值仿射线性分析统计量 $T_{\text{aff}}(D)$ 的计算分成 2^s 个类别。由 λ_1 的取值, 每个类别是两个子集之间的差异。如果 $\lambda(x) = (\lambda_1(x), \lambda_2(x))$ 在 $\{0, 1\} \times F_2^s$ 上均匀分布, 差异的期望为零。两个子集之间的差异越大越易区分出正确与错误密钥。

可以证明仿射检验统计量 $T_{\text{aff}}(D)$ 服从卡方分布(见文献[12]):

$$T_{\text{aff}}(D) \sim \begin{cases} \chi_{2^s}^2, & p_{(b, \eta)}(D) \text{ 服从均匀分布,} \\ \chi_{2^s}^2(N\delta), & \text{otherwise.} \end{cases}$$

其中, $\delta = (C_\lambda - C_{\lambda_2}) = 2^s \sum_{\eta \in F_2^s} (p_{(1, \eta)} - p_{(0, \eta)})^2$, C_λ 、

C_{λ_2} 分别表示线性多维线性逼近 λ 、 λ_2 的分布的平均容量。

由(1)式知, 利用 χ^2 方法仿射线性攻击的数据复杂度表达式为

$$N_{\text{aff}} = \frac{2\sqrt{a \cdot 2^s} + 4\Phi^{-2}(2P_s - 1)}{\delta} \quad (2)$$

3 Nyberg 猜想

本节介绍并证明了 Nyberg^[12]猜想, 并给出它的一个应用方法。

3.1 Nyberg 猜想的证明

为了证明 Nyberg^[12]猜想, 我们需要以下引理。

引理 2. 使用上面的记号, 由 λ 确定的 t 维线性攻击统计量 $T_\lambda(D)$ 可以表示为

$$T_\lambda(D) = X^T (A^T A) X$$

由 λ_2 确定的 $t-1$ 维线性攻击统计量 $T_{\lambda_2}(D)$ 可以表示为

$$T_{\lambda_2}(D) = 2^{-1} X^T (A^T B A) X$$

这里 $A = \begin{pmatrix} I_{M-1} \\ -1_{M-1}^T \end{pmatrix}$, $B = \begin{pmatrix} I_{M/2} & I_{M/2} \\ I_{M/2} & I_{M/2} \end{pmatrix}$, $M = 2^{s+1}$,

I_{M-1} 是 $M-1$ 阶单位矩阵, 1_{M-1} 是所有元素都是 1 的 $M-1$ 阶列向量。令 η 为 λ_2 的取值: $\eta = 0, 1, \dots, 2^s - 1$, 记

$$X_{\eta+1} = \sqrt{N2^{s+1}} (\hat{p}_{(0,\eta)} - 2^{-s-1}),$$

$$X_{\eta+2^s+1} = \sqrt{N2^{s+1}} (\hat{p}_{(1,\eta)} - 2^{-s-1}),$$

满足 $\sum_{i=1}^{2^{s+1}} X_i = 0$ 。随机向量 $X = (X_1, X_2, \dots, X_{M-1})^T$

近似地服从多元正态分布 $N_{M-1}(\theta, \Sigma)$, 其中期望向量为

$$\theta = \sqrt{N2^{s+1}} (N^{-1} \cdot \mu - 2^{-s-1} \cdot 1_{M-1}),$$

$\mu = (Np_0, Np_2, \dots, Np_{M-1})^T$, 协方差阵为

$$\Sigma = I_{M-1} - 2^{-s-1} 1_{M-1} 1_{M-1}^T。$$

证明. 先介绍几个符号。记 $t = s+1$, 把 t 维的线性逼近 λ 划分成两个部分:

$$\lambda_1: F_2^n \rightarrow F_2, \lambda_2: F_2^n \rightarrow F_2^s,$$

满足 $\lambda(x) = (\lambda_1(x), \lambda_2(x))$ 。 λ_1 是一维线性逼近, λ_2 是 s 维线性逼近。

$V(\xi)$ 表示 N 对明密文使得 $s+1$ 维线性逼近 λ 取值为 ξ 的频数, $\xi = 0, 1, \dots, 2^{s+1} - 1$, 相应的频率为

$$\hat{p}_\xi = V(\xi)/N。$$

记 $M = 2^{s+1}$, $V = (V(0), V(1), \dots, V(M-1))^T$, 随机向量 V 服从多项式分布, 其期望向量记为 μ , 这里 $\mu = (Np_0, Np_2, \dots, Np_{M-1})^T$, 又由于

$$\text{Var}(V(\xi)) = Np_\xi(1 - p_\xi) \approx N2^{-s-1}(1 - 2^{-s-1}),$$

$\text{Cov}(V(i), V(j)) = -Np_i p_j \approx -N2^{-(s+1)}, i \neq j$, 故协方差阵近似为 $N\Sigma_0$, $\Sigma_0 = 2^{-s-1} \cdot (I_{M-1} - 2^{-s-1} 1_{M-1} 1_{M-1}^T)$, I_{M-1} 是 $M-1$ 阶单位矩阵, 1_{M-1} 是所有元素都是 1 的 $M-1$ 阶列向量。

记 λ_2 取值为 η , 令 $X_{\eta+1} = \sqrt{N2^{s+1}} (\hat{p}_{(0,\eta)} - 2^{-s-1})$,

$X_{\eta+2^s+1} = \sqrt{N2^{s+1}} (\hat{p}_{(1,\eta)} - 2^{-s-1})$, $\eta = 0, 1, \dots, 2^s - 1$, 满足 $\sum_{i=1}^{2^{s+1}} X_i = 0$ 。由多项式分布的渐近正态性, 可以知随机向量 $X = (X_1, X_2, \dots, X_{M-1})^T$ 近似地服从多元正态分布 $N_{M-1}(\theta, \Sigma)$, 其中期望向量为

$$\theta = \sqrt{N2^{s+1}} (N^{-1} \cdot \mu - 2^{-s-1} \cdot 1_{M-1}),$$

协方差阵为: $\Sigma = 2^{s+1} \cdot \Sigma_0 = I_{M-1} - 2^{-s-1} 1_{M-1} 1_{M-1}^T$ 。

记 $X = (X_1, X_2, \dots, X_M)^T$, 因为 $\sum_{i=1}^M X_i = 0$, 所以

$$X = (X_1, X_2, \dots, X_{M-1}, -\sum_{i=1}^{M-1} X_i)^T = A \cdot (X_1, X_2, \dots, X_{M-1})^T$$

这里 A 是 $M \times (M-1)$ 矩阵, $A = \begin{pmatrix} I_{M-1} \\ -1_{M-1}^T \end{pmatrix}$ 。下面使用

以上符号表示仿射检验统计量。

先表示由 λ 确定的 $s+1$ 维线性模型:

$$T_\lambda(D) = N2^{s+1} \cdot \sum_{b=0,1;\eta \in F_2^s} (\hat{p}_{(b,\eta)}(D) - 2^{-s-1})^2$$

$$= X^T X = X^T (A^T A) X$$

这里 $A^T A = I_{M-1} + 1_{M-1} 1_{M-1}^T$ 。

再表示由 λ_2 确定的 s 维线性模型:

$$T_{\lambda_2}(D) = N2^s \cdot \sum_{\eta \in F_2^s} (\hat{p}_{(1,\eta)}(D) + \hat{p}_{(0,\eta)}(D) - 2^{-s})^2$$

$$= 2^{-1} [(X_1 + X_{2^s+1})^2 + \dots + (X_{2^s} + X_{2^s+1})^2]$$

$$= 2^{-1} \cdot X^T (A^T B A) X$$

这里 B 是 $M \times M$ 矩阵, $B = \begin{pmatrix} I_{M/2} & I_{M/2} \\ I_{M/2} & I_{M/2} \end{pmatrix}$ 。

证毕。

下面介绍 Nyberg^[12]猜想, 观察仿射检验统计量:

$$T_{\text{aff}}(D) = N2^s \sum_{\eta \in F_2^s} (\hat{p}_{(1,\eta)}(D) - \hat{p}_{(0,\eta)}(D))^2,$$

如果对某些类别 $\eta \in F_2^s$, 得分 $\hat{p}_{(1,\eta)}(D) - \hat{p}_{(0,\eta)}(D)$ 接近于零, 那么它们对仿射统计量的贡献较小。因此考虑舍弃统计量 $T_{\text{aff}}(D)$ 中的这些项。为了简便, 假设统计量中剩余 $\eta_1, \dots, \eta_k \in F_2^s$ 对应的 k 项, 仿射检验统计量简化为

$$N2^s \sum_{i=1}^k (\hat{p}_{(1,\eta_i)}(D) - \hat{p}_{(0,\eta_i)}(D))^2,$$

记作 $T_k(D)$ 。Nyberg^[12]猜想利用上述剩余项构造的统计量 $T_k(D)$ 近似地服从自由度为 k , 非中心参数为 $N\delta'$ 的卡方分布, 这里 $\delta' = 2^s \sum_{i=1}^k (p_{(1,\eta_i)} - p_{(0,\eta_i)})^2$ 。

统计量 $T_k(D)$ 与仿射统计量 $T_{\text{aff}}(D)$ 相比较, 两者的均值之间差异较小, 但 $T_k(D)$ 的方差比 $T_{\text{aff}}(D)$ 的方差小很多, 因此该剩余的和统计量比仿射统计量更有效的。下面我们给出该猜想的证明。

定理 2[Nyberg 猜想]. 使用上面的记号, 检验统计量 $T_k(D)$ 近似服从自由度为 k , 非中心参数为 $N\delta'$ 的卡方分布, 即

$$T_k(D) \xrightarrow{d} \chi_k^2(N\delta').$$

其中, $\delta' = 2^s \sum_{i=1}^k (p_{(1,\eta_i)} - p_{(0,\eta_i)})^2$ 。

证明. 由引理 2, 仿射检验统计量可以表示为

$$T_{\text{aff}}(D) = T_k(D) - T_{\lambda_2}(D) = X^T CX$$

这里 $C = 2^{-1} \cdot [A^T(2I_M - B)A]$ 。可以验证 $(C\Sigma_0)^2 = C\Sigma_0$ (见附录 A), 对称阵 C 的秩为 2^s 。

我们把二次型 $(X_l - X_{2^s+l})^2$ 表示为 $X^T A_l X$, $l=1, \dots, M/2$ 。这里 A_l 是二次型对应的 $M-1$ 阶对称阵 (A_l 见附录 B)。

因为 $\sqrt{N2^{s+1}}(\hat{p}_{(0,\eta)}(D) - \hat{p}_{(1,\eta)}(D)) = X_{\eta+1} - X_{\eta+2^{s+1}}$, 所以仿射检验统计量可以表示为

$$\begin{aligned} T_{\text{aff}}(D) &= N2^s \sum_{\eta \in F_2^s} (\hat{p}_{(1,\eta)}(D) - \hat{p}_{(0,\eta)}(D))^2 \\ &= 2^{-1} [(X_1 - X_{2^s+1})^2 + \dots + (X_{2^s} - X_{2^{s+1}})^2] \\ &= 2^{-1} \cdot X^T \left(\sum_{j=1}^{M/2} A_j \right) X. \end{aligned}$$

记 $C_j = 2^{-1} \cdot A_j, j=1, \dots, M/2, C = \sum_{j=1}^{M/2} C_j$, 则

$$T_{\text{aff}} = X^T CX = X^T \left(\sum_{j=1}^{M/2} C_j \right) X.$$

矩阵 C_j 的秩为 1, 且满足 $\sum_{j=1}^{M/2} \text{rank}(C_j) = \text{rank}(C) = 2^s, (C\Sigma_0)^2 = C\Sigma_0$ 。又因为 $X \sim N_{M-1}(\theta, \Sigma)$, 由科克伦定理[见附录 C]得到以下结论:

(1) $T_{\text{aff}} = X^T CX$ 服从自由度为 2^s , 非中心参数为 $N\delta = \theta^T C\theta = N2^s \sum_{\eta \in F_2^s} (p_{(1,\eta)} - p_{(0,\eta)})^2$ 的卡方分布。

(2) $X^T C_j X$ 服从自由度为 1, 非中心参数为 $\theta^T C_j \theta$ 的卡方分布, $j=1, \dots, M/2$, 且相互独立, 即仿射统计量和式的每一项 $(\hat{p}_{(0,\eta)}(D) - \hat{p}_{(1,\eta)}(D))^2$ 服从自由度为 1 的卡方分布, 且相互独立。

因为有限的独立卡方分布随机变量的和仍服从卡方分布, 其自由度为各个卡方分布的自由度的和, 其非中心参数为各个卡方分布非中心参数的和。所以, $T_k(D)$ 服从自由度为 k 的卡方分布, 非中心参数为 $N\delta'$ 。

证毕。

3.2 Nyberg 猜想的应用

仿射统计量 $T_{\text{aff}}(D)$ 的表达式是一个关于样本数据的平方和。Nyberg 猜想就是根据和式中的各个平方项数值的大小考虑舍弃某些取值较小的项, 构造统计量 $T_k(D)$ 。在实际应用中, 我们要舍弃哪些项? 舍弃的标准是什么呢?

下面我们分析各项的概率分布, 给出一个舍弃标准。按照舍弃方法构造统计量 $T_k(D)$, 确定攻击的数据复杂度的表达式。

假设: 在一个 $s+1$ 维线性攻击中, $s+1$ 维线性逼近的概率分布 $p_\xi, \xi=0, 1, \dots, 2^{s+1}-1$ 服从独立同分布的正态分布 $N(2^{-s-1}, \sigma^2)$, 这里 $\sigma^2 = 2^{-s-1} \cdot \text{ELP}$ 。

由以上假设近似地有 $\hat{p}_{(b,\eta)}(D) \sim N(2^{-s-1}, \sigma_1^2)$, 其中方差 $\sigma_1^2 = 2^{-s-1} \cdot (\frac{1}{N} + \text{ELP})$ 假设 $\hat{p}_{(1,\eta)}(D), \hat{p}_{(0,\eta)}(D)$ 相互独立, 那么有,

$$\hat{p}_{(1,\eta)}(D) - \hat{p}_{(0,\eta)}(D) \sim N(0, 2\sigma_1^2).$$

可见, $\hat{p}_{(1,\eta)}(D) - \hat{p}_{(0,\eta)}(D)$ 的分布是均值为 0, 方差为 $2\sigma_1^2$ 的对称分布, 应该考虑舍弃落在均值 0 附近的项。我们给出如下舍弃标准。

舍弃标准: 舍弃 $\hat{p}_{(1,\eta)}(D) - \hat{p}_{(0,\eta)}(D)$ 落在一个

标准差内的项, 即舍弃 $|\hat{p}_{(0,\eta)}(D) - \hat{p}_{(1,\eta)}(D)| \leq \sqrt{2}\sigma_1$ 的项。

根据正态分布可以计算概率:

$$\Pr[|\hat{p}_{(0,\eta)}(D) - \hat{p}_{(1,\eta)}(D)| \leq \sqrt{2}\sigma_1] = 0.6826$$

那么, 剩余和的自由度大约为 $2^s \cdot (1 - 0.6826) = 2^{s-1.66}$ 。

下面考虑统计量 $T_k(D)$ 的非中心参数。由假设可知, $p_{(1,\eta)} - p_{(0,\eta)} \sim N(0, 2\sigma^2)$, 仿射统计量 $T_{\text{aff}}(D)$ 分布的非中心参数为

$$N\delta = N2^s \sum_{\eta \in F_2^s} (p_{(1,\eta)} - p_{(0,\eta)})^2,$$

舍弃该和式中满足 $|p_{(0,\eta)} - p_{(1,\eta)}| \leq \sqrt{2}\sigma$ 的项。

根据 $p_{(1,\eta)} - p_{(0,\eta)}$ 的概率分布可知, 仿射统计量 $T_{\text{aff}}(D)$ 的和式中平方项的均值为

$$\text{Exp}_K[(p_{(0,\eta)} - p_{(1,\eta)})^2] = 2\sigma^2$$

满足 $|p_{(0,\eta)} - p_{(1,\eta)}| \leq \sqrt{2}\sigma$ 条件的平方项的均值为

$$\begin{aligned} & \text{Exp}_K[(p_{(0,\eta)} - p_{(1,\eta)})^2 \cdot I(|p_{(0,\eta)} - p_{(1,\eta)}| \leq \sqrt{2}\sigma)] \\ &= \sigma^2 \int_{-1}^1 x^2 \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \approx 0.4\sigma^2. \end{aligned}$$

故按照舍弃标准, 在平均意义下非中心参数 $N\delta$ 舍弃的比例为 $0.4\sigma^2/2\sigma^2 = 0.2$ 。从而, 仿射统计量剩余部分的非中心参数为: $(1 - 0.2)\delta = 2^{-0.32} \cdot \delta$ 。

由以上分析, 按照我们的舍弃标准, 在正确密钥猜测下, 统计量 $T_k(D)$ 服从自由度为 $2^{s-1.66}$, 非中心参数为 $2^{-0.32} \cdot \delta$ 的卡方分布, 即

$$T_k(D) \sim \chi_{2^{s-1.66}}^2(2^{-0.32} N\delta)$$

根据多维线性密码分析的 χ^2 方法(见文献[7])的数据复杂度表达式(1), 知在 Nyberg 猜想下舍弃一个标准差内的数据时, 仿射线性攻击的 Nyberg 猜想下数据复杂度为

$$N_k = \frac{2\sqrt{2^{s-1.66}}a + 4\Phi^{-2}(2P_S - 1)}{2^{-0.32}\delta} \quad (3)$$

4 PRESENT 和 Serpent 的仿射线性攻击

因为 PRESENT 算法的 S 盒对线性分析很敏感, 我们为了验证仿射线性密码分析的有效性, 根据 PRESENT 算法 S 盒的 1-比特线性逼近的性质选择线性逼近, 利用仿射线性分析进行了 26、27 轮密钥恢复攻击。为了进一步验证 Nyberg 猜想的检验统计量

有效性, Serpent 算法比较了 4 轮 7 维线性分析与仿射线性密码分析的数据复杂度。

4.1 PRESENT 仿射线性攻击

我们由文献[19]可以知道, PRESENT 算法的 S 盒对线性分析很敏感, 它的某些单比特线性逼近是显著的。这些单比特线性逼近起始和结束的 S 盒 S_i 满足 $i \in \{5, 6, 7, 9, 10, 11, 13, 14, 15\}$ 。我们选择具有较大平均线性概率(ELP)的线性逼近(见图 2)进行仿射线性攻击。线性逼近的输入掩码为

$$\alpha = 2^{4i+3}, i \in \{5, 6, 7, 9, 10, 11, 13, 14\},$$

输出掩码为

$$\beta_1 = 2^{4i+2}, \beta_2 = 2^{4i+3}, i \in \{5, 6, 7, 9, 11, 13, 14, 15\}$$

我们的目标是利用上述 128 个 23 轮线性逼近构造仿射线性检验统计量, 对 PRESENT-80 进行密钥恢复攻击。

4.1.1 26 轮仿射线性密钥恢复攻击

本节应用上述的线性逼近进行 26 轮仿射线性密钥恢复攻击。首先构造仿射检验统计量, 计算数据复杂度; 接着给出攻击的步骤, 并分析攻击的时间复杂度; 最后考虑 Nyberg 猜想的应用。对给定的输入掩码 α , 输出掩码 $\beta_1 = 2^{4i+2}$ 和 $\beta_2 = 2^{4i+3}$, 假设线性逼近 $(\alpha, 2^{4i+2}), (\alpha, 2^{4i+3})$ 的联合概率分布为

$$p_{ij}, i, j = 0, 1,$$

$$p_{ij} = \Pr(\alpha \cdot x \oplus \beta_1 \cdot F(x) = i, \alpha \cdot x \oplus \beta_2 \cdot F(x) = j)$$

$\hat{p}_{ij}$ 表示由样本数据确定的概率 p_{ij} 的估计。

这两个线性逼近 $(\alpha, 2^{4i+2}), (\alpha, 2^{4i+3})$ 可以构成一个二维线性子空间, 利用仿射线性分析方法舍弃容量为零的形如 $(0, b)$ 的线性逼近。利用仿射线性分析

构造自由度为 2 的 χ^2 检验统计量:

$$T_{\text{aff}}^{(\alpha, \beta)} = 2N[(\hat{p}_{00} - \hat{p}_{11})^2 + (\hat{p}_{01} - \hat{p}_{10})^2],$$

也可表示为: $T_{\text{aff}}^{(\alpha, \beta)} = N(\hat{c}_{(\alpha, \beta_1)}^2(D) + \hat{c}_{(\alpha, \beta_2)}^2(D))$ 。

形如这样的仿射统计量一共可以构造 64 个, 假设这些统计量相互独立, 把这些统计量加在一块得到检验统计量:

$$T = \sum_{(\alpha, \beta)} T_{\text{aff}}^{(\alpha, \beta)}$$

在正确密钥猜测下 T 服从自由度为 128, 非中心参数为 $N\delta$, $\delta = \sum_{j=1}^{128} \text{ELP}_j$ 为 64 个二维线性子空间容量的期望和, $\text{ELP}_j, j = 1, \dots, 128$, 是第 j 个线性逼近的平均线性概率。我们利用信噪分解方法^[17]估计 128

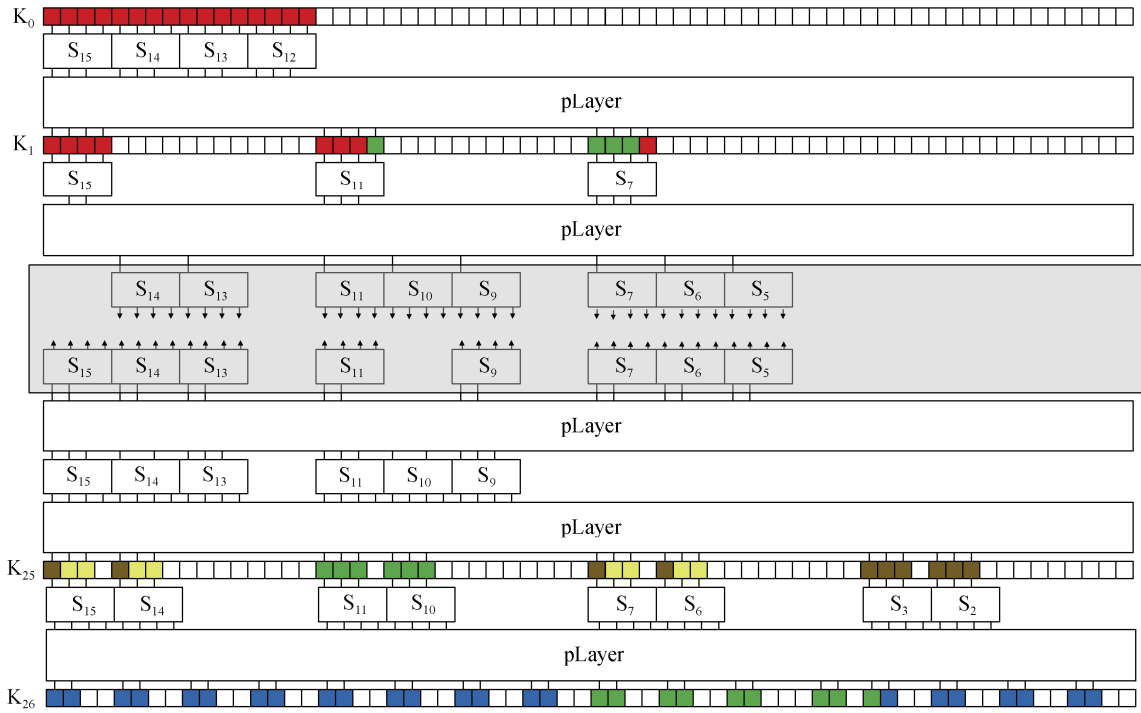


图 2 PRESENT 的 26 轮线性攻击

Figure 2 The 26-round affine linear attack on PRESENT

(注: 红色、蓝色和灰色密钥比特是需要猜测的, 共 57 比特, 绿色轮密钥比特是由红色轮密钥比特确定的, 黄色密钥比特是可以由蓝色猜测密钥确定的轮密钥比特。)

条线性逼近的 ELP, 得到统计量 T 在正确密钥猜测下的非中心参数为 $N\delta = N2^{-55.3}$, 故 $T \sim \chi^2_{128}(N\delta)$ 。

取 $a=16, P_S=0.95$ 可得数据复杂度为

$$N_{\text{aff}} = \frac{2\sqrt{16 \cdot 128} + 4 \cdot 1.3^2}{2^{-55.3}} \approx 2^{61.9}$$

我们在图 2 标记出了输入输出掩码的比特位置, 也标记出了计算线性逼近相关系数需要加密/解密的 S 盒位置和需要猜测的轮密钥。通过分析密钥编排算法, 根据轮密钥之间的关系减小 4 轮部分加密/解密需要猜测密钥比特。我们先猜测以下 24 比特的主密钥(图 2 中用红色标记):

$$k_i, i \in [0, 2] \cup [15, 18] \cup [63, 79]$$

由这 24 比特主密钥可以确定 K_1 所需要猜测的 4 比特和 K_{25}, K_{26} 所需要猜测的 15 比特(图 2 中用绿色标记)。然后猜测攻击涉及的 K_{26} 的其他 23 比特(图 2 中用蓝色标记), 利用这些密钥比特可以确定 K_{25} 所需要猜测的 8 比特(图 2 中用黄色标记)。最后猜测所涉及的 K_{25} 的 10 比特(图 2 中用灰色标记)。综上, 一共需要猜测 57 比特的密钥。

攻击步骤:

1、过滤阶段

获取加解密所需的 N 对部分明密文 (p_i, c_i) , 这

里 p_i 是 16 比特的明文, c_i 是 32 比特的密文, 生成一个大小为 2^{48} 的向量 t , 这里 $t[s \parallel t] = \{i \mid p_i = s, c_i = t\}$ 。

2、分析阶段

对需要猜测的 57 比特密钥的每个可能猜测, 计算攻击涉及的 128 个线性逼近的相关系数, 步骤如下。

(1) 给定 24 比特的猜测主密钥 K_M ,

(a) 对每个输入掩码 α 计算两个大小为 2^{16} 向量 $t_{\alpha_1}^{K_M}$ 和 $t_{\alpha_2}^{K_M}$ 。这里,

$$t_{\alpha_x}^{K_M}[j] = \#\{(p_i, c_i) \mid G_x(c_i) = j, \alpha \cdot E_{K_M}(p_i) = 0\},$$

$E_{K_M}(p)$ 是在密钥 K_M 下两轮加密 p , $G_x(c_i)$ 表示为了计算 $(\alpha, \beta_x), x \in [1, 2]$ 的相关系数而选择的 16 比特密文。

(b) 利用 FFT 技术(见文献[20])计算相关系数。对每个输出掩码 β , 给定所需要猜测的 K_{25} 的 4 比特密钥, 记为 K_I 。计算 $A_{\beta}^{K_I}[i, j] = \beta \cdot D_{K_I}(i \oplus j)$, 这里 $D_{K_I}(c)$ 是用密钥 K_I 两轮部分解密 16 比特密文 c , 不包含第一个密钥异或, 得到 $2^{16} \times 2^{16}$ 矩阵 $A_{\beta}^{K_I} = (A_{\beta}^{K_I}[i, j])$ 。计算所有 128 条线性逼近的相关系数。对每条线性逼近猜测所需的 K_{26} 的 16 比特密钥,

计算矩阵向量乘积:

$$C_{(\alpha, \beta)} = \frac{2}{N} A_{\beta}^{K_I} \cdot t_{\alpha_x}^{K_M} - 1$$

(c) 对所有可能的 K_I 取值, 重复步骤(b), 得到所有线性逼近的相关系数。

(d) 计算 χ^2 检验统计量 $T = \sum_{(\alpha, \beta)} T_{\text{aff}}^{(\alpha, \beta)}$, 保留

$T > \tau$ 的所有密钥。

(2) 对所有可能的 K_M , 重复以上步骤。

3、穷搜阶段

对每个候选密钥, 实施加密找到主密钥剩余的 80-57=23 比特。

攻击的计算复杂度

用单轮加密数来度量计算复杂度, 依次计算攻击各个阶段的计算复杂度。

1、过滤阶段需要 $O(N)$ 操作。

2、分析阶段

—step(1)(a)需要 $O(2 \cdot 2^{48})$ 操作。

—step(1)(b)利用 FFT 技术计算相关系数, 需要计算每个 $A_{\beta}^{K_I}$ 的第一列, 花费 $2 \cdot 2^{16}$ 操作。接着计算 $C_{(\alpha, \beta)}$ (给定 β , 对所有的 α 计算), 时间复杂度为 $O((2 \cdot 8 + 1) \cdot 16 \cdot 2^{16})$ 。

—step(1)(c) K_I 有 2^4 可能取值, 有 16 个输出掩码, 需要时间为

$$O(16 \cdot 2^4 \cdot (2 \cdot 2^{16} + (2 \cdot 8 + 1) \cdot 16 \cdot 2^{16})) \approx O(2^{32.1})$$

综上, 分析阶段的时间复杂度为

$$O(2^{24} (2^{49} + 2^{32.1} + 2^{33} + 2^{34.58})) \approx O(2^{73})$$

3、穷搜阶段

共有 2^{57-a} 候选密钥, 需要对每一个候选密钥进行 2^{k-57} 次全轮加密, 共需要 $O(2^{k-a})$ 次全轮加密。

给定 $a=16, P_S=0.95$, PRESENT 的 26 轮密钥恢复攻击的数据复杂度为 $N_{\text{aff}} = 2^{61.9}$, 时间复杂度为 $O(2^{80-16} + 2^{73}/26) = O(2^{68.37})$ 。Bogdanov 等人^[21]在 $a=4.7, P_S=0.95$ 时, 得到 PRESENT 的 26 轮密钥恢复攻击数据复杂度为 $N = 2^{61.9}$, 时间复杂度为 $O(2^{75.3})$ 次加密。比较可见, 我们的攻击更优。

Nyberg 猜想的应用

我们舍弃 $\hat{p}_{(1, \eta)}(D) - \hat{p}_{(0, \eta)}(D)$ 落在一个标准差内的项, 仿射统计量的自由度减小为 $2^{s-1.66}$ 。非中心参数舍弃的比例为 0.2, 剩余的非中心参数为 $(1-0.2)$

$N\delta = N2^{-0.32-55.3}$ 。取 $a=16, P_S=0.95$, 数据复杂度为

$$N_k = \frac{2\sqrt{16 \cdot 128 \cdot 2^{-1.66}} + 4 \cdot 1.3^2}{2^{-0.32-55.3}} \approx 2^{61.5}。$$

4.1.2 27 轮仿射线性密钥恢复攻击

利用相同的 23 轮线性逼近, 可以进行 27 轮的密钥恢复攻击, 此时非中心参数为 $N\delta = N2^{-57.08}$ 。取 $a=16, P_S=0.95$ 时, 可得数据复杂度为 $N_{\text{aff}} \approx 2^{63.68}$ 。由攻击实施的方法可知, 分析阶段的时间复杂度基本不受 K_{27} 密钥猜测的影响, 时间复杂度基本与 26 轮攻击相同, 即时间复杂度约为 $O(2^{68.37})$ 次加密。舍弃仿射统计量中落在一个标准差内的项时, 数据复杂度为 $N_k = 2^{63.25}$ 。

取 $a=2.73, P_S=0.95$ 时, 可得数据复杂度为 $N_{\text{aff}} \approx 2^{62.54}$, 时间复杂度约为 $O(2^{80-2.73}) = O(2^{77.27})$ 次加密。舍弃仿射统计量中一个标准差内的项时, 由式(3)得数据复杂度为 $N_k = 2^{62.2}$ 。Bogdanov 等人^[21]给定 $a=2.73, P_S=0.95$ 时, 得到 PRESENT 的 27 轮密钥恢复攻击数据复杂度为 $N = 2^{63.83}$, 时间复杂度为 $O(2^{77.27})$ 次加密。比较可见, 在相同的优势和成功率情况下, 我们的攻击数据复杂度更小。

4.2 Serpent 仿射线性攻击

Hermelin 等人^[7]分别用 χ^2 方法和 LLR 方法讨论了多维线性分析, 用 Serpent 算法验证它们的有效性。

考虑文献[6]中 Serpent 的 7 维线性逼近, 7 个不同的输入掩码, 一个相同的输出掩码。4 轮 Serpent 的 7 维线性逼近的容量为 2^{-22} , 优势 $a=12$, 成功概率为 $P_S=0.95$ 时, 利用文献[7] χ^2 方法计算的数据复杂度为

$$N_{\chi^2} = \frac{2\sqrt{12 \cdot (2^7 - 1)} + 4 \cdot 1.3^2}{2^{-22}} \approx 2^{28.4}$$

舍弃 7 维线性子空间中容量为零的一半的线性逼近, 仿射 χ^2 统计量的自由度减小为 2^6 , 非中心参数不变。在优势 $a=12$, 成功概率为 $P_S=0.95$ 下, 此时数据复杂度为

$$N_{\text{aff}} = \frac{2\sqrt{12 \cdot 2^6} + 4 \cdot 1.3^2}{2^{-22}} \approx 2^{27.96}$$

利用 Nyberg 猜想, 舍弃一个标准差的项, 此时在相同的条件下的数据复杂度为

$$N_k = \frac{2\sqrt{12 \cdot 2^{6-1.66}} + 4 \cdot 1.3^2}{2^{-0.32} 2^{-22}} \approx 2^{27.57}$$

文献[5]中给出了相同条件下 LLR 方法的数据复杂度 $N_{\text{LLR}} = 2^{27.5}$ 。可见此种方法和 LLR 方法效率相当。

5 结论

多维线性密码分析利用某个线性子空间上除了零之外的所有线性逼近进行攻击, 仿射线性密码分析舍弃线性子空间中对容量贡献较小或没有贡献的一半线性逼近构造了仿射 χ^2 统计量, 从而达到减小自由度的目的, 进而减小 χ^2 统计量的方差、提高攻击效率的目的。为了进一步减小自由度, Nyberg 猜想舍弃仿射统计量中贡献较小的项, 利用剩余项构造的统计量仍服从 χ^2 分布。本文证明了该猜想是正确的, 并用 PRESENT 和 Serpent 验证了仿射线性分析的有效性。我们利用仿射线性分析和 Nyberg 猜想对 PRESENT 进行了 26、27 轮密钥恢复攻击, 得到的攻击结果是当前最好的。对于 4 轮 Serpent 算法 7 维线性分析, 我们在相同的优势和成功概率下, 比较了各种分析方法的数据复杂度。结果发现 χ^2 方法数据复杂度最大, 仿射方法次之, Nyberg 猜测方法再次之, 它与 LLR 方法的效率相当。

参考文献

- [1] Matsui M. Linear Cryptanalysis Method for DES Cipher[M]. Advances in Cryptology — EUROCRYPT '93. Berlin, Heidelberg: Springer Berlin Heidelberg, 1994: 386-397.
- [2] Matsui M. The First Experimental Cryptanalysis of the Data Encryption Standard[M]. Advances in Cryptology — CRYPTO '94. Berlin, Heidelberg: Springer Berlin Heidelberg, 1994: 1-11.
- [3] Kaliski B S Jr, Robshaw M J B. Linear Cryptanalysis Using Multiple Approximations[M]. Advances in Cryptology — CRYPTO '94. Berlin, Heidelberg: Springer Berlin Heidelberg, 1994: 26-39.
- [4] Biryukov A, De Cannière C, Quisquater M. On Multiple Linear Approximations[M]. Advances in Cryptology — CRYPTO 2004. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004: 1-22.
- [5] Baignères T, Junod P, Vaudenay S. How far can we Go beyond Linear Cryptanalysis[M]. Advances in Cryptology - ASIACRYPT 2004. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004: 432-450.
- [6] Hermelin M, Cho J Y, Nyberg K. Multidimensional Linear Cryptanalysis of Reduced Round Serpent[C]. Information Security and Privacy, 13th Australasian Conference, ACISP, 2008, 5107: 203-215.
- [7] Hermelin M, Cho J Y, Nyberg K. Multidimensional Extension of Matsui's Algorithm 2[C]. Fast Software Encryption, 16th International Workshop, 2009: 209-227.
- [8] Blondeau C, Nyberg K. Joint Data and Key Distribution of Simple, Multiple, and Multidimensional Linear Cryptanalysis Test Statistic and Its Impact to Data Complexity[J]. Designs, Codes and Cryptography, 2017, 82(1): 319-349.
- [9] Bogdanov A., Wang M. Zero correlation linear cryptanalysis with reduced data complexity[C]. FSE, 2012, 7549: 29-48.
- [10] Bogdanov A, Leander G, Nyberg K, et al. Integral and Multidimensional Linear Distinguishers with Correlation Zero[M]. Advances in Cryptology — ASIACRYPT 2012. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012: 244-261.
- [11] Selçuk A A. On Probability of Success in Linear and Differential Cryptanalysis[J]. Journal of Cryptology, 2008, 21(1): 131-147.
- [12] Nyberg K. Affine Linear Cryptanalysis[J]. Cryptography and Communications, 2019, 11(3): 367-377.
- [13] Bogdanov A, Knudsen L R, Leander G, et al. PRESENT: An Ultra-Lightweight Block Cipher[M]. Cryptographic Hardware and Embedded Systems - CHES 2007. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007: 450-466.
- [14] Anderson R, Biham E, Knudsen L, Serpent: A Proposal for the Advanced Encryption Standard[N]. NIST AES Proposal, 1998.
- [15] Daemen J, Rijmen V. Probability Distributions of Correlation and Differentials in Block Ciphers[J]. Journal of Mathematical Cryptology, 2007, 1(3): 221-242.
- [16] Blondeau C, Nyberg K. Improved Parameter Estimates for Correlation and Capacity Deviates in Linear Cryptanalysis[J]. IACR Transactions on Symmetric Cryptology, 2017: 162-191.
- [17] Cho J Y. Linear cryptanalysis of reduced-round PRESENT[C]. The Cryptographers' Track at the RSA Conference, 2010: 302-317.
- [18] The Cryptographers' Track at the RSA Conference, 2010: 302-317. Bogdanov A., Tischhauser E. On the wrong key randomisation and key equivalence hypotheses in Matsui's Algorithm 2[C]. FSE, 2013, 8424: 19-38.
- [19] Ohkuma K. Weak keys of reduced-round PRESENT for linear cryptanalysis[C]. Selected Areas in Cryptography, 16th Annual International Workshop, 2009: 249-265.
- [20] Collard B, Standaert F, Quisquater J. Improving the time complexity of Matsui's linear cryptanalysis[C]. Information Security and Cryptology, 2007: 77-88.
- [21] Bogdanov A, Tischhauser E, Vejre P S. Multivariate Profiling of Hulls for Linear Cryptanalysis[J]. IACR Transactions on Symmetric Cryptology, 2018: 101-125.



曹文芹 于 2005 年在兰州大学概率论与数理统计专业获得硕士学位。现在中国科学院信息工程研究所网络空间安全专业攻读博士学位。研究领域为对称密码算法安全性分析。Email: caowenqin@sdut.edu.cn



张文涛 于 2004 年在中国科学院软件研究所获得博士学位。现为中国科学院信息工程研究所国家重点实验室研究员。研究领域为对称密码算法的设计与分析。Email: zhangwentao@iie.ac

附录 A 验证: $(C\Sigma_0)^2 = C\Sigma_0$ 。

证明: $A^T A$ 可以表示为

$$A^T A = I_{M-1} + 1_{M-1} 1_{M-1}^T = \begin{pmatrix} I_{M/2} + 1_{M/2} 1_{M/2}^T & 1_{(M/2)} 1_{(M/2)-1}^T \\ 1_{(M/2)-1} 1_{(M/2)}^T & I_{(M/2)-1} + 1_{(M/2)-1} 1_{(M/2)-1}^T \end{pmatrix},$$

$A^T(2I_M - B)A$ 可以表示为

$$\begin{aligned} & A^T(2I_M - B)A \\ &= 2A^T A - \begin{pmatrix} I_{M/2} & (0: -1_{M/2}) \\ 0 & (I_{(M/2)-1} : -1_{M/2}) \end{pmatrix} \begin{pmatrix} I_{M/2} & I_{M/2} \\ I_{M/2} & I_{M/2} \end{pmatrix} \begin{pmatrix} I_{M/2} & (0: -1_{M/2}) \\ 0 & (I_{(M/2)-1} : -1_{M/2}) \end{pmatrix}^T \\ &= \begin{pmatrix} I_{(M/2)-1} + 1_{(M/2)-1} 1_{(M/2)-1}^T & 2 \cdot 1_{(M/2)-1} & 1_{(M/2)-1} 1_{(M/2)-1}^T - I_{(M/2)-1} \\ 2 \cdot 1_{(M/2)-1}^T & 4 & 2 \cdot 1_{(M/2)-1}^T \\ 1_{(M/2)-1} 1_{(M/2)-1}^T - I_{(M/2)-1} & 2 \cdot 1_{(M/2)-1} & I_{(M/2)-1} + 1_{(M/2)-1} 1_{(M/2)-1}^T \end{pmatrix}, \end{aligned}$$

Σ_0 表示为

$$\Sigma_0 = \begin{pmatrix} I_{(M/2)-1} - 2^{-s-1} \cdot 1_{(M/2)-1} 1_{(M/2)-1}^T & -2^{-s-1} \cdot 1_{(M/2)-1} & -2^{-s-1} \cdot 1_{(M/2)-1} 1_{(M/2)-1}^T \\ -2^{-s-1} \cdot 1_{(M/2)-1}^T & 1 - 2^{-s-1} & -2^{-s-1} \cdot 1_{(M/2)-1}^T \\ -2^{-s-1} \cdot 1_{(M/2)-1} 1_{(M/2)-1}^T & -2^{-s-1} \cdot 1_{(M/2)-1} & I_{(M/2)-1} - 2^{-s-1} \cdot 1_{(M/2)-1} 1_{(M/2)-1}^T \end{pmatrix}$$

计算得

$$A^T(2I_M - B)A \cdot \Sigma_0 = \begin{pmatrix} \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 1 \\ 0 & 1 & 0 & \dots & 0 & 1 \\ 0 & 0 & 1 & \dots & 0 & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & 1 \\ 0 & 0 & 0 & \dots & 0 & 2 \end{pmatrix}_{\frac{M}{2} \times \frac{M}{2}} & \begin{pmatrix} -1 & 0 & 0 & \dots & 0 & 0 \\ 0 & -1 & 0 & \dots & 0 & 0 \\ 0 & 0 & -1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & -1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}_{\frac{M}{2} \times (\frac{M}{2}-1)} \\ \begin{pmatrix} -1 & 0 & 0 & \dots & 0 & 1 \\ 0 & -1 & 0 & \dots & 0 & 1 \\ 0 & 0 & -1 & \dots & 0 & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & -1 & 1 \end{pmatrix}_{(\frac{M}{2}-1) \times \frac{M}{2}} & \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}_{(\frac{M}{2}-1) \times (\frac{M}{2}-1)} \end{pmatrix}$$

为简便, $A^T(2I_M - B)A\Sigma_0$ 表示为

$$A^T(2I_M - B)A\Sigma_0 = \begin{pmatrix} I_{(M/2)-1} & 1_{(M/2)-1} & -I_{(M/2)-1} \\ 0 & 2 & 0 \\ -I_{(M/2)-1} & 1_{(M/2)-1} & I_{(M/2)-1} \end{pmatrix},$$

因为 $C = 2^{-1} \cdot [A^T(2I_M - B)A]$, 所以 $C\Sigma_0 = 2^{-1} \cdot \begin{pmatrix} I_{(M/2)-1} & 1_{(M/2)-1} & -I_{(M/2)-1} \\ 0 & 2 & 0 \\ -I_{(M/2)-1} & 1_{(M/2)-1} & I_{(M/2)-1} \end{pmatrix}$, 计算得,

$$(C\Sigma_0)^2 = 2^{-2} \cdot \begin{pmatrix} 2I_{(M/2)-1} & 21_{(M/2)-1} & -2I_{(M/2)-1} \\ 0 & 4 & 0 \\ -2I_{(M/2)-1} & 21_{(M/2)-1} & 2I_{(M/2)-1} \end{pmatrix},$$

即 $(C\Sigma_0)^2 = C\Sigma_0$, 证毕。

附录 B 沿用文中记号, 我们利用向量 X 把二次型 $(X_l - X_{2^s+l})^2$ 表示成二次型 $X^T A_l X$, A_l 是 $M-1$ 阶对称阵, $l=1, \dots, M/2$ 。当 $l=1, 2, \dots, M/2-1$ 时, $A_l = (a_{i,j}^{(l)})_{(M-1) \times (M-1)}$, 对称阵 A_l 的元素满足:

$a_{l,l}^{(l)} = 1, a_{l,M/2+l}^{(l)} = -1, a_{M/2+l,l}^{(l)} = -1, a_{M/2+l,M/2+l}^{(l)} = 1$, 其余元素均为零。当 $l=M/2$ 时,

$$A_{M/2} = \begin{pmatrix} 1_{(M/2)-1} 1_{(M/2)-1}^T & 2 \cdot 1_{(M/2)-1} & 1_{(M/2)-1} 1_{(M/2)-1}^T \\ 2 \cdot 1_{(M/2)-1}^T & 4 & 2 \cdot 1_{(M/2)-1}^T \\ 1_{(M/2)-1} 1_{(M/2)-1}^T & 2 \cdot 1_{(M/2)-1} & 1_{(M/2)-1} 1_{(M/2)-1}^T \end{pmatrix}$$

附录 C 定理: [科克伦定理] 假设 M 维随机向量 $X \sim N_M(\mu, \Sigma)$, $\Sigma > 0$, A_i 为对称矩阵, $rk(A_i) = r_i, i=1, \dots, k$, 又

$A = \sum_{i=1}^k A_i$, $rk(A) = r$, 则 $X^T A_i X \sim \chi_{r_i}^2(\lambda_i), i=1, \dots, k$, 其中 $\lambda_i = \mu^T A_i \mu$, 而且 $X^T A X \sim \chi_r^2(\lambda), \lambda = \sum_{i=1}^k \lambda_i$ 的充要条件是下列 (a),(b),(c) 中任何两者, 或者 (c) 和 (d):

$$(a) (A_i \Sigma)^2 = A_i \Sigma, i=1, \dots, k;$$

$$(b) A_i \Sigma A_j \Sigma = 0, i \neq j, irk((A_i \Sigma)^2) = rk(A_i \Sigma);$$

$$(c) (A \Sigma)^2 = A \Sigma;$$

$$(d) r = \sum_{i=1}^k r_i.$$