论 文

演化密码对抗差分密码分析能力的研究

张焕国^{①②*}, 李春雷^{①②}, 唐明^{①②}^① 武汉大学计算机学院, 武汉 430072^② 空天信息安全与可信计算教育部重点实验室, 武汉 430072

* 通信作者. E-mail: liss@whu.edu.cn

收稿日期: 2011-01-02; 接受日期: 2011-06-28

国家自然科学基金 (批准号: 69973034, 90104005, 60373087, 60673071, 60970115, 60970116, 91018008, 60970006, 61003267)
资助项目

摘要 演化密码是我国学者提出的一种新型密码体制. 本文对演化密码对抗差分密码分析的能力进行了研究, 研究表明演化密码对抗传统差分攻击的能力高于普通固定算法密码. 本文在分析差分攻击的数据复杂度, 最佳差分特征概率, 比特优势以及预期成功率关系的基础上, 证明了比特优势和预期成功率相同时, 攻击演化密码的数据复杂度大于攻击固定算法密码的数据复杂度; 并证明了在数据复杂度和预期成功率相同的情况下, 攻击演化密码的时间复杂度明显高于攻击固定算法密码的时间复杂度. 这表明演化密码对抗传统差分攻击的能力高于固定算法密码.

关键词 信息安全 密码学 演化密码 差分密码分析 比特优势

1 引言

密码技术是保障信息安全的关键技术, 如何设计出高安全强度的密码是人们长期追求的目标. 对于一个密码系统 $\langle \mathcal{P}, \mathcal{C}, \mathcal{K}, E, D \rangle$, 频繁地更换密钥带来的不确定性能够提高密码系统的安全强度. 更一般地, 若密码系统中密钥是随机的, 密钥空间不小于明文空间, 并且每个密钥只使用一次, 即每次加密时都采用与已使用密钥相互独立的新密钥, 此系统为众所周知的“一次一密”密码系统, 可提供完善保密性. 由于“一次一密”密码系统给密钥的产生和管理都带来严重的问题, 故这种密码系统并不实用. 受“一次一密”更换密钥能够增强密码安全性的启发, 文献 [1,2] 提出了演化密码的概念.

迄今为止的密码体制都是加解密算法固定而密钥随机可变的密码体制, 设 E 为加密算法, $K = K_0 K_1 \cdots K_{t-1}$ 为密钥, 则对明文 $P = P_0 P_1 \cdots P_{t-1}$ 的加密过程可表示成

$$C_0 = E(P_0, K_0), C_1 = E(P_1, K_1), \dots, C_{t-1} = E(P_{t-1}, K_{t-1}),$$

其中对于 $i = 0, 1, \dots, t-1$, $P_i \in \mathcal{P}$, $K_i \in \mathcal{K}$. 如果在密码体制中, 不仅密钥随机可变, 而且加密算法也是动态可变的, 则称此密码体制为动态可变密码, 其加密过程可表示为

$$C_0 = E_0(P_0, K_0), C_1 = E_1(P_1, K_1), \dots, C_{t-1} = E_{t-1}(P_{t-1}, K_{t-1}),$$

引用格式: 张焕国, 李春雷, 唐明. 演化密码对抗差分密码分析能力的研究. 中国科学: 信息科学, 2013, 43: 545-554

进一步地, 记 $S(E)$ 表示加密算法 E 对抗已有某些攻击的安全强度的指标, 若动态可变密码中加密算法 $E_i, i = 0, 1, \dots, t-1$ 的安全强度逐渐增加, 即

$$S(E_0) \leq S(E_1) \leq \dots \leq S(E_{t-1}),$$

则称其为演化密码 [3].

一般而言, 加密算法 E 的一些部件在其对抗某种已知攻击时发挥核心作用, 如 AES 中非线性部件 S 盒和扩散层的 P 置换部件是 AES 对抗线性攻击和差分攻击的核心. 为了有效对抗已有的各种攻击, 密码算法中一些部件通常需要满足一系列密码准则. 这些密码准则被看成一种量化指标, 用来度量密码部件对抗已有攻击的安全强度. 以加密算法 AES 为例, 算法中 S 盒的非线性度被用来度量其抵抗线性攻击的能力, 而差分均匀性被用来度量其抵抗差分攻击的能力. 一般情况下, 演化密码可表示成一个六元组 $\langle E, \omega, \Omega, \Gamma, \Psi, f \rangle$, 其中:

1. E 表示一个加密算法;
2. ω 表示加密算法 E 中的核心密码部件, 如分组密码算法中的非线性部件 S 盒;
3. Ω 表示所有可能的核心部件 ω 组成的集合;
4. Γ 表示密码部件 ω 已有的一些密码准则, 简称规则集. 如密码部件为 S 盒时, 规则集中包含差分均匀性, 非线性度, 代数次数等;
5. Ψ 表示所有 Ω 到 Ω 的映射组成的集合;
6. f 表示 Ω 到正实数集 $\mathbb{R}^+$ 的映射, 被称为评估函数, 其函数值被称为适应值.

通过比较密码部件 ω 对应的适应值, 设计者可以得出一组密码部件 $\omega_0, \omega_1, \dots, \omega_{t-1}$, 使算法 $E_i, i = 0, 1, \dots, t-1$ 对抗已知某种攻击安全强度满足 $S(E_0) \leq S(E_1) \leq \dots \leq S(E_{t-1})$, 其中算法 E_i 表示密码部件被更换为 ω_i 时的算法 E . 为下面讨论的简便, 本文中我们将固定算法密码记为 $\langle E, \omega_0 \rangle$, 将演化密码记为 $\langle E, \omega_0, \omega_1, \dots, \omega_{t-1} \rangle$.

基于演化密码的思想, 文献 [1,2] 从 DES 对抗差分密码分析和线性密码分析的安全强度出发, 利用演化计算的方法设计出性能优良的非线性部件 S 盒, 并以此为基础设计出了演化 DES 密码体制. 文献 [4~6] 利用演化算法在 Bent 函数的设计和分类方面也取得了一些成果. 文献 [7] 设计了一种新的 Hash 函数, 这种 Hash 函数的安全性基于多变量非线性方程组的难解性. 实验表明新 Hash 函数的各项性能指标与其他 Hash 函数具有可比性. 文献 [8] 提出了一种新的加噪扰动算法, 用于增强多元公钥密码体制 (MPKCs) 的安全性. 以 Matsumoto-Imai 体制为例, 这种新的加噪扰动算法能破坏其公钥函数潜在的数学特性, 在不降低算法效率的前提下, 可有效抵抗差分攻击, 增加其安全性. 文献 [9] 研究了演化密码体制对抗多重线性密码分析的能力, 研究表明演化密码对抗多重线性密码分析的能力高于固定算法密码.

本文余下部分组织如下: 第 2 节回顾了差分密码分析的基本方法, 并考察了差分密码分析的复杂度水平. 第 3 节阐述了对抗差分攻击的演化密码体制 $\langle E, \omega_0, \omega_1, \dots, \omega_{t-1} \rangle$ 的设计方法, 在考察预期成功率、数据复杂度、最佳差分特征概率以及比特优势关系的基础上, 比较了差分攻击固定算法密码 $\langle E, \omega_0 \rangle$ 和演化密码 $\langle E, \omega_0, \omega_1, \dots, \omega_{t-1} \rangle$ 的复杂度水平. 第 4 节总结了本文的工作.

2 差分攻击简介

1990 年, Biham 和 Shamir 提出了攻击低轮 DES 类密码的差分密码分析方法, 其基本思想是通过分析明文对的差值对密文对差值的影响来恢复某些密钥比特 [10,11]. 随后, 他们对原有方法加以改进,

首次给出了攻击 16-轮 DES 的密码分析方法, 改进后算法攻击 16-轮 DES 密码的数据复杂度减少为 2^{47} [12]. 由于对迭代型分组密码的安全性构成威胁, 差分密码分析提出后受到广泛的关注, 并被看作是评估分组密码安全性的一个重要手段. 差分密码分析的有效性依赖于分组密码的高概率差分特征, 然而对于有些密码体制, 寻找高概率差分特征非常困难. 基于差分攻击的思想, 密码学者们陆续提出了一些改进和扩展差分密码分析的方法, 如截断差分密码分析 [13], 不可能差分密码分析 [14], 差分线性攻击 [15], 高阶差分密码分析 [16] 以及飞去来器攻击 [17] 等.

差分密码分析的基本思想是通过分析密码算法输入差分对输出差分的影响来恢复某些密钥比特. 对于一个 r -轮迭代型分组密码 E , 设初始密钥长度为 n_k , 轮函数为 F , 每轮密钥长度为 m . 令 X_0, X_0^* 是 E 的一对明文, 其差分为 $\Delta X_0 = X_0 \oplus X_0^*$, 经过加密后可得到一条差分序列 $\Delta X_0, \Delta X_1, \dots, \Delta X_r$, 其中 ΔX_i 是第 i ($1 \leq i \leq r$) 轮输出 $X_i = F(X_{i-1}, K_i)$ 和 $X_i^* = F(X_{i-1}^*, K_i)$ 的差分, K_i 表示第 i ($1 \leq i \leq r$) 轮的密钥. 研究表明, 如果 $\Delta X_{i-1}, X_i$ 和 X_i^* 已知, 则根据简单轮函数 F 可确定包含密钥 K_i 的密钥集合. 该集合中的元素通常比较少, 被称为密钥测试集. 在差分密码分析过程中, 攻击者先选择符合特定差分值 ΔX_0 的明文对 X_0, X_0^* , 并使最后一轮的输入差分以高概率取特定值 ΔX_{r-1} , 进而由轮函数 F 确定最后一轮的子密钥或部分子密钥的密钥测试集.

分组密码算法的差分特征是一条差分序列 $\Delta X_0, \Delta X_1, \dots, \Delta X_r$, 它的概率是指在明文和各轮子密钥相互独立的条件下, 明文差分为 ΔX_0 时, 第 i ($1 \leq i \leq r$) 轮差分为 ΔX_i 的概率. 特别地, 当明文差分为 $\Delta X_0 = 0$ 时, 显然各轮的差分也等于零, 此时差分特征概率等于 1, 被称为平凡差分特征. 差分特征的概率影响差分攻击的有效性, 概率越大对攻击者越有利, 我们称概率取得最大时的非平凡差分特征为最佳差分特征. 对于差分特征 $\Delta X_0, \Delta X_1, \dots, \Delta X_r$, 若明文对 (X_0, X_0^*) 的差分为 ΔX_0 , 且第 i ($1 \leq i \leq r$) 轮输出差分 ΔX_i , 则称明文对 (X_0, X_0^*) 是该差分特征的正确对, 否则称为该差分特征的错误对. 当尝试的明文对是符合最佳差分特征的“正确对”时, 正确子密钥一定包含在相应的密钥测试集中; 当尝试的明文对是“错误对”时, 相应的密钥测试集中由一些随机密钥组成.

通过尝试大量的明文对, 攻击者对出现在密钥测试集中的候选密钥进行计数, 正确子密钥出现的次数会明显高于错误子密钥出现的次数, 通过对候选子密钥按累计次数排序, 攻击者可以从累计次数较高的候选密钥中恢复出正确子密钥.

对 r -轮分组密码的差分分析大致分为如下 3 个阶段:

1. 密钥提取阶段: 攻击者设法获得一个最佳差分特征 $\Delta X_0, \Delta X_1, \dots, \Delta X_{r-1}$, 均匀随机地选择差分为 ΔX_0 的 N 个明文对, 对每个明文对 (X_0, X_0^*) , 加密获得密文对 (X_r, X_r^*) , 根据三元组 $(\Delta X_{r-1}, X_r, X_r^*)$ 和轮函数 F 确定相应的密钥测试集;

2. 密钥分析阶段: 最后一轮子密钥 K_r 有 2^m 个可能值, 对每个可能的 K_r^i , 设置相应的计数器 T_i , ($0 \leq i \leq 2^m - 1$), T_i 表示 K_r^i 在密钥提取阶段的密钥测试集中出现的次数, 对计数器 T_i , ($0 \leq i \leq 2^m - 1$) 按递增的顺序排列;

3. 密钥搜索阶段: 按计数器 T_i ($0 \leq i \leq 2^m - 1$) 的取值从大到小尝试候选密钥, 直至恢复出正确子密钥.

在统计密码分析中, 恢复密钥的统计分析方法通常情况下基于如下的错误密钥假设:

错误密码假设 [18] 在密钥的统计分析过程中, 存在两个不同的概率分布 p 和 q 满足: 对于正确的密钥, 与明密文相关的数据服从概率分布 p ; 而对于其他错误密钥, 与明密文相关的实验数据都服从概率分布 q .

下面对差分攻击的复杂度进行分析. 差分密码分析的复杂度主要取决于数据复杂度和时间复杂

表 1 差分密码分析的复杂度

Table 1 The complexity of differential cryptanalysis on E

	Data	Time	Memory
Distillation phase	N	$\vartheta_1^{(E_0)} + N \cdot \vartheta_2$	$\mathcal{O}(N)$
Analysis phase	—	$2^m \vartheta_3 + \vartheta_4^{(m)}$	$\mathcal{O}(2^m)$
Search phase	—	$2^{n_k-a} \vartheta_5^{(E)}$	$\mathcal{O}(2^m)$

度^[11]. 在理想的统计分析模型下, 所有候选密钥对应的计数器排列中, 正确密钥对应的计数器应取值最大. 如果正确密钥的计数器取值不是最大, 则攻击者需要依次尝试其他候选密钥, 直至找出正确的密钥. 文献 [19] 提出密码分析的“比特优势 (advantage)”指标, 用来考察差分攻击密钥搜索阶段的复杂度水平.

定义 1 设一个密钥分析方法预期恢复的密钥长度为 m . 在所有 2^m 个候选密钥中, 若根据该方法确定的密钥顺序, 正确密钥排在前 $r = 2^{m-a}$ 个候选密钥中, 则称该密钥分析方法相对于穷举攻击的比特优势为 a .

本文的研究重点是比较差分攻击固定算法密码和攻击演化密码的复杂度, 由于差分攻击的复杂度与密码算法 E 紧密相关, 故下面的复杂度分析都是基于某些基本运算执行次数的讨论, 并非精确的复杂度分析.

在文献 [20] 中, Matsui 给出了搜索 DES 型分组密码的最佳差分特征的算法, 此算法的复杂度依赖于具体的算法 E . 在密钥提取阶段, 分别令 $\vartheta_1^{(E)}$ 表示搜索算法 E 的最佳差分特征的时间复杂度, ϑ_2 表示三元组 $(\Delta X_{r-1}, X_r, X_r^*)$ 确定密钥测试集的时间复杂度, 故密钥提取阶段的时间复杂度为 $\vartheta_1^{(E)} + N \cdot \vartheta_2$; 在密钥分析阶段, 令 ϑ_3 表示确定计数器 T_i 取值的时间复杂度, $\vartheta_4^{(m)}$ 表示利用快速排序算法对计数器 T_i ($0 \leq i \leq 2^m - 1$) 排序的时间复杂度, 从而此阶段的时间复杂度为 $2^m \vartheta_3 + \vartheta_4^{(m)}$, 在密钥分析阶段, 设差分攻击方法的比特优势为 a , 执行一次加密算法 E 的时间复杂度为 $\vartheta_5^{(E)}$, 则搜索阶段的时间复杂度为 $2^{n_k-m} \cdot 2^{m-a} \cdot \vartheta_5^{(E)} = 2^{n_k-a} \vartheta_5^{(E)}$. 根据上面的分析, 差分分析的复杂度如表 1 所示.

在差分攻击的密钥分析和搜索阶段, 文献 [19] 考察了密钥测试集中候选密钥对应的计数器的样本分布, 以及按递增顺序排列后的样本分布. 设 $\xi_1, \xi_2, \dots, \xi_n$ 是取自相同分布的 n 个独立的样本, 将这组样本按递增的顺序排列, 重排后的样本为 $\xi_{\tau(1)} \leq \xi_{\tau(2)} \leq \dots \leq \xi_{\tau(n)}$, 则样本 $\xi_{\tau(i)}$ 被称为是 $\xi_1, \xi_2, \dots, \xi_n$ 的第 i 个统计量. 重排后的样本 $\xi_{\tau(1)}, \xi_{\tau(2)}, \dots, \xi_{\tau(n)}$ 的概率分布可由下面的命题 1 确定.

命题 1^[21] 设 $\xi_1, \xi_2, \dots, \xi_n$ 是取自相同分布 $F(x)$ 的 n 个独立的变量, 设其密度函数 $f(x) = F'(x)$ 在区间 $[a, b]$ 上连续且取正值. 设 $0 < F(a) < q < F(b) < 1$, 序列 $i(n)$ 是一条整数序列, 满足 $\lim_{n \rightarrow \infty} \sqrt{n} \left| \frac{i(n)}{n} - q \right| = 0$, $\xi_{\tau(i)}$ 是 $\xi_1, \xi_2, \dots, \xi_n$ 的第 i 个统计量, 则

$$\lim_{n \rightarrow \infty} \Pr \left(\frac{\xi_{\tau(i(n))} - \mu_q}{\sigma_q} < x \right) = \Phi(x),$$

其中 $\mu_q = F^{-1}(q)$, $\sigma_q = \frac{1}{f(\mu_q)} \sqrt{\frac{q(1-q)}{n}}$, $\Phi(x)$ 表示标准正态分布的分布函数.

根据命题 1, 若取 $i(n) = \lfloor qn \rfloor + 1$, 当 n 取值较大时, 则样本 $\xi_{\tau(i(n))}$ 近似地服从期望 $\mu_q = F^{-1}(q)$, 标准差 $\sigma_q = \frac{1}{f(\mu_q)} \sqrt{\frac{q(1-q)}{n}}$ 的正态分布 $\mathcal{N}(\mu_q, \sigma_q^2)$. 在此基础上, 文献 [19] 研究了差分攻击成功率 P_S , 最佳差分特征概率 p , 明文对数量 N 以及比特优势 a 之间的关系.

命题 2 对算法 E 进行差分攻击时, 设攻击者利用 N 个选择明文分析 m 比特密钥, 最佳差分特征的概率为 p , 差分攻击的比特优势等于 a , 当 2^m 个候选密钥计数器统计独立且错误密钥的计数器服

从相同分布时, 差分攻击的成功率

$$P_S = \Phi \left(\frac{\sqrt{p^2 N} - \sqrt{p_r} \Phi(1 - 2^{-a})}{\sqrt{p + p_r}} \right), \quad (1)$$

其中 p_r 表示一个随机明文对确定的密钥测试集包含某特定错误密钥的平均概率.

下面讨论差分攻击演化密码时数据复杂度, 最佳差分概率和比特优势的关系.

3 演化密码对抗差分密码分析

对于 r -轮迭代型密码算法 E , 令 ω 表示算法 E 的核心部件, 如 E 表示 AES 加密算法, ω 表示算法 E 中的轮函数. 设 γ 表示算法的某个密码指标, 在对抗差分密码分析中, 我们令指标 γ 表示算法 E 的 $(r-1)$ -轮最佳差分特征 $\Delta X_0, \Delta X_1, \dots, \Delta X_{r-1}$ 的概率 p , 算法 E 中所有可能的 γ 组成集合 $\Gamma \subset \mathbb{R}^+$. 对于加密算法 E , 更换其核心部件 ω 将导致其密码指标相应地发生变化. 若利用某智能算法对部件 ω_0 进行作用, 可得到一些新的密码部件 $\omega_1, \omega_2, \dots, \omega_{t-1}$, 使对应的最佳差分特征的概率 $p^{(0)}, p^{(1)}, \dots, p^{(t-1)}$ 满足递减关系:

$$p^{(0)} \geq p^{(1)} \geq \dots \geq p^{(t-1)}.$$

基于此智能算法可以得到一个演化密码体制 $E_0, E_1, \dots, E_{t-1}$, 其中对于 $i = 0, 1, \dots, t-1$, 算法 E_i 对应于核心部件为 ω_i 的算法 E . 我们将上面的演化密码体制记为

$$\langle E, \omega_0, \omega_1, \dots, \omega_{t-1} \rangle,$$

其中算法对应的密码指标 $p^{(i)}$ 满足不等式

$$p^{(0)} \geq p^{(1)} \geq \dots \geq p^{(t-1)}. \quad (2)$$

需要指出的是对于分组算法 E , 算法中有些密码部件和轮密钥的产生有关, 有些部件和轮密钥的产生无关, 我们仅讨论部件更换不会改变轮密钥的情形. 换句话说, 本文假设算法 $\langle E, \omega_0 \rangle$ 和 $\langle E, \omega_0, \omega_1, \dots, \omega_{t-1} \rangle$ 具有相同的轮密钥.

对于算法 $\langle E, \omega_0 \rangle$ 和 $\langle E, \omega_0, \omega_1, \dots, \omega_{t-1} \rangle$, 下面比较差分攻击这两个算法的复杂度水平. 先给出如下一些攻击假设:

1. 攻击者在时间周期 T 内可以截获 N 个明密文对;
2. 攻击者的数据截获能力与算法相互独立, 即对算法 $\langle E, \omega_0 \rangle$ 和 $\langle E, \omega_0, \omega_1, \dots, \omega_{t-1} \rangle$ 攻击者的数据截获能力相当;
3. 演化密码 $\langle E, \omega_0, \omega_1, \dots, \omega_{t-1} \rangle$ 每隔一个时间周期 T/t , 更换一次核心部件 ω ;
4. 攻击者在每个时间周期内的数据截获能力基本相当, 换句话说, 攻击者攻击每个算法 $E_i, i = 0, 1, \dots, t-1$ 时可利用的明密文数量约为 N/t ;
5. 攻击者每次攻击算法 $E_i, (0 \leq i \leq t-1)$ 时都利用 $(r-1)$ -轮差分特征分析第 r 轮的轮密钥.

差分密码分析的复杂度主要取决于数据复杂度以及密钥搜索阶段的时间复杂度, 我们先对差分攻击固定密码算法 $\langle E, \omega_0 \rangle$ 和演化密码算法 $\langle E, \omega_0, \omega_1, \dots, \omega_{t-1} \rangle$ 的数据复杂度进行比较.

定理 1 对于算法结构相同的两个分组密码算法 E 和 E' , 设它们对应的最佳差分特征的概率为 p 和 p' , 且 $p > p'$. 当差分攻击算法 E 和 E' 时, 若攻击算法 E 和 E' 的预期成功率和比特优势相同, 则两次差分攻击所需的明文对数量 N 和 N' 满足 $N < N'$.

证明 设攻击算法 E 和 E' 时的比特优势为 a , 由于攻击算法 E 和 E' 的预期成功率和比特优势相同, 根据式 (1),

$$\Phi\left(\frac{\sqrt{p^2N} - \sqrt{p_r}\Phi(1-2^{-a})}{\sqrt{p+p_r}}\right) = P_S = P'_S = \Phi\left(\frac{\sqrt{p'^2N'} - \sqrt{p'_r}\Phi(1-2^{-a})}{\sqrt{p'+p'_r}}\right).$$

由于算法 E 和 E' 的算法结构相同, 故一个随机明文对确定某特定错误密钥的平均概率基本相同, 即 $p_r = p'_r$. 从而有

$$\frac{\sqrt{p^2N} - \sqrt{p_r}\Phi(1-2^{-a})}{\sqrt{p+p_r}} = \frac{\sqrt{p'^2N'} - \sqrt{p_r}\Phi(1-2^{-a})}{\sqrt{p'+p_r}}.$$

进一步地, 由于 $p > p'$, 我们有

$$\frac{\sqrt{p^2N}}{\sqrt{p+p_r}} < \frac{\sqrt{p'^2N'}}{\sqrt{p'+p_r}},$$

故有

$$N < \frac{p'^2(p+p_r)}{p^2(p'+p_r)}N' = \frac{p'^2p + p'^2p_r}{p^2p' + p^2p_r}N' < N'.$$

命题得证.

定理 2 对于算法 $\langle E, \omega_0 \rangle$ 和 $\langle E, \omega_0, \omega_1, \dots, \omega_{t-1} \rangle$, 设算法 $\langle E, \omega_0, \omega_1, \dots, \omega_{t-1} \rangle$ 的最佳差分特征概率满足式 (2), 若差分攻击算法 $\langle E, \omega_0 \rangle$ 和 $\langle E, \omega_0, \omega_1, \dots, \omega_{t-1} \rangle$ 的预期成功率相同且比特优势相同, 则差分攻击算法 $\langle E, \omega_0 \rangle$ 和 $\langle E, \omega_0, \omega_1, \dots, \omega_{t-1} \rangle$ 所需的明文对数量 N 和 $\hat{N}$ 满足 $N \leq \hat{N}$.

证明 类似于定理 1 的证明, 我们通过考察差分攻击算法 $\langle E, \omega_0 \rangle$ 和 $\langle E, \omega_0, \omega_1, \dots, \omega_{t-1} \rangle$ 过程中对应的差分特征概率, 比特优势和成功率的关系式来证明上面的结论. 对于攻击算法 $\langle E, \omega_0 \rangle$ 的情形, 命题 1 确定了差分特征, 比特优势和成功率的关系式. 下面我们讨论攻击算法 $\langle E, \omega_0, \omega_1, \dots, \omega_{t-1} \rangle$ 的情形. 对于 $i = 0, 1, \dots, t-1$, 差分攻击算法 E_i 可利用的明文对数量仅为 N/t , 攻击者每次攻击 E_i 时都试图恢复出算法的第 r 轮密钥, 相应的密钥长度为 m . 在 2^m 个候选密钥中, 记 k_0 表示正确密钥, k_j ($1 \leq j \leq 2^m - 1$) 表示错误密钥. 在对算法 E_i ($0 \leq i \leq t-1$) 进行差分攻击过程中, 令 $p_j^{(i)}$ 表示一个明文对确定的密钥测试集包含密钥 k_j 的概率, $p_r^{(i)}$ 表示一个随机明文对确定的密钥测试集包含某错误密钥的概率, 则有 $p_0^{(i)} = p^{(i)} + (1-p^{(i)})p_r^{(i)} \approx p^{(i)} + p_r^{(i)}$; 根据错误密钥假设, 对于 $1 \leq j \leq 2^m - 1$, 我们有 $p_j^{(i)} = p_r^{(i)}$.

对于 $i = 0, 1, \dots, t-1$, 记 $T_j^{(i)}$ 表示差分攻击 E_i 时密钥 k_j 出现的次数, $\mu^{(i)} = p^{(i)} \cdot N/t$ 表示正确的期望值. 根据错误密钥假设, 密钥计数器 $T_j^{(i)}$ ($1 \leq j \leq 2^m - 1$) 相互独立且服从相同的概率分布. 不难看出计数器 $T_0^{(i)}$ 服从二项分布 $\mathcal{B}(N/t, p_0^{(i)})$, 而 $T_j^{(i)}$ ($1 \leq j \leq 2^m - 1$) 服从二项分布 $\mathcal{B}(N/t, p_r^{(i)})$, 当 N/t 充分大时, 计数器 $T_0^{(i)}$ 近似服从正态分布 $\mathcal{N}(\mu_0^{(i)}, (\sigma_0^{(i)})^2)$, $T_j^{(i)}$ ($1 \leq j \leq 2^m - 1$) 近似服从正态分布 $\mathcal{N}(\mu_j^{(i)}, (\sigma_j^{(i)})^2)$, 其中

$$\begin{aligned} \mu_0^{(i)} &= p_0^{(i)}N/t, & (\sigma_0^{(i)})^2 &= p_0^{(i)}(1-p_0^{(i)})N/t \approx p_0^{(i)}N/t, \\ \mu_j^{(i)} &= p_r^{(i)}N/t, & (\sigma_j^{(i)})^2 &= p_r^{(i)}(1-p_r^{(i)})N/t \approx p_r^{(i)}N/t, \quad 1 \leq j \leq 2^m - 1. \end{aligned} \quad (3)$$

为了下面讨论简便, 我们记 $\hat{p} = \sum_{i=0}^{t-1} p^{(i)}/t$ 表示这 t 个最佳差分特征概率的均值, 并定义统计量 $U_j = \sum_{i=0}^{t-1} T_j^{(i)}$, $j = 0, 1, \dots, 2^m - 1$. 根据正态分布的性质, 统计量 U_j 近似服从正态分布 $\mathcal{N}(\mu_j, \sigma_j^2)$, 其中

$$\begin{aligned} \mu_0 &= \sum_{i=0}^{t-1} \mu_0^{(i)} = \sum_{i=0}^{t-1} p_0^{(i)} N/t = (\hat{p} + p_r)N, & \sigma_0^2 &= \sum_{i=0}^{t-1} (\sigma_0^{(i)})^2 \approx (\hat{p} + p_r)N, \\ \mu_j &= \sum_{i=0}^{t-1} \mu_j^{(i)} = \sum_{i=0}^{t-1} p_r^{(i)} N/t = p_r N, & \sigma_j^2 &= \sum_{i=0}^{t-1} (\sigma_j^{(i)})^2 \approx p_r N, \quad 1 \leq j \leq 2^m - 1. \end{aligned} \quad (4)$$

对统计量 $U_0, U_1, \dots, U_{2^m-1}$ 按递增的顺序重排, 记重排后的次序统计量为 $U_{\tau(0)}, U_{\tau(1)}, \dots, U_{\tau(2^m-1)}$, 设差分攻击的比特优势为 a , 记 $\varrho = 2^m - 2^a$, 在前 2^{m-a} 个候选密钥中恢复出正确密钥 k_0 的成功率 $\tilde{P}_S = \Pr(U_0 > U_{\tau(\varrho)})$. 根据命题 2, 由于统计量 U_j ($1 \leq j \leq 2^m - 1$) 近似服从正态分布 $\mathcal{N}(p_r N, p_r N)$, 故 $U_{\tau(\varrho)}$ 近似服从正态分布 $\mathcal{N}(\mu_q, \sigma_q^2)$, 其中

$$\begin{aligned} \mu_q &= F^{-1}(1 - 2^{-a}) = p_r N + p_r N \Phi^{-1}(1 - 2^{-a}), \\ \sigma_q &= \frac{1}{f(\mu_q)} 2^{-\frac{m+a}{2}} = \frac{p_r N}{\phi(\Phi^{-1}(1 - 2^{-a}))} 2^{-\frac{m+a}{2}}, \end{aligned}$$

进而统计量 $U_0 - U_{\tau(\varrho)}$ 近似服从正态分布 $\mathcal{N}(\mu_0 - \mu_q, \sigma_0^2 + \sigma_q^2)$, 我们记此正态分布的密度函数为 f' . 因此,

$$\tilde{P}_S = \Pr(U_0 - U_{\tau(\varrho)} > 0) = \int_0^\infty f'(x) dx = \int_{-\frac{\mu_0 - \mu_q}{\sqrt{\sigma_0^2 + \sigma_q^2}}}^\infty \phi(x) dx,$$

其中 $\phi(x)$ 表示标准正态分布的密度函数. 由于 $\sigma_q^2 \ll \sigma_0^2$, 故有

$$\tilde{P}_S = \int_{-\infty}^{\frac{\mu_0 - \mu_q}{\sigma_0}} \phi(x) dx = \Phi\left(\frac{\mu_0 - \mu_q}{\sigma_0}\right), \quad (5)$$

其中 $\mu_0 = \sigma_0^2 = (\hat{p} + p_r)N$, $\mu_q = p_r N + p_r N \Phi^{-1}(1 - 2^{-a})$.

由于攻击算法 $\langle E, \omega_0 \rangle$ 和 $\langle E, \omega_0, \dots, \omega_{t-1} \rangle$ 的预期成功率相同, 根据式 (1) 和 (5), 可知

$$\Phi\left(\frac{\sqrt{(p^{(0)})^2 N} - \sqrt{p_r} \Phi(1 - 2^{-a})}{\sqrt{p^{(0)} + p_r}}\right) = P_S = \tilde{P}_S = \Phi\left(\frac{\sqrt{\hat{p}^2 \hat{N}} - \sqrt{p_r} \Phi(1 - 2^{-a})}{\sqrt{\hat{p} + p_r}}\right).$$

类似于定理 1 的证明, 由于 $\hat{p} = \sum_{i=0}^{t-1} p^{(i)}/t \leq p^{(0)}$, 我们有

$$N \leq \frac{\hat{p}^2(p^{(0)} + p_r)}{(p^{(0)})^2(\hat{p} + p_r)} \hat{N} = \frac{\hat{p}^2 p^{(0)} + \hat{p}^2 p_r}{(p^{(0)})^2 \hat{p} + (p^{(0)})^2 p_r} \hat{N} \leq \hat{N}.$$

命题得证.

定理 2 比较了在预期成功率和比特优势相同的情况下, 差分攻击算法 $\langle E, \omega_0 \rangle$ 和 $\langle E, \omega_0, \dots, \omega_{t-1} \rangle$ 的数据复杂度水平. 下面我们在数据复杂度相当的情况下, 比较差分攻击算法 $\langle E, \omega_0 \rangle$ 和 $\langle E, \omega_0, \dots, \omega_{t-1} \rangle$ 的时间复杂度. 假设攻击者在时间 T 内可截获的明密文数量为 N , 差分攻击算法 $\langle E, \omega_0 \rangle$ 的情形前面已经讨论. 对于算法 $\langle E, \omega_0, \dots, \omega_{t-1} \rangle$ 的情形, 在密钥提取阶段, 攻击者需搜索每个算法 E_i ($0 \leq i \leq t-1$) 的 $(r-1)$ -轮最佳差分特征, 记 $\vartheta_1^{(E_i)}$ 表示利用文献 [20] 中的算法搜索 E_i 的 $(r-1)$ -轮最

表 2 $\langle E, \omega_0 \rangle$ 和 $\langle E, \omega_0, \omega_1, \dots, \omega_{t-1} \rangle$ 的复杂度比较Table 2 Comparison of the complexity for differential cryptanalysis on $\langle E, \omega_0 \rangle$ and $\langle E, \omega_0, \omega_1, \dots, \omega_{t-1} \rangle$

	$\langle E, \omega_0 \rangle$		$\langle E, \omega_0, \omega_1, \dots, \omega_{t-1} \rangle$	
	Time compl.	Memory compl.	Time compl.	Memory compl.
Distillation phase	$\vartheta_1^{(E_0)} + N \cdot \vartheta_2$	$\mathcal{O}(N)$	$\sum_{i=0}^{t-1} \vartheta_1^{(E_i)} + N \cdot \vartheta_2$	$\mathcal{O}(N)$
Analysis phase	$2^m \vartheta_3$	$\mathcal{O}(2^m)$	$2^m \vartheta_3$	$\mathcal{O}(2^m)$
Search phase	$2^{n_k - a} \vartheta_5^{(E)}$	$\mathcal{O}(2^m)$	$2^{n_k - \hat{a}} \vartheta_5^{(E)}$	$\mathcal{O}(2^m)$

佳差分特征的时间复杂度, 沿用上节中的符号, 对于算法 E_i ($0 \leq i \leq t-1$), 计算三元组 $(\Delta X_{r-1}, X_r, X_r^*)$ 的密钥测试集的时间复杂度为 $N\vartheta_2/t$, 故密钥提取阶段的时间复杂度为 $\sum_{i=0}^{t-1} \vartheta_1^{(E_i)} + N\vartheta_2$, 空间复杂度为 $\mathcal{O}(N)$. 在密钥分析阶段, 分析算法 $\langle E, \omega_0 \rangle$ 和 $\langle E, \omega_0, \omega_1, \dots, \omega_{t-1} \rangle$ 的情形类似, 时间复杂度和空间复杂度基本相当. 在密钥搜索阶段, 当攻击 $\langle E, \omega_0 \rangle$ 和 $\langle E, \omega_0, \omega_1, \dots, \omega_{t-1} \rangle$ 的预期成功率相同时, 由于核心部件改变引起最佳差分特征概率发生变化, 两次攻击的比特优势度不同, 从而导致搜索阶段的时间复杂度也不同. 根据上面的分析, 在数据复杂度为 N , 预期成功率相同的情况下, 表 2 比较了差分攻击算法 $\langle E, \omega_0 \rangle$ 和 $\langle E, \omega_0, \omega_1, \dots, \omega_{t-1} \rangle$ 的复杂度水平. 其中攻击算法 $\langle E, \omega_0 \rangle$ 和 $\langle E, \omega_0, \omega_1, \dots, \omega_{t-1} \rangle$ 的比特优势分别为

$$\begin{aligned} a &= -\log_2 \left(1 - \Phi \left(\sqrt{(p^{(0)})^2 N / p_r} - \sqrt{p^{(0)} / p_r + 1} \Phi^{-1}(P_S) \right) \right), \\ \hat{a} &= -\log_2 \left(1 - \Phi \left(\sqrt{\hat{p}^2 N / p_r} - \sqrt{\hat{p} / p_r + 1} \Phi^{-1}(P_S) \right) \right). \end{aligned} \quad (6)$$

由于 $\hat{p} = \sum_{i=0}^{t-1} p^{(i)} / t \leq p^{(0)}$, 成功率 P_S 和数据复杂度 N 相同, 不难推出 $a \geq \hat{a}$. 因此在密钥提取阶段和密钥搜索阶段, 差分攻击 $\langle E, \omega_0, \omega_1, \dots, \omega_{t-1} \rangle$ 的时间复杂度高于攻击 $\langle E, \omega_0 \rangle$ 的时间复杂度.

4 结论

本文阐述了演化密码的基本思想, 分析了差分攻击演化密码 $\langle E, \omega_0, \omega_1, \dots, \omega_{t-1} \rangle$ 过程中, 数据复杂度, 最佳差分特征概率, 比特优势以及成功率的关系. 在此基础上, 证明了在比特优势和预期成功率相同的条件下, 差分攻击演化密码 $\langle E, \omega_0, \omega_1, \dots, \omega_{t-1} \rangle$ 的数据复杂度大于攻击固定不变的算法 $\langle E, \omega_0 \rangle$ 的数据复杂度; 并证明了数据复杂度和预期成功率相同时, 差分攻击演化密码 $\langle E, \omega_0, \omega_1, \dots, \omega_{t-1} \rangle$ 的时间复杂度明显高于攻击固定算法 $\langle E, \omega_0 \rangle$ 的情形. 这表明演化密码在对抗传统的差分攻击的安全性高于固定算法密码.

参考文献

- 1 Zhang H G, Feng X T, Qin Z P, et al. Evolutionary cryptosystems and evolutionary design for DES. J China Institute Commun, 2002, 23: 57-64 [张焕国, 冯秀涛, 覃中平, 等. 演化密码与 DES 密码的演化设计. 通信学报, 2002, 23: 57-64]
- 2 Zhang H G, Feng X T, Qin Z P, et al. Research on evolutionary cryptosystems and evolutionary DES. Chin J Comput, 2003, 26: 1678-1684 [张焕国, 冯秀涛, 覃中平, 等. 演化密码与 DES 的演化研究. 计算机学报, 2003, 26: 1678-1684]
- 3 Zhang H G, Qin Z P. Introduction to Evolutionary Cryptography. Wuhan: Wuhan University Publishing Company, 2010 [张焕国, 覃中平. 演化密码引论. 武汉: 武汉大学出版社, 2010]
- 4 Meng Q S, Zhang H G, Wang Z Y, et al. Designing Bent functions using evolving method. Chin J Eletron, 2004, 32: 1901-1903 [孟庆树, 张焕国, 王张宜, 等. Bent 函数的演化设计. 电子学报, 2004, 32: 1901-1903]

- 5 Meng Q S, Zhang H G, Yang M, et al. Analysis of affinely equivalent Boolean functions. *Sci China Ser F-Inf Sci*, 2007, 37: 151–158 [孟庆树, 张焕国, 杨敏, 等. 仿射等价布尔函数的分析. *中国科学 F 辑: 信息科学*, 2007, 37: 151–158]
- 6 Meng Q S, Tang M, Zhang H G. Evolutionary design of trace form Bent function. Available <http://eprint.iacr.org.2005/332>
- 7 Wang H Z, Zhang H G, Wu Q H, et al. Design theory and method of multivariate hash function. *Sci China Inf Sci*, 2010, 40: 1299–1311 [王后珍, 张焕国, 伍前红, 等. 多变量 Hash 函数的构造理论与方法. *中国科学: 信息科学*, 2010, 40: 1299–1311]
- 8 Wang H Z, Zhang H G, Guan H M, et al. A new perturbation algorithm and enhancing security of SFLASH signature scheme. *Sci China Inf Sci*, 2010, 40: 393–402 [王后珍, 张焕国, 管海明, 等. 一种新的加噪扰动算法及其对 SFLASH 签名方案安全性的增强. *中国科学: 信息科学*, 2010, 40: 393–402]
- 9 Zhang H G, Li C L, Tang M. Research on evolutionary cryptography against multidimensional linear cryptanalysis. *Sci China Inf Sci*, 2012, 42: 603–616 [张焕国, 李春雷, 唐明. 演化密码对抗多重线性密码分析能力的研究, *中国科学: 信息科学*, 2012, 42: 603–616]
- 10 Biham E, Shamir A. Differential cryptanalysis of DES-like cryptosystems. In: Menezes A J, Vanstone S A, eds. *Advances in Cryptology-Crypto'90*. LNCS 537. Berlin: Springer-Verlag, 1991. 2–21
- 11 Biham E, Shamir A. *Differential Cryptanalysis of the Data Encryption Standard*. New York: Springer-Verlag, 1993
- 12 Biham E, Shamir A. Differential cryptanalysis of the full 16-round DES. In: Brickell E F, ed. *Advances in Cryptology-Crypto'92*, LNCS 740. Heidelberg: Springer-Verlag, 1993. 487–496
- 13 Knudsen L R. Truncated and higher order differentials. In: Preneel B, ed. *FSE 1994*, LNCS 1008. Heidelberg: Springer-Verlag, 1995. 196–211
- 14 Biham E, Biryukov A, Shamir A. Cryptanalysis of Skipjack reduced to 31 rounds using impossible differentials. In: Stern J, ed. *Advances in Cryptology-Eurocrypt'99*, LNCS 1592. Heidelberg: Springer-Verlag, 1999. 12–23
- 15 Hellman M E, Langford S K. Differential-linear cryptanalysis. In: Stern J, ed. *Advances in Cryptology-Crypto'94*, LNCS 839. Heidelberg: Springer-Verlag, 1994. 26–39
- 16 Lai X J. Higher order derivatives and differential cryptanalysis. In: *Proceedings of Symposium on Communication, Coding and Cryptography in Honor of James L. Massey on the Occasion of His 60th Birthday*. Berlin: Springer-Verlag, 1994. 227–233
- 17 Wagner D. The boomerang attack. In: Knudsen L R, ed. *FSE 1999*, LNCS 1636. Heidelberg: Springer-Verlag, 1999. 156–170
- 18 Junod P, Vaudenay S. Optimal key ranking procedures in a statistical cryptanalysis. In: Johansson T, ed. *FSE 2003*, LNCS 2887. Berlin: Springer-Verlag, 2003. 235–246
- 19 Selcuk A. On probability of success in linear and differential cryptanalysis. *J Cryptol*, 2008, 21: 131–147
- 20 Matsui M. On correlation between the order of S-boxes and the strength of DES. In: De Santis A, ed. *Advances in Cryptology-Eurocrypt'93*, LNCS 950. Berlin: Springer-Verlag, 1995. 366–375
- 21 Renyi A. *Probability Theory*. New York: American Elsevier Publishing Company, Inc., 1970

Capability of evolutionary cryptosystems against differential cryptanalysis

ZHANG HuanGuo^{1,2*}, LI ChunLei^{1,2} & TANG Ming^{1,2}

¹ School of Computer, Wuhan University, Wuhan 430072, China;

² Key Laboratory of Aerospace Information Security and Trusted Computing of Ministry of Education of China, Wuhan University, Wuhan 430072, China

*E-mail: liss@whu.edu.cn

Abstract The evolutionary cryptosystem is a new cryptosystem. This paper studies its security level resisting

against differential cryptanalysis. It is shown that the evolutionary cryptosystem possesses higher resistance than its initial fixed cryptosystem against differential cryptanalysis. On the basis of the relationship among the data complexity, the bit advantage and the success rate of differential cryptanalysis, it is proven that more data is needed for attacking the evolutionary cryptosystem when the bit advantage and success rate are identical. Moreover, it is shown that the time complexity for attacking the evolutionary cryptosystem is higher than that of differential attacking its initial fixed cryptosystem with the same amount of plaintext-ciphertext pairs. The research indicates that the evolutionary cryptosystem is more robust than its initial fixed cryptosystem against differential cryptanalysis.

Keywords information security, cryptology, evolutionary cryptosystem, differential cryptanalysis, bit advantage

ZHANG HuanGuo was born in 1945. He was graduated from XiDian University in 1970. He is now Professor and Ph.D. Director of Computer School in Wuhan University. His research interests include information security, cryptography, trusted computing, cloud computing, fault tolerant computing and computer application. He is a Senior Researcher of CCF. His professional services at present comprise co-chair of Didactic Committee of Information Security Major of Ministry of Education of China as well as a director of Chinese Association for Cryptologic Research.

LI ChunLei was born in 1984. He received the M.S. in mathematics from Hubei University, Wuhan in 2008. Currently, he is a Ph.D. candidate in information security in Wuhan University. His research interests include cryptography, evolutionary cryptosystem, sequence design and coding theory.