基于卡方统计量的多差分攻击方法

高海英,金晨辉,张军琪
(解放军信息工程大学,河南郑州 450001)

摘 要: 为了精确地估计分组密码算法抵抗差分攻击的能力,在已知多个具有高概率差分特征的条件下,提出了基于卡方统计量的多差分攻击方法.分析了正确密钥和错误密钥对应的统计量的分布规律,给出了多差分攻击方法的成功率、数据复杂度和计算复杂度的关系.在分组密码算法的差分特征概率未知的条件下,该方法仍然是适用的.

关键词: 分组密码;多差分密码分析;差分特征;数据复杂度;成功率
中图分类号: TN918.1 **文献标识码:** A **文章编号:** 0372-2112 (2014)09-1775-06
电子学报 URL: http://www.ejournal.org.cn **DOI:** 10.3969/j.issn.0372-2112.2014.09.017

Multiple Differential Cryptanalysis Using Chi-Square Statistics

GAO Hai-ying, JIN Chen-hui, ZHANG Jun-qi
(*Information Engineering University*, *Zhengzhou*, *Henan* 450001, *China*)

Abstract: In order to evaluate the capability that block ciphers resist differential attack accurately, a multiple differential cryptanalysis method is proposed in which chi-square statistics is constructed based on multiple differential characteristics with high probablities. We analyze the probability distribution of statistics corresponding correct key and incorrect key, and give the relation of data complexity, computational complexity and success probability. We point that the multiple differential cryptanalysis method can be applied to the the instance when the probabilities of differential characteristics are unknown.

Key words: block cipher; multiple differential cryptanalysis; differential characteristic; data complexity; success probability

1 引言

差分密码分析方法是由 Biham 和 Shamir[1]提出的一种针对分组密码算法的选择明文攻击方法,该方法已经成为一种常用的密码分析方法,不仅应用于分组密码算法,而且应用于序列密码、杂凑函数的安全性分析.基于该方法已扩展出一系列不同的分析方法,例如不可能差分分析[2]、条件差分分析[3]、截断差分[4]和高阶差分分析方法[5]等,并且利用该分析方法已成功地攻击了多个分组密码算法[6~10].

最初提出的差分分析方法只使用了一个差分特征进行攻击,为了更精确地估计分组密码算法抵抗差分攻击的能力,密码分析者研究利用多个差分特征对分组密码算法进行攻击的方法,1991 年,文献[11]提出的差分分析方法中综合利用多个具有相同输出差分的差分特征;1994 年,Knudsen[9]提出了截断差分分析方法,该方法用到的多个差分特征中的输出差分必须构成一个线性空间.2011 年,Blondeau 等[12]提出了综合利用多个输入差分、多个输入差分的差分特征的多差分攻击方法,但该方法没有最大限度地利用差分特征分布不均匀性这个信息泄漏.为了寻找更优的多差分攻击方法,2012 年,Blondeau 等[13]利用 LLR 统计量提出了针对单个输入差分、多个输出差分情况的最优的多差分攻击方法.但是,文献[13]的多差分攻击方法的应用存在一定的局限性,即统计量的计算都需要用到差分特征概率.针对多个输入差分、多个输出差分的情况,若这些差分特征的概率未知,如何对分组密码算法进行多差分攻击?为了解决该问题,本文采用与均匀分布进行拟合的思想提出了一种多差分攻击方法,分别分析了正确密钥和错误密钥对应的统计量的概率分布,给出了多差分攻击方法的成功率和数据复杂度之间的关系.

2 相关定义

令 E 表示一个分组密码算法,设该算法的圈数是 r,密钥是 K,分组规模是 m 比特,$E: Z_2^m \to Z_2^m$,$x \to y = E_K(x)$,$E_K(x) = F_{K_r} \circ \cdots \circ F_{K_1}(x)$,其中 F 表示圈函数,K_r 表示第 r 圈的圈子密钥.

下面给出与多差分密码分析方法相关的定义.

收稿日期:2013-04-13;修回日期:2013-07-27;责任编辑:梅志强
基金项目:国家自然科学基金(No.61272488,No.60272041,No.61202491)

定义 1[12] 设(δ_0,δ_{r-1})是分组密码算法 E 的$(r-1)$圈差分特征,该差分特征的概率为

$$P[\delta_0\rightarrow\delta_{r-1}]$$
$$=P[F_{K_r}^{-1}(E_K(X))\oplus F_{K_r}^{-1}(E_K(X\oplus\delta_0))=\delta_{r-1}]$$

若 K_r 是错误密钥,令 $P[\delta_0\rightarrow\delta_{r-1}]=2^{-m}$.

针对一个 r 圈分组密码算法,假设攻击者找到了多个$(r-1)$圈的具有较大概率的差分特征,将这些差分特征构成集合令为 $\Delta=\{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})\mid i=1,\cdots,|\Delta_0|,j=1,\cdots,|\Delta_{r-1}^{(i)}|\}$,并且令输入差分构成的集合为 $\Delta_0\overset{\text{def}}{=}\{\delta_0,\exists\,\delta_{r-1},(\delta_0,\delta_{r-1})\in\Delta\}=\{\delta_0^{(1)},\cdots,\delta_0^{(|\Delta_0|)}\}$,对 Δ_0 中的每个输入差分 $\delta_0^{(i)}$,对应的输出差分集合令为 $\Delta_{r-1}^{(i)}\overset{\text{def}}{=}\{\delta_{r-1}\mid(\delta_0^{(i)},\delta_{r-1})\in\Delta\}=\{\delta_{r-1}^{(i,1)},\cdots,\delta_{r-1}^{(i,|\Delta_{r-1}^{(i)}|)}\}$.

多差分密码攻击方法需要解决两个问题:一是针对每个实验密钥,如何利用多个差分特征构造统计量,使得正确密钥对应的统计量和错误密钥对应的统计量服从不同的概率分布;二是如何分析多差分密码攻击方法的各项性能指标.文中第 3、4 节分别解决这两个问题.

3 多差分攻击方法的具体描述

假设已知 N_i 个明文对构成的序列$\{(x_k,x_k\oplus\delta_0^{(i)})\}_{k=1}^{N_i}$,以及对应的密文对序列$\{(y_k,y_k')\}_{k=1}^{N_i}$,$i=1,2,\cdots,|\Delta_0|$,设第 r 圈的子密钥 K_r 有 2^{n_k}个可能值 $K_r^t(1\leqslant t\leqslant 2^{n_k})$,利用实验密钥 $K_r^t(1\leqslant t\leqslant 2^{n_k})$解密$\{(y_k,y_k')\}_{k=1}^{N_i}$,得到$\{(z_k,z_k')\}_{k=1}^{N_i}$,令序列 $\overrightarrow{\delta_{r-1}^{i}(K_r^t)}=\{z_k\oplus z_k'\}_{k=1}^{N_i}$,令 $N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^t)}=\#\{z_k\oplus z_k'=\delta_{r-1}^{(i,j)}\mid(x_k,x_k\oplus\delta_0^{(i)}),k=1,\cdots,N_i\}$.下面分析 K_r^t 是正确密钥和错误密钥时计数器 $N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^t)}$的分布规律.

定理 1 已知输入差分是 $\delta_0^{(i)}$ 的明密对的个数是 N_i,并且对于 $\forall\,j\in\{1,2,\cdots,|\Delta_{r-1}^{(i)}|\}$,已知 $P[\delta_{r-1}^{(i)}=\delta_{r-1}^{(i,j)}]=p^{(i,j)}$.

(1) 当 K_r^t 是正确密钥时,则

$$P[N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^t)}=d]=N_i(p^{(i,j)})^d(1-p^{(i,j)})^{N_i-d}$$
$$\mathrm{E}(N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^t)})=N_i p^{(i,j)}$$
$$\mathrm{D}(N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^t)})=N_i p^{(i,j)}(1-p^{(i,j)})$$

(2) 当 K_r^t 是错误密钥时,则

$$P[N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^t)}=d]=N_i(2^{-m})^d(1-2^{-m})^{N_i-d}$$
$$\mathrm{E}(N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^t)})=N_i2^{-m}$$
$$\mathrm{D}(N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^t)})=N_i2^{-m}(1-2^{-m})$$

证明 已知序列$\overrightarrow{\delta_{r-1}^{i}(K_r^t)}=\{z_k\oplus z_k'\}_{k=1}^{N_i}$,令

$$\xi_{k,\delta_0^{(i)},\delta_{r-1}^{(i,j)}}=\begin{cases}1, & x_k\oplus x_k'=\delta_0^{(i)}\text{ 且 }z_k\oplus z_k'=\delta_{r-1}^{(i,j)}\\0, & x_k\oplus x_k'=\delta_0^{(i)}\text{ 且 }z_k\oplus z_k'\neq\delta_{r-1}^{(i,j)}\end{cases}$$

当 K_r^t 是正确密钥时,

$$P[\xi_{k,\delta_0^{(i)},\delta_{r-1}^{(i,j)}}=1]=p^{(i,j)}$$
$$P[\xi_{k,\delta_0^{(i)},\delta_{r-1}^{(i,j)}}=0]=1-p^{(i,j)}$$

当 K_r^t 是错误密钥时,

$$P[\xi_{k,\delta_0^{(i)},\delta_{r-1}^{(i,j)}}=1]=2^{-m}$$
$$P[\xi_{k,\delta_0^{(i)},\delta_{r-1}^{(i,j)}}=0]=1-2^{-m}$$

由于 $N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^t)}=\sum_{k=1}^{N_i}\xi_{k,\delta_0^{(i)},\delta_{r-1}^{(i,j)}}$,因此 $N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^t)}$服从二项分布,由二项分布的相关结论可得定理 1. 证毕

对于 $\forall\,(\delta_0^{(i)},\delta_{r-1}^{(i,j)})\in\Delta$,令

$$\eta_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^t)}=\frac{N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^t)}-N_i2^{-m}}{\sqrt{N_i2^{-m}(1-2^{-m})}}$$

当 N_i 较大且 K_r^t 是错误密钥时,根据定理 1 的结论,可知 $\eta_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^t)}$近似服从 $\mathcal{N}(0,1)$分布.令统计量 $T_{K_r^t}=\sum_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})\in\Delta}\left(\eta_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^t)}\right)^2$,则错误密钥 K_r^F 对应的统计量 $T_{K_r^F}$应小于正确密钥 K_r^T 对应的统计量 $T_{K_r^T}$.

根据上述分析,在已知 N_i 个明文对构成的序列$\{(x_k,x_k\oplus\delta_0^{(i)})\}_{k=1}^{N_i}$和对应密文对序列$\{(y_k,y_k')\}_{k=1}^{N_i}$($i=1,2,\cdots,|\Delta_0|$)的条件下,本文提出了基于 χ^2 统计量的多差分攻击方法,如下所述:

Step1 利用每个实验密钥 $K_r^t(1\leqslant t\leqslant 2^{n_k})$解密$\{(y_k,y_k')\}_{k=1}^{N_i}$,得到$\{(z_k,z_k')\}_{k=1}^{N_i}$,令序列$\overrightarrow{\delta_{r-1}^{i}(K_r^t)}=\{z_k\oplus z_k'\}_{k=1}^{N_i}$,$i=1,2,\cdots,|\Delta_0|$.令 $N_{(\delta_0^{(i)}\rightarrow\delta_{r-1}^{(i,j)})}^{(K_r^t)}=\#\{z_k\oplus z_k'=\delta_{r-1}^{(i,j)}\mid(x_k,x_k\oplus\delta_0^{(i)}),k=1,\cdots,N_i\}$,计算统计量 $T_{K_r^t}=\sum_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})\in\Delta}\left(\eta_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^t)}\right)^2$,其中 $\eta_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^t)}=\frac{N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^t)}-N_i2^{-m}}{\sqrt{N_i2^{-m}(1-2^{-m})}}$.

Step2 计算每个实验密钥 K_r^t 对应的统计量 $T_{K_r^t}$,将 $T_{K_r^t}(1\leqslant t\leqslant 2^{n_k})$按从大到小的顺序排列,将前 l 个统计量对应的实验密钥作为候选密钥.

4 多差分攻击方法的指标分析

本节分析基于 χ^2 统计量的多差分攻击方法的成

功率和数据复杂度之间的关系.首先分析错误密钥的统计量 $T_{K_r^{'}}$的分布规律.

定理 2　设 $\eta_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^{'})}((\delta_0^{(i)},\delta_{r-1}^{(i,j)})\in\Delta)$ 相互独立,则当 $N_i(1\leqslant i\leqslant|\Delta_0|)$较大时,错误密钥的统计量 $T_{K_r^{'}}$近似服从$\mathcal{N}(|\Delta|,2|\Delta|)$分布.

证明　由第 3 节的分析可知,N_i 较大且 K_r^t 是错误密钥时,$\eta_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^{'})}$ 近似服从 $\mathcal{N}(0,1)$ 分布,假设 $\eta_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^{'})}$相互独立,则 $T_{K_r^{'}}$服从自由度为$|\Delta|$的 χ^2 分布,该分布的期望是$|\Delta|$,方差是 $2|\Delta|$.因此,当 $N_i(1\leqslant i\leqslant|\Delta_0|)$较大时,$T_{K_r^{'}}$近似服从 $\mathcal{N}(|\Delta|,2|\Delta|)$分布.

证毕

为了研究 $T_{K_r^{'}}$ 的分布规律,我们需要计算 $\mathrm{E}(N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)})^2$、$\mathrm{E}(N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)})^3$ 和 $\mathrm{E}(N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)})^4$ 的值.下面给出相关结论,首先介绍结论中用到的符号,记 P_n^r 表示 n 元集合的 r 元排列数,C_n^r 表示 n 元集合的 r 元组合数.

定理 3

$$\mathrm{E}\left(N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)}\right)^2=N_i\cdot p^{(i,j)}+N_i(N_i-1)(p^{(i,j)})^2$$

证明　已知序列$\overline{\delta_{r-1}^i(K_r^T)}=\{z_k\oplus z_k^{'}\}_{k=1}^{N_i}$,令

$$\xi_{k,\delta_0^{(i)},\delta_{r-1}^{(i,j)}}=\begin{cases}1, & x_k\oplus x_k^{'}=\delta_0^{(i)}\ \text{且}\ z_k\oplus z_k^{'}=\delta_{r-1}^{(i,j)}\\ 0, & x_k\oplus x_k^{'}=\delta_0^{(i)}\ \text{且}\ z_k\oplus z_k^{'}\neq\delta_{r-1}^{(i,j)}\end{cases}$$

由定理 1 的证明过程可知,当实验密钥是正确密钥时,

$$P[\xi_{k,\delta_0^{(i)},\delta_{r-1}^{(i,j)}}=1]=p^{(i,j)},\ P[\xi_{k,\delta_0^{(i)},\delta_{r-1}^{(i,j)}}=0]=1-p^{(i,j)}$$

因此,计算出

$$\mathrm{E}(\xi_{k,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})=\mathrm{E}(\xi_{k,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})^2=\mathrm{E}(\xi_{k,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})^3=\mathrm{E}(\xi_{k,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})^4=p^{(i,j)}\quad(1)$$

假设 $\xi_{k,\delta_0^{(i)},\delta_{r-1}^{(i,j)}}(1\leqslant k\leqslant N_i)$ 相互独立,由于 $N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)}=\sum_k\xi_{k,\delta_0^{(i)},\delta_{r-1}^{(i,j)}}$,因此

$$\begin{aligned}&\mathrm{E}\left(N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)}\right)^2=\mathrm{E}\left(\sum_k\xi_{k,\delta_0^{(i)},\delta_{r-1}^{(i,j)}}\right)^2\\&=\mathrm{E}\Big(\sum_k(\xi_{k,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})^2+2\sum_{\substack{k1,k2\\k1\neq k2}}(\xi_{k1,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})(\xi_{k2,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})\Big)\\&=\sum_k\mathrm{E}(\xi_{k,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})^2+2\sum_{\substack{k1,k2\\k1\neq k2}}\left(\mathrm{E}(\xi_{k1,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})\cdot\mathrm{E}(\xi_{k2,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})\right)\\&=N_ip^{(i,j)}+2C_{N_i}^2(p^{(i,j)})^2\\&=N_ip^{(i,j)}+N_i(N_i-1)(p^{(i,j)})^2\end{aligned}$$

证毕

定理 4

$$\mathrm{E}\left(N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)}\right)^3=N_i\cdot p^{(i,j)}+3P_{N_i}^2(p^{(i,j)})^2+6C_{N_i}^3(p^{(i,j)})^3$$

证明　已知多项式$(a+b+c+\cdots+l)^n$ 的展开式的通项为:

$$\frac{n!}{p!\ q!\ \cdots s!}a^pb^qc^r\cdots l^s(\text{其中 } p+q+r+\cdots+s=n)$$

因此,

$$\begin{aligned}&(N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)})^3=\left(\sum_k\xi_{k,\delta_0^{(i)},\delta_{r-1}^{(i,j)}}\right)^3\\&=\sum_k(\xi_{k,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})^3+\sum_{k1\neq k2}\frac{3!}{1!2!}(\xi_{k1,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})\cdot(\xi_{k2,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})^2\\&+\sum_{k1\neq k2\neq k3}3!(\xi_{k1,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})\cdot(\xi_{k2,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})\cdot(\xi_{k3,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})\end{aligned}$$

由于 $\xi_{k,\delta_0^{(i)},\delta_{r-1}^{(i,j)}}(1\leqslant k\leqslant N_i)$相互独立,并且根据式(1)可得:

$$\begin{aligned}&\mathrm{E}(N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)})^3\\&=\sum_k\mathrm{E}(\xi_{k,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})^3+\sum_{k1\neq k2}3\mathrm{E}(\xi_{k1,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})\cdot\mathrm{E}(\xi_{k2,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})^2\\&+\sum_{k1\neq k2\neq k3}6\mathrm{E}(\xi_{k1,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})\cdot\mathrm{E}(\xi_{k2,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})\cdot\mathrm{E}(\xi_{k3,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})\\&=N_i\cdot p^{(i,j)}+3P_{N_i}^2(p^{(i,j)})^2+6C_{N_i}^3(p^{(i,j)})^3\end{aligned}$$

证毕

定理 5

$$\begin{aligned}&\mathrm{E}(N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)})^4\\&=N_i\cdot p^{(i,j)}+(4P_{N_i}^2+6C_{N_i}^2)(p^{(i,j)})^2+6P_{N_i}^3(p^{(i,j)})^3\\&+24C_{N_i}^4(p^{(i,j)})^4\end{aligned}$$

证明

$$\begin{aligned}&(N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)})^4=\sum_k(\xi_{k,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})^4\\&+\sum_{k1\neq k2}\frac{4!}{1!3!}(\xi_{k1,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})(\xi_{k2,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})^3\\&+\sum_{k1\neq k2}\frac{4!}{2!2!}(\xi_{k1,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})^2(\xi_{k2,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})^2\\&+\sum_{k1\neq k2\neq k3}\frac{4!}{1!1!2!}(\xi_{k1,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})(\xi_{k2,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})\cdot(\xi_{k3,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})^2\\&+\sum_{k1\neq k2\neq k3\neq k4}4!(\xi_{k1,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})(\xi_{k2,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})(\xi_{k3,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})(\xi_{k4,\delta_0^{(i)},\delta_{r-1}^{(i,j)}})\end{aligned}$$

由于 $\xi_{k,\delta_0^{(i)},\delta_{r-1}^{(i,j)}}(1\leqslant k\leqslant N_i)$相互独立,根据式(1)可得:

$$
\begin{aligned}
&\mathrm{E}\left(N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)}\right)^4\\
&=\sum_k \mathrm{E}\left(\xi_{k,\delta_0^{(i)},\delta_{r-1}^{(i,j)}}\right)^4+\\
&\sum_{k1\neq k2}\frac{4!}{1!3!}\mathrm{E}\left(\xi_{k1,\delta_0^{(i)},\delta_{r-1}^{(i,j)}}\right)\mathrm{E}\left(\xi_{k2,\delta_0^{(i)},\delta_{r-1}^{(i,j)}}\right)^3+\\
&\sum_{k1\neq k2}\frac{4!}{2!2!}\mathrm{E}\left(\xi_{k1,\delta_0^{(i)},\delta_{r-1}^{(i,j)}}\right)^2\mathrm{E}\left(\xi_{k2,\delta_0^{(i)},\delta_{r-1}^{(i,j)}}\right)^2\\
&+\sum_{k1\neq k2\neq k3}\frac{4!}{1!1!2!}\mathrm{E}\left(\xi_{k1,\delta_0^{(i)},\delta_{r-1}^{(i,j)}}\right)\mathrm{E}\left(\xi_{k2,\delta_0^{(i)},\delta_{r-1}^{(i,j)}}\right)\\
&\cdot\mathrm{E}\left(\xi_{k3,\delta_0^{(i)},\delta_{r-1}^{(i,j)}}\right)^2\\
&+\sum_{k1\neq k2\neq k3\neq k4}4!\left[\mathrm{E}\left(\xi_{k1,\delta_0^{(i)},\delta_{r-1}^{(i,j)}}\right)\mathrm{E}\left(\xi_{k2,\delta_0^{(i)},\delta_{r-1}^{(i,j)}}\right)\right.\\
&\left.\cdot\mathrm{E}\left(\xi_{k3,\delta_0^{(i)},\delta_{r-1}^{(i,j)}}\right)\mathrm{E}\left(\xi_{k4,\delta_0^{(i)},\delta_{r-1}^{(i,j)}}\right)\right]\\
&=N_i\cdot p^{(i,j)}+(4P_{N_i}^2+6C_{N_i}^2)(p^{(i,j)})^2+6P_{N_i}^3(p^{(i,j)})^3\\
&\quad+24C_{N_i}^4(p^{(i,j)})^4
\end{aligned}
$$

证毕

利用定理 3 至定理 5 的结论,可求出 $T_{K_r^T}$的期望与方差.详见定理 6 和定理 7.

定理 6 设 $\eta_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)}((\delta_0^{(i)},\delta_{r-1}^{(i,j)})\in\Delta)$ 相互独立,则

$$
\begin{aligned}
\mathrm{E}(T_{K_r^T})&\approx 2^m\sum_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})\in\Delta}\left[(N_i-1)(p^{(i,j)})^2\right.\\
&\left.+(1-2^{-m+1}N_i)p^{(i,j)}+2^{-2m}N_i\right]
\end{aligned}
$$

证明

$$
\begin{aligned}
\mathrm{E}(T_{K_r^T})&=\sum_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})\in\Delta}\mathrm{E}\left(\eta_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)}\right)^2\\
&=\frac{1}{N_i2^{-m}(1-2^{-m})}\sum_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})\in\Delta}\mathrm{E}\left(N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)}-2^{-m}N_i\right)^2\\
&\approx\frac{2^m}{N_i}\sum_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})\in\Delta}\mathrm{E}\left(\left(N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)}\right)^2\right.\\
&\quad\left.-2^{-m+1}N_i\cdot N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)}+(2^{-m}N_i)^2\right)\\
&\approx\frac{2^m}{N_i}\sum_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})\in\Delta}\left[(N_i\cdot p^{(i,j)}+N_i(N_i-1)(p^{(i,j)})^2)\right.\\
&\quad\left.-2^{-m+1}N_ip^{(i,j)}+2^{-2m}N_i^2\right]\\
&\approx 2^m\sum_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})\in\Delta}\left[(N_i-1)(p^{(i,j)})^2\right.\\
&\quad\left.+(1-2^{-m+1}N_i)p^{(i,j)}+2^{-2m}N_i\right]
\end{aligned}
$$

证毕

定理 7 $\forall(\delta_0^{(i)},\delta_{r-1}^{(i,j)})\in\Delta$,

$$
\begin{aligned}
\mathrm{D}\left(\eta_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)}\right)^2&=(10N_i^2-4N_i^3-6N_i)(p^{(i,j)})^4\\
&+\left(4N_i^3-16N_i^2+12N_i+\frac{23N_i^3-5N_i^4-3N_i^2}{2^{m-2}}\right)(p^{(i,j)})^3\\
&+\left(6N_i^2-7N_i+\frac{3N_i^2-2N_i^3}{2^{m-2}}-\frac{2N_i^3}{2^{2m-1}}\right)(p^{(i,j)})^2\\
&+\left(N_i-\frac{N_i^2}{2^{m-2}}+\frac{4N_i^3}{2^{2m}}\right)\cdot p^{(i,j)}
\end{aligned}
$$

假设 $\eta_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)}((\delta_0^{(i)},\delta_{r-1}^{(i,j)})\in\Delta)$是相互独立的,则 $\mathrm{D}(T_{K_r^T})=\sum_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})\in\Delta}\mathrm{D}\left(\eta_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)}\right)^2$.

证明 由于 $\mathrm{D}\left(\eta_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)}\right)^2=\mathrm{E}\left(\eta_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)}\right)^4-\left(\mathrm{E}\left(\eta_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)}\right)^2\right)^2=\frac{1}{(N_i2^{-m}(1-2^{-m}))^2}$

$$
\cdot\left[\mathrm{E}\left(N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)}-N_i2^{-m}\right)^4-\left(\mathrm{E}\left(N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)}-N_i2^{-m}\right)^2\right)^2\right]
$$

因此,需要计算 $\mathrm{E}\left(N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)}-N_i2^{-m}\right)^4$ 和 $\mathrm{E}\left(N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)}-N_i2^{-m}\right)^2$.

$$
\begin{aligned}
&\mathrm{E}\left(N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)}-N_i2^{-m}\right)^4\\
&=\mathrm{E}\left(N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)}\right)^4-4(N_i2^{-m})\cdot\mathrm{E}\left(N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)}\right)^3\\
&\quad-4(N_i2^{-m})^3\cdot\mathrm{E}\left(N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)}\right)\\
&\quad+6(N_i2^{-m})^2\cdot\mathrm{E}\left(N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)}\right)^2\\
&\quad+(N_i2^{-m})^4\left(\mathrm{E}\left(N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)}-N_i2^{-m}\right)^2\right)^2\\
&=\left(\mathrm{E}\left(N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)}\right)^2-N_i2^{-m+1}\mathrm{E}\left(N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)}\right)\right.\\
&\quad\left.+(N_i2^{-m})^2\right)^2
\end{aligned}
$$

利用定理 1 和定理 3、4、5 可求出

$$
\begin{aligned}
&\mathrm{D}\left(\eta_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)}\right)^2\\
&=(10N_i^2-4N_i^3-6N_i)(p^{(i,j)})^4\\
&\quad+\left(4N_i^3-16N_i^2+12N_i+\frac{23N_i^3-5N_i^4-3N_i^2}{2^{m-2}}\right)(p^{(i,j)})^3\\
&\quad+\left(6N_i^2-7N_i+\frac{3N_i^2-2N_i^3}{2^{m-2}}-\frac{2N_i^3}{2^{2m-1}}\right)(p^{(i,j)})^2\\
&\quad+\left(N_i-\frac{N_i^2}{2^{m-2}}+\frac{4N_i^3}{2^{2m}}\right)\cdot p^{(i,j)}
\end{aligned}
$$

由于 $T_{K_r^T}=\sum_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})\in\Delta}\left(\eta_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)}\right)^2$,且假设 $\eta_{(\delta_0^{(i)}\to\delta_{r-1}^{(i,j)})}^{(K_r^T)}((\delta_0^{(i)},\delta_{r-1}^{(i,j)})\in\Delta)$ 相互独立,因此 $\mathrm{D}(T_{K_r^T})=\sum_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})\in\Delta}\mathrm{D}\left(\eta_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^T)}\right)^2$.

证毕

在已知数据复杂度的条件下,为了求出基于 χ^2 统计量的多差分攻击方法的成功率,需要用到文献[14]的结论.该结论描述如下.

引理 1[14] 设攻击的密钥是 n 比特,已知正确密

钥 K_T 对应的统计量 $S(K_T)$服从 $\mathcal{N}(\mu_T,\sigma_T^2)$分布,错误密钥 K_F 对应的统计量 $S(K_F)$服从 $\mathcal{N}(\mu_F,\sigma_F^2)$分布.将每个实验密钥对应的统计量 $S(K_i)$按从大到小排序,取前 l 个实验密钥作为候选密钥,记攻击成功的概率为 P_s,则

$$P_s\approx\varphi_{0,1}\left(\frac{\mu_T-\mu_F-\sigma_F\varphi_{0,1}^{-1}(1-2^{-a})}{\sigma_T}\right)$$

其中,$a=n-\log_2(l)$.

根据定理 6 和定理 7,可以求出 $T_{K_r^t}$的期望和方差,为了描述方便,记 $\mathrm{E}(T_{K_r^t})=\mu_T$,$\mathrm{D}(T_{K_r^t})=\sigma_T^2$,当样本量较大时,即 $N_i(i=1,2,\cdots,|\Delta_0|)$较大时,$T_{K_r^t}$近似服从 $\mathcal{N}(\mu_T,\sigma_T^2)$分布.又由定理 2 可知,$T_{K_r^F}$服从 $\mathcal{N}(|\Delta|,2|\Delta|)$分布.因此,基于引理 1[14],下面给出基于 χ^2 统计量的多差分分析方法的指标分析结果.

定理 8　针对基于 χ^2 统计量的多差分攻击方法,若选择前 l 个统计量对应的实验密钥作为候选密钥,令 $a=n_k-\log_2(l)$,在已知 $\sum_{i=1}^{|\Delta_0|}N_i+\max_{i=1,\cdots,|\Delta_0|}\{N_i\}$个明密对的条件,该攻击方法的计算复杂度约为 $O\left(\left(\sum_{i=1}^{|\Delta_0|}N_i+\max_{i=1,\cdots,|\Delta_0|}\{N_i\}\right)2^{n_k}\right)$次第 r 圈的圈函数解密,$P_s\approx\varphi_{0,1}\left(\frac{\mu_T-|\Delta|-\sqrt{2|\Delta|}\,\varphi_{0,1}^{-1}(1-2^{-a})}{\sigma_T}\right)$,$P_s$ 为成功率.

证明　已知正确密钥统计量 $T_{K_r^t}$近似服从 $\mathcal{N}(\mu_T,\sigma_T^2)$分布,错误密钥的统计量 $T_{K_r^F}$近似服从 $\mathcal{N}(|\Delta|,2|\Delta|)$分布,根据引理 1[14]可直接计算出成功率 P_s.

基于 χ^2 统计量的多差分攻击方法需要已知 N_i 个明文对构成的序列 $\{(x_k,x_k\oplus\delta_0^{(i)})\}_{k=1}^{N_i}$,以及对应的密文对序列 $\{(y_k,y_k')\}_{k=1}^{N_i}$,$i=1,2,\cdots,|\Delta_0|$,因此,该攻击方法的数据复杂度是 $\sum_{i=1}^{|\Delta_0|}N_i+\max_{i=1,\cdots,|\Delta_0|}\{N_i\}$个明密对.

该多差分攻击方法的计算复杂度包括:$\left(\sum_{i=1}^{|\Delta_0|}N_i+\max_{i=1,\cdots,|\Delta_0|}\{N_i\}\right)2^{n_k}$次第 r 圈的圈函数解密、2^{n_k}个统计量的计算以及 2^{n_k}个统计量的排序的计算复杂度,相对于圈函数解密的计算复杂度,统计量和排序的计算复杂度可忽略,因此,该算法的计算复杂度约为 $O\left(\left(\sum_{i=1}^{|\Delta_0|}N_i+\max_{i=1,\cdots,|\Delta_0|}\{N_i\}\right)2^{n_k}\right)$次第 r 圈的圈函数解密.　　证毕

特别说明:文献[12、13]中构造的统计量需要用到每个差分特征的差分特征概率,例如,文献[13]中构造的针对单个输入差分 δ_0、多个输出差分为 $\delta_{r-1}^{(j)}(j=1,2,\cdots,k)$的 χ^2 统计量

$$\chi^2(K_r^t)=N\sum_{j=1,2,\cdots,k}\frac{(P[\delta_0\to\delta_{r-1}^{(j)}]-2^{-m})^2}{2^{-m}}$$

其中 N 表示明文差分为 δ_0 的明密对的个数.

本文提出的多差分攻击方法中构造的 χ^2 统计量

$$T_{K_r^t}=\sum_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})\in\Delta}\left(\frac{N_{(\delta_0^{(i)},\delta_{r-1}^{(i,j)})}^{(K_r^t)}-N_i2^{-m}}{\sqrt{N_i2^{-m}(1-2^{-m})}}\right)^2$$

统计量 $\chi^2(K_r^t)$中需要用到概率 $P[\delta_0\to\delta_{r-1}^{(j)}]$,而统计量 $T_{K_r^t}$中却未用概率 $P[\delta_0\to\delta_{r-1}^{(j)}]$,因此,基于统计 $T_{K_r^t}$的多差分攻击方法在未知差分特征概率的情况下仍然是适用的,这也是本文提出的多差分攻击方法的一个优势.

5　结束语

针对每个实验密钥,采用与均匀分布拟合的思想,构造了 χ^2 统计量,在此基础上提出了针对分组密码算法的多差分密码攻击方法,分析了正确密钥和错误密钥对应的统计量的概率分布,从而给出了多差分分析方法的各项性能指标分析结果.与已有的多差分攻击方法相比,该方法适用性更广,不仅适用于差分特征概率已知的情况,也适用于差分特征概率未知的情况.

参考文献

[1] Biham E, Shamir A. Differential cryptanalysis of DES-like cryptosystems[A]. CRYPTO 1990[C]. Santa Barbara, California, USA: Springer-Verlag, 1990. 2 – 21.

[2] E Biham, A Biryukov, A Shamir. Cryptanalsis of skipjack reduced to 31 rounds using impossible differentials[A]. EUROCRYPT 1999[C]. Prague, Czech Republic: Springer-Verlag, 1999. 12 – 23.

[3] Simon Knellwolf, Willi Meier, Naya Plasencia. Conditional differential cryptanalysis of NLFSR-based cryptosystems[A]. ASIACRYPT 2010[C]. Swissôtel Merchant Court, Singapore: Springer-Verlag, 2010. 130 – 145.

[4] L R Knudsen. Truncated and higher order differentials[A]. FSE' 94[C]. Leuven, Belgium: Springer-Verlag, 1994. 196 – 211.

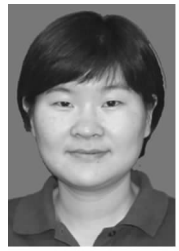
[5] 胡豫濮,蔡勉,肖国镇.一类高阶差分密码分析[J].电子学报,1999,27(10):74 – 78.
Hu Yu Pu, Cai Mian, Xiao Guo-zhen. A class of high-order differential cryptanalysis[J]. Acta Electronica Sinica, 1999, 27(10): 74 – 78. (in Chinese)

[6] Wang W. Differential cryptanalysis of reduced-round PRESENT[A]. AFRICACRYPT 2008[C]. Africa, Casablanca, Morocco: Springer-Verlag, 2008. 40 – 49.

[7] Gaoli Wang. Improved differential cryptanalysis of serpent[A]. Computational Intelligence and Security (CIS), 2010[C]. Nan-

ning: Springer-Verlag, 2010. 367 – 371.
[8] Lei Zang, Wen Tao Zhang, Wen Ling Wu. Cryptanalysis of reduced-round SMS4 block cipher[A]. Information Security and Privacy[C]. Wollongong, Australia: Springer-Verlag, 2008. 216 – 229.
[9] 李超,沈静. Camellia 的差分和线性迭代特征[J]. 电子学报,2005,33(8):1345 – 1348.
Li Chao, Shen Jing. Differential andlinear iterative characteristics of camellia[J]. Acta Electronica Sinica, 2005, 33(8): 1345 – 1348. (in Chinese)
[10] Nicolas T Corutois. An Improved Differential Attack on Full GOST[EB/OL]. http://eprint.iacr.org, 2013-02.
[11] Biham E, Shamir A. Differential cryptanalysis of DES-like cryptosystems[J]. Journal of Cryptology, 1991, 4(1): 3 – 72.
[12] Céline Blondeau, Benoît Gérard. Multiple differential cryptanalysis: Theory and practice[A]. Fast Soft Encryption 2011 [C]. Lyngby, Denmark: Springer-Verlag, 2011. 35 – 54.
[13] Céline Blondeau, Benoît Gérard, Kaisa Nyberg. Multiple Differential Cryptanalysis Using LLR and χ^2 Statistics[EB/OL]. http://eprint.iacr.org, 2012 – 09.
[14] Selcuk A A. On probability of success in linear and differential cryptanalysis[J]. Journal of Cryptology, 2008, 21(1): 131 – 147.

作者简介



高海英 女,1978 年 7 月出生,河南沈丘人. 2006 年获北京邮电大学密码学专业博士学位,现为解放军信息工程大学副教授,研究方向为密码理论.
E-mail: ghyyueyue@126.com

金晨辉 男,1965 年 3 月出生,河南扶沟人. 解放军信息工程大学教授,博士生导师,研究方向为密码理论.

张军琪 男,1991 年 5 月出生,河南商丘人. 现为信息工程大学硕士研究生,研究方向为密码算法分析.