概率积分密码分析

李晓千[1)] 吴文玲[1)] 李 宝[1)] 于晓丽[2)]
[1)](中国科学院信息工程研究所 北京 100093)
[2)](中国科学院软件研究所 北京 100190)

摘 要 在传统的积分密码分析中,积分区分器都是以概率1成立的.虽然Knudsen等学者提到过:"就像差分一样,积分也可以是概率的",但是,没有文献报道过进一步的研究.文中对此问题进行了探讨,提出了概率积分密码分析方法,并从理论和实验两方面验证了概率积分分析方法的有效性.对于采用S盒设计的分组密码,文中证明了如果S盒的差分均匀性越接近随机概率,则分组密码抵抗概率积分密码分析的能力就越强.同时,文中指出高阶积分分析的某些技巧对于概率积分分析是行不通的,主要原因是随着求和变量个数的增加,积分特征概率趋近于随机概率.最后,文中通过对AES和LBlock这两个算法的概率积分分析实例,说明目前广泛使用的分组密码算法对于概率积分密码分析方法都是免疫的.

关键词 积分密码分析;差分密码分析;概率积分密码分析;轻量级密码;AES;LBlock
中图法分类号 TP309 **DOI号**:10.3724/SP.J.1016.2012.01897

Probabilistic Integral Cryptanalysis

LI Xiao-Qian[1)] WU Wen-Ling[1)] LI Bao[1)] YU Xiao-Li[2)]
[1)](*Institute of Information Engineering, Chinese Academy of Sciences, Beijing* 100093)
[2)](*Institute of Software, Chinese Academy of Sciences, Beijing* 100190)

Abstract In traditional integral cryptanalysis, the integral distinguisher was with probability 1. Knudsen once mentioned: "Integrals can be probabilistic just like differentials", but there was no further research afterwards. In this paper, we study deeply into this problem. Firstly, we introduce the method of Probabilistic Integral Cryptanalysis, and then we testify the validity of this method both in theory and experiment. For block cipher which contains S-boxes, it is proved that the closer the differential uniformity of S-boxes gets to random probability, the stronger the resistance of the block cipher to probabilistic integral cryptanalysis is. Finally, we point out that the techniques of higher order differential cryptanalysis do not work in probabilistic integral cryptanalysis, for the reason that the probability of integral characteristic regresses toward the random probability with increment of the number of sum variables. Furthermore, after the experimental probabilistic integral cryptanalysis of two block ciphers LBlock and AES, we come to a conclusion that most of existing modern block ciphers is immune to probabilistic integral cryptanalysis.

Keywords integral cryptanalysis; differential cryptanalysis; probabilistic integral cryptanalysis; lightweight block cipher; AES; LBlock

---

收稿日期:2012-05-11;最终修改稿收到日期:2012-07-18.本课题得到国家自然科学基金(61070171)、中国科学院战略性先导专项基金(XDA06010702)资助.**李晓千**,女,1987年生,博士研究生,目前主要从事分组密码分析方面的研究.E-mail:xqli@is.ac.cn.**吴文玲**,女,1966年生,博士,研究员,博士生导师,主要研究领域为分组密码、杂凑函数的设计与分析、分组密码的工作模式及可证明安全的密码方案.**李 宝**,男,1962年生,博士,研究员,博士生导师,主要研究领域为可证明安全方法的理论与应用、安全协议的设计与分析及椭圆曲线密码学.**于晓丽**,女,1986年生,博士研究生,目前主要从事分组密码分析方面的研究.

1 引 言

在过去的几十年里,对分组密码分析方法的研究有了很大的进展,其中应用最广泛的分析方法有差分密码分析[1]、线性密码分析[2]和积分密码分析[3]等. 1990 年,Biham 等密码学家提出了差分密码分析. 差分分析的基本思想是:通过分析两个中间状态的"差"经过几轮密码变换后的变化情况来恢复某些密钥比特. 与差分分析类似,积分分析的基本思想是:通过分析多个中间状态的"和"经过几轮密码变换后的变化情况来恢复某些密钥比特. 这两种分析方法的不同之处在于,差分分析关注的是两个中间状态的异或值,而积分分析关注的是多个中间状态的异或值以及这些状态之间的关系.

差分分析经扩展得到不可能差分分析[4],进一步地,不可能差分分析经扩展得到不太可能差分分析[5]. 差分分析利用概率高于随机概率的差分特征筛选正确密钥,即 $p>p_r$;不可能差分分析利用概率为 0 的差分特征排除错误密钥,即 $p=0$;而不太可能差分利用概率低于随机概率的差分特征排除错误密钥,即 $0<p<p_r$.

类似地,在传统的积分分析中,我们利用的积分特征都必须满足概率为 1,即 $p=1$. 比如对 Square[3]、AES[6]、Crypton[7]、Hierocrypt[8]、IDEA[9]、Skipjack[10]、Camellia[11]、MISTY1[6]、FOX[12]、SAFER+[13]、Twofish[14]等算法的积分攻击. 那么,能否构造概率不为 1 的有效概率积分特征,即 $p_r<p<1$ 或 $0\leq p<p_r$?在文献[6]中,Knudsen 等学者提到过:"就像差分一样,积分也可以是概率的",这点明了概率积分特征的存在性,但是没有任何文献做进一步的研究. 本文对此问题进行了探讨,提出了概率积分密码分析方法,并从理论和实验两方面验证了概率积分分析方法的有效性. 本文的主要结论有:对于采用 S 盒设计的分组密码,如果 S 盒的差分均匀性越接近随机概率,则分组密码抵抗概率积分密码分析的能力就越强;并且,随着求和变量个数的增加,积分特征概率趋近于随机概率,因此高阶积分分析的某些技巧对于概率积分分析是行不通的. 最后,通过对 AES 和 LBlock 这两个算法的概率积分分析实例,说明目前广泛使用的分组密码算法对于概率积分密码分析方法是免疫的.

本文第 2 节简要介绍积分密码分析中的符号定义;第 3 节提出采用 S 盒的分组密码抵抗概率积分分析的充分条件,并用实验证明,随着求和变量个数的增加,积分特征概率趋近于随机概率 p_r;第 4 节通过对 AES、LBlock 的概率积分密码分析实例,验证第 3 节的结论,说明目前广泛使用的分组密码算法对于概率积分密码分析方法是免疫的.

2 基本记号和定义

本文沿用文献[6]中的记号. 如非特别说明,下文的"集合"均指多重集,即集合中的元素可以重复,例如 $S=\{1\}$ 与 $T=\{1,1\}$ 表示两个不同的集合.

定义 1[10]. 在积分分析中,定义活跃集

$$A=\{(a_0,a_1,\cdots,a_{m-1})\in\mathbb{F}_{2^n}^m \mid a_i\neq a_j,\ \forall i\neq j\},$$

稳定集

$$C=\{(a_0,a_1,\cdots,a_{m-1})\in\mathbb{F}_{2^n}^m \mid \forall a_i=a_0,0\leq i\leq m-1\},$$

平衡集

$$B=\left\{(a_0,a_1,\cdots,a_{m-1})\in\mathbb{F}_{2^n}^m \,\middle|\, \sum_{i=0}^{m-1}a_i=0\right\},$$

在本文中,定义

$$V_x=\left\{(a_0,a_1,\cdots,a_{m-1})\in\mathbb{F}_{2^n}^m \,\middle|\, \sum_{i=0}^{m-1}a_i=x\right\},$$

特别的,$V_0=B$. 其中 m 称作求和变量个数.

性质 1[6].

(1) A/C 经过一个双射后,仍然是 A/C;

(2) A 的线性组合为 B;

(3) A 与 C 的对应元素异或求和仍然为 A,记作 $A\oplus C=A$;

(4) $V_x\oplus V_y=V_{x\oplus y}$.

由性质 1(1)得,A/C 经过 S 盒仍然是 A/C,但是 B 经过 S 盒后性质无法确定,如何确定 B 经过 S 盒后的性质便成为积分分析的瓶颈,概率积分分析便是针对这一问题引入的.

定义 2[6]. 几个集合的直积称作一个结构,例如

$$CCCCABBB=C\times C\times C\times C\times A\times B\times B\times B$$

就是一个结构.

与差分特征的定义类似,下面给出积分特征的定义.

定义 3. r-轮积分特征 Ω 是一个结构序列:$\alpha_0,\alpha_1,\cdots,\alpha_r$,其中 α_0 是明文满足的结构,$\alpha_i(1\leq i\leq r)$ 是第 i 轮输出满足的结构.

定义 4. r-轮积分特征 $\Omega=\alpha_0,\alpha_1,\cdots,\alpha_r$ 的概率是指,在满足输入积分特征的结构独立、均匀随机选取时,该结构经过 i 轮加密,满足第 i 轮$(1\leq i\leq r)$输出积分特征 α_i 的概率,记作 p^{Ω}.

定义 5[15]. 对于一个 $n\times n$ 的 S 盒，令

$$\delta_S=2^{-n}\max_{\alpha\in\mathbb{F}_{2^n}\setminus\{0\}}\max_{\beta\in\mathbb{F}_{2^n}}\#\{(a_0,a_1)\in\mathbb{F}_{2^n}^2\mid a_0\oplus a_1=\alpha,S(a_0)\oplus S(a_1)=\beta\},$$

称 δ_S 为 S 盒的差分均匀性.

定义 6. 对于一个 $n\times n$ 的 S 盒，记

$$Prob^m[\alpha\xrightarrow{S}\beta]=\frac{\#\left\{(a_0,a_1,\cdots,a_{m-1})\in\mathbb{F}_{2^n}^m\mid\sum_{i=0}^{m-1}a_i=\alpha,\sum_{i=0}^{m-1}S(a_i)=\beta\right\}}{\#\left\{(a_0,a_1,\cdots,a_{m-1})\in\mathbb{F}_{2^n}^m\mid\sum_{i=0}^{m-1}a_i=\alpha\right\}}$$

$$=2^{-n(m-1)}\#\left\{(a_0,a_1,\cdots,a_{m-1})\in\mathbb{F}_{2^n}^m\mid\sum_{i=0}^{m-1}a_i=\alpha,\sum_{i=0}^{m-1}S(a_i)=\beta\right\}$$

为 S 盒的 m 次积分特征概率，其中 $\alpha,\beta\in\mathbb{F}_{2^n}$.

如果积分特征 (α,β) 满足 $\forall(\alpha',\beta')\neq(\alpha,\beta)$，

$$|Prob^m[\alpha\xrightarrow{S}\beta]-2^{-n}|\geq|Prob^m[\alpha'\xrightarrow{S}\beta']-2^{-n}|,$$

令 $\delta_S^m=Prob^m[\alpha\xrightarrow{S}\beta]$，称 δ_S^m 为 S 盒的 m 次积分均匀性，即与随机概率差距最大的 m 次积分特征.

定义 7. 如果一个差分特征概率/积分特征概率为 p，并且一个满足该输入差分特征/输入积分特征的明文对/结构，经过一个随机置换后，仍然满足该特征概率 p，则称此概率为随机概率，记作 p_r.

3 抵抗概率积分密码分析

3.1 抵抗概率积分密码分析的充分条件

这一节我们给出采用 S 盒的分组密码抵抗概率积分密码分析的充分条件：S 盒的差分均匀性越接近随机概率 p_r，分组密码抵抗概率积分密码分析的能力就越强.

定理 1. 如果一个 $n\times n$ S 盒的任意非零差分特征概率均不大于 p，那么任意积分特征概率也不大于 $p+2^{-n}(1-p)$.

证明. 记 $Prob^m[\alpha\xrightarrow{S}\beta]$ 为一个 S 盒的积分特征 (α,β) 的概率，其中 m 是指求和变量个数. 由已知得：$m=2$ 时，$\forall\alpha,\beta\in\mathbb{F}_{2^n}\setminus\{0\}$，$Prob^2[\alpha\xrightarrow{S}\beta]\leq p$，其中 $Prob^2[\alpha\xrightarrow{S}\beta]$ 指的是 S 盒的差分特征概率.

下面证明求和变量个数 $m>2$ 时，

$$\forall\alpha,\beta\in\mathbb{F}_{2^n},\ Prob^m[\alpha\xrightarrow{S}\beta]\leq p+2^{-n}(1-p).$$

(1) $m=2$ 时，由已知得

$$\forall\alpha,\beta\in\mathbb{F}_{2^n}\setminus\{0\},$$

$$Prob^2_{\alpha,\beta\neq0}[\alpha\xrightarrow{S}\beta]=\frac{\#\{(a_0,a_1)\in\mathbb{F}_{2^n}^2\mid a_0\oplus a_1=\alpha,S(a_0)\oplus S(a_1)=\beta\}}{\#\{(a_0,a_1)\in\mathbb{F}_{2^n}^2\mid a_0\oplus a_1=\alpha\}}$$

$$=2^{-n}\cdot\#\{(a_0,a_1)\in\mathbb{F}_{2^n}^2\mid a_0\oplus a_1=\alpha,S(a_0)\oplus S(a_1)=\beta\}\leq p.$$

(2) $m=3$ 时，

$$\forall(\alpha,\beta)\in\mathbb{F}_{2^n}^2\setminus\{(\alpha,\beta)\in\mathbb{F}_{2^n}^2\mid\beta=S(\alpha)\},$$

$$Prob^3_{\beta\neq S(\alpha)}[\alpha\xrightarrow{S}\beta]=\frac{\#\{(a_0,a_1,a_2)\mid a_0\oplus a_1\oplus a_2=\alpha,S(a_0)\oplus S(a_1)\oplus S(a_2)=\beta\}}{\#\{(a_0,a_1,a_2)\in\mathbb{F}_{2^n}^3\mid a_0\oplus a_1\oplus a_2=\alpha\}}$$

$$=2^{-2n}\cdot\sum_{a_2=0}^{2^n-1}\#\{(a_0,a_1)\in\mathbb{F}_{2^n}^2\mid a_0\oplus a_1=\alpha\oplus a_2,S(a_0)\oplus S(a_1)=\beta\oplus S(a_2)\}$$

$$=2^{-2n}\cdot\sum_{a_2=0}^{2^n-1}2^n\cdot Prob^2_{\alpha\oplus a_2\neq0,\beta\oplus S(a_2)\neq0}[\alpha\oplus a_2\to\beta\oplus S(a_2)]$$

$$\leq2^{-2n}\cdot2^n\cdot2^n\cdot p=p<p+2^{-n}(1-p).$$

$$\forall(\alpha,\beta)\in\{(\alpha,\beta)\in\mathbb{F}_{2^n}^2\mid\beta=S(\alpha)\},$$

$$Prob^3_{\beta=S(\alpha)}[\alpha\xrightarrow{S}\beta]$$

$$=2^{-2n}\cdot\sum_{a_2=0}^{2^n-1}\#\{(a_0,a_1)\in\mathbb{F}_{2^n}^2\mid a_0\oplus a_1=\alpha\oplus a_2,S(a_0)\oplus S(a_1)=\beta\oplus S(a_2)\}$$

$$=2^{-2n}\cdot\Big(\sum_{a_2\neq\alpha}\#\{(a_0,a_1)\in\mathbb{F}_{2^n}^2\mid a_0\oplus a_1=\alpha\oplus a_2,S(a_0)\oplus S(a_1)=\beta\oplus S(a_2)\}+\#\{(a_0,a_1)\in\mathbb{F}_{2^n}^2\mid a_0\oplus a_1=0,S(a_0)\oplus S(a_1)=0\}\Big)$$

$$\leq2^{-2n}\cdot((2^n-1)\cdot2^n\cdot p+2^n)=p+2^{-n}(1-p).$$

因此，$m=3$ 时，$\forall\alpha,\beta\in\mathbb{F}_{2^n}$，$Prob^3[\alpha\xrightarrow{S}\beta]\leq p+2^{-n}(1-p)$ 成立.

(3) 假设 $m-1$ 时定理成立，即 $\forall\alpha,\beta\in\mathbb{F}_{2^n}$，$Prob^{m-1}[\alpha\xrightarrow{S}\beta]\leq p+2^{-n}(1-p)$，那么求和变量个数为 m 时，有

$$\forall\alpha,\beta\in\mathbb{F}_{2^n},$$

$$Prob^m[\alpha\xrightarrow{S}\beta]=\frac{\#\left\{(a_0,a_1,\cdots,a_{m-1})\in\mathbb{F}_{2^n}^m\mid\sum_{i=0}^{m-1}a_i=\alpha,\sum_{i=0}^{m-1}S(a_i)=\beta\right\}}{\#\left\{(a_0,a_1,\cdots,a_{m-1})\in\mathbb{F}_{2^n}^m\mid\sum_{i=0}^{m-1}a_i=\alpha\right\}}$$

$$=2^{-n(m-1)}\cdot\sum_{a_{m-1}=0}^{2^n-1}\#\left\{(a_0,a_1,\cdots,a_{m-2})\in\mathbb{F}_{2^n}^{m-1}\mid\sum_{i=0}^{m-1}a_i=\alpha\oplus a_{m-1},\sum_{i=0}^{m-1}S(a_i)=\beta\oplus S(a_{m-1})\right\}$$

$$=2^{-n(m-1)}\sum_{a_{m-1}=0}^{2^n-1}2^{n(m-2)}\cdot$$

$$Prob^{m-1}[\alpha\oplus a_{m-1}\xrightarrow{S}\beta\oplus S(a_{m-1})]$$

$$\leq 2^{-n(m-1)}\cdot 2^n\cdot 2^{n(m-2)}\cdot(p+2^{-n}(1-p))$$

$$=p+2^{-n}(1-p).$$

综上所述，如果一个 $n\times n$ S 盒的任意非零差分特征概率均不大于 p，那么任意积分特征概率也不大于 $p+2^{-n}(1-p)$. 证毕.

由定理 1 我们得到：如果一个 S 盒的差分均匀性 δ 与随机概率 $p_r=2^{-n}$ 很接近，那么 S 盒的任意积分均匀性也与随机概率很接近，该算法抵抗概率积分密码分析的能力也就越强.

这个证明也从一个侧面反应了概率积分分析和差分分析的联系. 差分分析关注的是两个中间状态的异或值，而概率积分分析关注的是多个中间状态的异或值.

因为 S 盒的 m 次积分特征概率是多个 $m-1$ 次积分特征概率的均值，因此当求和变量个数增大时，积分特征概率趋近于随机概率. 下面我们用实验验证这一观察.

3.2 S 盒积分特征概率随求和变量个数的变化

从 3.1 节的证明过程我们可以看出：当求和变量个数增大时，积分特征概率趋近于随机概率. 这一节分别以 AES 的 8×8 S 盒、LBlock 的 4×4 S 盒为例验证这一观察.

图 1 表示 AES 中 S 盒积分均匀性随求和变量个数的变化规律. 横坐标表示求和变量个数 m，纵坐标表示 m 次积分均匀性. 从图 1 可以看出：对于 AES 的 S 盒，当求和变量个数增大时，积分均匀性趋近于随机概率 p_r，在这里 $p_r=2^{-8}$. 实验数据显示，当 $m\geq 11$ 时，m 次积分均匀性与 2^{-8} 的差距小于 10^{-24}.

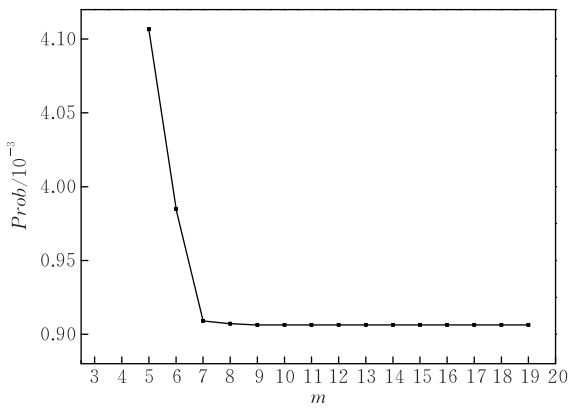

图 1 AES 中 S 盒的积分均匀性随求和变量个数的变化

图 2 表示 LBlock 中 S 盒积分均匀性随求和变量个数增长的变化规律. 图 2 只显示了 LBlock 加密算法中一个 S 盒 S5 的积分均匀性随求和变量个数增长的变化情况，其余 7 个 S 盒积分均匀性的变化情况类似. 从图 2 可以看出：对于 LBlock 的 S 盒 S5，当求和变量个数增大时，积分均匀性趋近于随机概率 p_r，在这里 $p_r=2^{-4}$. 实验数据显示，当 $m\geq 29$ 时，对于任意加密算法中的 S 盒 $S_i(0\leq i\leq 7)$，m 次积分均匀性与 2^{-4} 的差距小于 10^{-24}.

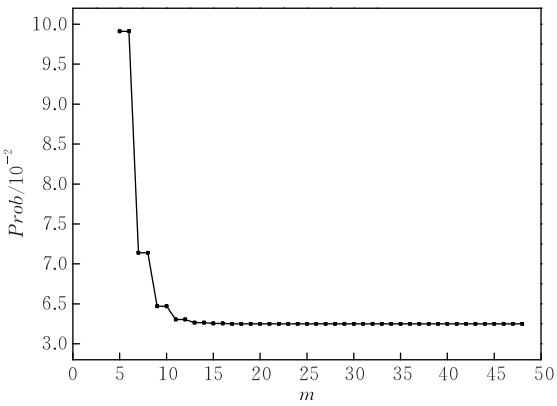

图 2 LBlock 中 S 盒的积分均匀性随求和变量个数的变化

由实验我们可以得到：对于 AES 和 LBlock 的 S 盒，当求和变量个数增大时，积分特征概率趋近于随机概率.

实际上，现在绝大多数采用 S 盒的分组密码算法，抵抗概率积分分析的能力都很强，因为：(1) 现代分组密码算法设计时，要求 S 盒必须具有良好的差分均匀性；(2) S 盒的 m 次积分特征概率是多个 $m-1$ 次积分特征概率的均值. 所以，如果 S 盒的均匀性越好，求和变量个数 m 越大，那么 S 盒的 m 次积分特征概率就越接近随机概率 p_r，利用此概率积分特征构造的区分器就越难把密码算法和随机置换区分开. 由于同样的原因，高阶积分分析的某些技巧对于概率积分分析是行不通的，这一点将在下一节的实例中详细阐述.

4 AES、LBlock 的概率积分密码分析

这一节通过对 AES、LBlock 进行概率积分密码分析，说明目前广泛使用的分组密码算法对于概率积分密码分析方法是免疫的.

4.1 计算积分特征概率

在差分密码分析中，我们假设每一轮子密钥统计独立且均匀分布；在概率积分密码分析中，对于积

分特征 $\Omega=\alpha_0,\alpha_1,\cdots,\alpha_r$，我们给出以下假设.

假设 1. 对于任意一个满足输入积分特征 α_0 的结构 T_0，经过 $i(i\geq 1)$ 轮加密得到第 $i+1$ 轮的输入结构 T_i，我们把这个积分特征中满足 α_i 的 T_i 组成的空间记作 $Space_i^{\Omega}(\alpha_i)$；另外，把任意满足 α_i 的结构组成的空间记作 $Space(\alpha_i)$. 假设条件概率

$$\underset{T_i,k_{i+1}}{Prob}[F_{k_{i+1}}(T_i)=\alpha_{i+1}\mid T_i\in Space_i^{\Omega}(\alpha_i)]=$$

$$\underset{T_i,k_{i+1}}{Prob}[F_{k_{i+1}}(T_i)=\alpha_{i+1}\mid T_i\in Space(\alpha_i)],$$

其中 k_{i+1} 表示第 $i+1$ 轮的子密钥，$F_{k_{i+1}}$ 表示第 $i+1$ 轮的轮函数.

这个假设比子密钥独立均匀的假设强. 在对任何一个算法进行概率积分密码分析时，都有必要验证假设 1 的正确性.

在假设 1 下，计算积分特征概率的问题可以归约到计算 S 盒的积分特征概率. 因为对于任意一个线性变换 $P:x\rightarrow P\cdot x$，给定输入积分特征 α，易得输出积分特征 $\beta=P\cdot\alpha$ 以概率 1 成立. 所以计算积分特征概率的问题可以归约到计算 S 盒的 m 次积分特征概率 $Prob^m[V_x\xrightarrow{S}V_y]$，$x,y\in\mathbb{F}_{2^n}$. 下面我们用归纳的方法给出一个计算 $Prob^m[V_x\xrightarrow{S}V_y]$ 的一般方法：构造表 $\{T_{x,y}^m\}$，$0\leq x,y\leq 2^n-1$，其中 m 指求和的变量个数.

(1) 令 $m=2$，计算表 $\{T_{x,y}^2\}(0\leq x,y\leq 2^n-1)$：$T_{x,y}^2=2^{-n}\cdot\#\{(a_0,a_1)\mid a_0\oplus a_1=x,S(a_0)\oplus(a_1)=y,a_i\in\mathbb{F}_{2^n},0\leq i\leq 1\}$，实际上这就是 S 盒的差分分布表.

(2) 假设已经计算得到求和变量个数为 $m-1$ 的表 $\{T_{x,y}^{m-1}\}(0\leq x,y\leq 2^n-1)$：

$$T_{x,y}^{m-1}=2^{-n(m-1)}\cdot\#\left\{(a_0,a_1,\cdots,a_{m-2})\,\middle|\,\sum_{i=0}^{m-2}a_i=x,\ \sum_{i=0}^{m-2}S(a_i)=y\right\}.$$

那么由 $\{T_{x,y}^{m-1}\}$ 可以得到 $\{T_{x,y}^m\}$：

$$\begin{aligned}T_{x,y}^m&=2^{-nm}\cdot\#\left\{(a_0,a_1,\cdots,a_{m-1})\,\middle|\,\sum_{i=0}^{m-1}a_i=x,\sum_{i=0}^{m-1}S(a_i)=y\right\}\\&=2^{-nm}\sum_{a_{m-1}=0}^{2^n-1}\#\left\{(a_0,a_1,\cdots,a_{m-2})\,\middle|\,\sum_{i=0}^{m-2}a_i=x\oplus a_{m-1},\sum_{i=0}^{m-2}S(a_i)=y\oplus S(a_{m-1})\right\}\\&=2^{-nm}\cdot 2^{n(m-1)}\sum_{a_{m-1}=0}^{2^n-1}T_{x\oplus a_{m-1},y\oplus S(a_{m-1})}^{m-1}\\&=2^{-n}\sum_{i=0}^{2^n-1}T_{x\oplus i,y\oplus S(i)}^{m-1}.\end{aligned}$$

4.2 对 AES 的概率积分密码分析

在这一节里，我们先在假设 1 的条件下对 AES 进行概率积分密码分析，然后通过实验验证假设 1 的正确性. AES 是 SP 结构的分组密码，明文分块为 128 bit、192 bit 和 256 bit. 本文只考虑对 AES-128 的概率积分分析.

我们将利用文献[6]中概率为 1 的积分特征. 在这个概率为 1 的积分特征中，输入积分特征由 1 个活跃字节 A 和 15 个稳定字节 C 组成；经过第 1 轮轮函数的作用，变为 1 列 4 个活跃字节 A 和 3 列 12 个稳定字节 C；经过第 2 轮轮函数的作用，全部变成 16 个活跃字节 A；经过第 3 轮轮函数的作用，全部变成 16 个稳定字节 B；经过第 4 轮子密钥异或，得到 16 个稳定字节 B. 接下来我们考虑在概率为 1 积分特征的基础上，向后加一轮，再利用高阶积分分析的技巧向前加几轮，构成概率 $p<1$ 的积分特征.

先考虑向后加一轮. 经过第 4 轮 S 盒的变换，设其中一个字节的积分特征为 V_x，令 $p_x=Prob[B\xrightarrow{S}V_x]$，易知 $p^{\Omega}=p$. 在假设 1 的条件下，由 3.1 节的实验数据得：当 $m\geq 11$ 时，AES 的 S 盒 m 次积分均匀性与 2^{-8} 的差距小于 10^{-24}. 因为在假设 1 的条件下，

$$\begin{aligned}p_x&=Prob[B\xrightarrow{S}V_x]\\&=2^{-8\cdot(2^8-1)}\cdot\#\left\{(a_0,a_1,\cdots,a_{2^8-1})\,\middle|\,\sum_{i=0}^{2^8-1}a_i=0,\sum_{i=0}^{2^8-1}S(a_i)=x\right\},\end{aligned}$$

又由积分均匀性的定义得 $|p_x-2^{-8}|\leq|\delta_S^{2^8}-2^{-8}|$，这里求和变量个数 $m=2^8>11$，所以我们有 $|p_x-2^{-8}|\leq|\delta_S^{2^8}-2^{-8}|<10^{-18}$，即 $p_x\approx 2^{-8}$. 利用这样的概率积分特征构成的区分器无效.

再考虑使用文献[6]中的方法，利用高阶积分分析的技巧向前加轮数——通过增加输入积分特征中活跃字节的个数，增加积分特征的轮数. 但是这个方法在概率积分分析中行不通，因为增加输入积分特征中活跃字节的个数 k，必然导致求和变量个数 $m=2^{8\cdot k}$ 增加. 在假设 1 的条件下，随着 m 的增大，积分特征概率 p 趋近于随机概率 $p_r=2^{-8}$. 实验数据显示，对 AES 的 S 盒，当 $m\geq 11$ 时，m 次积分均匀性与 2^{-8} 的差距小于 10^{-24}. 当输入积分特征中活跃字节个数 $k\geq 1$ 时，求和变量个数 $m\geq 2^8$，m 次积分均匀性几乎与 2^{-4} 相等，利用这样的概率积分特征构造的区分器无效.

我们通过实验验证了假设 1 在 AES 中的正确性. 随机选取 1000 个密钥、1000 个满足输入积分特征的结构,用 AES 加密. 对于每一个结构,在经过第 4 轮第一个 S 盒后,对这个结构中所有中间状态值求和得到 x. 我们对所有的 x 做统计,得到的实验结果与我们在假设 1 的条件下得到的理论结果吻合,证明了假设 1 对 AES 算法的正确性.

综上所述,AES 对概率积分密码分析是免疫的.

4.3 对 LBlock 的概率积分密码分析

在这一节里,我们对 LBlock 进行概率积分密码分析,在这之前,我们先验证假设 1 的正确性. LBlock 是 Feistel 结构的轻量级分组密码,由 32 轮迭代算法组成,明文分块为 64 bit,密钥为 80 bit. 本文沿用 LBlock 设计文档[16]中的记号.

4.3.1 验证假设 1

对于任意一个给定的 LBlock 的积分特征,我们先在假设 1 的条件下给出相应的概率分布,再在实际情况下给出相应的概率分布,然后用统计方法检验这两个分布的拟合程度. 这里只详细阐述对一个积分特征验证假设 1 的过程,其余积分特征的验证过程类似.

在概率为 1 的 9 轮积分特征后加 1 轮,构成一个 10 轮 LBlock 概率积分特征. 图 3(a)与 LBlock 设计文档中给出的 9 轮积分特征稍有不同. 图 3(b)是本文给出的 10 轮积分特征中的 9～10 轮,1～8 轮与图 3(a)的 1～8 轮积分特征一样. 在这个 10 轮 LBlock 积分特征中,1～8 轮的概率为 1. 在第 9 轮中, S_5 的输入积分特征为 B,令 S_5 的输出积分特征为 V_x,则第 10 轮的输出积分特征中有一个半字节为 V_x,这就构成了一个 10 轮概率积分特征. 我们记 $p_x=Prob[B\xrightarrow{S_5}V_x]$,易知这个 10 轮积分特征的概率为 p_x. 下面考虑在没有假设的条件下积分特征概率 p_x^* 的分布.

第 i 轮的输出为 (X_{i+1},X_i),其中 $X_i=X_{i,7}\parallel X_{i,6}\parallel\cdots\parallel X_{i,1}\parallel X_{i,0}$. 第 i 轮轮函数 F 的输出为 U_i, $U_i=U_{i,7}\parallel U_{i,6}\parallel\cdots\parallel U_{i,1}\parallel U_{i,0}$. 由加密算法得

$$U_{9,4}=S_5(X_{9,5}\oplus K_{9,5}),$$

$$X_{9,5}=U_{8,5}\oplus X_{7,3},$$

$$U_{8,5}=S_7(X_{8,7}\oplus K_{8,7})=S_7(S_7(S_7(X_{0,5}\oplus W_0)\oplus W_1)\oplus W_2),$$

$$X_{7,3}=S_2(S_0(S_1(X_{0,5}\oplus W_3)\oplus W_4)\oplus W_5)\oplus W_6,$$

其中,当密钥固定时, $W_i(0\leq i\leq 6)$是明文 X_i 的函数:

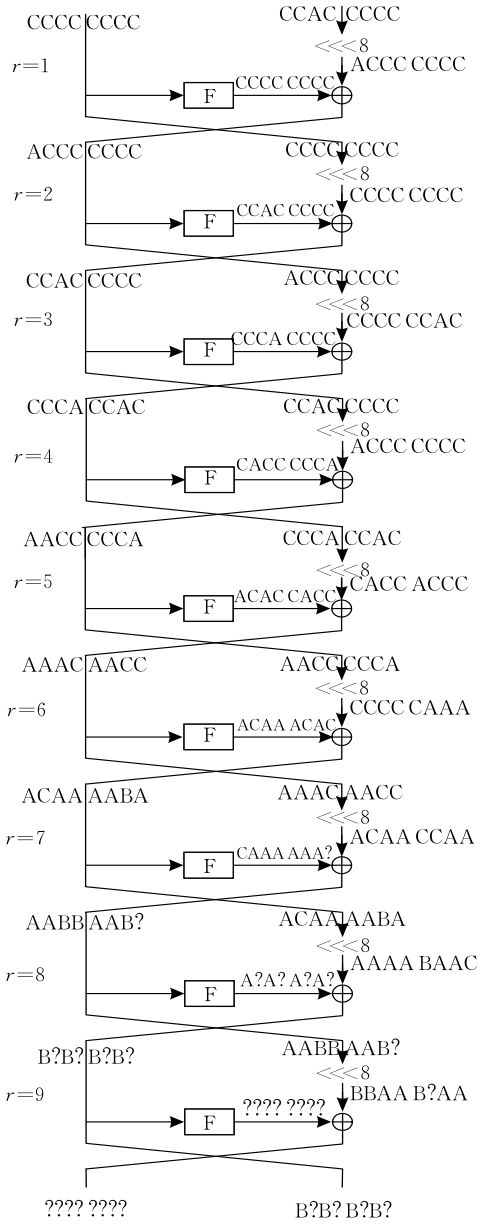

图 3 10 轮 LBlock 概率积分特征

$$W_0=W_0(X_{1,6}),$$
$$W_1=W_1(X_{0,2},X_{0,3},X_{1,2},X_{1,3},X_{1,5},X_{1,7}),$$
$$W_2=W_2(X_{0,0},X_{0,1},X_{0,2},X_{0,3},X_{0,4},X_{0,6},X_{0,7},$$
$$X_{1,0},X_{1,1},X_{1,2},X_{1,3},X_{1,4},X_{1,5},X_{1,6},X_{1,7}),$$
$$W_3=W_3(X_{0,0},X_{1,0},X_{1,1},X_{1,6}),$$
$$W_4=W_4(X_{0,2},X_{1,4},X_{1,5}),$$
$$W_5=W_5(X_{0,1},X_{0,4},X_{1,2},X_{1,4},X_{1,7}),$$
$$W_6=W_6(X_{0,4},X_{0,6},X_{0,7},X_{1,0},X_{1,1},X_{1,3},$$
$$X_{1,4},X_{1,5}).$$

容易证明，当明文独立、均匀随机选取时，对任意固定密钥，$W_i(0\leq i\leq 6)$都是均匀随机独立的变量，因此有

$$U_{9,4}=U_{9,4}(X_{0,5},W_0,W_1,\cdots,W_6).$$

根据积分特征概率的定义，

$$p_x^*=Prob[B\xrightarrow{S_5}V_x]$$
$$=2^{(-4)\times 7}\cdot\#\left\{(W_0,W_1,\cdots,W_6)\in\mathbb{F}_{2^4}^7\mid \sum_{X_{0,5}\in\mathbb{F}_{2^4}}U_{9,4}(X_{0,5},W_0,W_1\cdots W_6)=x\right\}.$$

把 p_x 的分布与 p_x^* 的分布做比较，得到图 4. 为了检测 p 和 p^* 的拟合优度测试得

$$\chi^2(p;p^*)=2^4\sum_{x\in\mathbb{F}_{2^4}}\frac{(p_x-p_x^*)^2}{p_x}=2.70\times10^{-4}.$$

从测试结果我们可以看出，在假设 1 的条件下积分特征概率的分布 p 和没有任何假设条件下的分布 p^* 十分接近.

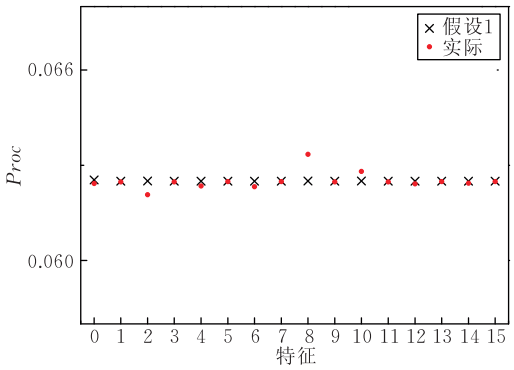

图 4　假设 1 与实际情况比较

我们对 LBlock 的其它积分特征也做了检验，结果都验证了假设 1 的正确性. 因此假设 1 对 LBlock 算法是合理的.

4.3.2　对 16 轮 LBlock 的概率积分密码分析

在这一节里，我们在概率为 1 的积分特征的基础上，向后加两轮，再利用高阶积分分析的技巧向前加几轮，构成概率 $p<1$ 的积分特征；然后利用这个积分特征构造区分器，对 16 轮 LBlock 做恢复密钥攻击；最后把概率积分密码分析的复杂度与传统积分分析作比较，证明 LBlock 对概率积分密码分析是免疫的.

先向后加两轮，见图 5. 图 5 是本文给出的 11 轮积分特征中的 8～11 轮，1～7 轮与图 3(a)的 1～7 轮积分特征一样. 这个积分特征是我们能找到的概率最大的积分特征，其概率 $p^a=p^3+(1-p^3)p_r\approx 2^{-3.9992\times 3}+2^{-4}$，其中 p^3 表示一个满足输入积分特征的结构加密时，满足每一轮输出积分特征的概率；$(1-p^3)p_r$ 表示一个满足输入积分特征的结构加密时，不满足所有中间轮输出积分特征的概率、满足最后一轮输出积分特征的概率.

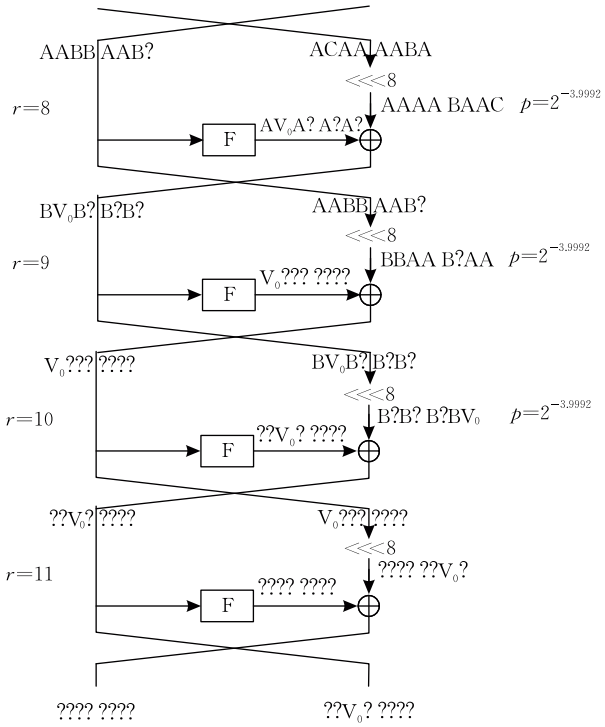

图 5　11 轮 LBlock 积分特征(8～11 轮)

利用图 5 的 11 轮积分特征，我们构造 11 轮概率积分区分器，对 16 轮 LBlock 进行恢复密钥攻击. 与文献[16]中的方法完全一样，我们猜测 12～16 轮中的 12 个半字节密钥，对每一个猜测密钥值建立一个计数器. 随机选择 2^4 个明文构成一个满足输入积分特征的结构，经过 16 轮加密，得到相应的 2^4 个密文. 用每一个猜测密钥值对 2^4 个密文进行部分解密，得到第 11 轮的输出值 $X_{11,5}$，检验 2^4 个 $X_{11,5}$ 的异或之和是否为 0，如果为 0，则相应密钥的计数器加 1. 再随机选一个满足输入积分特征的结构，重复以上步骤，直到某个密钥的计数远远大于其余密钥为止，这个密钥就是我们要找的正确密钥. 这样就恢复了 48 bit 密钥值，剩下的比特可以通过穷搜的方法得到.

对于概率为 $2^{-3.9992\times 3}+2^{-4}$ 的积分特征，用文

献[17]计算复杂度的方法，攻击成功至少需要 $2^{26.3}$ 个结构，成功概率是 0.99. 因此对 16 轮 LBlock 的概率积分密码分析的时间复杂度是 $2^{30.3}$，需要存储 2^{48} 个密钥对应的计数器，其中 1 个单位时间复杂度是指运行 16 轮 LBlock 加密算法一次需要的时间.

再考虑利用高阶积分分析的技巧向前加轮数，和在 AES 中一样，这个方法在概率积分分析中行不通，因为向前加轮数必然会使得输入积分特征中活跃字节的个数 k 增加，进而导致求和变量个数 $m=2^{4\cdot k}$ 增加. 在假设 1 的条件下，随着 m 的增大，积分特征概率 p 趋近于随机概率 $p_r=2^{-4}$. 实验数据显示，对 LBlock 的 S 盒 $S_i(0\leq i\leq 7)$，当 $m\geq 29$ 时，m 次积分均匀性与 2^{-4} 的差距小于 10^{-24}. 当输入积分特征中活跃字节个数 $k\geq 2$ 时，求和变量个数 $m\geq 2^8$，m 次积分均匀性几乎与 2^{-4} 相等，利用这样的概率积分特征构成的区分器无效.

因此概率积分密码分析对 LBlock 能攻击的最大轮数是 16，时间复杂度为 $2^{30.3}$；而传统积分密码分析对 LBlock 能攻击的最大轮数是 22. 由此看来，概率积分密码分析和传统积分密码分析相比，攻击能力较弱.

目前广泛使用的分组密码算法，在设计 S 盒时，都要求 S 盒必须具有良好的差分均匀性，并且在算法的安全性评估中，都考虑了算法抵抗传统积分密码分析的强度. 因此我们在概率为 1 积分特征的基础上，构造的积分特征的概率和随机概率很接近. 利用这样的概率积分特征构造的区分器，能够攻击的轮数也小于传统积分特征. 总的来说，目前广泛使用的分组密码算法对于概率积分密码分析方法是免疫的.

5 结 论

本文提出了一种新的分组密码攻击方法——概率积分密码分析. 通过理论和实验两方面的研究，验证了概率积分分析方法的有效性. 对于采用 S 盒设计的分组密码，本文的研究从一个侧面反映了概率积分密码分析与差分密码分析的关系：如果 S 盒的差分均匀性越接近随机概率，则分组密码抵抗概率积分密码分析的能力就越强. 并且，本文指出：高阶积分分析的某些技巧对于概率积分分析是行不通的，主要原因是随着求和变量个数的增加，积分特征概率趋近于随机概率. 通过对 AES、LBlock 的概率积分分析，说明目前广泛使用的分组密码算法对于概率积分密码分析方法是免疫的.

参 考 文 献

[1] Biham E, Shamir A. Differential cryptanalysis of DES-like cryptosystems. Journal of Cryptology, 1991, 4: 3-72

[2] Matsui M. Linear cryptanalysis method for DES cipher//Proceedings of the Advances in Cryptology—EUROCRYPT'93 Workshop on the Theory and Application of Cryptographic Techniques. Lofthus, Norway, 1993. Berlin: Springer Verlag, 1994: 386-397

[3] Daemen J, Knudsen L, Rijmen V. The block cipher Square//Proceedings of the Fast Software Encryption 4th International Workshop, FSE'97. Haifa, Israel, 1997. Berlin: Springer Verla, 1997: 149-165

[4] Biham E, Biryukov A, Shamir A. Cryptanalysis of Skipjack reduced to 31 rounds using impossible differentials//Proceedings of the Advances in Cryptology-EUROCRYPT'99 International Conference on the Theory and Application of Cryptographic Techniques. Prague, Czech Republic, 1999. Berlin: Springer Verlag, 1999: 12-23

[5] Tezcan C. The improbable differential attack: Cryptanalysisof reduced round CLEFIA//Proceedings of the Progress in Cryptology-INDOCRYPT 2010 11th International Conference on Cryptology in India. Hyderabad, India, 2010. Berlin: Springer Verlag, 2010: 197-209

[6] Knudsen L, Wagner D. Integral cryptanalysis (extended abstract)//Proceedings of the Fast Software Encryption 9th International Workshop, FSE 2002. Leuven, Belgium, 2002. Berlin: Springer Verlag, 2002: 629-632

[7] Halluin C D, Bijnens G, Rijmen V, Preneel B. Attack on six rounds of Crypton//Proceedings of the Fast Software Encryption 6th International Workshop, FSE'99. Rome, Italy, 1999. Berlin: Springer Verlag, 1999: 46-59

[8] Barreto P S L M, Rijmen V, Nakahara J, Preneel B, Vandewalle J, Kim H Y. Improved Square attacks against reduced-round hierocrypt//Proceedings of the Fast Software Encryption 8th International Workshop, FSE 2001. Yokohama, Japan, 2001. Berlin: Springer Verlag, 2002: 61-80

[9] Demirci H. Square-like attacks on reduced rounds of IDEA//Proceedings of the Selected Areas in Cryptography 9th Annual International Workshop, SAC 2002, St. John's. Newfoundland, Canada, 2002. Berlin: Springer Verlag, 2003: 147-159

[10] Hwang Kyungdeok, Lee Wonil, Lee Sungjae, Lee Sangjin, Lim Jongin. Saturation attacks on reduced round Skipjack//Proceedings of the Fast Software Encryption 9th International Workshop, FSE 2002. Leuven, Belgium, 2002. Berlin: Springer Verlag, 2002: 15-23

[11] Yeom Y, Park S, Kim I. On the security of CAMELLIA against the Square attack//Proceedings of the Fast Software Encryption 9th International Workshop, FSE 2002. Leuven, Belgium, 2002. Berlin: Springer Verlag, 2002: 189-99

[12] Wu W, Zhang W, Feng D. Integral cryptanalysis of reduced FOX block cipher//Proceedings of the Information Security

and Cryptology-ICISC 2005 8th International Conference. Seoul, Korea, 2005. Berlin: Springer Verlag, 2006: 229-241

[13] Hu Y. Integral cryptanalysis of SAFER+. Electronics Letters, 1999, 35: 1458-1459

[14] Lucks S. The saturation attack — A bait for Twofish//Proceedings of the Fast Software Encryption 8th International Workshop, FSE 2001. Yokohama, Japan, 2001. Berlin: Springer Verlag, 2002: 187-205

[15] Wu Wen-Ling, Feng Deng-Guo, Zhang Wen-Tao. The Design and Cryptanalysis of Block Ciphers. 2nd Edition. Beijing: Tsinghua University Press, 2009(in Chinese)
(吴文玲，冯登国，张文涛. 分组密码的设计与分析. 第 2 版. 北京：清华大学出版社，2009)

[16] Wu W, Zhang L. LBlock: A lightweight block cipher//Proceedings of the Applied Cryptography and Network Security 9th International Conference, ACNS 2011. Nerja, Spain, 2011. Berlin: Springer Verla, 2011: 327-344

[17] Blondeau C, Gérard B. On the data complexity of statistical attacks against block ciphers//Kholosha A, Rosnes E, Parker M G eds. Workshop on Coding and Cryptography-WCC 2009. Berlin: Springer Verlag, 2009: 469-488

[18] Knudsen L R. Truncated and higher order differentials//Proceedings of the Fast Software Encryption 2nd International Workshop. Leuven, Belgium, 1994. Berlin: Springer Verlag, 1995: 196-211

[19] Knudsen L R, Berson T A. Truncated differentials of SAFER//Proceedings of the Fast Software Encryption 3rd International Workshop. Cambridge, UK, 1996. Berlin: Springer Verlag, 1996: 15-26

[20] Langford S K, Hellman M E. Differential-linear cryptanalysis//Proceedings of the Advances in Cryptology-CRYPTO'94 14th Annual International Cryptology Conference. Santa Barbara, California, USA, 1994. Berlin: Springer Verlag, 1994: 17-25

[21] Biham E, Dunkelman O, Keller N. Enhancing differential-linear cryptanalysis//Proceedings of the Advances in Cryptology-ASIACRYPT 2002 8th International Conference on the Theory and Application of Cryptology and Information Security. Queenstown, New Zealand, 2002. Berlin: Springer Verlag, 2002: 587-592

[22] Wu Wen-Ling, Fan Wei-Jie, Zhang Lei. Advances in lightweight block ciphers//Proceedings of the Chinese Association for Cryptologic Research. Beijing: Publishing House of Electronics Industry, 2011: 140-159(in Chinese)
(吴文玲，范伟杰，张蕾. 轻量级密码//中国密码学发展报告 2010. 中国，北京，2010. 北京：电子工业出版社，2011：140-159)

LI Xiao-Qian, born in 1987, Ph. D. candidate. Her research interest is cryptanalysis of block cipher.

WU Wen-Ling, born in 1966, Ph. D., researcher, Ph. D. supervisor. Her main research interests include design and cryptanalysis of block ciphers and Hash functions, modes of operation for block ciphers, and the theory of provable security.

LI Bao, born in 1962, Ph. D., researcher, Ph. D. supervisor. His main research interests include the theory and application of provable security, design and analysis of cryptographic protocols, and elliptic curve cryptography.

YU Xiao-Li, born in 1986, Ph. D. candidate. Her research interest is cryptanalysis of block cipher.

Background

Integral cryptanalysis is a chosen-plaintext attack, which was first used to attack block cipher Square, so integral cryptanalysis is also called Square attack. In 2002, Knudsen and Wagner first proposed the definition of integral cryptanalysis and used it to attack block cipher Rijndeal. In 2008, Muhammad et al. presented bit-pattern based integral attack and used it to attack Noekeon, Serpent and PRESENT. So far, the integral attack applied to many kinds of block ciphers, such as IDEA, Camellia, MISTY1 and so on. Integrals have many interesting features. They are especially well-suited in analyzing ciphers designed with primarily bijective components. Moreover, they exploit the simultaneous relationship between many encryptions, in contrast to differential cryptanalysis where only pairs of encryptions are considered. Consequently, integrals apply to a lot of ciphers which are not vulnerable to differential and linear cryptanalysis. These features have made integral an increasingly popular tool in recent cryptanalysis work.

In traditional integral cryptanalysis, the integral distinguisher was with probability 1. Knudsen once mentioned: "Integrals can be probabilistic just like differentials", but there was no further research afterwards.

In this paper, we study deeply enough into this problem. We introduce the method of Probabilistic Integral Cryptanalysis, and testify the validity of this method both in theory and experiment. For block cipher which contains S-boxes, we prove that the closer the differential uniformity of S-boxes gets to random probability, the stronger the resistance of the block cipher to probabilistic integral cryptanalysis is. We also prove that the techniques of higher order differential cryptanalysis do not work in probabilistic integral cryptanalysis. Furthermore, we examine the resistance of LBlock and AES against probabilistic integral cryptanalysis, and come to a conclusion that most of existing modern block ciphers is immune to probabilistic integral cryptanalysis.