差分分析中的特征概率计算问题研究 [1]

李 贞 吕述望 王永传 王安胜

(中国科学技术大学研究生院 信息安全国家重点实验室 北京 100039)

摘 要 该文指出了长期以来在差分密码分析中所采用的差分特征概率计算方法与差分分析基本原理不相符合的矛盾，对这一问题进行了深入研究，给出了二者等价的充分条件，力图解决差分分析方法的理论基础问题.

关键词 差分密码分析，特征概率计算，等价条件

中图号 TN918

1 引 言

1991 年， Eli Biham 和 Adi Shamir 提出了对 DES(Data Encryption Standard) 类密码的差分密码分析方法 [1]，从此差分密码分析成为迄今为止对迭代型分组密码进行攻击和安全评估的最有力手段之一。然而从差分密码分析方法诞生之日起，就一直存在着实际攻击方法与基本原理不相符合、相互脱节的问题，至今，人们只是使用该方法，而对于该方法背后存在的极不可靠的理论基础问题一直没有正面提出和给予研究。虽然对几个算法的实际攻击结果表明该方法与实际结果拟合得较好 [2]，但这是个别现象还是普遍现象仍然令人产生极大的疑问。本文立足于解决实际差分分析中存在的这一疑问，深入研究对于差分分析可用的差分特征概率计算过程，给出了一个计算等价充分条件，力图回答差分分析的理论基础问题。

2 差分密码分析方法简介

一个迭代型分组密码往往设计成对一个密码学弱的轮函数进行多次迭代的结构。对一个分组长度为 n 的 r 轮迭代密码，将两个 n 比特串 Y_i 和 Y_i^* 的差分定义为

$$\Delta Y_i = Y_i \otimes Y_i^{*-1}$$

其中 $\otimes$ 表示 n 比特串集上的一个特定群运算， Y_i^{*-1} 表示 Y_i^* 在此群中的逆元。

定义 1 r 轮特征 Ω 是一个差分序列 $a_0, a_1, \cdots, a_r$，其中 a_0 是明文对 Y_0 和 Y_0^* 的差分，$a_i (1 \le i \le r)$ 是第 i 轮输出 Y_i 和 Y_i^* 的差分。 R 轮特征 $\Omega = a_0, a_1, \cdots, a_r$ 的概率 p^Ω 是指在子密钥 $K_1, K_2, \cdots, K_r$ 独立、均匀随机时，明文对 Y_0 和 Y_0^* 的差分为 a_0 的条件下，第 i 轮 $(1 \le i \le r)$ 输出 Y_i 和 Y_i^* 的差分为 a_i 的概率。

可以证明以下定理：

定理 1 在密钥独立随机假设下，若分组密码轮函数的结构形式如图 1：

则当差分运算定义为 $\otimes$ 时 (即定义的差分运算与密钥结合算法所采用的群运算一致时)，R 轮特征 $\Omega = a_0, a_1, \cdots, a_r$ 的特征概率 p^Ω 是 r 个单轮特征概率的乘积 $p^\Omega = \prod_{i=1}^r p_i^\Omega$，其中 $p_i^\Omega = P(a_{i-1} \to a_i) = P(\Delta F(Y_{i-1}) = a_i | \Delta Y_{i-1} = a_{i-1})$。

可见，在定理 1 条件下 $a_{i-1} \to a_i$ 的概率与本轮的输入无关，即和前面轮的作用无关，序列 $a_0, a_1, \cdots, a_r$ 形成一条马尔科夫链，因此， p^Ω 可由 $\prod_{i=1}^r p_i^\Omega$ 来计算。

[1] 2002-01-26 收到， 2002-08-29 改回
国家 973 项目资助 (No:G1999035808)

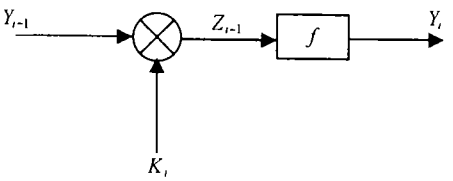

图 1　密码轮函数的结构

上述结构具有广泛的代表性. 就目前现有的迭代型分组密码而言，很多密码在进行差分分析时，可以上述结构作为基本差分分析模型进行研究.

算法 1[3]　差分攻击的基本算法如下：

第 1 步　定义一个差分运算，找出一个 $(r-1)$ 轮特征 $\Omega_{r-1}=a_0,a_1,\cdots,a_{r-1}$，使得它的概率尽量大；

第 2 步　均匀随机地选择明文 Y_0 并计算 Y_0^*，使得 Y_0 和 Y_0^* 的差分为 a_0，找出 Y_0 和 Y_0^* 在实际密钥加密下所得的密文 Y_r 和 Y_r^*. 若最后一轮的子密钥 K_r(或 K_r 的部分比特) 有 2^m 个可能值 $K_r^j(1\le j\le 2^m)$，设置相应的 2^m 个计数器 $A^j(1\le j\le 2^m)$；由 (a_{r-1},Y_r,Y_r^*) 确定可能的密钥值，并对相应的计数器加 1；

第 3 步　重复第 2 步，直到一个或几个计数器的值明显高于其他计数器的值，输出它们所对应的子密钥 (或部分比特).

攻击复杂度 [4]：若攻击中所利用的 $(r-1)$ 轮特征 Ω_{r-1} 的概率为 p^Ω，则攻击复杂度 $\mathrm{Comp}(r)\ge 2/p^\Omega-[1/(2^n-1)]$.

3 差分密码分析中的概率计算问题

差分分析利用的是密码体制的高概率差分特征，特征概率的计算是实现差分攻击的基础. 然而，根据差分分析的基本原理以及实现差分攻击的具体过程 (算法 1)，差分分析中所使用的特征概率应该在明文随机、密钥固定的条件下计算，但根据定义 1 所确定的特征定义以及在定理 1 保证下所确定的计算特征概率的方法，研究的对象都是在明文固定、密钥独立随机假设下的特征概率，二者概率计算的基础根本不同，它们的计算结果总是可以互相代替的吗？或者在什么条件下二者计算结果一致或近似？这正是差分分析方法中所存留的问题. 以下将详细对比两种概率计算的过程，对这一问题进行研究. 为叙述方便，以下将两种概率分别称为随机明文特征概率和随机密钥特征概率并以轮函数具有图 1 结构的密码为研究模型且假定 f 是置换. 当需要区分随机变量和确定量时，用大写字母表示随机变量，用小写字母表示一个确定的值.

3.1 随机明文特征概率的计算

第 1 轮

输入：(Y_0,Y_0^*), $\Delta Y_0=Y_0\otimes Y_0^{*-1}=a_0$，用 S_0 表示随机变量 Y_0 的样本空间，用 S_0^* 表示随机变量 Y_0^* 的样本空间，用 $(S_0\times S_0^*)a_0$ 表示差为 a_0 的有序对 (Y_0,Y_0^*) 的样本空间，则在第 1 轮有 $(Y_0,Y_0^*)\in(S_0\times S_0^*)a_0=(Z_2^n\times Z_2^n)a_0$；

子密钥：$k_1\in Z_2^n$

输出：(Y_1,Y_1^*)

求输出 $Y_1\otimes Y_1^{*-1}=a_1$ 的概率.

求解过程：

易证，此时 f 函数的输入是所有可能的 2^n 个差为 a_0 的对．记该轮 f 函数所有可能输入的集合为 $f^{1,I'}$，则 $f^{1,I'}=(Z_2^n\times Z_2^n)a_0$．记在 $f^{1,I'}$ 集合中那些使得输出差为 a_1 的对形成的子集合为 $f^{1,I'}(a_0,a_1)$，这一子集合在 f 函数下的输出对的集合为 $f^{1,O'}(a_0,a_1)$，则 $|f^{1,O'}(a_0,a_1)|=|f^{1,I'}(a_0,a_1)|$，且集合 $f^{1,O'}(a_0,a_1)$ 中的每一对数据的差为 a_1．记在随机明文下 $a_0\to a_1$ 的转移概率为 $P(a_0\to a_1)'$，则有

$$P(a_0\to a_1)'=\frac{|f^{1,I'}(a_0,a_1)|}{|(Z_2^n\times Z_2^n)a_0|}=\frac{|f^{1,I'}(a_0,a_1)|}{2^n}\ .$$

第 2 轮

输入：$(Y_1,Y_1^*),\ Y_1\otimes Y_1^{*-1}=a_1,\ \ (Y_1,Y_1^*)\in f^{1,O'}(a_0,a_1)$；

子密钥：$k_2\in Z_2^n$

输出：(Y_2,Y_2^*)

求输出 $Y_2\otimes Y_2^{*-1}=a_2$ 的概率。

求解过程：

此时 f 函数所有可能输入的集合 $f^{2,I'}=f^{1,O'}(a_0,a_1)\otimes k_2$(注：$f^{1,O'}(a_0,a_1)\otimes k_2$ 表示对集合 $f^{1,O'}(a_0,a_1)$ 中的每一数据 (y,y^*) 进行 $\otimes k_2$ 运算：$(y,y^*)\otimes k_2=(y\otimes k_2,y^*\otimes k_2)$)，且该集合中每对数据差均为 a_1．由于该集合的元素与 a_0,a_1,k_2 有关，将 $f^{2,I'}$ 更显式地记为 $f^{2,I'}_{[a_0,a_1,k_2]}$．在 f 函数的控制下，记 $f^{2,I'}_{[a_0,a_1,k_2]}$ 中那些使得输出差为 a_2 的对形成的子集合为 $f^{2,I'}_{[a_0,a_1,k_2]}(a_1,a_2)$，这一子集合在 f 函数的作用下得到的输出对的集合为 $f^{2,O'}_{[a_0,a_1,k_2]}(a_1,a_2)$，则 $|f^{2,O'}_{[a_0,a_1,k_2]}(a_1,a_2)|=|f^{2,I'}_{[a_0,a_1,k_2]}(a_1,a_2)|$，且集合 $f^{2,O'}_{[a_0,a_1,k_2]}(a_1,a_2)$ 中每一对数据的差为 a_2．这样，$P(a_1\to a_2)'=\dfrac{|f^{2,I'}_{[a_0,a_1,k_2]}(a_1,a_2)|}{|f^{2,I'}_{[a_0,a_1,k_2]}|}$，与 a_0,a_1,a_2,k_2 有关。

以下各轮关于差分转移概率的计算过程同第 2 轮。可见，在随机明文、固定密钥条件下特征 $a_0,a_1,\cdots,a_{r-1}$ 的概率不仅与 $a_0,a_1,\cdots,a_{r-1}$ 的取值有关，而且与 $k_2,k_3,\cdots,k_{r-1}$ 的值有关。

3.2 随机密钥特征概率的计算

第 1 轮

输入：$(y_0,y_0^*),\ \ \Delta y_0=y_0\otimes y_0^{*-1}=a_0$；

子密钥：$K_1\in Z_2^n$

输出：(Y_1,Y_1^*)

求输出 $Y_1\otimes Y_1^{*-1}=a_1$ 的概率。

求解过程：

易证，此时 f 函数的输入是所有可能的 2^n 个差为 a_0 的对．记该轮 f 函数所有可能输入的集合为 $f^{1,I}$，则 $f^{1,I}=(Z_2^n\times Z_2^n)a_0$．在 f 函数的控制下，记那些使得输出差为 a_1 的对形成的子集合为 $f^{1,I}(a_0,a_1)$，这一子集合在 f 函数的作用下得到的输出对的集合为 $f^{1,O}(a_0,a_1)$，则 $|f^{1,O}(a_0,a_1)|=|f^{1,I}(a_0,a_1)|$，且集合 $f^{1,O}(a_0,a_1)$ 中的每一对数据的差为 a_1，出现的概率均等．这样，$P(a_0\to a_1)=\dfrac{|f^{1,I}(a_0,a_1)|}{|(Z_2^n\times Z_2^n)a_0|}=\dfrac{|f^{1,I}(a_0,a_1)|}{|2^n|}$．

第 2 轮

输入：$(Y_1,Y_1^*),\ \ Y_1\otimes Y_1^{*-1}=a_1,\ \ (Y_1,Y_1^*)\in f^{1,O}(a_0,a_1)$；

子密钥：$K_2\in Z_2^n$

输出：(Y_2,Y_2^*)

求输出 $Y_2\otimes Y_2^{*-1}=a_2$ 的概率。

求解过程：

此时 f 函数所有可能输入的集合 $f^{2,I}=\bigcup_{k_2} f^{1,O}(a_0,a_1)\otimes k_2$, $k_2\in Z_2^n$,$=\bigcup_s s\otimes Z_2^n$, $s\in f^{1,O}(a_0,a_1)$,$=\bigcup_s(Z_2^n\times Z_2^n)a_1$, $s\in f^{1,O}(a_0,a_1)$,$=(Z_2^n\times Z_2^n)a_1$ 。记 $f^{2,I}$ 中那些使得输出差为 a_2 的对形成的子集合为 $f^{2,I}(a_1,a_2)$，这一子集合在 f 函数的作用下得到的输出对的集合为 $f^{2,O}(a_1,a_2)$，则 $|f^{2,O}(a_1,a_2)|=|f^{2,I}(a_1,a_2)|$，且集合 $f^{2,O}(a_1,a_2)$ 中的每一对数据的差为 a_2 。这样，$P(a_1\to a_2)=\frac{|f^{2,I}(a_1,a_2)|}{|f^{2,I}|}=\frac{|f^{2,I}(a_1,a_2)|}{|(Z_2^n\times Z_2^n)a_1|}=\frac{|f^{2,I}(a_1,a_2)|}{2^n}$ 只与 a_1,a_2 有关。

以下各轮关于差分转移概率的计算过程同第 2 轮。可见，在随机密钥、固定明文条件下特征 $a_0,a_1,\cdots,a_{r-1}$ 的概率仅与 $a_0,a_1,\cdots,a_{r-1}$ 的取值有关，且形成一条马尔科夫链。

那么，在什么条件下随机明文特征概率可以由随机密钥特征概率代替计算呢？以下对这一问题进行讨论，给出一个充分条件。

3.3 特征概率计算等价充分条件

符号：$x\in_R X$：表示从集合 X 中随机均匀地选取 x；

定义 2　对于 X 的一个子集 X'，若对于任意一个统计测试 $T: X\to\{0,1\}$，都有 $|P(T[x]=1)-P(T[x']=1)|<\varepsilon$，这里 $x\in_R X$, $x'\in_R X'$, ε 表示足够小的数，则称 X' 是 X 的一个随机子集。

该定义的含义是指一个集合的随机子集与该集合具有完全相似的统计特性。

可以证明随机子集的如下性质：

性质 1　若 B 是 A 的随机子集，f 是一个双射，则 $f(B)$ 是 $f(A)$ 的随机子集。

性质 2　若 B 是 A 的随机子集，f 是一个双射，R 是一个规则，若 $f(A)$ 中符合规则 R 的元素组成的集合 $Rf(A)$ 是集合 C 的随机子集，$Rf(A)$ 的原像集 $f^{-1}(Rf(A))$ 是 A 的随机子集，则 $f(B)$ 中符合规则 R 的元素组成的集合 $Rf(B)$ 是集合 C 的随机子集。

定义 3　设 f 是 $Z_2^n\to Z_2^n$ 上的一个置换，若在 f 函数的控制下，在 $(Z_2^n\times Z_2^n)a_{i-1}$ 集合中，使得输出差为 a_i 的子集合记为 $X^{i-1,I}(a_{i-1},a_i)$，该子集合的输出集合记为 $X^{i-1,O}(a_{i-1},a_i)$，则若 $X^{i-1,I}(a_{i-1},a_i)$ 是 $(Z_2^n\times Z_2^n)a_{i-1}$ 的一个随机子集，$X^{i-1,O}(a_{i-1},a_i)$ 是 $(Z_2^n\times Z_2^n)a_i$ 的一个随机子集，就称 f 对于差分 $a_{i-1}\to a_i$ 是一个随机映射。

定义 4　对于 r 轮特征 $\Omega=a_0,a_1,\cdots,a_r$，若 f 函数对于差分 $a_{i-1}\to a_i(1\le i\le r)$ 均为随机映射，则称 f 函数对于特征 Ω 是随机映射。

定理 2　(特征概率计算等价充分条件)　f 函数对于特征 $\Omega=a_0,a_1,\cdots,a_r$ 是随机映射，则该特征的随机明文特征概率在足够小的误差范围内近似等于随机密钥特征概率。

证明　对于第 1 轮，有 $f^{1,I'}(a_0,a_1)=f^{1,I}(a_0,a_1)$，因此 $P(a_0\to a_1)'=P(a_0\to a_1)$

对于第 2 轮，由于 f 对于 $a_0\to a_1$ 是随机映射，因此 $X^{1,O}(a_0,a_1)$ 是 $(Z_2^n\times Z_2^n)a_1$ 的一个随机子集，这样，$f^{2,I'}=f^{2,I'}_{[a_0,a_1,k_2]}=f^{1,O'}(a_0,a_1)\otimes k_2=X^{1,O}(a_0,a_1)\otimes k_2$ 。

由性质 1 得它也是 $(Z_2^n\times Z_2^n)a_1$ 的一个随机子集且对任意 k_2 都成立。

根据定义 2，对于任意一个统计测试 $T:(Z_2^n\times Z_2^n)a_1\to\{0,1\}$，都有 $|P(T[(z,z^*)]=1)-P(T[(s,s^*)]=1)|<\varepsilon$，这里 $(z,z^*)\in_R(Z_2^n\times Z_2^n)a_1$, $(s,s^*)\in_R f^{2,I'}$，现在定义一个测试 T_f 为：若一对差为 a_1 的输入 (y,y^*)，在 f 函数下的输出差为 a_2，则 $T_f(y,y^*)=1$；否则，$T_f(y,y^*)=0$ 。

由此得到 $|P(T_f[(z,z^*)]=1)-P(T_f[(s,s^*)]=1)|<\varepsilon$，其中 $(z,z^*)\in_R(Z_2^n\times Z_2^n)a_1$, $(s,s^*)\in_R f^{2,I'}$，而

$$P(T_f[(z,z^*)]=1)=\frac{|X^{2,I}(a_1,a_2)|}{|(Z_2^n\times Z_2^n)a_1|}=\frac{|f^{2,I}(a_1,a_2)|}{|(Z_2^n\times Z_2^n)a_1|}=P(a_1\to a_2),$$

$$P(T_f[(s,s^*)]=1)=\frac{|f^{2,I'}_{[a_0,a_1,k_2]}(a_1,a_2)|}{|f^{2,I'}_{[a_0,a_1,k_2]}|}=P(a_1\to a_2)',\text{且对于任意 } k_2 \text{ 都成立。}$$

由此得到 $|P(a_1 \to a_2)' - P(a_1 \to a_2)| < \varepsilon$, 且 $P(a_1 \to a_2)'$ 与 k_2 无关.

对于第 3 轮, 有 f 是双射, 定义规则 R 为两数据具有差 a_2 , 则有 $(Z_2^n \times Z_2^n)a_1 \xrightarrow{Rf} X^{2,O}(a_1, a_2)$, $f_{[a_0,a_1,k_2]}^{2,I'} \xrightarrow{Rf} f_{[a_0,a_1,k_2]}^{2,O'}(a_1, a_2)$.

由于 $(Z_2^n \times Z_2^n)a_1$ 在 Rf 下的像集 $X^{2,O}(a_1, a_2)$ 是集合 $(Z_2^n \times Z_2^n)a_2$ 的随机子集, $X^{2,O}(a_1, a_2)$ 在 Rf 下的原像集 $X^{2,I}(a_1, a_2)$ 是 $(Z_2^n \times Z_2^n)a_1$ 的随机子集, 而 $f_{[a_0,a_1,k_2]}^{2,I'}$ 是 $(Z_2^n \times Z_2^n)a_1$ 的随机子集, 根据性质 2 , $f_{[a_0,a_1,k_2]}^{2,O'}(a_1, a_2)$ 是集合 $(Z_2^n \times Z_2^n)a_2$ 的随机子集. 因此 $f^{3,I'} = f_{[a_0,a_1,k_2]}^{2,O'}(a_1, a_2) \otimes k_3$ 是 $(Z_2^n \times Z_2^n)a_2$ 的随机子集, 且对于任意 k_3 都成立. 以下证明完全类同于第 2 轮, 可以得到 $|P(a_2 \to a_3)' - P(a_2 \to a_3)| < \varepsilon$, 且 $P(a_2 \to a_3)'$ 与 k_2, k_3 无关.

其余各轮的推证同第 3 轮。 证毕

上述证明针对 f 是置换时的情况, 其实从证明过程看, 当 f 具有这样的性质: 每一个像所具有的原像个数都相同, 则上述结论仍然成立. 同时注意这样一个问题: 上述证明基于随机集合和随机映射的概念, 这些概念建立在统计学性质上, 必须有量的保证. 因此, 定理 2 给出的充分条件适用于那些具有高概率的差分特征, 而差分攻击中要寻找的恰恰是这些高概率差分特征.

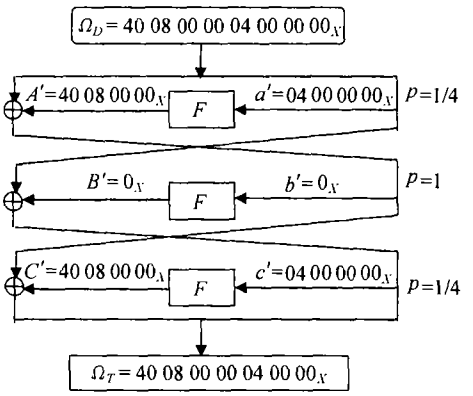

图 2 攻击 6 圈 DES 利用的高概率特征
注: 图中数据均表示差分值

由定理 2 , 还可以得到这样一个结论: 当满足定理中给出的充分条件时, 不需要进行密钥独立随机假设, 就可以将序列 $a_0, a_1, \cdots, a_r$ 看做马尔科夫链, 用 $\prod_{i=1}^{r} p_i^{\Omega}$ 来近似计算随机明文特征概率. 对于一个设计良好的密码, f 函数对于高概率特征的随机映射特性应该是可以得到一定程度的满足的, 如考察文献 [2] 中攻击 DES 所利用的高概率特征, 以攻击缩减为 6 圈的 DES 所利用的高概率特征为例, 攻击中利用了图 2 这样一种特征:

对于这一特征, 显然只要 F 对于差分 $a' \to A'$ 是随机映射, 则该特征满足定理 2 的计算等价充分条件. 使用文献 [5] 中给出的随机性检测方法, 对那些满足 $a' \to A'$ 的输入对集合和输出对集合进行随机性检测, 它们的随机特性与它们所在的全集 $(Z_2^n \times Z_2^n)a'$ 和 $(Z_2^n \times Z_2^n)A'$ 基本没有差别. 对于文献 [2] 中所利用的其他高概率特征可作类似的分析. 这说明利用目前给出的随机性检测方法, F 对于这些高概率差分的随机映射特性是好的. 这也许是在进行实际的差分攻击中, 在子密钥根本不满足独立随机假设的条件下, 用 $\prod_{i=1}^{r} p_i^{\Omega}$ 来计算高概率特征仍与实际拟合得较好的原因.

4 结束语

本文指出了长期以来存在于差分密码分析中的实际攻击方法与基本原理不相符合的矛盾, 研究了二者之间的关系, 给出了二者相互一致的条件, 使长期流于经验学的差分密码分析方法

在理论上得以解释和有所依靠。这对于指导实际的差分分析工作将十分有意义。然而本文给出的是一个充分条件，如何对该条件进一步弱化使之更加精确将是我们今后继续研究的方向。

参 考 文 献

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

[2] E. Biham, A. Shamir, Differential Cryptanalysis of the Data Encryption Standard, Berlin, Springer-Verlag 1993, Chapter 1–7.

[3] 冯登国，密码分析学，北京，清华大学出版社，广西科学技术出版社，2000, 15–32.

[4] Xuejia Lai, J. L. Massey, S. Murphy, Markov ciphers and differential cryptanalysis, Advances in cryptology-EUROCRYPT'91, Berlin, Springer-Verlag, 1992, 17–38.

[5] A. Rukhin, J. Soto, J. Nechvatal, A statistical test suite for random and pseudorandom number generators for cryptographic applications, http://csrc.nist.gov/cryptval/,2001, Chapter 1–2.

STUDY ON THE COMPUTATION OF DIFFERENTIAL CHARACTERISTIC PROBABILITY IN DIFFERENTIAL CRYPTANALYSIS

Li Zhen　Lü Shuwang　Wang Yongchuan　Wang Ansheng

(*The State Key Laboratory of Information Security, Graduate School at Beijing, University of Science & Technology of China, Beijing* 100039, *China*)

Abstract This paper points out the contradiction between the way of computing the differential characteristic probability and the principle of differential cryptanalysis, gives a deep reaserch on it and proposes a equivalent sufficient condition. The goal of the work is trying to solve the problem of the basic theoretic foundation for differential cryptanalysis.

Key words Differential cryptanalysis, Computation of differential characteristic probability, Equivalent conditions

李　贞：　女，1970 年生，硕士生，主要研究方向为密码算法分析、信号分析与处理.

吕述望：　男，1941 年生，教授，博士生导师，主要研究方向为网络与信息安全、密码学理论与密码算法设计与实现等.

王永传：　男，1969 年生，博士后，主要研究方向为网络与信息安全、密码算法设计与分析.

王安胜：　男，1973 年生，硕士生，主要研究方向为网络与信息安全、信号分析与处理.