Keccak 类非线性变换的差分性质研究

李倩男[1]，李云强[1]，蒋淑静[2]，路遥[3]

（1. 信息工程大学 电子技术学院，河南 郑州 450004；2. 中国科学院 光电研究院，北京 100094；
3. 国防科学技术大学 计算机学院，湖南 长沙 410073）

摘 要：通过对 Keccak 中非线性环节的分析，提出了 n 元 Keccak 类非线性变换模型，研究了这类变换的差分性质。证明了对于 n 元 Keccak 类非线性变换，差分转移概率关于循环移位是不变的，当输入差确定时其非零差分转移概率都相等，给出了其差分转移概率不等于 0 和 1 时的取值范围；通过研究输出差的差分布尔函数表达式，得到了非平凡最大差分转移概率和非零最小差分转移概率的差分结构，给出了一种相邻变元 Keccak 类非线性变换间的差分传递概率相关性。

关键词：Keccak；Keccak 类非线性变换；杂凑算法；差分分析

中图分类号：TN918.1 **文献标识码**：A **文章编号**：1000-436X(2012)09-0140-07

Research on differential properties of Keccak-like nonlinear transform

LI Qian-nan[1], LI Yun-qiang[1], JIANG Shu-jing[2], LU Yao[3]

(1. Institute of Electronic Technology, the Information Engineering University, Zhengzhou 450004, China;
2. Academy of Opto-Electronics, Chinese Academy of Sciences, Beijing 100094, China;
3. College of Computer, National University of Defense Technology, Changsha 410073, China)

Abstract: By analyzing the nonlinear transformation in Keccak, n-gram Keccak-like nonlinear transform model was proposed firstly, and the differential properties of this kind of transformation was studied. They are proved that to the n-gram Keccak-like nonlinear transform, the differential transition probability about cycle shift transform is unchanged, and nonzero differential transition probabilities are equal when the input difference was confirmed. The range of the differential transition probability was calculated when it wasn't 0 and 1. By analyzing the Boolean expressions of the output difference, the differential construction for largest nontrivial and smallest nonzero differential transition probability were obtained. At the end, one correlation between the differential probabilities of adjacent variable in n-gram Keccak-like nonlinear transform was given.

Key words: Keccak; Keccak-like nonlinear transform; hash algorithm; differential analysis

1 引言

2010 年 12 月 10 日，美国国家标准与技术协会（NIST）公布进入 SHA3 竞选算法[1~3]最后一轮的 5 个杂凑算法[4]，由 Bertoni G、Daemen J 和 Peeters M 等人设计的 Keccak 杂凑算法就是其中之一[5]。Keccak 杂凑算法蕴涵着杂凑算法的最新设计理念[6,7]，对其编码环节的系统分析具有重要的理论和应用价值。c 变换是 Keccak 压缩函数中唯一的非线性环节，也是文献[8~10]的算法所选用的非线性环节，文献[11]中通过对 Keccak 的非线性环节进行改造，构造了新的非线性变换 MiniKeccak。密码算法抵抗差分分析的强度也一直是人们关注的问题，对密码算法中的重要环节

收稿日期：2011-11-02；修回日期：2012-03-30

进行差分分析有助于分析整个密码算法的抗差分分析攻击的强度，也有助于对算法中的编码环节进行更深刻的认识和把握。为了更好地应用 Keccack 类非线性变换编码环节，本文将对其差分性质进行系统研究。

2　基本概念

定义 1　设 $X,Y\in Z_2^n$，$X=(x_0,x_1,\cdots,x_{n-1}),Y=(y_0,y_1,\cdots,y_{n-1})$，如果函数 $Y=f(X)$ 满足：

$$y_i=x_i\oplus\overline{x_{(i+1)\bmod n}}\,x_{(i+2)\bmod n},\quad 0\le i<n,\quad n\ge 3,$$

则称 $Y=f(X)$ 为 n 元 Keccak 类非线性变换。

显然，Keccak 压缩函数中 c 变换是 5 元 Keccak 类非线性变换，MiniKeccak 中的非线性变换是 3 元 Keccak 类非线性变换。

定义 2　设 $(X,+)$，$(Y,+)$ 是有限交换群，$f:X\to Y$，$a\in X,b\in Y$，

$$p_f(a\to b)=\frac{1}{|X|}\#\{x\in X:f(x+a)-f(x)=b\}$$

则称 $p_f(a\to b)$ 为 f 在输入差为 a 条件下，输出差为 b 的差分转移概率，并称 $a\to b$ 为 f 的一个差分对应，称 $p_f(a\to b)$ 为该差分对应的转移概率。其中，$|\bullet|$ 和 $\#\{\bullet\}$ 均表示集合 $\bullet$ 中的元素个数。

定义 3　设 $X=(x_0,x_1,\cdots,x_{n-1})\in \mathrm{GF}^n(2)$ 为 n 维布尔向量，称 $X=(x_0,x_1,\cdots,x_{n-1})$ 中的不为零的分量的个数为 X 的汉明重量，记做 $W_{\mathrm{H}}(X)$。

3　差分性质分析

定理 1　对于 n 元 Keccak 类非线性变换 f，如果 $a,a',b,b'\in Z_2^n$，且 $b=f(x\oplus a)\oplus f(x)$，$<a',b'>\in\{<a,b>|a=(a<<j)\text{且}b=(b<<j),\ 0\le j<n,\ j\in Z\}$，那么差分转移概率 $p_f(a'\to b')=p_f(a\to b)$。

证明　当 $j=1$，即 $a'=(a<<1)$ 时，输入、输入差和输出差从高位比特到低位比特依次可以表示为：$x=(x_0,x_1,\cdots,x_{n-1})$，$a=(a_0,a_1,\cdots,a_{n-1})$，$b=(b_0,b_1,\cdots,b_{n-1})$，$x'=(x_0',x_1',\cdots,x_{n-1}')$，$a'=(a_1,a_2,\cdots,a_{n-1},a_0)$，$b'=(b_0',b_1',\cdots,b_{n-1}')$，其中，$x_i,a_i,b_i,b_i',x_i'\in\{0,1\}$，$i\in\{0,1,\cdots,n-1\}$。那么输出差 b 和 b' 的每一个比特的差分变换布尔函数表达式可以写为

$$\begin{aligned}b_i&=(x_i\oplus(x_{(i+1)\bmod n}\times x_{(i+2)\bmod n})\oplus x_{(i+2)\bmod n})\oplus\\&((x_i\oplus a_i)\oplus((x_{(i+1)\bmod n}\oplus a_{(i+1)\bmod n})\times\\&(x_{(i+1)\bmod n}\oplus a_{(i+2)\bmod n}))\oplus(x_{(i+2)\bmod n}\oplus a_{(i+2)\bmod n}))\\&=a_i\oplus a_{(i+2)\bmod n}x_{(i+1)\bmod n}\oplus a_{(i+1)\bmod n}x_{(i+2)\bmod n}\oplus\\&a_{(i+1)\bmod n}a_{(i+2)\bmod n}\oplus a_{(i+2)\bmod n};\end{aligned}$$

$$b_i'=a_{(i+1)\bmod n}\oplus a_{(i+3)\bmod n}x'_{(i+2)\bmod n}\oplus a_{(i+2)\bmod n}x'_{(i+3)\bmod n}\oplus a_{(i+2)\bmod n}a_{(i+3)\bmod n}\oplus a_{(i+3)\bmod n}$$

记 $L_b^a=\{x|f(x\oplus a)\oplus f(x)=b\}$，$L_{b'}^{a'}=\{x'|f(x'\oplus a')\oplus f(x')=b'\}$。

由 b 和 b' 的每一个比特的差分变换布尔函数表达式可知：当 $x\in L_b^a$，$x'=(x<<1)=(x_1,x_2,\cdots,x_{n-1},x_0)$ 时：

$$\begin{aligned}b_i'&=a_{(i+1)\bmod n}\oplus a_{(i+3)\bmod n}x_{(i+2)\bmod n}\oplus\\&a_{(i+2)\bmod n}x_{(i+3)\bmod n}\oplus a_{(i+2)\bmod n}a_{(i+3)\bmod n}\oplus a_{(i+3)\bmod n}\\&=b_{(i+1)\bmod n}\end{aligned}$$

可知，当 $a'=(a<<1)$，$x'=(x<<1)$ 时，有 $b'=(b<<1)$，$f((x<<1)\oplus(a<<1))\oplus f((x<<1))=(b<<1)$，即 $f(x'\oplus a')\oplus f(x')=b'$。所以，$x'=(x<<1)\in L_{b'}^{a'}$。

记 $x'=f(x):x'=(x<<1)$，其中，$x\in L_b^a$，$x'\in L_{b'}^{a'}$。下面说明 $x'=f(x)$ 为一一映射。

当 $x,y\in L_b^a$，$x\ne y$ 时，$x'=f(x)$，$y'=f(y)$，$x',y'\in L_{b'}^{a'}$，且 $x'\ne y'$。所以 $x'=f(x)$ 为单射。

当 $x'\in L_{b'}^{a'}$，$x\in L_b^a$ 时，存在 $x=j(x'):x=(x'>>1)$，使 $x=(x'_{n-1},x'_0,x'_1,\cdots,x'_{n-2})$，

$$b_i=a_i\oplus a_{(i+2)\bmod n}x_{(i+1)\bmod n}\oplus a_{(i+1)\bmod n}x_{(i+2)\bmod n}\oplus a_{(i+1)\bmod n}a_{(i+2)\bmod n}\oplus a_{(i+2)\bmod n}=b'_{(i-1)\bmod n}$$

可知，当 $x=(x'>>1)$ 时，有 $f((x'>>1)\oplus a)\oplus f((x'>>1)=(b'<<1)=b$，即 $f(x\oplus a)\oplus f(x)=b$。所以，$x=(x'>>1)\in L_b^a$。

综上，当 $x'\in L_{b'}^{a'}$，$x\in L_b^a$ 时，$x'=f(x)$ 为一一映射。所以 $|L_{b'}^{a'}|=|L_b^a|$。

$$p_f(a'\to b')=\frac{|L_{b'}^{a'}|}{2^n}=\frac{|L_b^a|}{2^n}=p_f(a\to b)$$，所以 $p_f(a'\to b')=p_f(a\to b)$。

假设 $j=m-1,(1<m<n,m\in Z)$，$p_f(a'\to b')=p_f(a\to b)$ 成立。$j=m$ 时的情况就是在 $j=m-1$

的基础上继续向左循环移动 1 位，由 $j=1$ 时的结论，$p_f(a'\to b')=p_f(a\to b)$ 也成立。所以，由数学归纳法可知定理 1 成立。

定理 2 对于 Keccak 类非线性变换 f，如果 $a,b,g\in Z_2^n$，$p_f(a\to b)\neq 0$，$p_f(a\to g)\neq 0$，那么差分转移概率 $p_f(a\to b)=p_f(a\to g)$。

证明 记 $L_b^a=\{x\mid f(x\oplus a)\oplus f(x)=b\}$，当 $p_f(a\to b)\neq 0$ 时，分析 L_b^a 中元素个数 $|L_b^a|$。将输入、输入差和输出差从高位比特到低位比特依次可以表示为：$x=(x_0,x_1,\cdots,x_{n-1})$，$a=(a_0,a_1,\cdots,a_{n-1})$，$b=(b_0,b_1,\cdots,b_{n-1})$，其中，$x_i,a_i,b_i\in\{0,1\}$，$i\in\{0,1,\cdots,n-1\}$。$b$ 的每一个比特的差分变换布尔函数表达式为

$$b_i=a_i\oplus a_{(i+2)\bmod n}x_{(i+1)\bmod n}\oplus a_{(i+1)\bmod n}x_{(i+2)\bmod n}\oplus a_{(i+1)\bmod n}a_{(i+2)\bmod n}\oplus a_{(i+2)\bmod n}$$；

$$b_i\oplus a_i\oplus a_{(i+1)\bmod n}a_{(i+2)\bmod n}\oplus a_{(i+2)\bmod n}=a_{(i+2)\bmod n}x_{(i+1)\bmod n}\oplus a_{(i+1)\bmod n}x_{(i+2)\bmod n}$$。

令 $y_i=b_i\oplus a_i\oplus a_{(i+1)\bmod n}a_{(i+2)\bmod n}\oplus a_{(i+2)\bmod n}$，$i\in\{0,1,\cdots,n-1\}$，$y_i\in\{0,1\}$。

则 $y_i=a_{(i+2)\bmod n}x_{(i+1)\bmod n}\oplus a_{(i+1)\bmod n}x_{(i+2)\bmod n}$，$i\in\{0,1,\cdots,n-1\}$。又因为 $x_i,a_i,y_i\in\{0,1\}$，所以当 a_i、b_i 确定，$i\in\{0,1,\cdots,n-1\}$ 时，$y_i=a_{(i+2)\bmod n}x_{(i+1)\bmod n}\oplus a_{(i+1)\bmod n}x_{(i+2)\bmod n}$ 为 F_2 上的非齐次线性方程组。显然该方程组的解 $x=(x_0,x_1,\cdots,x_{n-1})\in L_b^a$，$L_b^a$ 中的元素也满足该方程组。$y_i=a_{(i+2)\bmod n}x_{(i+1)\bmod n}\oplus a_{(i+1)\bmod n}x_{(i+2)\bmod n}$，$i\in\{0,1,\cdots,n-1\}$，可以表示为矩阵形式：

$$\begin{pmatrix}y_0\\y_1\\\vdots\\y_{n-3}\\y_{n-2}\\y_{n-1}\end{pmatrix}=\begin{pmatrix}0&a_2&a_1&0&\cdots&0\\0&0&a_3&a_2&0&\cdots\\&&\cdots&\cdots&&\\0&\cdots&\cdots&0&a_{n-1}&a_{n-2}\\a_{n-1}&0&\cdots&\cdots&0&a_0\\a_1&a_0&0&\cdots&\cdots&0\end{pmatrix}\begin{pmatrix}x_0\\x_1\\\vdots\\x_{n-3}\\x_{n-2}\\x_{n-1}\end{pmatrix},$$

$$\boldsymbol{A}=\begin{pmatrix}0&a_2&a_1&0&\cdots&0\\0&0&a_3&a_2&0&\cdots\\&&\cdots&\cdots&&\\0&\cdots&\cdots&0&a_{n-1}&a_{n-2}\\a_{n-1}&0&\cdots&\cdots&0&a_0\\a_1&a_0&0&\cdots&\cdots&0\end{pmatrix}$$ 为系数矩阵，设系数矩阵 $\boldsymbol{A}$ 的秩 $R(\boldsymbol{A})=r$，X_0 为一个特解，那么所有的解可以表示为

$$X=X_0\oplus k_1h_1\oplus k_2h_2\oplus\cdots\oplus k_{n-r}h_{n-r}$$

其中，$k_l\in\{0,1\}$，$l\in\{1,2,\cdots,n-r\}$；

$$\begin{cases}h_1=(c_{11},\cdots,c_{1r},1,0,\cdots,0),\\h_2=(c_{21},\cdots,c_{2r},0,1,\cdots,0),\\\cdots\\h_{n-r}=(c_{n-r,1},\cdots,c_{n-r,r},0,0,\cdots,1).\end{cases}$$，$c_{m,n}\in\{0,1\}$，其中，$m\in\{1,2,\cdots,n-r\}$，$n\in\{1,2,\cdots,r\}$。$h_1,h_2,\cdots,h_{n-r}$ 为线性无关的基础解系。

由于 $k_l\in\{0,1\}$，$l\in\{1,2,\cdots,n-r\}$，所以方程组解的个数为 2^{n-r} 个，$|L_b^a|=2^{n-r}$。

记 $L_g^a=\{x\mid f(x\oplus a)\oplus f(x)=g\}$，由于输入差 a 确定，$p_f(a\to g)\neq 0$；同理可以求得，$|L_g^a|=2^{n-r}$。

又 $p_f(a\to b)=\dfrac{|L_b^a|}{2^n}=\dfrac{|L_g^a|}{2^n}=p_f(a\to g)$，所以定理 2 成立。

定理 3 对于 n $(n\geq 3)$ 元 Keccak 类非线性变换 f，如果输入差 a，输出差 b，使 $b=f(x\oplus a)\oplus f(x)$ 成立，且 $p_f(a\to b)\neq 0$ 和 1，那么 $\dfrac{1}{2^{n-1}}\leq p_f(a\to b)\leq\dfrac{1}{4}$。

证明 按照定理 2 的证明方法，$p_f(a\to b)\neq 0$ 时，将输入、输入差和输出差从高位比特到低位比特依次可以表示为：$x=(x_0,x_1,\cdots,x_{n-1})$，$a=(a_0,a_1,\cdots,a_{n-1})$，$b=(b_0,b_1,\cdots,b_{n-1})$，令 $y_i=b_i\oplus a_i\oplus a_{(i+1)\bmod n}a_{(i+2)\bmod n}\oplus a_{(i+2)\bmod n}$，其中，$x_i,a_i,b_i,y_i\in\{0,1\}$，$i\in\{0,1,\cdots,n-1\}$。$L_b^a=\{x\mid f(x\oplus a)\oplus f(x)=b\}$，$L_b^a$ 中元素个数 $|L_b^a|$ 就是 $x_i,y_i\in\{0,1\}$，$i\in\{0,1,\cdots,n-1\}$ 的线性方程组解的个数。线性方程组的矩阵形式为

$$\begin{pmatrix}y_0\\y_1\\\vdots\\y_{n-3}\\y_{n-2}\\y_{n-1}\end{pmatrix}=\begin{pmatrix}0&a_2&a_1&0&\cdots&0\\0&0&a_3&a_2&0&\cdots\\&&\cdots&\cdots&&\\0&\cdots&\cdots&0&a_{n-1}&a_{n-2}\\a_{n-1}&0&\cdots&\cdots&0&a_0\\a_1&a_0&0&\cdots&\cdots&0\end{pmatrix}\begin{pmatrix}x_0\\x_1\\\vdots\\x_{n-3}\\x_{n-2}\\x_{n-1}\end{pmatrix}$$

由定理 2 可知，方程组解的个数为 $|L_b^a|=2^{n-r}$，其中，r 为系数矩阵 $\boldsymbol{A}=\begin{pmatrix}0&a_2&a_1&0&\cdots&0\\0&0&a_3&a_2&0&\cdots\\&&\cdots&\cdots&&\\0&\cdots&\cdots&0&a_{n-1}&a_{n-2}\\a_{n-1}&0&\cdots&\cdots&0&a_0\\a_1&a_0&0&\cdots&\cdots&0\end{pmatrix}$ 的秩。

矩阵 A 的每一行和每一列都有输入差的 2bit，且同 1bit 出现在矩阵不同的行和列中，所以矩阵 A 中非零的行数越少，r 越小；当矩阵 A 的每一行都有非零的元素时，r 为最大值。显然，当 $a=0$，即 A 为零矩阵时，$r=0$；当 a 有 1bit 非零时，由于非零的比特出现在矩阵不同的行和列中，所以 $r=2$；

当 a 有 $n-1$ 和 n bit 非零时，矩阵 A 的所有行均非零。当 a 有 $n-1$ bit 非零时，不妨设 $a_i=0$，$0\le i\le n-1$，矩阵 $A=\begin{pmatrix} 0 & a_2 & a_1 & \cdots & 0 & \cdots & 0 \\ & & \vdots & & \vdots & & \\ & \cdots & 0 & a_{(i-1) \bmod n} & 0 & \cdots & \\ & \cdots & 0 & a_{(i+1) \bmod n} & 0 & \cdots & \\ & & \vdots & & \vdots & & \\ a_1 & a_0 & 0 & \cdots & \cdots & 0 & \end{pmatrix}$

$=\begin{pmatrix} 0 & 1 & 1 & 0 & \cdots & 0 \\ & & \vdots & \vdots & & \\ & \cdots & 0 & 1 & 0 & \cdots \\ & \cdots & 0 & 1 & 0 & \cdots \\ & & \vdots & \vdots & & \\ 1 & 1 & 0 & \cdots & \cdots & 0 \end{pmatrix}$，经过 F_2 上的初等行变换，化为阶梯型矩阵为 $\begin{pmatrix} 1 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 1 & 0 & \cdots & \\ & & \vdots & \vdots & & \\ & & \vdots & \vdots & & \\ 0 & \cdots & \cdots & 0 & 1 & 1 \\ 0 & 0 & \cdots & \cdots & 0 & 0 \end{pmatrix}$，矩阵 A 中非零行数为 $n-1$，所以 $r=n-1$。

当 a 有 n bit 全为 1 时，矩阵 $A=\begin{pmatrix} 0 & 1 & 1 & 0 & \cdots & \cdots \\ \cdots & 0 & 1 & 1 & 0 & \cdots \\ & & \vdots & \vdots & & \\ & & \vdots & \vdots & & \\ 1 & 0 & \cdots & \cdots & 0 & 1 \\ 1 & 1 & 0 & \cdots & \cdots & 0 \end{pmatrix}$，对矩阵 A 在 F_2 上进行初等行变换，化为阶梯型矩阵为 $\begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 1 \\ 0 & 1 & 1 & 0 & 0 & \cdots \\ \cdots & 0 & 1 & 1 & 0 & \cdots \\ & & \vdots & \vdots & & \\ 0 & 0 & \cdots & \cdots & 1 & 1 \\ 0 & 0 & \cdots & \cdots & 0 & 0 \end{pmatrix}$，矩阵 A 中非零行数为 $n-1$，所以 $r=n-1$。

当 $r=0$ 时，$|L_b^a|=2^n$，$p_f(a\to b)=1$。当 $r\ne 0$ 时，$2\le r\le n-1$，$2\le|L_b^a|\le 2^{n-2}$，其中，$n\ge 3$。

因为 $p_f(a\to b)=\dfrac{|L_b^a|}{2^n}$，所以 $\dfrac{2}{2^n}\le p_f(a\to b)\le\dfrac{2^{n-2}}{2^n}$，$\dfrac{1}{2^{n-1}}\le p_f(a\to b)\le\dfrac{1}{4}$。

综上，定理 3 成立。

定理 4　对于 n 元 Keccak 类非线性变换 f，如果 $<a,b>\in\{<(a\ll j),(b\ll j)>|\ a=(0,0,\underbrace{1,4,2,\cdots,8}_{n-1\text{个}0},1),\ b=(0,0,\underbrace{1,4,2,\cdots,8}_{n-3\text{个}0},*,*,1)\ \ j\in\{0,1,\cdots,n-1\}\}$，其中*表示这个比特可以任意取 0 或 1，那么 $p_f(a\to b)=\dfrac{1}{4}$。

证明　由定理 1，只需要证明 $p_f((0,0,\underbrace{1,4,2,\cdots,8}_{n-1\text{个}0},1)\to(0,0,\underbrace{1,4,2,\cdots,8}_{n-3\text{个}0},*,*,1))=\dfrac{1}{4}$ 成立，便可得到 $p_f((0,0,\underbrace{1,4,2,\cdots,8}_{n-1\text{个}0},1\ll j)\to(0,0,\underbrace{1,4,2,\cdots,8}_{n-3\text{个}0},*,*,1\ll j))=\dfrac{1}{4}$ 成立，其中，$0\le j<n$，$j\in Z$。

当输入差 $a=(0,0,\underbrace{1,4,2,\cdots,8}_{n-1\text{个}0},1)$ 时，输出差 b 的各个比特的差分变换的布尔函数表达式为：

$b_0=0\oplus 0x_1\oplus 0x_2\oplus 0\oplus 0=0$；

$b_1=0\oplus 0x_2\oplus 0x_3\oplus 0\oplus 0=0$；

…

$b_{n-3}=0\oplus x_{n-2}\oplus 0x_{n-1}\oplus 0\oplus 1=x_{n-2}\oplus 1$；

$b_{n-2}=0\oplus 0x_{n-1}\oplus 1x_0\oplus 0\oplus 0=x_0$；

$b_{n-1}=1\oplus 0\oplus 0\oplus 0\oplus 0\oplus 0=1$。

输出差 $b=(0,0,\underbrace{1,4,2,\cdots,8}_{n-3\text{个}0},x_{n-2}\oplus 1,x_0,1)$。$(x_{n-2}\oplus 1)$ 和 x_0 可以取 0 或 1，b 的取值只与 $x_{n-2}\oplus 1$ 和 x_0 的取值有关，与输入 x 的其他 $n-2$ bit 无关。b 共有 $(0,0,\underbrace{1,4,2,\cdots,8}_{n-1\text{个}0},1)$，$(0,0,\underbrace{1,4,2,\cdots,8}_{n-2\text{个}0},1,1)$，$(0,0,\underbrace{1,4,2,\cdots,8}_{n-3\text{个}0},1,0,1)$，$(0,0,\underbrace{1,4,2,\cdots,8}_{n-3\text{个}0},1,1,1)$ 4 种情形。由定理 2 可知，当 $a=(0,0,\underbrace{1,4,2,\cdots,8}_{n-1\text{个}0},1)$ 时，

$$\begin{aligned} & p_f(a\to(0,0,\underbrace{1,4,2,\cdots,8}_{n-1\text{个}0},1)) \\ &= p_f(a\to(0,0,\underbrace{1,4,2,\cdots,8}_{n-2\text{个}0},1,1)) \\ &= p_f(a\to(0,0,\underbrace{1,4,2,\cdots,8}_{n-3\text{个}0},1,0,1)) \end{aligned}$$

$= p_f(a \to (0,0,\underbrace{1,\cdots,1}_{n-3\text{个}0},1,1,1)) = \frac{1}{4}$。

再由定理 1，

$p_f((0,0,\underbrace{1,\cdots,1}_{n-1\text{个}0},1<<j) \to (0,0,\underbrace{1,\cdots,1}_{n-3\text{个}0},*,*,1<<j))$

$= p_f((0,0,\underbrace{1,\cdots,1}_{n-1\text{个}0},1) \to (0,0,\underbrace{1,\cdots,1}_{n-3\text{个}0},*,*,1)) = \frac{2^{n-2}}{2^n} = \frac{1}{4}$。

综上，定理 4 成立。

定理 5 对 n 元 Keccak 类非线性变换 f，如果 $<a,b>\in\{<(a<<j),(b<<j)>|a=(0,0,1,\underbrace{1,\cdots,1}_{n-2\text{个}1},1),b=(*,*,\underbrace{1,\cdots,1}_{n-1\text{个}*},1)$；或 $a=(0,1,\underbrace{1,\cdots,1}_{n-1\text{个}1},1)$，$b=(*,\underbrace{1,\cdots,1}_{n-2\text{个}*},\bullet,(\bullet\oplus1))$；或 $W_{\mathrm{H}}(a)=n,\begin{cases}W_{\mathrm{H}}(b)\text{为奇数，}n\text{为奇数；}\\W_{\mathrm{H}}(b)\text{为偶数，}n\text{为偶数}\end{cases}$，其中，*表示这个比特可以任意取 0 或 1，$\bullet$ 和 $(\bullet\oplus1)$ 表示这 2bit 是互补的，$0<j\le n$ ，$j\in Z\}$，则差分转移概率 $p_f(a\to b)=\frac{1}{2^{n-1}}$。

证明 1) 输入差 $a=(0,0,1,\underbrace{1,\cdots,1}_{n-2\text{个}1},1)$ 时，b 的每一个比特的布尔函数表达式为：

$b_0 = 0\oplus1x_1\oplus0x_2\oplus0\oplus1 = x_1\oplus1$；

$b_1 = 0\oplus1x_2\oplus1x_3\oplus0\oplus0 = x_2\oplus x_3$；

…

$b_{n-3} = 1\oplus1x_{n-2}\oplus1x_{n-1}\oplus1\oplus1 = x_{n-2}\oplus x_{n-1}\oplus1$；

$b_{n-2} = 1\oplus0x_{n-1}\oplus1x_0\oplus0\oplus0 = x_0\oplus1$；

$b_{n-1} = 1\oplus0x_0\oplus0x_1\oplus0\oplus0 = 1$。

输出差 b 的结构为 $(*,*,\underbrace{1,\cdots,1}_{n-1\text{个}*},1)$，*表示这个比特可以取 0 或 1，共有 2^{n-1} 个不同的输出差 b，由定理 2，差分转移概率 $p_f(a\to b)=\frac{2}{2^n}=\frac{1}{2^{n-1}}$。

2) 输入差 $a=(0,1,\underbrace{1,\cdots,1}_{n-1\text{个}1},1)$ 时，b 的每一个比特的布尔函数表达式为

$b_0 = 0\oplus1x_1\oplus1x_2\oplus1\oplus1 = x_1\oplus x_2$；

$b_1 = 1\oplus1x_2\oplus1x_3\oplus1\oplus1 = x_2\oplus x_3\oplus1$；

…

$b_{n-3} = 1\oplus1x_{n-2}\oplus1x_{n-1}\oplus1\oplus1 = x_{n-2}\oplus x_{n-1}\oplus1$；

$b_{n-2} = 1\oplus0x_n\oplus1x_0\oplus0\oplus0 = x_0\oplus1$；

$b_{n-1} = 1\oplus1x_0\oplus0x_2\oplus0\oplus1 = x_0$。

输出差 b 的结构为 $(*,*,\underbrace{1,\cdots,1}_{n-2\text{个}*},\bullet,(\bullet\oplus1))$，*表示这个比特可以取 0 或 1，$\bullet$ 和 $(\bullet\oplus1)$ 表示后 2bit 是互补的；由输入差的结构可知，共有 2^{n-1} 个不同的输出差 b。由定理 2 可知，差分转移概率 $p_f(a\to b)=\frac{2}{2^n}=\frac{1}{2^{n-1}}$。

3) 输入差 a 的汉明重量 $W_{\mathrm{H}}(a)=n$，$p_f(a\to b)\ne0$ 时，按照定理 2 的证明方法，将输入、输入差和输出差从高位比特到低位比特依次可以表示为：$x=(x_0,x_1,\cdots,x_{n-1})$，$a=(a_0,a_1,\cdots,a_{n-1})$，$b=(b_0,b_1,\cdots,b_{n-1})$，令 $y_i=b_i\oplus a_i\oplus a_{(i+1)\bmod n}\ a_{(i+2)\bmod n}\oplus a_{(i+2)\bmod n}$，其中，$x_i,a_i,b_i,y_i\in\{0,1\}$，$i\in\{0,1,\cdots,n-1\}$。$L_b^a=\{x\mid f(x\oplus a)\oplus f(x)=b\}$，则方程组解的个数即为 $W_{\mathrm{H}}(a)=n$ 时，L_b^a 中元素的个数 $|L_b^a|$。

$$\begin{pmatrix}y_0\\y_1\\\vdots\\y_{n-3}\\y_{n-2}\\y_{n-1}\end{pmatrix}=\begin{pmatrix}0&1&1&0&\cdots&0\\0&0&1&1&0&\cdots\\&&\cdots&\cdots&&\\0&\cdots&\cdots&0&1&1\\1&0&\cdots&\cdots&0&1\\1&1&0&\cdots&\cdots&0\end{pmatrix}\begin{pmatrix}x_0\\x_1\\\vdots\\x_{n-3}\\x_{n-2}\\x_{n-1}\end{pmatrix},$$

$$\boldsymbol{A}=\begin{pmatrix}0&1&1&0&\cdots&0\\0&0&1&1&0&\cdots\\&&\cdots&\cdots&&\\0&\cdots&\cdots&0&1&1\\1&0&\cdots&\cdots&0&1\\1&1&0&\cdots&\cdots&0\end{pmatrix}$$

为系数矩阵，系数矩阵的秩 $R(\boldsymbol{A})=r=n-1$；所以 $|L_b^a|=2^{n-(n-1)}=2$，$p_f(a\to b)=\frac{|L_b^a|}{2^n}=\frac{1}{2^{n-1}}$。

下面证明差分转移概率 $p_f(a\to b)=\frac{1}{2^{n-1}}$ 时，对应的输出差 b 的结构。

当 $W_{\mathrm{H}}(a)=n$ 时，b 的每个比特的布尔函数表达式为：

$b_i = x_{(i+1)\bmod n}\oplus x_{(i+2)\bmod n}\oplus1$，其中，$x_i\in\{0,1\}$，$i\in\{0,1,\cdots,n-1\}$。

当 n 为偶数时：

$b_0\oplus b_1\oplus b_2\oplus\cdots\oplus b_{n-1}$

$= x_1\oplus x_2\oplus x_2\oplus x_3\oplus x_3\oplus x_4\oplus\cdots\oplus$

$x_{n-1}\oplus x_0\oplus x_0\oplus x_1\oplus(1\oplus1\oplus\cdots\oplus1)$

$= x_1\oplus x_2\oplus x_2\oplus x_3\oplus x_3\oplus x_4\oplus\cdots\oplus$

$x_{n-1}\oplus x_0\oplus x_0\oplus x_1 = 0$

所以，此时输出差 b 的汉明重量为偶数。

因为差分转移概率 $p_f(a \to b)=\frac{1}{2^{n-1}}$，由定理 2 可知，不同的输出差个数为 2^{n-1} 个。n 为偶数时，全体 $W_{\mathrm{H}}(b)$ 为偶数的 b 个数恰好为 2^{n-1} 个。假设存在一个 $W_{\mathrm{H}}(b')$ 为偶数，但 $p_f(a \to b') \neq \frac{1}{2^{n-1}}$，则一定存在另一个输出差 b^* 的 $W_{\mathrm{H}}(b^*)$ 为奇数，在 $W_{\mathrm{H}}(a)=n$ 的前提下，使 $p_f(a \to b^*)=\frac{1}{2^{n-1}}$，这与 $b_0 \oplus b_1 \oplus b_2 \oplus \mathrm{L} \oplus b_{n-1}=0$ 相矛盾，所以假设不成立。当 $W_{\mathrm{H}}(a)=n$，n 为偶数时，对于所有的 $W_{\mathrm{H}}(b)$ 为偶数的输出差 b，都有 $p_f(a \to b)=\frac{1}{2^{n-1}}$。

同理可以证明，当 $W_H(a)=n$，n 为奇数时，对于所有的 $W_H(b)$ 为奇数的输出差 b，都有 $p_f(a \to b)=\frac{1}{2^{n-1}}$。

综上，定理 5 成立。

定理 6　n 元的 Keccak 类非线性变换 f 和 $n+1$ 元的 Keccak 类非线性变换 g 有如下关系：对于 n 元 Keccak 类非线性变换 f，输入差 $a=(0,0,a_0,a_1,\mathrm{L},a_{n-3})$，输出差 $b=(b_0,b_1,\mathrm{L},b_{n-1})$，对应的差分转移概率为 $p_f(a \to b)$；$n+1$ 元 Keccak 类非线性变换 g，输入差 $a'=(0,0,0,a_0,a_1,\mathrm{L},a_{n-3})$，输出差 $b'=(0,b_0,b_1,\mathrm{L},b_{n-1})$，差分转移概率为 $p_f(a' \to b')$；那么 $p_g(a' \to b')=p_f(a \to b)$。其中，$a_i,b_j \in \{0,1\},\ i \in \{0,1,\mathrm{L},n-3\},\ j \in \{0,1,\mathrm{L},n-1\}$。

证明　当输入 $x=(x_0,x_1,\mathrm{L},x_{n-1})$、$a=(0,0,a_0,a_1,\mathrm{L},a_{n-3})$ 时，n 元 Keccak 类非线性变换 f 对应的输出差 b 每个比特的布尔函数表达式为：

$b_0=0 \oplus a_0x_1 \oplus 0x_2 \oplus 0a_0 \oplus a_0$；

$b_1=0 \oplus a_1x_2 \oplus a_0x_3 \oplus a_0a_1 \oplus a_1$；

…

$b_{n-3}=a_{n-5} \oplus a_{n-3}x_{n-2} \oplus a_{n-4}x_{n-1} \oplus a_{n-4}a_{n-3} \oplus a_{n-3}$；

$b_{n-2}=a_{n-4} \oplus 0x_{n-1} \oplus a_{n-3}x_0 \oplus a_{n-3}0 \oplus 0$；

$b_{n-1}=a_{n-3} \oplus 0x_0 \oplus 0x_1 \oplus 0 \oplus 0$。

当输入 $x'=(x'_0,x'_1,x'_2,\mathrm{L},x'_n)$，$a=(0,0,0,a_0,a_1,\mathrm{L},a_{n-3})$ 时，$n+1$ 元 Keccak 类非线性变换 g 对应的输出差 b' 每一个比特的布尔函数表达式为：

$b'_0=0 \oplus 0x'_1 \oplus 0x'_2 \oplus 0 \oplus 0$；

$b'_1=0 \oplus a_0x'_2 \oplus 0x'_3 \oplus 0 \oplus a_0$；

…

$b'_{n-3}=a_{n-6} \oplus a_{n-4}x'_{n-2} \oplus a_{n-5}x'_{n-1} \oplus a_{n-5}a_{n-4} \oplus a_{n-4}$；

$b'_{n-2}=a_{n-5} \oplus a_{n-3}x'_{n-1} \oplus a_{n-4}x'_n \oplus a_{n-4}a_{n-3} \oplus a_{n-3}$；

$b'_{n-1}=a_{n-4} \oplus 0x'_n \oplus a_{n-3}x'_0 \oplus a_{n-3}0 \oplus 0$；

$b'_n=a_{n-3} \oplus 0x'_0 \oplus 0x'_1 \oplus 0 \oplus 0$。

由表达式可知，当 $x_0=x'_0$，$x_i=x'_{i+1}$ $(1 \le i \le n-1)$ 时，$b=(b_0,b_1,\mathrm{L},b_{n-1})$，$b'=(0,b_0,b_1,\mathrm{L},b_{n-1})$；$x'_1$ 取值 0 或 1 对 b' 的值没有影响。记 $f_b^a=\{x \mid f(x \oplus a) \oplus f(x)=b\}$，$g_{b'}^{a'}=\{x' \mid g(x' \oplus a') \oplus g(x')=b'\}$，那么 $2 \mid f_b^a \mid = \mid g_{b'}^{a'} \mid$。

$$p_g(a' \to b')=\frac{|g_{b'}^{a'}|}{2^{n+1}}=\frac{2|f_b^a|}{2^{n+1}}=\frac{|f_b^a|}{2^n}=p_f(a \to b)$$

所以，$p_g(a' \to b')=p_f(a \to b)$。

综上，定理 6 成立。

4　结束语

杂凑函数 Keccak 中的压缩函数的非线性变换已经被广泛应用到很多密码算法中，文献[11]通过对其进行改造，提出了 Minikeccak 的非线性变换，并取得了良好的效果。为了更好地应用这一类非线性变换，本文建立了 n 元 Keccak 类非线性变换模型，并且分析了它的差分转移概率性质，给出了最大的非平凡差分转移概率和最小的非平凡差分转移概率的结构和计数，给出了这类非线性变换相邻变元间取相同差分转移概率的结构。但是，还有许多 Keccak 类非线性变换模型的差分性质例如次大差分转移概率、差分转移概率为 0 的结构等情况，本文还没有去研究分析。另外，对 Keccak 类非线性变换模型的 Walsh 谱值特性的研究将是下一个工作重点。

参考文献：

[1] NIST. Announcing request for candidate algorithm nominations for a new cryptographic hash algorithm (SHA-3) family[J]. Federal Register Notices 72, 2007, 212: 62212-62220.

[2] ANDREW R, RAY P, CHANG S J. Status Report on the First Round of the SHA-3 Cryptographic Hash Algorithm Competition[R]. Information Technology Laboratory National Institute of Standards and Technology, Gaithersburg, 2009.

[3] MELTEM S T, RAY P, LAWRENCE E B, *et al*. Status Report on the Second Round of the SHA-3 Cryptographic Hash Algorithm Competition. Computer Security Division[R]. Information Technology Labor-

atory National Institute of Standards-and Technology, Gaithersburg, 2011.

[4] NIST. The SHA-3 Finalists candidates U S department of commerce national information service[EB/OL]. http://csrc. nist.gov. /groups/ ST/hash/sha-3/Round3/submissions_round3. html.

[5] GUIDO B, JOAN D, MICHAEL P, *et al*. Keccak sponge function family main document[EB/OL]. http://csrc. nist.gov /groups/ ST /hash /sha-3 /Round1/submissions_round1. Html.

[6] **薛宇**, **吴文玲**, **王张宜**. SHA3 **杂凑密码候选算法简评**[J]. **中国科学院研究生院学报**,2009, 26(5): 577-586.
XUE Y,WU W L,WANG Z Y. Remarks of candidates hash algorithms for SHA3[J].Journal of CAS Postgraduates, 2009, 26(5): 577-586.

[7] **罗岚**,**叶娅兰**,**许春香等**.**在信念网模型下的** SHA3 **前五名算法注记**[EB/OL].http://www.scienceet.cn/upload/blog/f-ile/2010/12/2010121592436256375.pdf.
LUO L, YE Y L, XU C X, *et al*. A note finalist5 of SHA3 at faithful network[EB/OL].http://www.scienceet.cn/upload/blog/file/2010/12/2010121592436256375.pdf.

[8] GUIDO B, JOAN D, MICHAEL P, *et al*. A belt-and-mill hash function[EB/OL]. http://radiogatun.noekeon.org.

[9] JOAN D, CLAPP C S K. Fast hashing and stream encryption with PANAMA[A]. Fast Software Encryption 1998 (S Vaudenay, ed)[C]. 1998. 60-74.

[10] JOAN D. Cipher and Hash Function Design Strategies Based on Linear and Differential Cryptanalysis[D]. Belgium: Katholieke Universities Leuven, 1995.

[11] EPHRAIM A. Sharing Nonlinear Gates in the Presence of Glitches[D]. Enschede, Holland: University of Twente, 2010.

作者简介：

李倩男（1988-），**男，河南柘城人，信息工程大学硕士生，主要研究方向为密码理论和应用数学等。**

李云强（1968-），**男，河南尉氏人，信息工程大学教授，主要研究方向为密码理论和应用数学。**

蒋淑静（1986-），**女，河南叶县人，中国科学院硕士生，主要研究方向为信号与信息处理、应用数学和计算机理论等。**

路遥（1987-），**男，河南柘城人，国防科学技术大学硕士生，主要研究方向为计算机理论和应用等。**