文章编号:1002-1175(2011)02-0266-08

# 6 轮 ARIA 的最优不可能差分分析

张 磊†,郭建胜

(解放军信息工程大学电子技术学院, 郑州 450004)

(2010 年 3 月 23 日收稿; 2010 年 5 月 28 日收修改稿)

**Zhang L, Guo J S. Best impossible differential cryptanalysis of 6-round ARIA [J]. Journal of the Graduate School of the Chinese Academy of Sciences, 2011, 28(2):266 – 273.**

**摘 要** 研究了 ARIA 在不可能差分分析下的安全性. 通过对算法扩散层的分析,给出了 ARIA 中间状态在加密过程的差分传递性质. 在此基础上证明了 6 轮 ARIA 不存在使得输入输出差分重量小于 10 的不可能差分路径,同时证明了在输入输出差分重量为 10 的情况下 6 轮 ARIA 只存在 2 类形式的不可能差分路径. 利用构造出的这 2 类不可能差分路径,从理论上证明了 6 轮 ARIA 不可能差分攻击的最优结果为:$2^{120}$个选择明文和$2^{94.5}$次 6 轮加密.

**关键词** 分组密码,不可能差分分析,ARIA,数据复杂性,时间复杂性

**中图分类号** TP309

ARIA[1]是韩国国家安全研究所(National Security Research Institute)提出的一个 SPN 结构的分组密码算法,于 2004 年被韩国商业部、工业部和能源部确认为韩国分组密码算法标准. ARIA 的明文分组长度是 128bit,密钥长度是 128、192 和 256bit 的可变密钥长度分组密码算法. 在不同的密钥规模下,算法相应的轮数分别为 12、14、16 轮. 与常规的 SPN 结构算法不同,ARIA 在连续 2 轮加密的非线性层采用 S 与 $S^{-1}$结构,扩散层变换 $P$ 采用对合变换,从而形成 $SPS^{-1}$结构,这样其解密算法与加密算法结构一致,节约算法实现成本.

不可能差分[2-5]最早由 Knudsen 和 Biham 分别独立提出,可以看作是差分分析的一种变体. 在不可能差分分析中,目标是寻找概率为 0 的差分路径来排除错误的候选密钥,从而恢复正确的密钥. 在针对 ARIA 的不可能差分分析方面,其主要结果有:2003 年 Daesung Kwon 等[6]在 ARIA 算法的提交报告中称 ARIA 不存在 4 轮的不可能差分;2007 年 Wu 等[7]根据“中间相遇原理”首先指出了 ARIA 存在 4 轮不可能差分,并给出了对 6 轮 ARIA 的攻击,其数据量约为$2^{121}$个选择明文,计算量约为$2^{112}$次 6 轮加密;2008 年 Zhang 等[8]改进了 Wu 的不可能差分,给出一系列 4 轮不可能差分,在攻击 6 轮 ARIA 时,给出 2 类算法,其中算法 1 的数据量约为$2^{120.5}$个选择明文,计算量约为$2^{104.5}$次 6 轮加密;算法 2 的数据量约为$2^{113}$个选择明文,计算量约为$2^{121.6}$次 6 轮加密;2008 年李申华等[9]在 Zhang 的工作基础上改进了其分析方法,减少了子密钥的猜测量,得到 6 轮攻击所需的数据量约$2^{120}$个选择明文,计算量约为$2^{96}$次 6 轮加密. 然而此攻击算法在分析过程中存在错误,本文将在下面的分析中给出改进后的攻击结果.

由上述研究工作可以看出,更深入地分析 ARIA 算法在不可能差分分析下的安全性,并给出其安全性评估是目前的研究热点之一. 本文基于 ARIA 加密算法的扩散层 $P$ 盒的设计,通过对 4 轮不可能差分路径的分析,证明给出了 6 轮 ARIA 所能达到最优的 2 类形式不可能差分路径,从而从理论上给出了 6 轮 ARIA 不可能差分攻击的最优结果.

† E-mail:zzll2000@163.com

本文的主要内容如下:第 1 节介绍了 ARIA 分组密码算法及针对 ARIA 的不可能差分研究现状;第 2 节分析研究了 ARIA 扩散层变换 $P$ 的相关性质及对密码变换的影响;第 3 节证明并给出了所有达到最优的 6 轮 ARIA 的不可能差分路径;最后总结全文.

## 1　相关研究

ARIA 是分组长度为 128bit,密钥长度是 128、192 和 256bit 的可变分组密码. 算法采用 SPN 结构,易于软硬件实现.

### 1.1　ARIA 分组密码算法

ARIA 密码算法的各类变换均作用在字节上,可以将 128bit 的明文写成 4×4 的状态矩阵的形式,如图 1 所示. 算法每轮的轮函数包括轮密钥加、置换层变换、扩散层变换 3 个操作,最后一轮只包括轮密钥加变换和置换层变换,不含扩散层变换.

| | | | |
|---|---|---|---|
| 0 | 4 | 8 | 12 |
| 1 | 5 | 9 | 13 |
| 2 | 6 | 10 | 14 |
| 3 | 7 | 11 | 15 |

**图 1　中间状态的字节矩阵表示形式**

1)轮密钥加(RKA)　将中间状态与 128bit 的轮子密钥($RK_i$)作模 2 加.

2)置换层变换(SL)　SL 包含 2 种类型 $SL_1$ 和 $SL_2$,涉及 4 个 8×8 的 S 盒($S_1$,$S_2$ 以及它们的逆 $S_1^{-1}$ 和 $S_2^{-1}$),其中,

$$SL_1 = S_1S_2S_1^{-1}S_2^{-1}\ S_1S_2S_1^{-1}S_2^{-1}\ S_1S_2S_1^{-1}S_2^{-1}\ S_1S_2S_1^{-1}S_2^{-1}$$

$$SL_2 = S_1^{-1}S_2^{-1}S_1S_2\ S_1^{-1}S_2^{-1}S_1S_2\ S_1^{-1}S_2^{-1}S_1S_2\ S_1^{-1}S_2^{-1}S_1S_2$$

加密过程中交替使用这 2 种类型,在奇数轮使用 $SL_1$,偶数轮使用 $SL_2$,亦即:在奇数轮时,在第 0,4,8,12 个字节使用 $S_1$,在 1,5,9,13 个字节使用 $S_2$,在 2,6,10,14 个字节使用 $S_1^{-1}$,在 3,7,11,15 个字节使用 $S_2^{-1}$;在偶数轮时,在第 0,4,8,12 个字节使用 $S_1^{-1}$,在 1,5,9,13 个字节使用 $S_2^{-1}$,在 2,6,10,14 个字节使用 $S_1$,在 3,7,11,15 个字节使用 $S_2$.

3)扩散层变换(DL)　扩散层 DL 是 2 元域上面向字节的线性变换,其具体操作可以定义为 $GF(2^8)^{16}\rightarrow GF(2^8)^{16}$的一个映射,即

$$(x_0,x_1,\cdots,x_{15}) \mapsto (y_0,y_1,\cdots,y_{15})$$

$$y_0 = x_3 \oplus x_4 \oplus x_6 \oplus x_8 \oplus x_9 \oplus x_{13} \oplus x_{14},\quad y_8 = x_0 \oplus x_1 \oplus x_4 \oplus x_7 \oplus x_{10} \oplus x_{13} \oplus x_{15},$$

$$y_1 = x_2 \oplus x_5 \oplus x_7 \oplus x_8 \oplus x_9 \oplus x_{12} \oplus x_{15},\quad y_9 = x_0 \oplus x_1 \oplus x_5 \oplus x_6 \oplus x_{11} \oplus x_{12} \oplus x_{14},$$

$$y_2 = x_1 \oplus x_4 \oplus x_6 \oplus x_{10} \oplus x_{11} \oplus x_{12} \oplus x_{15},\quad y_{10} = x_2 \oplus x_3 \oplus x_5 \oplus x_6 \oplus x_8 \oplus x_{13} \oplus x_{15},$$

$$y_3 = x_0 \oplus x_5 \oplus x_7 \oplus x_{10} \oplus x_{11} \oplus x_{13} \oplus x_{14},\quad y_{11} = x_2 \oplus x_3 \oplus x_4 \oplus x_7 \oplus x_9 \oplus x_{12} \oplus x_{14},$$

$$y_4 = x_0 \oplus x_2 \oplus x_5 \oplus x_8 \oplus x_{11} \oplus x_{14} \oplus x_{15},\quad y_{12} = x_1 \oplus x_2 \oplus x_6 \oplus x_7 \oplus x_9 \oplus x_{11} \oplus x_{12},$$

$$y_5 = x_1 \oplus x_3 \oplus x_4 \oplus x_9 \oplus x_{10} \oplus x_{14} \oplus x_{15},\quad y_{13} = x_0 \oplus x_3 \oplus x_6 \oplus x_7 \oplus x_8 \oplus x_{10} \oplus x_{13},$$

$$y_6 = x_0 \oplus x_2 \oplus x_7 \oplus x_9 \oplus x_{10} \oplus x_{12} \oplus x_{13},\quad y_{14} = x_0 \oplus x_3 \oplus x_4 \oplus x_5 \oplus x_9 \oplus x_{11} \oplus x_{14},$$

$$y_7 = x_1 \oplus x_3 \oplus x_6 \oplus x_8 \oplus x_{11} \oplus x_{12} \oplus x_{13},\quad y_{15} = x_1 \oplus x_2 \oplus x_4 \oplus x_5 \oplus x_8 \oplus x_{10} \oplus x_{15}.$$

ARIA 算法加解密算法结构相似,差别仅在于轮子密钥的使用. 在本文的分析中不涉及密钥生成算法,这里不再叙述,具体描述参见文献[1].

### 1.2　ARIA 的不可能差分性质研究现状

**定义 1**　设 $\boldsymbol{X}=(x_0,x_1,\cdots,x_{15})$,这里 $x_i\in(GF(2^8),0\leqslant i\leqslant 15$. 令 $\boldsymbol{X}'=\boldsymbol{X}\oplus\boldsymbol{X}^*$ 代表差分值,差分重量为差分值中所有非 0 的状态字节的个数. 输入输出差分重量为输入差分重量与输出差分重量之和.

2007 年 Wu 等[2]首先发现 ARIA 存在 4 轮不可能差分路径,其找到的不可能差分路径输入差分重量为 1,输出差分重量为 5. 输入差分经过 2 轮变换,输出差分经过 2 轮逆变换后,在中间产生矛盾,从而形成 4 轮 ARIA 的不可能差分路径. 在此基础上,前后各加一轮,完成对 6 轮 ARIA 的不可能差分攻击.

此时其输入差分重量为 7,输出差分重量为 5,输入输出差分重量为 12,记此不可能差分路径为 ID – I.

Zhang 等[8]提出了 ARIA 的另外 2 类不可能差分路径,其中一类改进了 Wu 的不可能差分路径,使得 4 轮 ARIA 输入差分重量为 1,输出差分重量为 4,其对应的 6 轮 ARIA 不可能差分攻击,输入差分重量为 7,输出差分重量为 4,输入输出重量为 11,记此不可能差分路径为 ID-II;另一类 4 轮 ARIA 的输入差分重量为 2,输出差分重量为 4,其对应的 6 轮 ARIA 不可能差分攻击,输入差分重量为 10,输出差分重量为 4,输入输出重量为 14,记此不可能差分路径为 ID-III;

李申华等[9]改进了 Zhang 提出的第二类不可能差分路径,使得 4 轮 ARIA 输入差分重量为 2,输出差分重量为 4,其对应的 6 轮 ARIA 不可能差分攻击,输入差分重量为 6,输出差分重量为 4,输入输出重量为 10,记此不可能差分路径为 ID-IV,这也是目前针对 ARIA 不可能差分分析的最好结果.

**定义 2** 记表达式 $a\rightarrow\times\leftarrow b$ 表示由输入差分重量为 $a$,输出差分重量为 $b$ 构成的一类不可能差分路径;$c\rightarrow a\rightarrow\times\leftarrow b\leftarrow d$ 表示上述不可能差分路径前后各扩展一轮后的不可能差分路径,其输入差分重量为 $c$,输出差分重量为 $d$.

根据定义 2,上述 4 类情况具体见表 1.

**表 1 ARIA 的不可能差分路径比较**

| | 4 轮不可能差分路径 | 扩展至 6 轮不可能差分攻击 | 输入输出差分重量 |
|---|---|---|---|
| ID-I | $1\rightarrow\times\leftarrow 5$ | $7\rightarrow 1\rightarrow\times\leftarrow 5\leftarrow 5$ | 12 |
| ID-II | $1\rightarrow\times\leftarrow 4$ | $7\rightarrow 1\rightarrow\times\leftarrow 4\leftarrow 4$ | 11 |
| ID-III | $2\rightarrow\times\leftarrow 4$ | $10\rightarrow 2\rightarrow\times\leftarrow 4\leftarrow 4$ | 14 |
| ID-IV | $2\rightarrow\times\leftarrow 4$ | $6\rightarrow 2\rightarrow\times\leftarrow 4\leftarrow 4$ | 10 |

目前,针对 ARIA 的不可能差分分析所达到的最好结果即猜测的输入输出差分重量为 10. 我们将在分析扩散层变换的基础上,给出在输入输出重量为 10 情况下的所有不可能差分路径,并证明不存在输入输出重量小于 10 的不可能差分路径.

## 2 ARIR 扩散层变换的相关性质

ARIA 扩散层变换 $P$ 是一个对合变换,将加密中间状态写成如图 1 所示的 4 × 4 的字节矩阵的形式. 由于扩散层变换的分支数为 8,则可以得到性质 1.

**性质 1** 加密变换的中间状态经过一次线性 $P$ 变换,任选其输出的 3 块,则它们至少与输入的 5 块相关;任选其输出的 4 块,则它们至少与输入的 4 块相关;任选其输出的 5 块,则它们至少与输入的 3 块相关.

观察扩散层变换 $P$ 的性质,任意一个输出均表示为输入 7 块的模 2 和. 当输入差分值相等时,其求和之后,出现偶数次数的输入将抵消. 应用分支数理论,此时输入输出涉及的非 0 块数最少. 为了降低输入输出差分重量,下面讨论时所选取的输入差分与输出差分均指各差分值相等的情况,当其差分值不相等时经过加密变换其所涉及的 S 盒数目一定大于上述讨论的情况. 这里仅考虑最优的状况.

通过计算机程序验证可得,满足输入差分重量为 1 输出差分重量为 7 的变量共计 16 个;满足输入差分重量为 2 输出差分重量为 6 的变量共计 48 个;满足输入差分重量为 3 输出差分重量为 5 的变量共计 144 个;满足输入差分重量为 4 输出重量为 4 的变量共计 204 个. 根据 $P$ 的对合性,可以得到满足输入差分重量为 5 输出差分重量为 3 的变量共计 144 个,满足输入差分重量为 6 输出差分重量为 2 的变量共计 48 个;满足输入差分重量为 7 输出差分重量为 1 的变量共计 16 个.

更进一步,利用计算机程序我们将从上述变量筛选出符合要求的所有不可能差分路径,可以得到如下性质.

**性质 2** 输入差分重量为 3 输出差分重量为 5 的变量经过一次 SL 和 DL 变换后,其输出的 16 个字节块中会在 4 个位置差分值为 0,并且这 4 个位置的分布有 36 种情况(附录中表 3 具体给出了这 36 个分布情况),可归为 3 类:

1) 从列上看有 2 列差分重量为 2,2 列差分重量为 0,从行上看有 4 行差分重量为 1,即附录表 3 中的 I,此类分布有 12 个;

2) 从列上看 4 列差分重量为 1,从行上看有 2 行差分重量为 2,2 行差分重量为 0,即附录表 3 中的 II,此类分布有 12 个;

3) 从列上看有 4 列差分重量为 1,从行上看有 4 行差分重量为 1,即附录表 3 中的 III,此类分布有 12 个.

举例说明,当输入差分在第 1、2、12 处不为 0,其他位置为 0 时,经过一次 SL 和 DL 变换后,其分布如图 2 所示.

输入差分为(1,2,12)时,经过一轮加密后在(0,3,13,14)这 4 个位置的差分值为 0,从列上看其分布为(2,0,0,2),从行上其分布为(1,1,1,1),符合性质 2 中 1)的情况.

**性质 3**　输入差分重量为 4 输出差分重量为 4 的变量经过一次 SL 和 DL 变换后,其值分布有以下 3 种情况:

1) 其输出的 16 个字节块中不含为 0 的差分值,此类情况有 24 个;

2) 其输出的 16 个字节块中含有 1 个为 0 的差分值,此类情况有 144 个;

3) 其输出的 16 个字节块中含有 4 个为 0 的差分值,这 4 个为 0 的差分值位置分布与性质 3 中 4 个为 0 的差分值位置分布相同,此类情况有 36 个.

**性质 4**　考虑在性质 2 与性质 3 中出现的 36 个有 4 处差分值为 0 分布,当其符合性质 2 中 1)时( 即附录表 3 中的 I ),经过一次 SL 和 DL 变换后,在输入列重复为 2 的那 2 列中存在位于不同行的 2 个差分字节,使得这 4 个差分字节模 2 和为 0;当其符合性质 2 中 2)时( 即附录表 3 中的 II ),经过一次 SL 和 DL 变换后,在输出的不同列不同行存在 4 个差分字节它们的模 2 和为 0;当其符合性质 2 中 3)时( 即附录表 3 中的 III ),经过一次 DL 和 SL 变换后,在输出的不同列 2 行中存在 4 个差分字节,它们的模 2 和为 0.

举例说明,当输入差分在第 0,3,13,14 处不为 0 时,经过一轮变换结果如图 3 所示.

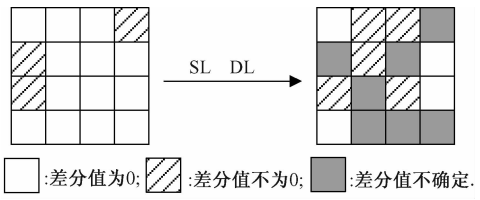


**图 2　输入差分为(1,2,12)时的一轮变换差分分布**

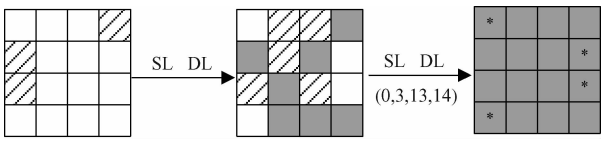


**图 3　输入差分为(1,2,12)时的 2 轮变换差分分布**

当在(0,3,13,14)这 4 个位置的差分值为 0 时,再经过一轮变换,各个字节的差分值均不能确定,但(0,3,13,14)这 4 个位置的差分值模 2 和为 0.

**性质 5**　满足输入差分重量为 5 输出重量为 3 的变量,其输入差分重量无论从列上还是行上看其分布只有 2 种情况:2 列(行)重量为 2,1 列(行)重量为 1,1 列(行)重量为 0 或者 1 列(行)重量为 3,2 列(行)重量为 1,1 列(行)重量为 0,亦即不存在使得各列(行)差分重量均不为 0 的输入.

## 3　ARIR 的不可能差分分析

对 6 轮 ARA 的不可能差分攻击需要寻找到一条 4 轮不可能差分路径,然后前后扩展一轮,构成 6 轮的不可能差分分析. 攻击的复杂性由第 1 轮与最后 1 轮所需猜测的密钥字节个数决定,即输入输出差分的重量决定,当输入输出差分重量越大时,其攻击的复杂性越高. 同时,对给定 4 轮 ARIA 不可能差分路径时,当其向后扩展 1 轮时,只经过 SL 变换,从而不改变其输出差分重量,当其向前扩展时需要经过一层 DL 变换,这里可以利用 ARIA 扩展层的分支数,得到向前一轮的最优的输入差分重量. 从而研究 6

轮 ARIA 的不可能差分输入输出差分重量,可以等价为研究寻找在此重量下的 4 轮 ARIA 的不可能差分路径.

考虑对 6 轮 ARIA 的不可能差分,在输入输出差分重量为 10 的情况下,可能出现的不可能差分路径如表 2 所示.

下面我们将分别讨论这 10 类可能的不可能差分路径,并证明不存在输入输出差分重量 < 10 的不可能差分路径. 为便于说明,我们去掉前后 2 轮,直接考虑 4 轮不可能差分路径的情况.

**表 2 输入输出差分重量为 10 的情况下所有可能的不可能差分路径**

| | 可能的不可能差分路径 | | 可能的不可能差分路径 |
|---|---|---|---|
| (1) | 7→1→ × ←3←3 | (2) | 6→2→ × ←4←4 |
| (3) | 5→3→ × ←5←5 | (4) | 4→4→ × ←6←6 |
| (5) | 4→6→ × ←6←6 | (6) | 3→5→ × ←7←7 |
| (7) | 2→6→ × ←8←8 | (8) | 1→7→ × ←9←9 |
| (9) | 8→? → × ←2←2 | (10) | 9→? → × ←1←1 |

**引理 1** 当输入差分重量为 1 时,不存在输出差分重量 ≤ 3 的 4 轮 ARIA 的不可能差分路径.

证明 当输入差分重量为 1 时,经过一次 SL 和 DL 变换,会在 7 个位置产生非 0 的差分值,其他位置差分值为 0;再经过一次 SL 和 DL 变换,各位置差分值不能确定,即正向可传递 1.5 轮.

当输出差分重量为 3 时,经过一次 $DL^{-1}$和 $SL^{-1}$变换,会在 5 个位置产生非 0 的差分值,再经过一次 $DL^{-1}$和 $SL^{-1}$变换,有一处差分值非零,其他位置差分值不能确定. 显然从单个字节上考虑前后 2 轮变换不能构成矛盾;考虑几个字节求和的情况. 由性质 1 正向 2 轮的几个位置求和得到确定的值,但逆向求和值仍不确定,从而不能构成 4 轮不可能差分路径.

当输出差分重量为 2 时,经过一次 $DL^{-1}$和 $SL^{-1}$变换,会在 6 个位置产生非 0 的差分值;再经过一次 $DL^{-1}$和 $SL^{-1}$变换,有 2 个位置的差分值为 0,其他位置差分值不能确定. 因为正向 2 轮变换后各个位置值均无法确定,从而此种情况也不能构成 4 轮不可能差分路径.

当输出差分重量为 1 时,经过一次 $DL^{-1}$和 $SL^{-1}$变换,会在 7 个位置产生非 0 的差分值,再经过一次 $DL^{-1}$和 $SL^{-1}$变换,各位置差分值不能确定,即逆向传递 1.5 轮. 从而不能构成 4 轮不可能差分路径.

综上,当输入差分重量为 1 时,不存在输出差分重量 ≤ 3 的 4 轮 ARIA 的不可能差分路径.

**引理 2** 当输入差分重量为 2 时,不存在输出差分重量 ≤ 3 的 4 轮 ARIA 的不可能差分路径.

证明 当输入差分重量为 2 时,经过一次 SL 和 DL 变换后,会产生 6 个位置差分值为 0,6 个位置差分值非 0,其他 4 个位置差分值不能确定;再经过一次 SL 和 DL 变换后,各处差分值均不能确定,但可以找到有 2 处差分值相等.

当输出差分重量为 3 时,经过一次 $DL^{-1}$和 $SL^{-1}$变换,会在 5 个位置产生非 0 的差分值,再经过一次 $DL^{-1}$ $SL^{-1}$变换,有 1 个位置差分值非 0,其他位置差分值不能确定. 这与输入有 2 个位置差分值相等不能找到矛盾. 考虑几个位置求和的情况,由性质 1 及性质 5,可知逆向几个字节的求和值不能确定,从而不能构成 4 轮不可能差分路径.

当输出差分重量为 2 和 1 时,其变化形式如引理 1 中的证明描述,它们也不能与输入差分重量为 2 的输入构成 4 轮不可能差分路径.

综上,当输入差分重量为 2 时,不存在输出差分重量 ≤ 3 的 4 轮 ARIA 的不可能差分路径.

**引理 3** 当输入差分重量为 2 时,存在一类输出差分重量为 4 的不可能差分路径(其形式为 2→ × ←4),并且其输出差分重量分布只能为性质 3 中 2)的形式.

证明 当输入差分重量为 2 时,经过一次 SL 和 DL 变换后,会产生 6 个位置差分值为 0,6 个位置差分值非 0,其他 4 个位置差分值不能确定;再经过一次 SL 和 DL 变换后,各处差分值均不能确定,但可以找到有 2 处差分值相等.

当输出差分重量为 4 时,由性质 3 知,经过一次 $DL^{-1}$ 和 $SL^{-1}$ 变换,其输出分为 3 种情况. 当为性质 3 中 1)时,再经过一轮变换,各字节差分值不确定,无法构成不可能差分路径;当为性质 3 中 2)时,再经过一次 $DL^{-1}$ 和 $SL^{-1}$ 变换,输出在 1 个位置差分值为 0,在 4 个位置差分值非 0,在其他位置的差分值不确定,从而可以找到 2 处差分值不相等. 经过程序验证,正向找到的使得差分值相等的 2 个位置与这里找到的使得差分值不相等的位置相同,从而构成矛盾,这样就证明了存在形式为 $2\rightarrow\times\leftarrow 4$ 的 4 轮不可能差分路径. 当为性质 3 中 3)时,再经过一轮变换,各字节差分值不确定,不能直接构成不可能差分路径. 考虑几个位置的差分值求和,由性质 1,性质 4,性质 5 知,前后 2 轮不能构成矛盾,从而不存在不可能差分路径.

综上,当输入差分重量为 2 时,存在一类形式如 $2\rightarrow\times\leftarrow 4$ 的 4 轮不可能差分路径,并且其输出差分只能为性质 3 中 2)的形式.

特别地,李申华等找到的不可能差分路径即这种形式的几个特例,其证明过程参见文献[9].

**引理 4**　当输入差分重量为 3 时,不存在输出差分重量 $\leqslant 5$ 的 4 轮 ARIA 的不可能差分路径.

证明　当输入差分重量为 3 时,由性质 2 知,经过一次 SL 和 DL 变换,会在 4 个位置差分值为 0;再经过一次 SL 和 DL 变换,各个字节的差分值均不能确定.

当输出差分重量为 5 时,经过一次 $DL^{-1}$ 和 $SL^{-1}$ 变换,会在 3 个位置产生非 0 的差分值;再经过一次 $DL^{-1}$ 和 $SL^{-1}$ 变换,会在 4 个位置差分值为 0,4 个位置差分值非 0,其他位置的差分值不确定. 由于正向单个字节的差分值均不确定,直接的矛盾不存在,只能考察它们作和的情况. 为便于说明,这里以图 2 中所示差分输入为证明,其他情况可类似给出. 正向经过 2 轮变换后,在(0,3,13,14)这 4 个位置的差分值和为 0,这 4 个位置来自 2 列不同的行,若要形成矛盾在逆向 2 轮变换后在这 4 个位置的和为非 0. 由性质 2 知,逆向 2 轮变换后在这 4 个位置的和只有 2 种情况:和为 0 或和不确定,从而不能与正向 2 轮变换构成不可能差分. 考虑 3 项求和的情况,由性质 5 知,正向 2 轮变换后,任选 3 顶之和都是不确定的值,不能与逆向变换构成不可能差分;考虑 5 项求和的情况,由于线性变换的对合性,当输出差分经过一次 $DL^{-1}$ 变换后,其形式与输入相同,此时逆向 2 轮变换后由性质 5 知,任 5 项之和都是不确定的值,不能与正向变换构成不可能差分. 至于其他求和的情况,由性质 1 易知都不能构成不可能差分路径.

当输出差分重量为 4 时,显然符合性质 3 中 1)、2)的差分变量不能与正向变换的输出构成不可能差分. 下面考虑符合性质 3 中 3)的差分变量. 经过一次 $DL^{-1}$ 和 $SL^{-1}$ 变换,由性质 3 知,其差分为 0 的字节分布与输出为 5 的变量经过一轮逆变换后的分布相同. 从而,类似上述证明过程,可以得出此时也不能构成不可能差分路径.

当输出差分重量为 3、2、1 时,经过 2 轮逆变换,各字节差分值均不确定. 考虑几个位置求和的情况,由性质 1 及性质 5,可知逆向字节求和值不能确定,从而不能构成 4 轮不可能差分路径.

综上,当输入差分重量为 3 时,不存在输出差分重量 $\leqslant 5$ 的 4 轮 ARIA 的不可能差分路径.

**引理 5**　当输入差分重量为 4 时,不存在输出差分重量 $\leqslant 5$ 的 4 轮 ARIA 的不可能差分路径.

证明　当输入差分重量为 4 时,由性质 3 知,经过一次 SL 和 DL 变换后,其输出分为 3 种情况. 当符合性质 3 中 1)时,输出有 4 个位置的差分不为 0,其他位置的差分值不确定,再经过一轮变换,各个位置的差分值都不确定;当符合性质 3 中 2)时,输出有 1 个位置的差分值为 0,4 个位置的差分值非 0,其他位置的差分值不确定,再经过一轮变换,各个位置的差分值都不确定;当符合性质 3 中 3)时,输出有 4 个位置的差分值为 0,其他位置的差分值不确定.

当输出差分重量为 5 时,其差分形式如引理 3 证明中的描述. 直接的矛盾不存在,考察它们作和的情况. 易知,当输入符合性质 3 中 1)、2)时,其求和值不确定,此时不能构成不可能差分. 当输入符合性质 3 中 3)时,其证明过程与引理 3 证明中输入差分重量为 3,输出差分重量为 4 情况类似. 可知此时也不能构成矛盾. 从而不存在输出差分重量为 5 的不可能差分路径.

当输出差分重量为 4 时,经过 2 轮逆变换后,各个位置的差分值都不确定. 当输入输出有符合性质 3 中 1)、2),单个字节差分值及几项求和值都不确定,从而不可能差分不存在;当输入输出符合性质 3

中 3)时,考虑 4 项求和的情况,在一侧的求和值为 0,在另一侧的求和值只有 2 种情况:求和值也为 0 或是不确定,从而不存在不可能差分. 至于其他的求和情况,由差分重量为 4 的分布,其求和值也不能构成矛盾. 从而,不存在输出差分重量为 4 的不可能差分路径.

当输出差分重量分别为 3、2、1 时,依据上述引理的证明,可以类似证明.

综上,当输入差分重量为 4 时,不存在输出差分重量 $\leqslant 5$ 的 4 轮 ARIA 的不可能差分路径.

**引理 6** 当输入差分重量为 4 时,存在输出差分重量为 6 的 4 轮不可能差分路径(其形式为 4→ ×←6),并且其输入差分重量分布只能为性质 3 中 2)的形式.

证明 当输入差分重量为 4,输出差分重量为 6 时,输入差分重量经过一次 SL 变换,其差分重量为 4,输出差分重量经过一次 $DL^{-1}$ 变换,其差分重量为 2. 通过分析算法结构知,由于算法各环节变换的对合性,此时寻找的不可能差分路径与引理 3 中证明的 4 轮不可能差分路径输入输出相反,即在引理 3 中得到的 4 轮不可能差分路径 2→ ×←4,前面加一次 DL 变换,后减少一次 $DL^{-1}$ 变换. 由引理 3 的证明,可以得出当输入差分重量为 4 时,存在一类形式如 4→ ×←6 的 4 轮不可能差分路径,并且其输入差分重量分布只能为性质 3 中 2)的形式.

**引理 7** 当输入差分重量为 5 时,不存在输出差分重量 $\leqslant 7$ 的 4 轮 ARIA 的不可能差分路径.

**引理 8** 当输入输出差分重量形如表 2 中的(7)、(8)、(9)、(10)时,不存在 4 轮 ARIA 的不可能差分路径.

根据算法的对合性,引理 7 和引理 8 的证明类似于上述引理的证明,这里不再一一具体给出.

由以上引理的证明可知,针对表 2 列出的 6 轮 ARIA 所有可能的输入输出差分重量为 10 的不可能差分路径仅有 2 类情况成立,并且不存在输入输出差分重量小于 10 的 6 轮 ARIA 的不可能差分路径,从而有以下定理成立.

**定理 1** 不存在输入输出差分重量小于 10 的 6 轮 ARIA 的不可能差分分析,当其输入输出差分重量为 10 时,只存在 2 类形式的不可能差分路径:4→4→ ×←6←6 和 6→2→ ×←4←4.

证明 根据 4 轮不可能差分路径,前后各加一轮即为 6 轮的不可能差分. 由引理 1 至引理 8 的证明,显然可以得到,不存在输入输出差分重量小于 10 的 6 轮 ARIA 的不可能差分分析. 根据引理 3 及引理 6 知,当输入输出差分重量为 10 时,存在这 2 类不可能差分路径;根据其他引理的证明知不存在其他形式的不可能差分路径. 从而,定理成立.

定理 1 给出了使得 ARIA 的不可能差分达到最优的 2 类不可能差分路径. 分析这 2 类路径,仅在差分重量符合性质 3 中 2)的条件成立. 通过程序可以验证得,符合条件的输入计 144 个,从而可以构造出达到最优条件的所有的不可能差分路径.

利用上述不可能差分,采用文献[9]的方法,可以实现 6 轮 ARIA 的不可能差分攻击. 通过分析,利用此路径需要猜测 10 个子密钥字节,与文献[9]所需猜测的密钥字节数目一致. 在文献[9]中,进行第 5 步计算的计算量应为 $2^{32}\times(2^{16}\times 2^{46}+2^{24}\times 2^{38}+\cdots+2^{48}\times 2^{14})=5\times 2^{94}$ 次一轮加密运算,而不是 $5\times 2^{95}$ 次一轮加密运算. 更为一般的结论,根据文献[8-9]的分析,第 4 步与第 5 步计算上没有严格的先后之分,这里我们首先对猜测较多的量进行计算,筛除一部分,而后再进行下一步的计算,这样使得第 5 步的计算量为 $3\times 2^{94}$,从而达到优化的结果,即对 6 轮 ARIA 的不可能差分攻击所需数据量约为 $2^{120}$ 个选择明文和 $2^{94.5}$ 次 6 轮加密运算.

## 4 总结

本文在分析研究 ARIA 扩散层性质基础上,从输入输出差分重量的角度研究了 ARIA 的不可能差分性质. 证明了在输入输出差分重量为 10 的条件下存在且仅存在 2 类形式的 6 轮 ARIA 的最优不可能差分路径,在些基础上,从理论上证明了 6 轮 ARIA 的不可能差分攻击的最优结果为:$2^{120}$ 个选择明文和 $2^{94.5}$ 次 6 轮加密. 在我们的分析过程中,没有考虑密钥的影响,是否存在相关密钥的不可能差分分析,还待作进一步的研究.

## 附录

输入差分重量为 3 时,经过一次 SL 和 DL 变换后,会在 4 个字节位置差分值为 0. 这 4 个位置的分布有 36 种情况,可分为 3 类,如表 3 所示.

**表 3　轮变换后差分值为 0 的分布**

| | | | | | | |
|---|---|---|---|---|---|---|
| Ⅰ | (0,1,10,11) | (0,2,5,7) | (0,3,13,14) | (1,2,12,15) | (1,3,4,6) | (2,3,8,9) |
| | (4,5,14,15) | (4,7,9,10) | (5,6,8,11) | (6,7,12,13) | (8,10,13,15) | (9,11,12,14) |
| Ⅱ | (0,4,10,14) | (0,5,8,13) | (0,7,11,12) | (1,4,9,12) | (1,5,11,15) | (1,6,10,13) |
| | (2,5,9,14) | (2,6,8,12) | (2,7,10,15) | (3,4,8,15) | (3,6,11,14) | (3,7,9,13) |
| Ⅲ | (0,5,10,15) | (0,6,11,13) | (0,7,9,14) | (1,4,11,14) | (1,6,8,15) | (1,7,10,12) |
| | (2,4,9,15) | (2,5,11,12) | (2,7,8,13) | (3,4,10,13) | (3,5,8,14) | (3,6,9,12) |

## 参考文献


[1] Daesung K, Jaesung K, Sangwoo P, et al. New block cipher: ARIA[C]// Proceedings of the Information Security and Cryptology, ICISC'03. Springer-Verlag, LNCS 2971, 2003: 432-445.

[2] Wu W L, Zhang L. The state-of-the-art of research on impossible differential cryptanalysis[J]. Journal of Systems Science and Mathematical Sciences, 2008, 28(8): 971-983(in Chinese).
吴文玲,张蕾. 不可能差分密码分析研究进展[J]. 系统科学与数学, 2008, 28(8):971-983.

[3] Kim J, Hong S, Sung J, et al. Impossible differential cryptanalysis for block cipher structures[C]// Proceedings of Indocrypt 2003. Springer-Verlag, LNCS 2904, 2003: 82-96.

[4] Zhang W T, Wu W L, Feng D G.. New results on impossible differential cryptanalysis of reduced AES[C]// Proceeding of ICISC 2007. LNCS 4817, 2007: 239-250.

[5] Tsunoo Y, Tsujihara E, Shigeri M, et al. Impossible differential cryptanalysis of CLEFIA [C]// FSE 2008. Springer-Verlag, LNCS 5086, 2008: 289-302.

[6] Alex B, Christophe D C, Joseph L, et al. Security and performance analysis of ARIA: Version 1.2[R/OL]. 2003[2010-03-15]. http://homes.esat.kuleuven.be/abiryuko/ARIA-COSICreport.pdf.

[7] Wu W L, Zhang W T, Feng D G. Impossible differential cryptanalysis of reduced-round ARIA and Camellia[J]. Journal of Computer Science and Technology, 2007, 22(3): 449-456.

[8] Peng Z, Ruilin L, Bing S, et al. New impossible differential cryptanalysis of ARIA[R/OL]. [2010-03-15]. Cryptology ePrint Archive, Report 2008. http://eprint.iacr.org/.

[9] Li S H. Cryptanalysis of two symmetric encryption algorithms ARIA and SALSA20[D]. Jinan: Institute of Mathematics and System Science, Shandong University, 2008(in Chinese).
李申华. 对称密码算法 ARIA 和 Salsa20 的安全性分析[D]. 济南:山东大学数学与系统科学学院, 2008.


# Best impossible differential cryptanalysis of 6-round ARIA


ZHANG Lei, GUO Jian-Sheng

(*Institute of Electronic Technology*, *PLA Information Engineering University*, *Zhengzhou* 450004, *China*)



**Abstract**　The security of the block cipher ARIA against impossible differential cryptanalysis is studied. First, we analyze the diffusion layer of ARIA and indicate some differential characters of the intermediate state through the encryption transformation. On the basis of these, we show that there is no 6-round impossible differential with the input-and-output differential weight less than ten and that there are only two kinds of 6-round impossible differential with the input-and-output differential weight of ten. Both kinds of the best impossible differentials can be found and can be used to attack the 6-round ARIA with the best results: the data complexity being $2^{120}$ chosen plaintexts and the time complexity being $2^{94.5}$ encryptions of 6-round ARIA.



**Key words**　block cipher, impossible differential cryptanalysis, ARIA, data complexity, time complexity