Improved Linear Cryptanalysis of Round-Reduced ARIA

Ahmed Abdelkhalek, Mohamed Tolba, and Amr M. Youssef(✉)

Concordia Institute for Information Systems Engineering,
Concordia University, Montréal, Québec, Canada
youssef@ciise.concordia.ca

Abstract. ARIA is an iterated SPN block cipher developed by a group of Korean cryptographers in 2003, established as a Korean standard in 2004 and added to the Transport Layer Security (TLS) supported cipher suites in 2011. It encrypts 128-bit blocks with either 128, 192, or 256-bit key. In this paper, we revisit the security of round-reduced ARIA against linear cryptanalysis and present a 5-round linear hull using the correlation matrix approach to launch the first 8-round key recovery attack on ARIA-128 and improve the 9 and 11-round attacks on ARIA-192/256, respectively, by including the post whitening key. Furthermore, sin all our attacks, we manage to recover the secret master key. The (data in known plaintexts, time in round-reduced encryption operations, memory in 128-bit blocks) complexities of our attacks are ($2^{122.61}$, $2^{123.48}$, $2^{119.94}$), ($2^{122.99}$, $2^{154.83}$, $2^{159.94}$), and ($2^{123.53}$, $2^{238.13}$, $2^{239.95}$) for ARIA-128, ARIA-192, and ARIA-256, respectively.

Keywords: Block cipher · Cryptanalysis · Linear cryptanalysis · ARIA · Key recovery · Linear hull · Correlation matrix

1 Introduction

ARIA is an iterated Substitution Permutation Network (SPN) block cipher that operates on 128-bit blocks with 128, 192 or 256-bit key. It was designed by a group of Korean cryptographers and published in ICISC 2003 [11]. When ARIA was published in ICISC, it had 10/12/14 rounds for key sizes of 128/192/256 bits, respectively, and used 4 distinct S-boxes. In 2004, it was adopted by the Korean Agency for Technology and Standards (KATS) as the Korean 128-bit block encryption algorithm standard after increasing the number of rounds to 12/14/16 and introducing some modifications in the key scheduling algorithm. The life span of ARIA has been extended since then and the latest extension was in December 2014 where its life span was extended for another 5 years (KS X 1213-1:2014) [9]. Since 2011, ARIA is also one of the ciphers that are supported in the Transport Layer Security (TLS) protocol [10].

Since its introduction, the security of ARIA was scrutinized by several cryptographers. After the initial analysis of ARIA by its designers, Biryukov *et al.* [4]

M. Bishop and A.C.A. Nascimento (Eds.): ISC 2016, LNCS 9866, pp. 18–34, 2016.
DOI: 10.1007/978-3-319-45871-7_2

evaluated the security of ARIA against many cryptanalytic techniques. The best attack they developed was based on a 7-round truncated differential. They have also put forward dedicated linear attacks on 7-round ARIA-128 and 10-round ARIA-192/256 in the weak-key setting, i.e., these attacks succeed for a limited number of weak keys. Apart from the cipher designers, Wu *et al.* [24] were the first to evaluate the security of ARIA against impossible differential cryptanalysis. They have proved, in contrast to the designers' expectations, that 4-round impossible differentials do exist and they can be used to mount a 6-round attack on ARIA. The impossible differential attack proposed by Wu *et al.* was independently enhanced by Li and Song [15] and Li *et al.* [21], and then it was extended to 7-round ARIA-256 by Du and Chen [7]. Li *et al.* [14] presented 3-round integral distinguishers that can be used to attack 4/5-round ARIA and 6-round ARIA-192/256. Afterwards, these 3-round integral distinguishers were modified by Li *et al.* [16] to 4-round integral distinguishers which improved the complexity of the 6-round integral attack and extended it to 7-round attack on ARIA-256. Boomerang attacks on 5/6-round ARIA and 7-round ARIA-256 were presented by Fleischmann [8]. Meet-in-the-Middle (MitM) attacks were applied to ARIA for the first time by Tang *et al.* [23], where they presented 5 & 6/7/8 MitM attacks on ARIA-128/192/256, respectively. The complexities of these MitM attacks were further improved by Bai and Yu [3] which enabled them to extend the MitM attacks to 7-round ARIA-128 and 9-round ARIA-256. The complexities of the 7/8-round MitM attacks on ARIA-192/256 were also enhanced by Akshima *et al.* [2] and they presented the first master key recovery attacks on ARIA. Although the designers of ARIA did not expect the existence of effective attacks on 8 or more rounds of ARIA with any key size using linear cryptanalysis, Liu *et al.* [17] managed to attack 7/9/11-round ARIA-128/192/256, respectively, by presenting a special kind of linear characteristics exploiting the diffusion layer employed in ARIA. However, the attacked rounds by Liu *et al.* [17] did not include the post whitening key. This means that if the post whitening key is considered, then the number of the reported rounds in their attacks will be reduced by one for all versions of ARIA. Finally, after the introduction of the Biclique cryptanalysis, it was applied on the full-round ARIA-256 [25], however, this class of attacks is considered as an optimized exhaustive search.

Linear cryptanalysis is one of the major cryptanalysis techniques used against symmetric-key ciphers. It was applied for the first time to FEAL and then to DES by Matsui [18,19]. In linear cryptanalysis, which is a known plaintext attack, the adversary tries to find a linear approximation between some bits from the plaintext, ciphertext and the secret key which can be used as a statistical distinguisher over several rounds of the cipher. Such linear distinguishers are then extended to key-recovery attacks on a few additional rounds using partial decryption and/or encryption. Subkeys of the appended rounds are guessed and the ciphertext is decrypted and/or plaintext is encrypted using these subkeys to calculate intermediate state value at the ends of the distinguisher. If the subkeys are correctly guessed then the distinguisher should hold and it fails, otherwise. After the introduction of linear cryptanalysis, many extensions and improvements have been

proposed. One particular improvement that we use in this paper is the introduction of the notion of linear hull by Nyberg [20]. A linear hull is a set of linear approximations that involve the same bits in the plaintext and ciphertext and each one involves different intermediate state bits. An equally important framework for the description and understanding of the mechanisms of linear cryptanalysis is the concept of correlation matrices of boolean functions which was introduced by Daemen *et al.* [5]. The elements of the correlation matrices of a boolean function F are all the correlation coefficients between linear combinations of input bits and that of output bits of F.

In this paper, we revisit the security of ARIA against linear cryptanalysis. Inspired by the work of Liu *et al.* [17], we first explore all the iterative patterns across ARIA's diffusion layer which have 8 active S-boxes in 2 rounds such as 3-5-3 and 4-4-4. Then, in order to have a good balance between the complexity of the analysis rounds and the number of S-boxes involved in the distinguisher, we focus our attention on the patterns that involve 4 S-boxes in each round, i.e., 4-4-4. Among these patterns, we found 2 patterns that involve only 2 distinct S-boxes (out of the 4 possible distinct S-boxes used in ARIA) in both the even and odd rounds. Then, to simplify our analysis, we focus on these 2 patterns and build their correlation potential matrices to estimate their linear hull effect. In a correlation potential matrix, every element of the correlation matrix is squared. One of these patterns provide a new 5-round linear hull distinguisher with correlation $2^{-114.93}$ which gives us one more round as compared to [17]. Based on this 5-round linear hull, we append 3/4/6 analysis rounds which enables us to mount the first attack on 8-round ARIA-128 and improve the 9 and 11-round attacks on ARIA-192/256, respectively, to include the post whitening key. Further, we use the recovered bytes of information from the round keys to recover the master key. Our results and all previous attacks are summarized in Table 1.

The rest of the paper is organized as follows. Section 2 provides a description of ARIA and the notations adopted in the paper. In Sect. 3, we briefly give the concepts required for the linear cryptanalysis of ARIA. In Sect. 4, we use the correlation potential matrix to establish a linear hull of ARIA and present our 8, 9 and 11-round attacks on ARIA-128/192/256. We also show how the master key can be recovered. Finally, we conclude the paper in Sect. 5.

2 Specification of ARIA

ARIA [12] is an iterative 128-bit block cipher that follows the SPN structure. It can be used with 3 different key lengths, i.e., 128, 192 and 256 bits. The number of rounds in ARIA differs by the key length, i.e., 12 rounds for ARIA-128, 14 rounds for ARIA-192 and 16 rounds for ARIA-256. Similar to AES, the internal state of ARIA can be represented as a 4×4 matrix, where each byte of the matrix is an element in $GF(2^8)$. An ARIA round applies the following three transformations to the state matrix:

– Add Key (AK): XORing a 128-bit round key with the internal state. The round keys are deduced from the master key via the key scheduling algorithm which is described later in this section.

Table 1. Summary of attacks on ARIA

Key size	Rounds	Attack type	Data	Time	Memory	Reference
128/192/256	4	IC	2^{25} CP	2^{25}	$*$	[14]
	5	IDC	$2^{71.3}$ CP	$2^{71.6}$	$2^{72\dagger}$	[21]
	5	IC	$2^{27.2}$ CP	$2^{76.7}$	$2^{27.5\dagger}$	[14]
	5	MitM	25 CP	$2^{65.4}$	2^{121}	[23]
	5	BA	2^{109} ACPC	2^{110}	2^{57}	[8]
	6	IDC	2^{121} CP	2^{112}	$2^{121\dagger}$	[24]
	6	IDC	2^{120} CP	2^{96}	$*$	[15]
	6	IDC	$2^{120.5}$ CP	$2^{104.5}$	$2^{121\dagger}$	[21]
	6	IDC	2^{113} CP	$2^{121.6}$	$2^{113\dagger}$	[21]
	6	MitM	2^{56} CP	$2^{121.5}$	2^{121}	[23]
	6	IC	$2^{99.2}$ CP	$2^{71.4}$	$*$	[16]
	6	BA	2^{128} KP	2^{108}	2^{56}	[8]
	7	TDC	2^{81} CP	2^{81}	2^{80}	[4]
	7	TDC	2^{100} CP	2^{100}	2^{51}	[4]
	7‡	LC	$2^{105.8}$ KP	$2^{100.99}$	$2^{79.73}$	[17]
	7	MitM	2^{121} CP	$2^{125.7}$	2^{122}	[3]
192/256	6	IC	$2^{124.4}$ CP	$2^{172.4}$	$2^{124.4\dagger}$	[14]
	7	MitM	2^{113} CP	2^{132}	2^{130}	[23]
	7	MitM	2^{96} CP	$2^{161.3}$	2^{185}	[23]
	9‡	LC	$2^{108.3}$ KP	$2^{154.83}$	$2^{159.77}$	[17]
	10 ^{wk}	LC	2^{119} KP	2^{119}	2^{63}	[4]
128	7 ^{wk}	LC	2^{77} KP	2^{88}	2^{61}	[4]
	8^{mk}	LC	$2^{122.61}$ KP	$2^{123.48}$	$2^{119.94}$	This paper
192	7	MitM	2^{113} CP	$2^{135.1}$	2^{130}	[2]
	9^{mk}	LC	$2^{122.99}$ KP	$2^{154.83}$	$2^{159.94}$	This paper
256	7	IC	$2^{100.6}$ CP	$2^{225.8}$	$*$	[16]
	7	IDC	2^{125} CP	2^{238}	$*$	[7]
	7	BA	2^{128} KP	2^{236}	2^{184}	[8]
	7^{mk}	MitM	2^{115} CP	$2^{136.1}$	2^{130}	[2]
	8	MitM	2^{56} CP	$2^{251.6}$	2^{250}	[23]
	8	MitM	2^{113} CP	$2^{244.61}$	2^{130}	[3]
	8^{mk}	MitM	2^{56} CP	$2^{251.6}$	2^{252}	[2]
	8^{mk}	MitM	2^{113} CP	$2^{245.9}$	2^{138}	[2]
	9	MitM	2^{121} CP	$2^{253.37}$	2^{250}	[3]
	11‡	LC	$2^{110.3}$ KP	$2^{218.54}$	$2^{239.8}$	[17]
	11^{mk}	LC	$2^{123.53}$ KP	$2^{238.13}$	$2^{239.95}$	This paper
	16^{mk}	BC	2^{80} CP	$2^{255.2}$	$*$	[25]

Time in round-reduced ARIA encryptions and memory in 128-bit blocks

BA: Boomerang Attack

BC: Biclique Cryptanalysis

IC: Integral Cryptanalysis

IDC: Impossible Differential Cryptanalysis

LC: Linear Cryptanalysis

MitM: Meet-in-the-Middle

TDC: Truncated Differential Cryptanalysis

ACPC: Adaptive Chosen Plaintexts and Ciphertext

CP: Chosen Plaintext — KP: Known Plaintext

mk: Recovers the master key — wk: Weak-key setting

$*$: Not given in the related paper — †: Estimated in [8]

‡: Without post whitening key

– SubBytes (SB): Applying non-linear invertible 8-bit to 8-bit S-box to each byte of the state. ARIA employs 4 distinct S-boxes, namely, S_1, S_2 and their inverses S_1^{-1}, S_2^{-1}. Moreover, the order in which the S-boxes are applied to the internal state differs between odd and even rounds. In the odd rounds, the S-boxes are applied, column-wise, in the order: $(S_1, S_2, S_1^{-1}, S_2^{-1})$ while in the even rounds, the order, for each column, is: $(S_1^{-1}, S_2^{-1}, S_1, S_2)$. Figure 1 depicts the order in which the S-boxes are applied in both odd (X_1) and even (X_2) rounds.
– MixState (MS): Multiplication of the internal state by an involutional binary matrix that has a branch number of 8. Given an input state Y, the output state Z of the MS operation is computed as:

$$\begin{pmatrix} Z[0] \\ Z[1] \\ Z[2] \\ Z[3] \\ Z[4] \\ Z[5] \\ Z[6] \\ Z[7] \\ Z[8] \\ Z[9] \\ Z[10] \\ Z[11] \\ Z[12] \\ Z[13] \\ Z[14] \\ Z[15] \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} Y[0] \\ Y[1] \\ Y[2] \\ Y[3] \\ Y[4] \\ Y[5] \\ Y[6] \\ Y[7] \\ Y[8] \\ Y[9] \\ Y[10] \\ Y[11] \\ Y[12] \\ Y[13] \\ Y[14] \\ Y[15] \end{pmatrix}$$

In the last round of ARIA, the MS linear transformation is replaced by an AK operation, which is referred to as the post whitening key. The full encryption function of an r-round ARIA is given in Fig. 1, where the ciphertext C is computed from the plaintext P via r rounds using $r + 1$ round keys.

Key Schedule. The key schedule algorithm of ARIA takes the master key and outputs 13, 15, or 17 128-bit round keys for ARIA-128/192/256, respectively. First, the master key is divided into 2 128-bit values KL and KR, where KL is the leftmost 128-bits of the master key and KR is the remaining bits, if any, of the master key, right-padded with zeros to a 128-bit value. Then, a 3-round, 256-bit Feistel structure, as shown in Fig. 2, is used to compute 4 128-bits words ($W0, W1, W2$, and $W3$), where F_o and F_e denote ARIA odd and even round functions replacing the AK operation with pre-defined constants addition. The

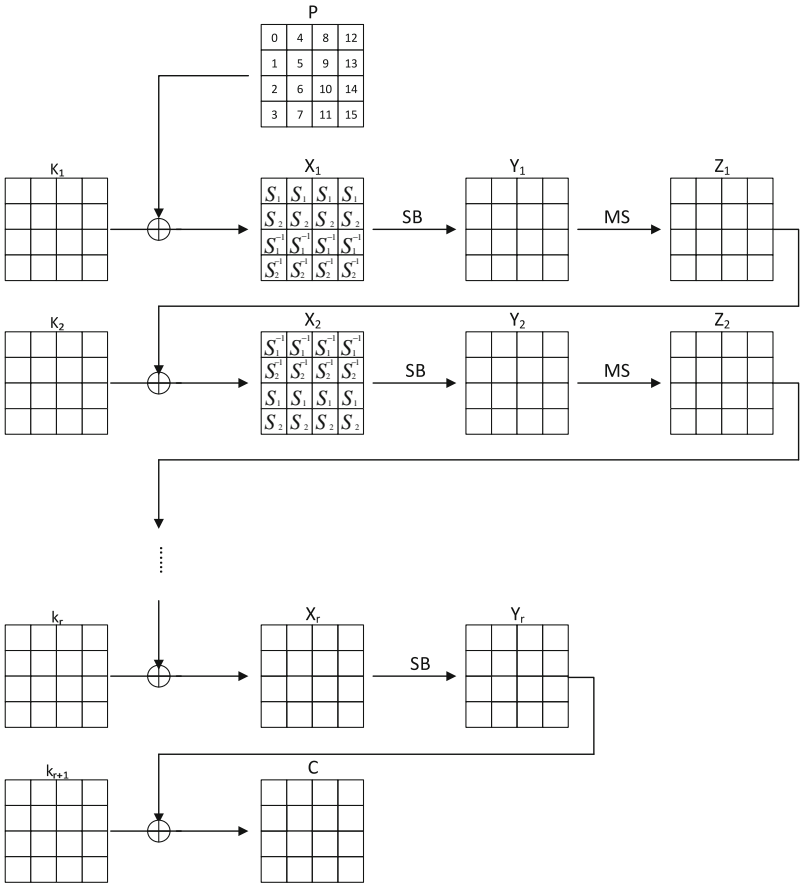

Fig. 1. r-round ARIA

round keys are deduced from $W0, W1, W2$, and $W3$ as follows:

$$
\begin{aligned}
K_1 &= W0 \oplus (W1 \ggg 19), & K_2 &= W1 \oplus (W2 \ggg 19),\\
K_3 &= W2 \oplus (W3 \ggg 19), & K_4 &= (W0 \ggg 19) \oplus W3,\\
K_5 &= W0 \oplus (W1 \ggg 31), & K_6 &= W1 \oplus (W2 \ggg 31),\\
K_7 &= W2 \oplus (W3 \ggg 31), & K_8 &= (W0 \ggg 31) \oplus W3,\\
K_9 &= W0 \oplus (W1 \lll 61), & K_{10} &= W1 \oplus (W2 \lll 61),\\
K_{11} &= W2 \oplus (W3 \lll 61), & K_{12} &= (W0 \lll 61) \oplus W3,\\
K_{13} &= W0 \oplus (W1 \lll 31), & K_{14} &= W1 \oplus (W2 \lll 31),\\
K_{15} &= W2 \oplus (W3 \lll 31), & K_{16} &= (W0 \lll 31) \oplus W3,\\
K_{17} &= W0 \oplus (W1 \lll 19), &&
\end{aligned}
$$

where $a \lll b$ and $a \ggg b$ denote that a is circularly rotated by b bit to the left and right, respectively.

For more detailed information regarding the S-boxes and the key schedule algorithm, the reader is referred to [12].

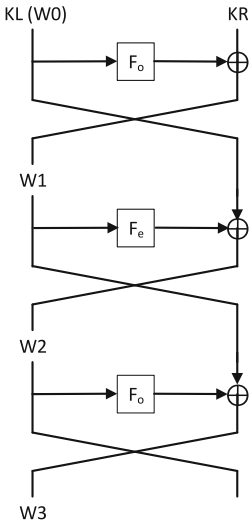

Fig. 2. ARIA key schedule - Initialization phase

2.1 Notations

The following notations are used throughout the rest of this paper:

- I_i: State value at the input of round i, where I_1 is the plaintext P.
- X_i: State value after the AK operation of round i, where X_{R+1} is the ciphertext C and R is 12 for ARIA-128, 14 for ARIA-192, and 16 for ARIA-256.
- Y_i: State value after the SB operation of round i.
- Z_i: State value after the MS operation of round i.
- O_i: State value at the output of round i, i.e., $O_i = I_{i+1}$.
- $S_i[j]$: The $(j+1)^{th}$ byte of state S at round i, where $0 \leq j \leq 15$, as numbered in P in Fig. 1.
- $S_i^k[j]$: The $(j+1)^{th}$ byte of state S^k at round i which corresponds to the plaintext/ciphertext pair (P^k, C^k).
- $K_i^{\{a,b,c,d\}}$: The XOR of 4 bytes of K_i, i.e., $K_i[a] \oplus K_i[b] \oplus K_i[c] \oplus K_i[d]$.

3 Linear Cryptanalysis

As mentioned above, linear cryptanalysis [18,19] is a known plaintext cryptanalysis technique, in which the adversary attempts to construct linear approximations for each round of a block cipher E, such that the output mask of a round equals the input mask of the next round. The concatenation of these linear approximations creates a linear trail (Ω) whose correlation is computed by multiplying the correlations of each round linear approximation. This results in a linear distinguisher covering several rounds of E that can be used to distinguish it from a random permutation. A linear approximation of a block cipher E is typically given by a plaintext mask α and a ciphertext mask β, such that the corresponding correlation $CO_E(\alpha, \beta)$ is non-negligible:

$$CO_E(\alpha, \beta) = |2 \times Pr[\alpha \bullet P \oplus \beta \bullet C = \gamma \bullet K] - 1| \gg 2^{-n/2},$$

where α, β, and γ denote the masks of the plaintext, ciphertext, and key, respectively, n denotes the block length of the cipher and $a \bullet b$ denotes the bitwise inner product of a and b. To distinguish E, the adversary gathers $N = \mathcal{O}(1/CO_E^2(\alpha, \beta))$ plaintexts and their corresponding ciphertexts and computes the empirical correlation $\hat{CO}_E(\alpha, \beta)$:

$$\hat{CO}_E(\alpha, \beta) = |2 \times \#\{i : \alpha \bullet P^i \oplus \beta \bullet C^i = 0\}/N - 1|.$$

The computed empirical correlation is close to $CO_E(\alpha, \beta)$ for the attacked block cipher E, and smaller than $1/\sqrt{N}$, with high probability, for a random permutation [13]. By adding more rounds, the so-called analysis rounds, at the bottom and/or the top of such linear distinguisher, it can be used to perform a key recovery attack using partial decryption and/or encryption. The attack proceeds by guessing the round keys used in the appended rounds, and computing an intermediate state value(s) from the guessed round keys, ciphertext and/or plaintext. The distinguisher is then applied to the deduced intermediate state value(s): if the round keys guess is correct, the distinguisher is expected to hold, and fail for wrong key guesses.

Linear Hulls. The notion of linear hulls was introduced by Nyberg [20], where an r-round linear hull of a block cipher E is a set of all linear trails having the same input mask α, output mask β and can differ in the intermediate masks. If we denote the square of a correlation by correlation potential, then the average correlation potential of a linear hull over r rounds of a key-alternating block cipher, averaged over all values of the expanded key (i.e. the concatenation of all round keys), is the sum of the correlation potentials of all individual trails that compose that linear hull, assuming independent round keys (Theorem 7.9.1 in [6]).

Correlation Matrices. High-probable linear hulls can be found by creating a correlation matrix, or rather a correlation potential matrix, a notion that was introduced by Daemen *et al.* [5]. For a key-alternating cipher of n-bit block

length, a correlation potential matrix M is an $2^n \times 2^n$ matrix where the element M_{ij} in row i and column j of the matrix corresponds to the correlation potential of an input mask α_i and an output mask β_i. Computing M^r gives the correlation potential after r rounds [1]. Constructing the correlation potential matrix for modern block ciphers is infeasible as n is quite large. An alternative approach, then, is to construct a submatrix of the correlation potential matrix that enables us to obtain a lower bound on the average correlation potential of a linear hull.

4 Linear Cryptanalysis of ARIA

Liu *et al.* [17] have proposed a special kind of linear characteristics for byte-oriented SPN block ciphers and applied it on ARIA. Their proposal exploited the MS linear transformation in ARIA by finding a linear relation between 4 bytes of its input and 4 bytes of its output. Then, the linear approximation over one round is formed by applying an input mask α and an output mask β to the XOR of these input/output bytes, i.e.,

$$\alpha \bullet \oplus_{i \in V} I_r[i] = \beta \bullet \oplus_{i \in V} O_r[i],$$

where V is the set of the input/output bytes positions. For example, in their attack $V = \{0, 3, 12, 15\}$.

Inspired by their work, we have first explored the space of all iterative patterns that have 8 active S-boxes in 2 rounds such as 3-5-3 and 4-4-4. We have found that, for 5-round distinguisher and 3 analysis rounds, there is a trade-off between the number of S-boxes involved in the linear characteristic, or rather the linear hull, and the number of key bytes to be guessed in the analysis rounds. On the one hand, the more S-boxes involved in the linear hull, the smaller the correlation potential of the linear hull will be and thus the higher data complexity of the attack will be. On the other hand, the more key bytes to be guessed in the analysis rounds, the higher time complexity will be. Therefore, such a trade-off can be thought of as a trade-off between the data complexity and the time complexity. As an example, in a 3-5-3-5-3-5-3-5 pattern, its first 5-round linear hull involves a total of 19 S-boxes and in its last three analysis rounds, there are 13 key bytes to be guessed as will be illustrated in our attacks later. If the same pattern is shifted by one round to be 5-3-5-3-5-3-5-3, then the number of S-boxes involved in the 5-round distinguisher increases to 21 while the number of key bytes to be guessed in the analysis rounds drops to 11 bytes. The pattern that achieves the balance between the number of S-boxes in the distinguisher and the number of guessed key bytes in the analysis rounds is the pattern 4-4-4.

We have automated the search for all the 4-4-4 patterns across ARIA's MS linear transformation and found that there are 204 such patterns. Among all these patterns, there are only 2 patterns that have 2 active distinct S-boxes even though the order of the application of the S-boxes alternates between the odd and even rounds. The first set of these two patterns is $V1 = \{8, 10, 12, 14\}$ which has S_1 and S_1^{-1} as the active S-boxes in both the odd and even rounds (the gray cells in Fig. 3). The other set is $V2 = \{9, 11, 13, 15\}$ which has S_2 and S_2^{-1} as the

active S-boxes, once again in both the odd and even rounds (the black-hatched cells in Fig. 3). Based on these two patterns, a 1-round linear trail of round i with input mask α and output mask β can be written as:

$$\begin{aligned} V1 :& \alpha \bullet (I_i[8] \oplus I_i[10] \oplus I_i[12] \oplus I_i[14]) = \\ & \beta \bullet (O_i[8] \oplus O_i[10] \oplus O_i[12] \oplus O_i[14]), \\ V2 :& \alpha \bullet (I_i[9] \oplus I_i[11] \oplus I_i[13] \oplus I_i[15]) = \\ & \beta \bullet (O_i[9] \oplus O_i[11] \oplus O_i[13] \oplus O_i[15]). \end{aligned}$$

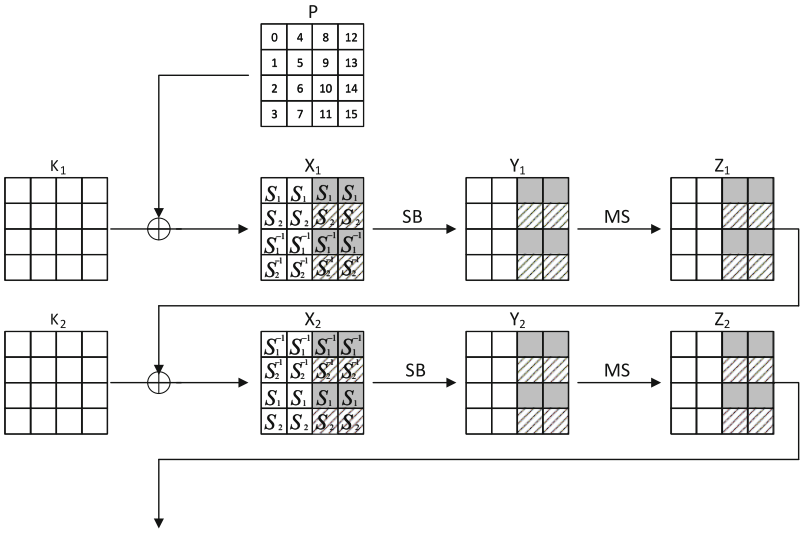

Fig. 3. ARIA 4-4-4 iterative patterns involving 2 distinct S-boxes, each. The gray cells represent pattern $V1$ while the black-hatched cells represent pattern $V2$.

Since both α and $\beta \in GF(2^8)$, the 1-round correlation potential matrix M for each pattern has a size of a $2^8 \times 2^8$ and as the S-boxes involved in these patterns do not change over the odd and even rounds, an M^r correlation potential matrix to get the average correlation potential after r rounds can be constructed by simply raising M to the power r. Such a correlation potential matrix can be regarded as a correlation potential submatrix of ARIA, restricting the inputs and outputs of the matrix to the values that follow our specific patterns. We have automatically constructed the 1-round correlation potential matrix for both patterns. We were not able to go for more than 5 rounds as the highest correlation potential starting M^6 exceeds 2^{-128}. So, for M^5 of pattern $V1$, the highest average correlation potential was found to be $2^{-114.93}$ when the input mask α is $0x09$ and the output mask β is $0x0E$ while for M^5 of pattern $V2$, the highest average correlation potential was found to be $2^{-115.63}$ when the input mask α is $0x24$ and the output mask β is $0xD3$.

4.1 Key Recovery Attacks on ARIA

As the highest correlation potential in $V1$ is greater than the highest one in $V2$, we have opted for using pattern $V1$. In our attacks, we have placed the 5-round linear hull to cover rounds 1–5, hence it is represented as:

$$0x09 \bullet (I_1[8] \oplus I_1[10] \oplus I_1[12] \oplus I_1[14]) = \\ 0x0E \bullet (O_5[8] \oplus O_5[10] \oplus O_5[12] \oplus O_5[14])$$

and since:

$$O_5[8] \oplus O_5[10] \oplus O_5[12] \oplus O_5[14] = \\ X_6[8] \oplus X_6[10] \oplus X_6[12] \oplus X_6[14] \oplus \\ K_6[8] \oplus K_6[10] \oplus k_6[12] \oplus K_6[14]$$

the 5-round linear hull can be re-written as:

$$0x09 \bullet (P[8] \oplus P[10] \oplus P[12] \oplus P[14]) = \\ 0x0E \bullet (X_6[8] \oplus X_6[10] \oplus X_6[12] \oplus X_6[14])$$

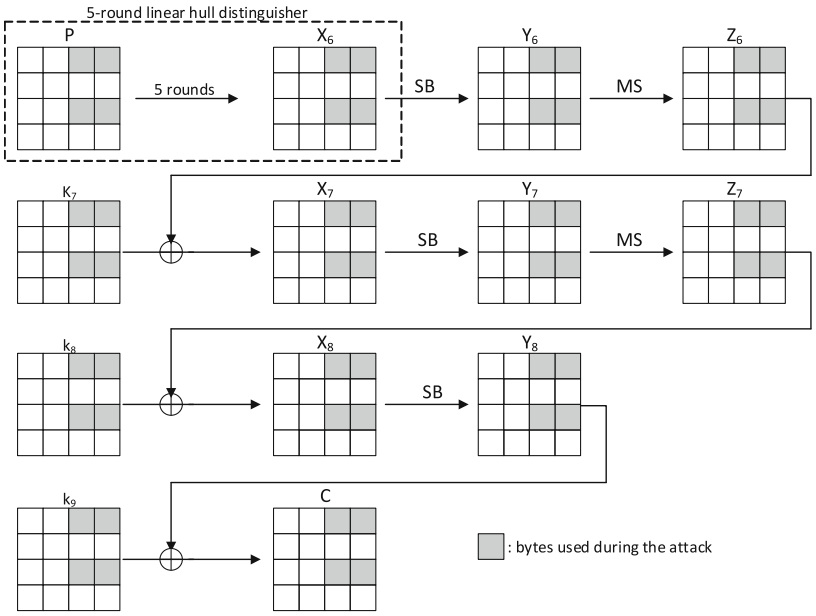

Fig. 4. Attack on 8-round ARIA

ARIA-128. The attack on 8-round ARIA-128 is based on the above 5-round linear hull and adding 3 more rounds at its end, as illustrated in Fig. 4. The attack proceeds as follows:

1. First, we gather N plaintexts and their corresponding ciphertexts (P^i, C^i), where $1 \leq i \leq N$.
2. Next, we initialize 2^{32} counters U_m, where the size of each counter is $\lceil log_2 N \rceil$ bits and $0 \leq m \leq 2^{32} - 1$. Then, for each plaintext/ciphertext pair (P^i, C^i), we increment (resp. decrement) the counter U_m by 1 if the parity of

$$0x09 \bullet (P^i[8] \oplus P^i[10] \oplus P^i[12] \oplus P^i[14])$$

   is 0 (resp. 1) and m equals the value of $C^i[8]\|C^i[10]\|C^i[12]\|C^i[14]$.
3. We initialize 2^{120} counters U_l, where the size of each counter is $\lceil log_2 N \rceil$ bits as well and $0 \leq l \leq 2^{120} - 1$ and l represents the possible value of the 15 bytes of $K_9^{\{8,10,12,14\}}\|K_8^{\{8,10,12,14\}}\|K_7^{\{8,10,12,14\}}\|Y_8[8]\ \|Y_8[10]\|Y_8[12]\|Y_8[14]\|Y_7[8]\|Y_7[10]\|Y_7[12]\|Y_7[14]\|Y_6[8]\|Y_6[10]\|Y_6[12]\|Y_6[14]$.
4. Then, for each possible value of $K_9^{\{8,10,12,14\}}$, $K_8^{\{8,10,12,14\}}$ and $K_7^{\{8,10,12,14\}}$, we do the following:
   (a) For each possible value of the 2^{32} values of m, compute $Y_8^m[8] \oplus Y_8^m[10] \oplus Y_8^m[12] \oplus Y_8^m[14] = m[0] \oplus m[1] \oplus m[2] \oplus m[3] \oplus K_9^{\{8,10,12,14\}}$ and denote this value as t_8^m.
   (b) For any value of the 2^{24} values of $Y_8[8]\|Y_8[10]\|Y_8[12]\|Y_8[14]$ satisfying t_8^m, we deduce $X_8^m[8], X_8^m[10], X_8^m[12], X_8^m[14]$ from the corresponding S-boxes. Then, using the guessed value of $K_8^{\{8,10,12,14\}}$, compute $Y_7^m[8] \oplus Y_7^m[10] \oplus Y_7^m[12] \oplus Y_7^m[14] = Z_7^m[8] \oplus Z_7^m[10] \oplus Z_7^m[12] \oplus Z_7^m[14] = X_8^m[8] \oplus X_8^m[10] \oplus X_8^m[12] \oplus X_8^m[14] \oplus K_8^{\{8,10,12,14\}}$ and denote this value as t_7^m.
   (c) Then, for any value of the 2^{24} values of $Y_7[8]\|Y_7[10]\|Y_7[12]\|\ Y_7[14]$ satisfying t_7^m, we deduce $X_7^m[8], X_7^m[10], X_7^m[12], X_7^m[14]$ from the corresponding S-boxes. Then, using the guessed value of $K_7^{\{8,10,12,14\}}$, compute $Y_6^m[8] \oplus Y_6^m[10] \oplus Y_6^m[12] \oplus Y_6^m[14] = Z_6^m[8] \oplus Z_6^m[10] \oplus Z_6^m[12] \oplus Z_6^m[14] = X_7^m[8] \oplus X_7^m[10] \oplus X_7^m[12] \oplus X_7^m[14] \oplus K_7^{\{8,10,12,14\}}$ and denote this value as t_6^m.
   (d) For any value of the 2^{24} values of $Y_6[8]\|Y_6[10]\|Y_6[12]\|Y_6[14]$ satisfying t_6^m, we deduce $X_6^m[8], X_6^m[10], X_6^m[12], X_6^m[14]$ from the corresponding S-boxes. Then, calculate the parity of:

$$0x0E \bullet (X_6^m[8] \oplus X_6^m[10] \oplus X_6^m[12] \oplus X_6^m[14])$$

   If the parity is 0 (resp. 1), increment (resp. decrement) the corresponding counter U_l by the value of U_m.
5. For l such that the value of U_l is maximal, output the value of the corresponding $K_9^{\{8,10,12,14\}}\|\ K_8^{\{8,10,12,14\}}\|K_7^{\{8,10,12,14\}}$ as the correct key information.

Attack complexity. The number of known plaintext/ciphertext pairs N required to perform the attack is estimated by the following formula, which is adopted from Corollary 1 in [22]:

$$N = \left(\frac{\Phi^{-1}(P_s) + \Phi^{-1}(1 - 2^{-a-1})}{2} \right)^2 \times \frac{4}{CO^2}, \tag{1}$$

where P_s is the probability of success, CO is the correlation of the linear hull, Φ^{-1} is the inverse cumulative function of the standard normal distribution, and a is the advantage of the adversary over the exhaustive search and equals $k - log_2\ d$ if the correct key was ranked among the top d candidates out of the 2^k possible candidates of an k-bit key.

In our attack, we guess 120 bits, set the advantage a to 120, i.e., the correct key information is the first one of the list of candidates and set the probability of success to 0.95. Then, the number of plaintext/ciphertext pairs N equals $2^{5.68} \times \frac{4}{2^{-114.93}} = 2^{122.61}$. The time complexity of the attack is dominated by steps 2 and 4.(d). Therefore, the time complexity of the attack equals $2^{122.61} \times \frac{4}{16\times 8} + 2^{24} \times 2^{32} \times 2^{24} \times 2^{24} \times 2^{24} \times \frac{4}{16\times 8} \approx 2^{123.03}$ 8-round ARIA encryptions. The memory complexity of the attack is attributed to storing the counters U_l, where the size of each counter is set to 123 bits. Hence, the memory complexity of the attack is $2^{120} \times \frac{123}{128} \approx 2^{119.94}$ 128-bit blocks.

ARIA-192/256. The attack on 8-round ARIA-128 can be extended to 9-round ARIA-192 (resp. 11-round ARIA-256) with the post whitening key by utilizing the same 5-round linear hull and having 4 (resp. 6) analysis rounds. The attack procedure is similar to the attack on ARIA-128, except that in step 3, we initialize 2^{160} (resp. 2^{240}) counters U_l, where in this case, $0 \leq l \leq 2^{160} - 1$ (resp. $0 \leq l \leq 2^{240} - 1$) and represents the possible value of the 20 (resp. 30) bytes of $K_{r+1}^{\{8,10,12,14\}} \| Y_r[8] \| Y_r[10] \| Y_r[12] \| Y_r[14]$, where $6 \leq r \leq 9$ (resp. $6 \leq r \leq 11$) and we add 1 (resp. 3) more sub-step(s) in step 4 to accommodate the additional round(s).

In this case, for an advantage a of 160 (resp. 240) and P_s of 0.95, the number of known plaintext/ciphertext pairs N is $2^{6.06} \times \frac{4}{2^{-114.93}} = 2^{122.99}$ for ARIA-192 and is $2^{6.6} \times \frac{4}{2^{-114.93}} = 2^{123.53}$ for ARIA-256. The time complexity of the attack is $2^{122.99} \times \frac{4}{16\times 9} + 2^{32} \times 2^{32} \times 2^{24} \times 2^{24} \times 2^{24} \times 2^{24} \times \frac{4}{16\times 9} \approx 2^{154.83}$ 9-round ARIA encryptions for ARIA-192 and is $2^{123.53} \times \frac{4}{16\times 11} + 2^{48} \times 2^{32} \times 2^{24} \times 2^{24} \times 2^{24} \times 2^{24} \times 2^{24} \times 2^{24} \times \frac{4}{16\times 11} \approx 2^{218.54}$ 11-round ARIA encryptions for ARIA-256. The size of each counter of U_l is set to 123 (resp. 124) bits, therefore, the memory complexity of the attack is $2^{160} \times \frac{123}{128} \approx 2^{159.94}$ 128-bit blocks for ARIA-192 and $2^{240} \times \frac{124}{128} \approx 2^{239.95}$ 128-bit blocks for ARIA-256.

4.2 Recovering the Master Key

In this subsection, we show how the recovered bytes of information from the round keys can be used to recover the master key in all versions of ARIA.

ARIA-128. In the attack on 8-round ARIA-128, we recover 3 bytes of information from K_9, K_8, and K_7. Recall that in ARIA-128, KR is all zeros and KL is the 128-bit master key and at the same time it is $W0$. In order to recover the master key, we do the following:

- First, we guess 15 bytes of $W0$, i.e., all the bytes except $W0[7]$. These bytes enable us to compute $W1[0], W1[2]$, $W1[4]$, with 6 other bytes, which gives us the first 5 bits of bytes $8, 10$, and 12 of $(W1 \lll 61)$.
- From the key schedule, we know that $K_9 = W0 \oplus (W1 \lll 61)$. As we recover $K_9^{\{8,10,12,14\}}$, i.e., $K_9[8] \oplus K_9[10] \oplus K_9[12] \oplus K_9[14]$, this means that we recover $W0[8] \oplus W0[10] \oplus W0[12] \oplus W0[14] \oplus (W1 \lll 61)[8] \oplus (W1 \lll 61)[10] \oplus (W1 \lll 61)[12] \oplus (W1 \lll 61)[14]$.
- As we guessed $W0[8], W0[10], W0[12]$ and $W0[14]$, recovered $K_9^{\{8,10,12,14\}}$ and computed the first 5 bits of bytes $8, 10$, and 12 of $(W1 \lll 61)$, we can deduce the first 5 bits of byte 14 of $(W1 \lll 61)$ which in turn enables us to deduce the last 5 bits of $SB(W0[7])$.
- Afterwards, we guess the 3 first bits of $SB(W0[7])$ which means that we have 2^{123} candidates for $W0$ or rather the master key.
- Then, we run the key schedule and use the remaining 3 bits of $K_9^{\{8,10,12,14\}}$ and the two bytes of $K_8^{\{8,10,12,14\}}$ and $K_7^{\{8,10,12,14\}}$ to discard the wrong guesses and so we end up with 2^{104} candidates for the master key which we can test using 2 plaintext/ciphertext pairs.

The time complexity of the master key recovery phase is dominated by the last step and equals $2^{123} \times \frac{3}{8} + 2 \times 2^{104} \approx 2^{121.59}$ 8-round ARIA encryptions as we need to compute 3 rounds of ARIA for the 2^{123} candidates to deduce $W2$ and $W3$ and then test the remaining 2^{104} candidates using 2 plaintext/ciphertext pairs. Therefore the total time complexity of the attack is $2^{123.03} + 2^{121.59} \approx 2^{123.48}$.

ARIA-192. In the attack on 9-round ARIA-192, we recover 4 bytes of information from K_{10}, K_9, K_8, and K_7. In order to recover the master key, we do the following:

- First, we guess the 16 bytes of $W0$ and calculate $F_o(W0, CK1)$. Then, to be able to compute bytes $8, 10, 12$ and 14 of $(W1 \lll 61)$, we guess 29 bits of KR as the 8 right bytes of KR are zeros.
- We use the recovered $K_9^{\{8,10,12,14\}}$ to discard the wrong guesses of $W0$ and the 29 bits guessed from KR and so we have 2^{149} for $W0$ along with the 29 bits of KR.
- Next, we guess the remaining 35 bits of the master key, i.e., the remaining 35 bits of KR so we have 2^{184} candidates for the master key.
- Then, we run the key schedule to compute $W1, W2$, and $W3$ and use the 3 bytes of $K_{10}^{\{8,10,12,14\}}$, $K_8^{\{8,10,12,14\}}$ and $K_7^{\{8,10,12,14\}}$ to discard the wrong guesses and we end up with 2^{160} candidates which we test using 2 plaintext/ciphertext pairs.

The time complexity of the master key recovery phase equals $2^{184} \times \frac{3}{9} + 2 \times 2^{160} \approx 2^{182.42}$ 9-round ARIA encryptions, hence the total time complexity of the attack is $2^{154.83} + 2^{182.42} \approx 2^{182.42}$.

ARIA-256. In the attack on 11-round ARIA-256, we recover 6 bytes of information from $K_{12}, K_{11}, K_{10}, K_9, K_8$, and K_7. In order to recover the master key, we do the following:

– First, we guess the 16 bytes of $W2$ and 14 bytes of $W3$ and use the recovered $K_7^{\{8,10,12,14\}}$ and $K_{11}^{\{8,10,12,14\}}$, both of them are deduced from $W2$ and $W3$, to calculate the remaining two bytes of $W3$ which means that we have 2^{240} candidates for both $W2$ and $W3$.
– Next, starting from $W2$ and $W3$, we run the key schedule to compute $W0$ and $W1$ and use the other 4 bytes of $K_{12}^{\{8,10,12,14\}}$, $K_{10}^{\{8,10,12,14\}}$, $K_9^{\{8,10,12,14\}}$ and $K_8^{\{8,10,12,14\}}$ to discard the wrong guesses and we end up with 2^{208} candidates for the master key which we test using 2 plaintext/ciphertext pairs.

The time complexity of the master key recovery phase is $2^{240} \times \frac{3}{11} + 2 \times 2^{208} \approx 2^{238.13}$ 11-round ARIA encryptions, therefore the total time complexity of the attack is $2^{218.54} + 2^{238.13} \approx 2^{238.13}$.

5 Conclusion

In this paper, we have revisited the security of round-reduced ARIA against linear cryptanalysis and presented the first 8-round attack on ARIA-128 and improved the previous 9 and 11-round attacks on ARIA-192/256 by including the post whitening key. We have achieved these results by constructing a 5-round linear hull on ARIA using the correlation matrix approach and exploiting the binary linear transformation layer in the analysis rounds. For all our attacks, we showed how the recovered bytes of information from the round keys can be used to recover the master key. This paper shows some weaknesses of reduced versions of ARIA, but the full round ARIA remains still secure.

References

1. Abdelraheem, M.A., Alizadeh, J., Alkhzaimi, H.A., Aref, M.R., Bagheri, N., Gauravaram, P.: Improved linear cryptanalysis of reduced-round SIMON-32 and SIMON-48. In: Biryukov, A., Goyal, V. (eds.) Progress in Cryptology - INDOCRYPT 2015. LNCS, vol. 9462, pp. 153–179. Springer, Cham (2015). http://dx.doi.org/10.1007/978-3-319-26617-6_9
2. Biryukov, A., Goyal, V. (eds.): Progress in Cryptology – INDOCRYPT 2015. LNCS, vol. 9462. Springer, Cham (2015). http://dx.doi.org/10.1007/978-3-319-26617-6_11
3. Bai, D., Yu, H.: Improved meet-in-the-middle attacks on round-reduced ARIA. In: Desmedt, Y. (ed.) ISC 2013. LNCS, vol. 7807, pp. 155–168. Springer, Heidelberg (2015). http://dx.doi.org/10.1007/978-3-319-27659-5_11
4. Biryukov, A., De Canniere, C., Lano, J., Ors, S.B., Preneel, B.: Security and performance analysis of ARIA, version 1.2. Technical report, Katholieke Universiteit Leuven, Belgium (2004).http://www.cosic.esat.kuleuven.be/publications/article-500.pdf
5. Daemen, J., Govaerts, R., Vandewalle, J.: Fast Software Encryption. LNCS, vol. 1008. Springer, Heidelberg (1995). http://dx.doi.org/10.1007/3-540-60590-8_21
6. Daemen, J., Rijmen, V.: The Design of Rijndael. Springer, New York (2002)

7. Du, C., Chen, J.: Impossible differential cryptanalysis of ARIA reduced to 7 rounds. In: Heng, S.-H., Wright, R.N., Goi, B.-M. (eds.) CANS 2010. LNCS, vol. 6467, pp. 20–30. Springer, Heidelberg (2010). http://dx.doi.org/10.1007/978-3-642-17619-7_2
8. Fleischmann, E., Forler, C., Gorski, M., Lucks, S.: New boomerang attacks on ARIA. In: Gong, G., Gupta, K.C. (eds.) INDOCRYPT 2010. LNCS, vol. 6498, pp. 163–175. Springer, Heidelberg (2010). http://dx.doi.org/10.1007/978-3-642-17401-8_13
9. Korean Agency for Technology and Standards (KATS): 128-bit Block Encryption Algorithm ARIA KS X 1213–1: December 2014 (in Korean)
10. Kim, W., Lee, J., Park, J., Kwon, D.: Addition of the ARIA cipher suites to Transport Layer Security (TLS). RFC 6209, RFC Editor, April 2011. http://www.rfc-editor.org/rfc/rfc6209.txt, http://www.rfc-editor.org/rfc/rfc6209.txt
11. Daesung, K., et al.: Information Security and Cryptology - ICISC 2003. LNCS, vol. 2971. Springer, Heidelberg (2004). http://dx.doi.org/10.1007/978-3-540-24691-6_32
12. Lee, J., Lee, J., Kim, J., Kwon, D., Kim, C.: A Description of the ARIA Encryption Algorithm. RFC 5794, RFC Editor, March 2010
13. Leurent, G.: Improved differential-linear cryptanalysis of 7-round chaskey with partitioning. Cryptology ePrint Archive, Report 2015/968 (2015). http://eprint.iacr.org/
14. Li, P., Sun, B., Li, C.: Integral cryptanalysis of ARIA. In: Bao, F., Yung, M., Lin, D., Jing, J. (eds.) Inscrypt 2009. LNCS, vol. 6151, pp. 1–14. Springer, Heidelberg (2010). http://dx.doi.org/10.1007/978-3-642-16342-5_1
15. Li, S., Song, C.: Improved impossible differential cryptanalysis of ARIA. In: A Description of the ARIA Encryption Algorithm. RFC 5794, RFC Editor International Conference on Information Security and Assurance, ISA 2008, pp. 129–132, April 2008
16. Li, Y., Wu, W., Zhang, L.: Integral attacks on reduced-round ARIA block cipher. In: Kwak, J., Deng, R.H., Won, Y., Wang, G. (eds.) ISPEC 2010. LNCS, vol. 6047, pp. 19–29. Springer, Heidelberg (2010). http://dx.doi.org/10.1007/978-3-642-12827-1_2
17. Liu, Z., Gu, D., Liu, Y., Li, J., Li, W.: Linear cryptanalysis of ARIA block cipher. In: Qing, S., Susilo, W., Wang, G., Liu, D. (eds.) ICICS 2011. LNCS, vol. 7043, pp. 242–254. Springer, Heidelberg (2011). http://dx.doi.org/10.1007/978-3-642-25243-3_20
18. Matsui, M.: Linear cryptanalysis method for DES cipher. In: Helleseth, T. (ed.) EUROCRYPT 1993. LNCS, vol. 765, pp. 386–397. Springer, Heidelberg (1994). http://dx.doi.org/10.1007/3-540-48285-7_33
19. Matsui, M., Yamagishi, A.: A new method for known plaintext attack of FEAL cipher. In: Rueppel, R.A. (ed.) EUROCRYPT 1992. LNCS, vol. 658, pp. 81–91. Springer, Heidelberg (1993). http://dx.doi.org/10.1007/3-540-47555-9_7
20. Nyberg, K.: Linear approximation of block ciphers. In: De Santis, A. (ed.) EUROCRYPT 1994. LNCS, vol. 950, pp. 439–444. Springer, Heidelberg (1995). http://dx.doi.org/10.1007/BFb0053460
21. Li, R., Bing Sun, P.Z., Li, C.: New Impossible Differential Cryptanalysis of ARIA. Cryptology ePrint Archive, Report 2008/227 (2008). http://eprint.iacr.org/2008/227.pdf
22. Selçuk, A.A.: On probability of success in linear and differential cryptanalysis. J. Cryptology **21**(1), 131–147 (2007). http://dx.doi.org/10.1007/s00145-007-9013-7

23. Tang, X., Sun, B., Li, R., Li, C., Yin, J.: A meet-in-the-middle attack on reduced-round ARIA. J. Syst. Softw. **84**(10), 1685–1692 (2011). http://www.sciencedirect.com/science/article/pii/S016412121100104X
24. Wu, W.L., Zhang, W.T., Feng, D.G.: Impossible differential cryptanalysis of reduced-round ARIA and Camellia. J. Comput. Sci. Technol. **22**(3), 449–456 (2007). http://dx.doi.org/10.1007/s11390-007-9056-0
25. zhen Chen Tian-min Xu, S.: Biclique Attack of the Full ARIA-256. Cryptology ePrint Archive, Report 2012/011 (2012). http://eprint.iacr.org/2012/011.pdf