A Deep Study of The Impossible Boomerang Distinguishers: New Construction Theory and Automatic Search Methods

Xichao Hu[1], Dengguo Feng[1], Lin Jiao[1], Yonglin Hao[1], Xinxin Gong[1], Yongqiang Li[2,3]

[1] State Key Laboratory of Cryptology, Beijing, China xchao_h@163.com
[2] Key Laboratory of Cyberspace Security Defense, Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China
[3] School of Cyber Security, University of Chinese Academy of Sciences, Beijing, China

Abstract. The impossible boomerang attack (IBA) is a combination of the impossible differential attack and boomerang attack, which has demonstrated remarkable power in the security evaluation of AES and other block ciphers. However, this method has not received sufficient attention in the field of symmetric cipher analysis. The only existing search method for impossible boomerang distinguishers (IBD), the core of IBAs, is the $\mathcal{UB}$-method, but it is considered rather rudimentary given current technological advancements and may result in missed opportunities for effective attacks. Therefore, this paper delves into a comprehensive study on the construction theory and automatic search method of IBDs.

Theoretically, we propose 5 IBD constructions aligned with the techniques of arbitrary S-box, boomerang distinguisher, Boomerang Connectivity Table, U/L/EBCT and mixed tables for differential propagation for SPN-network block ciphers, and 2 IBD constructions accompanied by state propagation for block ciphers with any structure. Furthermore, we investigate the relationship among these IBD constructions and demonstrate that the most superior IBD aligns precisely with the original definition. Technically, we develop a general SAT-based automatic search tool for IBDs by introducing optimized search strategies of the composite model method and the mixed model method. This tool not only considers the details of each operation but also takes into account the impact of key schedule in a single-key setting.

As applications, we first acquire 59584 4-round 1 active word truncated IBDs for AES-128, and 192 of those IBDs cannot be detected by the $\mathcal{UB}$-method. For Midori64, we first demonstrate the non-existence of 7-round 1 active word truncated IBDs, and obtain 7296 6-round 1 active word truncated IBDs, which is complementary to the finding that there are no existing 6-round 1 active word truncated IDs. For PRESENT-80, we get the first 6-round IBDs which cannot be detected by the $\mathcal{UB}$-method. Those results indicate that our method outperforms the $\mathcal{UB}$-method and offer an advantage over IDs. We believe that our work can bring new insights to symmetric cipher analysis.

Keywords: Impossible Boomerang Distinguishers · Propagation of S-tates · Composite Model Method · Mixed Model Method.

1 Introduction

The differential attack, proposed by Biham and Shamir [1], is considered one of the most crucial methods for analyzing the security of block ciphers. Its fundamental idea involves identifying a high-probability differential characteristic to be a differential distinguisher, and then adding specific rounds at the beginning and end of the distinguisher to recover the key. Provable security against a d-ifferential attack has become a significant consideration in the design of block ciphers. Numerous cryptanalytic techniques have been developed based on the principles of differential attacks, with two well-known approaches being impossible differential attacks and boomerang attacks.

The impossible differential attack was proposed by Biham et al. and Knudsen to attack Skipjack [2] and DEAL [3] respectively. This approach serves as a complement to differential attack, with the distinguisher of the impossible differential (ID) being its zero probability of occurrence. Additionally, the boomerang attack proposed by Wagner at the same time [4], represents a variant of differential attack. The main idea is to connect two (or more) short high-probability differential characteristics so as to generate a boomerang distinguisher (BD) to analyze more rounds of the primitive. There is no doubt that these two attacks have played a very important role in the security analysis of block ciphers [5,6,7,8,9].

The impossible boomerang attack (IBA) is proposed by J. Lu [10]. This attack combines the concepts of impossible differential attack and boomerang attacks, utilizing an impossible boomerang distinguisher (IBD). Similar to a boomerang attack, a block cipher E is treated as two sub-ciphers $E_0 \circ E_1$. Two (or more) differentials with probability 1 for E_0 and two (or more) differentials with probability 1 for E_1 are employed, where the XOR of the intermediate differences of these differentials is not equal to zero. In [10,11], the impossible boomerang attack was utilized to successfully break 6-round AES-128, 7-round AES-192 and 7-round AES-256 in a single key attack scenario, as well as 8-round AES-192 and 9-round AES-256 in a related-key attack scenario involving two keys, based on a 4-round IBD.

The automatic search method can search for distinguishers effectively, thoughtfully and precisely. Given the existing solvers for the Boolean Satisfiability Problem (SAT)/Satisfiability Modulo Theories (SMT) problem [12,13,14], the Mixed Integer Linear Programming (MILP) problem [15,16], and the Constraint Programming (CP) problem [17,18], cryptographers typically convert search problems into these mathematical problems to achieve the automatic search for distinguishers. Automatic tools for cryptanalysis are increasingly influential in both symmetric cipher design and analysis.

The initial work to search for the IDs using such automatic methods is documented in [19]. In their study, Cui et al. proposed a MILP-based tool to search for the IDs of lightweight block ciphers, considering the detailed propagation of differences through each operation. Subsequently, Sasaki and Todo [20] presented a further MILP-based tool for searching the IDs of SPN block ciphers, by introducing arbitrary S-box (AS) mode. That is, the large S-boxes are treated as permutations only to identify contradictions in the linear components, which demonstrated its applicability for block ciphers based on large S-boxes. In [21], Hu et al. presented a SAT/SMT-based tool to search for the IDs, by describing the state propagation simultaneously. It allows the search of IDs considering the specific key schedule in the single-key scenario.

In recent years, significant advancements have been made in the technique of constructing BDs. Initially, the original concepts of the boomerang attack postulated that the two sub-ciphers E_0 and E_1 were independent of each other. However, Murphy [22] highlighted that two independently chosen characteristics might be incompatible, leading to a probability of generating a right quartet of plaintext-ciphertext pairs being zero. Furthermore, numerous improvements considering the dependence between the two differential characteristics have been proposed, including techniques such as the middle round S-box trick [23], ladder switch, S-box switch and Feistel switch [7]. These insights can be encapsulated within the framework of the sandwich attack proposed by Dunkelman et al. [8,24]. It divides the block cipher E into three parts $E_1 \circ E_m \circ E_0$, where the upper part E_0 and the lower part E_1 are covered by ordinary differential distinguishers, while the middle part E_m is subject to a small boomerang distinguisher that connects the two parts by specified input difference and output difference. The main role of the middle part E_m is to take the dependency between E_0 and E_1 into account while computing the probability of the BD distinguisher. Recently, new insights on what exactly happens in the middle part E_m have been investigated. At Eurocrypt 2018, Cid et al. [25] presented the Boomerang Connectivity Table (BCT), a tool facilitating the straightforward evaluation of the BD's probability of the middle part E_m in the single-round scenario for SPN network. Subsequently, Wang et al. [26] proposed the Boomerang Difference Table (BDT) and its variant BDT', enabling systematic evaluation of boomerang switching effect in the multiple rounds involved scenario. Furthermore, in [27], Boukerrou et al. generalized the BCT and BDT to feistel network and proposed the concept of FBCT. Subsequently, Delaune et al. [28] proposed a CP-based method to search for BDs, and renamed BDT and BDT' as UBCT and LBCT for upper BCT and lower BCT, respectively. Additionally, they defined the EBCT for SPN network based on the definition of FBCT for Feistel network. In [29], a SAT-based tool was presented for discovering BDs in ARX ciphers.

As is demonstrated in [10,11,30], IBA exhibits great strength in the security evaluation of block ciphers, indicating its high potency as a cryptanalytic technique. Consequently, the number of rounds resistant to IBD reflects the security level of block ciphers. However, the only known method for searching for IBDs at present is the $\mathcal{UB}$-method. This method employs the concept of the miss-in-the-middle approach to construct IBDs. The core idea is to transform the differential propagation into the manipulation of a matrix, and to seek contradictions by defining certain criteria. The $\mathcal{UB}$-method has the following limitations.

- **Unable to take into account the details of S-box.** This method merely regards the S-box as a permutation, omitting the detailed properties of the differential propagation through the S-box.

- **Unable to take into account the details of linear layer.** In fact, this method roughly depicts the differential propagation via the Xor operation, and naturally it is unable to take into account the details of the linear layer as the differential propagation through the linear layer is based on the Xor in this method.
- **Unable to take into account the key schedule in the single-key setting.** This method constructs the IBDs through the differential propagation, and in this way it naturally counteracts the influence of the key schedule.

Our contributions. Motivated by both the strong threat of IBA method (e.g. powerful attacks on AES) and its significant lack of systematic theory and general search models, we initiate the comprehensive research work on its core, constructing IBDs, synchronously related to the theoretical development of boomerang attacks and impossible differential attacks.

Firstly, we have established a new theoretical framework for constructing IBDs. We define a series of IBDs from both the perspectives of differential propagation and state propagation as follows.

T_0**-IBD:** the IBD regarding the S-box appearing in a S-box based block cipher as only a permutation.

T_1**-IBD:** the IBD constructed based on DDT purely for S-box based block ciphers, which corresponds to the only existing method proposed in [10].

T_2**-IBD:** the IBD constructed based on our newly defined GBCT for SPN-network block ciphers, which is a generalization of BCT in context with BD.

T_3**-IBD:** the IBD constructed based on our newly defined GEBCT merely for SPN-network block ciphers, which is a generalization of EBCT in context with BD.

T_P**-IBD:** the IBD constructed by a pre-defined propagation rule P based on a mixed use of our newly defined DDT^2_{upper}, DDT^2_{lower}, GBCT, GUBCT, GLBCT, GEBCT for SPN-network block ciphers, which are a generalization of (U/L/E)BCT in context with BD.

T_4**-IBD:** the IBD constructed through the propagation of pure states in the case of where the round keys are mutually independent for a block cipher.

T_5**-IBD:** the IBD constructed through the propagation of pure states considering the key schedule for a block cipher.

We further prove the inclusion relations between these newly-defined IBDs. Let S_{T_i} be the set containing all T_i-IBDs, and then we derive that

$$S_{T_0} \subseteq S_{T_1} \subseteq S_{T_2} \subseteq S_{T_3} = S_{T_4} \subseteq S_{T_5}.$$
$$S_{T_P} \subseteq S_{T_3}$$

Specifically,

- For $0 \leq i \leq 4$, an r-round T_i-IBD is always an r-round T_{i+1}-IBD.
- $S_{T_3} = S_{T_4}$ means that T_3-IBD is equivalent with T_4-IBD within SPN-network block ciphers.

- $S_{T_P} \subseteq S_{T_3}$ means that any construction method based on the mixed use of DDT and G(U/L/E)BCT cannot be superior to T_3-IBD.

Moreover, we prove and choose the most superior IBD by the following conclusions.

- All r-round T_i-IBDs ($0 \leq i \leq 4$) and T_P-IBDs are r-round T_5-IBDs. That is, any construction method based on DDT and G(U/L/E)BCT will not be superior to T_5-IBD.
- The definition of T_5-IBD is equivalent to origin definition of IBD in [10]. That is, the construction of T_5-IBD is the tightest method for constructing IBDs.

Therefore, our ultimate objective is to construct and search for the T_5-IBD.

Secondly, we develop our general automatic search models so as to efficient tools for identifying IBDs based on our newly established theoretical framework. Specific, we devise the SAT-based automatic method to search for the IBDs. Although our ultimate objective is to search for the T_5-IBD, we realize IBD constructions both from the perspectives of differential propagation and state propagation, considering the following optimized search strategies.

Composite model method: The core idea of this method lies in constructing the IBD based on the propagation of both differences and states. Concretely, we define a composite unit consisting of two differences and four states, and investigate the propagation rules of the composite unit through each operation. Subsequently, we propose a new SAT-based method for searching for the T_5-IBD.

Mixed model method: Let C_{T_i} denote the cost time for determining T_i-IBD. Through our experiments, it turns out that $C_{T_0} < C_{T_4} < C_{T_5}$. Additionally, our theoretical analysis has demonstrated that the T_0-IBDs and T_4-IBDs are T_5-IBDs. Hence, rather than directly identifying T_5-IBD, we can first determine whether a quarter $\mathcal{D} = (\alpha, \alpha', \beta, \beta')$ is an r-round T_0-IBD or T_4-IBD. If $\mathcal{D}$ is not an T_4-IBD, then the information generated in the intermediate process can assist in demonstrating that $\mathcal{D}$ is not an T_5-IBD.

Our newly developed automatic search tool, which integrates these optimized search strategies, enables us to rapidly and efficiently search for IBDs.

Our method has the following new features that previous work did not have.

Able to search the IBDs by considering the impact of key schedule in the single-key setting. We consider the round key as variables influencing state propagation, and subsequently establish relationships between the round keys and the master key based on the key schedule. This approach allows us to identify IBDs while accounting for the impact of the key schedule in a single-key setting.

Able to search the IBDs by considering all the details of each operation. We investigate the propagation of states and composite units through each operation, and formalize the propagation rules in our automatic search model for IBDs. Specifically, in addition to small S-boxes based block ciphers, our methodology is also applicable to large S-boxes based block ciphers that are capable of considering all details comprehensively.

Finally, we apply our method to various block ciphers to verify its effectivity, including the large S-box based block AES [31], the lightweight block cipher Midori [32] which adopts the almost MDS matrix, and the lightweight block cipher PRESENT [33] which employs the bit permutation, and get the following results. To verify the correctness of these results, we also selected some of the results and manually verified them.

- For AES-128, we apply our method to search for all the 4-round 1 active word truncated IBDs. As a result, we first acquire 59584 such IBDs, and 192 of those IBDs cannot be detected through the propagation of pure differences, i.e. they cannot be detected by the $\mathcal{UB}$-method.
- For Midori64, we first demonstrate the non-existence of 7-round 1 active word truncated IBDs, and obtain 7296 6-round 1 active word truncated IBDs. This result is complementary to the finding that there are no existing 6-round 1 active word truncated IDs [21].
- For PRESENT-80, we search for the 1 active word IBDs, which only restrict the 0-th S-box of the input two differences to be active and the 0-th S-box of the output two differences to be active. As a result, we first obtain 58 6-round 1 active word IBDs. Note that, all those IBDs cannot be detected by the $\mathcal{UB}$-method.

All these findings suggest that our method outperforms the only $\mathcal{UB}$-method for searching IBDs currently in use. Furthermore, certain results indicate that IBDs offer an advantage over IDs in terms of the number of rounds of the distinguishers. Given the significance of impossible difference attacks as a crucial analysis method, it is imperative to take IBA seriously as a commonly employed cryptanalysis technique.

Outline. We introduce the notations and related work in Section 2. We establish our theoretical framework with a series constructions of IBDs and study their relationship in Section 3. The first automatic search method for IBDs is detailed in Section 4. In Section 5, we applied our method to various block ciphers. We conclude this paper in Section 6.

2 Preliminaries

2.1 Notation

The primary notations used hereafter are detailed as follows.

- Let k and k_i denote the master key and the i-th round key, respectively. The key schedule is denoted as KS, which generates $k_i = KS_i$ using k as input.
- Let $E_k^r(x)$ represent a r-round block cipher, encrypting the input $x \in \mathbb{F}_2^n$ under the master key $k \in \mathbb{F}_2^m$ to produce the output $y = E_k^r(x) \in \mathbb{F}_2^n$.
- Let $E_{i,k_i}(x_i)$ represent the i-th round of $E_k^r(x)$, encrypting the input $x_i \in \mathbb{F}_2^n$ under the round key $k_i \in \mathbb{F}_2^{m_i}$ to produce the output $x_{i+1} \in \mathbb{F}_2^n$. That is, $E_k^r(x) = E_{r-1,k_{r-1}} \circ \cdots \circ E_{0,k_0}(x)$. In unambiguous cases, $E_k^r(x)$ and $E_{i,k_i}(x_i)$ are abbreviated as E, E_k and E_{i,k_i}.
- Let S, SL, LL, and $AddKey$ denote a Sbox, a Sbox layer, a linear layer and the key xored layer respectively.

2.2 Boomerang distinguishers

We revisit the definitions corresponding to boomerang attacks.

Definition 1. *Basic definitions of differential analysis are as follows.*

1. *For a function $f\colon \mathbb{F}_2^n \to \mathbb{F}_2^m$, the probability that an input difference α propagates to an output difference β is given by $P_f(\alpha,\beta) = \#\left\{x \in \mathbb{F}_2^n \mid f(x) \oplus f(x \oplus \alpha) = \beta\right\}/2^n$. If $P_f(\alpha,\beta) \neq 0$, it is denoted as $\alpha \xrightarrow{f} \beta$. Define $DP_f(\alpha) = \{\beta | \alpha \xrightarrow{f} \beta\}$.*
2. *For a composite function $f : \mathbb{F}_2^n \to \mathbb{F}_2^n$, $f = f_{r-1} \circ \cdots \circ f_1 \circ f_0$, an r-round differential characteristic is defined as a series of differences $\Omega = (\alpha_0, \ldots, \alpha_r)$, where $\alpha_i \xrightarrow{f_i} \alpha_{i+1}, 0 \leq i \leq r-1$, and the probability of Ω is given by $P_f(\alpha) = \prod_{i=0}^{r-1} P_{f_i}(\alpha_i, \alpha_{i+1})$. Moreover, the probability of differential (α_0, α_r) is given by $P_f(\alpha_0, \alpha_r) = \sum_{\alpha_1,\ldots,\alpha_{r-1}} P_f(\Omega)$.*

Definition 2. *Given $\gamma, \theta, \delta \in \mathbb{F}_2^n$ three differences, the DDT and the BCT for an n-bit S-box are defined as*

$$\begin{aligned}
\mathrm{DDT}(\gamma,\theta) &= \#\left\{x \in \mathbb{F}_2^n \mid S(x) \oplus S(x \oplus \gamma) = \theta\right\},\\
\mathrm{BCT}(\gamma,\delta) &= \#\left\{x \in \mathbb{F}_2^n \mid S^{-1}(S(x) \oplus \delta) \oplus S^{-1}(S(x \oplus \gamma) \oplus \delta) = \gamma\right\}.
\end{aligned}$$

Definition 3. *Let $E = E^1 \circ E^m \circ E^0$ be an r-round SPN-network block cipher with $r = r_0 + r_1 + 1$, where E^0, E^m and E^1 denote the initial r_0 rounds, the middle 1 round and the final r_1 rounds of E, respectively. Suppose $\alpha \xrightarrow{E^0} \gamma$ and $\delta \xrightarrow{E^1} \beta$, then the probability*

$$\Pr(E^{-1}(E(x) \oplus \beta) \oplus E^{-1}(E(x \oplus \alpha) \oplus \beta) = \alpha) = (P_{E^0}(\alpha,\gamma))^2 (P_{E^1}(\delta,\beta))^2 P_m,$$

where $P_m = \prod_{i=0}^{t} \left(\mathrm{BCT}(\gamma_i,\delta_i)/2^n\right)$, assuming that there are t n-bit S-boxes in E_m with the input difference γ_i and the output difference δ_i.

To apply boomerang switch in multiple rounds, more tables have been proposed.

Definition 4. *Given $\gamma, \theta, \lambda, \delta \in \mathbb{F}_2^n$ four differences, the UBCT, LBCT and EBCT for an n-bit S-box are defined as*

$$\begin{aligned}
\mathrm{UBCT}(\gamma,\theta,\delta) &= \#\left\{ x \in \mathbb{F}_2^n \middle| \begin{array}{l} S(x) \oplus S(x \oplus \gamma) = \theta \\ S^{-1}(S(x) \oplus \delta) \oplus S^{-1}(S(x \oplus \gamma) \oplus \delta) = \gamma \end{array} \right\},\\
\mathrm{LBCT}(\gamma,\lambda,\delta) &= \#\left\{ x \in \mathbb{F}_2^n \middle| \begin{array}{l} S(x) \oplus S(x \oplus \lambda) = \delta \\ S^{-1}(S(x) \oplus \delta) \oplus S^{-1}(S(x \oplus \gamma) \oplus \delta) = \gamma \end{array} \right\},\\
\mathrm{EBCT}(\gamma,\theta,\lambda,\delta) &= \#\left\{ x \in \mathbb{F}_2^n \middle| \begin{array}{l} S(x) \oplus S(x \oplus \gamma) = \theta \\ S(x) \oplus S(x \oplus \lambda) = \delta \\ S^{-1}(S(x) \oplus \delta) \oplus S^{-1}(S(x \oplus \gamma) \oplus \delta) = \gamma \end{array} \right\}.
\end{aligned}$$

The properties present in one table have corresponding counterparts in the other tables. In [28], Delaune et al. proposed a method for establishing a BD with optimal probability using mixed tables.

2.3 The impossible boomerang attack

The IBA proposed by J. Lu [10] is recalled here. The core of IBA is an efficient IBD, originally defined as follows.

Definition 5 (IBD). *Given a block cipher $E : \mathbb{F}_2^n \times \mathbb{F}_2^m \to \mathbb{F}_2^n$ under a key $k \in \mathbb{F}_2^m$, if for $\alpha, \alpha', \beta, \beta'$ four differences, any pair of plaintexts (x_1, x_2) cannot satisfy*

$$E_k(x_1) \oplus E_k(x_2) = \beta,$$
$$E_k(x_1 \oplus \alpha) \oplus E_k(x_2 \oplus \alpha') = \beta'$$

at the same time, then $(\alpha, \alpha', \beta, \beta')$ is called an impossible boomerang distinguisher for E_k, denoted by $(\alpha, \alpha') \not\twoheadrightarrow (\beta, \beta')$.

Furthermore, they put forward a method for establishing the IBDs.

Theorem 1. *Let $E = E^1 \circ E^0$. Given $\alpha \xrightarrow{E^0} \gamma$, $\alpha' \xrightarrow{E^0} \gamma'$ and $\beta \xrightarrow{(E^1)^{-1}} \delta$, $\beta' \xrightarrow{(E^1)^{-1}} \delta'$ all with probability 1, if $\delta \oplus \delta' \oplus \gamma \oplus \gamma' \neq 0$, then $(\alpha, \alpha') \not\twoheadrightarrow (\beta, \beta')$.*

Let the block cipher $E : \mathbb{F}_2^n \times \mathbb{F}_2^m \to \mathbb{F}_2^n$ be a cascade of three sub-ciphers $E = E^c \circ E^b \circ E^a$, where E^b corresponds to the derived IBD $(\alpha, \alpha') \not\twoheadrightarrow (\beta, \beta')$, then the IBA recover the key as follows:

1. For each guess of k_a and k_c, the subkey used in E^a and E^c respectively, check whether a candidate quartet of plaintext/ciphertext pairs $((x_0, y_0), (x_1, y_1), (x_2, y_2), (x_3, y_3))$ satisfies the following four conditions:

$$E^a_{k_a}(x_0) \oplus E^a_{k_a}(x_1) = \alpha,$$
$$E^a_{k_a}(x_2) \oplus E^a_{k_a}(x_3) = \alpha',$$
$$(E^c_{k_c})^{(-1)}(y_1) \oplus (E^c_{k_c})^{(-1)}(y_2) = \beta,$$
$$(E^c_{k_c})^{(-1)}(y_0) \oplus (E^c_{k_c})^{(-1)}(y_3) = \beta'.$$

2. If the quartet does satisfy the above conditions, then discard the subkey guess, and go to the previous step until unique subkey remains.

2.4 The SAT method

The SAT problem [34] involves determining whether a given Boolean formula can be satisfied. A typical framework for the SAT-based automatic search method is to convert the search for a distinguisher into a SAT problem, and then solve this SAT problem by calling upon the available solvers. In this paper, we utilize STP[4] along with CryptoMiniSat[5] as backends.

3 New Theory for Constructing IBDs

In this section, we establish a new theoretical framework for constructing IBDs from both the perspectives of differential propagation and state propagation, and illustrating the relationships between all construction methods.

[4] https://stp.github.io

[5] https://github.com/msoos/cryptominisat

3.1 Constructing IBDs from the perspective of differential propagation

We propose five IBD definitions so as to construction methods in line with the techniques of AS, BD, BCT, U/L/EBCT and mixed tables describing differential propagation for SPN-network block ciphers as follows.

Firstly, we relax the definition of $DP_f(\alpha)$ by regarding the S-box appearing in f as only a permutation.

Definition 6. *Let $\overline{DP}_f(\alpha) = \{\beta | \alpha \xrightarrow{\overline{f}} \beta\}$, where $\alpha \xrightarrow{\overline{f}} \beta$ denotes that the input difference α propagates to the output difference β via f by considering all the details of operations of f expect S-boxes, then we drive a roughly propagation set*

$$\overline{DP}_S(\alpha) = \begin{cases} \{0\}, & \alpha = 0, \\ \mathbb{F}_2^n/\{0\}, & \textit{otherwise}. \end{cases}$$

Then we present two boomerang trail based on $\overline{DP}_f(\alpha)$ and $DP_f(\alpha)$.

Definition 7. *Given an r-round block cipher $E = E^1 \circ E^0$, for two input differences α, α' and two output differences β, β',*

- *if there exist $\gamma \in \overline{DP}_{E^0}(\alpha)$, $\gamma' \in \overline{DP}_{E^0}(\alpha')$, $\delta \in \overline{DP}_{(E^1)^{-1}}(\beta)$, and $\delta' \in \overline{DP}_{(E^1)^{-1}}(\beta')$, such that $\gamma \oplus \gamma' \oplus \delta \oplus \delta' = 0$, then*

  $$(\alpha, \alpha') \to \cdots \to \underbrace{(\gamma, \gamma')(\delta, \delta')}_{\gamma \oplus \gamma' \oplus \delta \oplus \delta' = 0} \to \cdots \to (\beta, \beta')$$

  is called an r-round T_0 boomerang trail.
- *if there exist $\gamma \in DP_{E^0}(\alpha)$, $\gamma' \in DP_{E^0}(\alpha')$, $\delta \in DP_{(E^1)^{-1}}(\beta)$, and $\delta' \in DP_{(E^1)^{-1}}(\beta')$, such that $\gamma \oplus \gamma' \oplus \delta \oplus \delta' = 0$, then*

  $$(\alpha, \alpha') \to \cdots \to \underbrace{(\gamma, \gamma')(\delta, \delta')}_{\gamma \oplus \gamma' \oplus \delta \oplus \delta' = 0} \to \cdots \to (\beta, \beta')$$

  is called an r-round T_1 boomerang trail.

Accordingly, we present the following two IBD construction methods, with T_0-IBD being a new construction and T_1-IBD corresponding to the only method for constructing IBDs currently as depicted in Theorem 1.

Construction 1 (T_0-IBD). *Given an r-round block cipher E, for two input differences α, α' and two output differences β, β', if there is no r-round T_0 boomerang trail, then $((\alpha, \alpha'), (\beta, \beta'))$ is an IBD, called an r-round T_0-IBD.*

Construction 2 (T_1-IBD). *Given an r-round block cipher E, for two input differences α, α' and two output differences β, β', if there is no r-round T_1 boomerang trail, then $((\alpha, \alpha'), (\beta, \beta'))$ is an IBD, called an r-round T_1-IBD.*

We prove the inclusion relationship between T_0-IBD and T_1-IBD. The proofs of the theorems in this section are given in Appendix A.

Theorem 2. *An r-round T_0-IBD $((\alpha, \alpha'), (\beta, \beta'))$ is an r-round T_1-IBD.*

The above construction methods are based on the original boomerang attack, assuming that the two sub-ciphers E^0 and E^1 are independent. However, it may not hold true for two selected differential characteristics as demonstrated in [22]. In other words, an r-round T_1 boomerang trail might not exist at all. Consequently, this method may overlook certain IBDs. In order to address this issue, we introduce and generalize the concept of BCT in the context of boomerang attack for constructing IBDs.

Definition 8. *Given $\mu, \mu', \varphi, \varphi' \in \mathbb{F}_2^n$ four differences, the GBCT for an n-bit S-box is defined as*

$$\mathrm{GBCT}(\mu, \mu', \varphi, \varphi') = \# \left\{ (u_1, u_2) \in \{0,1\}^{2n} \middle| \begin{array}{l} S(u_1) \oplus S(u_2) = \varphi \\ S(u_1 \oplus \mu) \oplus S(u_2 \oplus \mu') = \varphi' \end{array} \right\}.$$

Definition 9. *Let $E = E^1 \circ E^m \circ E^0$ be an r-round block cipher with $r = r_0 + r_1 + 1$, where E^0, E^m and E^1 denote the initial r_0 rounds, the middle 1 round and the final r_1 rounds of E, respectively. For two input differences α, α' and two output differences β, β', if there exist $\gamma \in DP_{E^0}(\alpha)$, $\gamma' \in DP_{E^0}(\alpha')$, and $\delta \in DP_{(E^1)^{-1}}(\beta)$, $\delta' \in DP_{(E^1)^{-1}}(\beta')$, such that $GBCT_{E^m}(\gamma, \gamma', \delta, \delta') \neq 0$, then*

$$(\alpha, \alpha') \to \cdots \to \underbrace{(\gamma, \gamma') \to (\delta, \delta')}_{(\gamma,\gamma') \xrightarrow{GBCT} (\delta,\delta')} \to \cdots \to (\beta, \beta')$$

is called an r-round T_2 boomerang trail. Here, $\xrightarrow{GBCT}$ represents that the propagation rule through S-boxes in E^m follows GBCT.

Accordingly, we present the following new IBD construction method, which is able to deal with the dependence problem of one round.

Construction 3 (T_2-IBD). *Given an r-round block cipher E, for two input differences α, α' and two output differences β, β', if there is no r-round T_2 boomerang trail, then $((\alpha, \alpha'), (\beta, \beta'))$ is an IBD, called an r-round T_2-IBD.*

We prove the inclusion relationship between T_1-IBD and T_2-IBD.

Theorem 3. *An r-round T_1-IBD $((\alpha, \alpha'), (\beta, \beta'))$ is an r-round T_2-IBD.*

However, as highlighted by Song et al. [35], it has been observed that the dependencies can exert a much greater influence over multiple rounds, e.g. up to 6 rounds for SKINNY. To eliminate the incompatibility in multiple rounds, we further introduce and generalize the concepts of the UBCT, LBCT, and EBCT in boomerang attacks to IBDs. This serves to compensate for the limitations of T_2-boomerang trails.

Definition 10. *Given $\mu, \mu', \rho, \rho', \theta, \theta', \varphi, \varphi' \in \mathbb{F}_2^n$ eight differences ($\rho' = \mu \oplus \mu' \oplus \rho$), the GUBCT, GLBCT and GEBCT for an n-bit S-box are defined as*

$$\mathrm{GUBCT}(\mu, \mu', \theta, \theta', \varphi, \varphi') = \# \left\{ (u_1, u_2) \in \mathbb{F}_2^{2n} \left| \begin{array}{l} S(u_1) \oplus S(u_1 \oplus \mu) = \theta \\ S(u_2) \oplus S(u_2 \oplus \mu') = \theta' \\ S(u_1) \oplus S(u_2) = \varphi \\ S(u_1 \oplus \mu) \oplus S(u_2 \oplus \mu') = \varphi' \end{array} \right. \right\},$$

$$\mathrm{GLBCT}(\mu, \mu', \rho, \rho', \varphi, \varphi') = \# \left\{ (u_1, u_2) \in \mathbb{F}_2^{2n} \left| \begin{array}{l} u_1 \oplus u_2 = \rho \\ S(u_1) \oplus S(u_2) = \varphi \\ S(u_1 \oplus \mu) \oplus S(u_2 \oplus \mu') = \varphi' \end{array} \right. \right\},$$

$$\mathrm{GEBCT}(\mu, \mu', \rho, \rho', \theta, \theta', \varphi, \varphi') = \# \left\{ (u_1, u_2) \in \mathbb{F}_2^{2n} \left| \begin{array}{l} u_1 \oplus u_2 = \rho \\ S(u_1) \oplus S(u_1 \oplus \mu) = \theta \\ S(u_2) \oplus S(u_2 \oplus \mu') = \theta' \\ S(u_1) \oplus S(u_2) = \varphi \\ S(u_1 \oplus \mu) \oplus S(u_2 \oplus \mu') = \varphi' \end{array} \right. \right\}.$$

A schematic diagram for these generalized BCTs is shown in Figure 1.

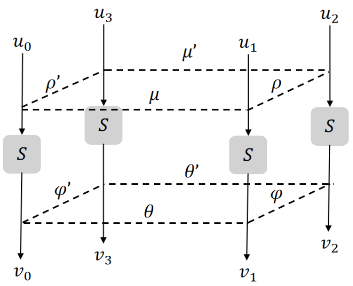

Fig. 1. The Generalized Boomerang Connectivity Tables

Definition 11. *Let $E = E_{r-1,k_{r-1}} \circ \cdots \circ E_{0,k_0}(x)$ be an r-round block cipher. Let $\epsilon_0^i, \epsilon_1^i, \epsilon_2^i, \epsilon_3^i$ be the four input differences and $\epsilon_0^{i+1}, \epsilon_1^{i+1}, \epsilon_2^{i+1}, \epsilon_3^{i+1}$ be the four output differences of the round function E_{i,k_i} for $i \in \{0, \ldots, r-1\}$. For two input differences α, α' and two output differences β, β' of the block cipher E, if there exits a trail*

$$(\epsilon_0^0 = \alpha, \epsilon_1^0, \epsilon_2^0 = \alpha', \epsilon_3^0) \xrightarrow{GEBCT} \cdots \xrightarrow{GEBCT} (\epsilon_0^r, \epsilon_1^r = \beta, \epsilon_2^r, \epsilon_3^r = \beta'),$$

then it is called an r-round T_3 boomerang trail. Here, $\xrightarrow{GEBCT}$ represents that the propagation rule through S-boxes in E_{i,k_i} follows GEBCT.

These definitions allow us to define new IBDs.

Construction 4 (T_3-IBD). *Given an r-round block cipher E, for two input differences α, α' and two output differences β, β', if there is no r-round T_3 boomerang trail, then $((\alpha, \alpha'), (\beta, \beta'))$ is an IBD, called an r-round T_3-IBD.*

We prove the inclusion relationship between T_2-IBD and T_3-IBD.

Theorem 4. *An r-round T_2-IBD $((\alpha,\alpha'),(\beta,\beta'))$ is an r-round T_3-IBD.*

Next, we consider whether the mixed use of DDT, GBCT, GUBCT, GLBCT, and GEBCT can discover more IBDs. We present two more notations for an n-bit S-box given the input and output difference pair $(\mu,\mu',\rho,\rho') - (\theta,\theta',\varphi,\varphi')$,

$$\mathrm{DDT}^2_{upper}(\mu,\mu',\theta,\theta') = \#\left\{ (u_1,u_2)\in\mathbb{F}_2^{2n} \middle| \begin{array}{l} S(u_1)\oplus S(u_1\oplus\mu)=\theta \\ S(u_2)\oplus S(u_2\oplus\mu')=\theta' \end{array} \right\},$$

$$\mathrm{DDT}^2_{lower}(\rho,\rho',\varphi,\varphi') = \#\left\{ (u_0,u_1)\in\mathbb{F}_2^{2n} \middle| \begin{array}{l} S(u_1)\oplus S(u_1\oplus\rho)=\varphi \\ S(u_0)\oplus S(u_0\oplus\rho')=\varphi' \end{array} \right\}.$$

Definition 12. *Let E be a block cipher which has totally t S-boxes $(S_0,\ldots,S_{t-1})$. Define $\mathcal{AP}_E = \{(p_0,\ldots,p_{t-1}) | p_i \in \{\mathrm{DDT}^2_{upper}, \mathrm{DDT}^2_{lower}, \mathrm{GBCT}, \mathrm{GUBCT}, \mathrm{GLBCT}, \mathrm{GEBCT}\}\}$ as a propagation rule set. Then for $P=(p_0,\ldots,p_{t-1})\in\mathcal{AP}_E$, the propagation rule through the i-th S-box follows p_i.*

Definition 13. *Let $E = E_{r-1,k_{r-1}}\circ\cdots\circ E_{0,k_0}(x)$ be an r-round block cipher. Let $P_i\in\mathcal{AP}_{E_{i,k_i}}$ be a propagation rule of E_{i,k_i} for $i\in\{0,\ldots,r-1\}$, which cascade a predefined propagation rule P of E. Let $\epsilon_0^i,\epsilon_1^i,\epsilon_2^i,\epsilon_3^i$ be the four input differences and $\epsilon_0^{i+1},\epsilon_1^{i+1},\epsilon_2^{i+1},\epsilon_3^{i+1}$ be the four output differences of the round function E_{i,k_i} for $i\in\{0,\ldots,r-1\}$. For two input differences α,α' and two output differences β,β' of the block cipher E, if there exits a trail*

$$(\epsilon_0^0=\alpha,\epsilon_1^0,\epsilon_2^0=\alpha',\epsilon_3^0)\xrightarrow{P_0}\cdots\xrightarrow{P_{r-1}}(\epsilon_0^r,\epsilon_1^r=\beta,\epsilon_2^r,\epsilon_3^r=\beta'),$$

then it is called an r-round T_P boomerang trail. Here, $\xrightarrow{P_i}$ represents that the propagation rule through S-boxes in E_{i,k_i} follows P_i.

Accordingly, we have the following construction.

Construction 5 (T_P-IBD). *Given an r-round block cipher E and a predefined rule $P\in\mathcal{AP}$, for two input differences α,α' and two output differences β,β', if there is no r-round T_P boomerang trail, then $((\alpha,\alpha'),(\beta,\beta'))$ is an IBD, called an r-round T_P-IBD.*

Next, we prove that, for any predefined rule P of the r-round block cipher E, Construction 5 cannot be superior to Construction 4.

Theorem 5. *For any predefined rule $P\in\mathcal{AP}_E$, an r-round T_P-IBD $((\alpha,\alpha'),(\beta,\beta'))$ is an r-round T_3-IBD.*

Therefore, we omit Construction 5 hereafter.

3.2 Constructing IBDs from the perspective of state propagation

Building on the concept proposed by Hu et al. [21] that constructs IDs by the propagation of two states, we construct IBDs by the propagation of four states adapt to block ciphers of any structure. Specifically, our approach considers both cases of independent keys and key relations in a single-key setting.

Definition 14. *Let $E = E_{r-1,k_{r-1}} \circ \cdots \circ E_{0,k_0}(x)$ be an r-round block cipher. Given $\alpha, \alpha', \beta, \beta'$ four differences, let $I = \{(x_0, x_1, x_2, x_3) \mid x_0 \oplus x_1 = \alpha, x_2 \oplus x_3 = \alpha'\}$ and $O = \{(y_0, y_1, y_2, y_3) \mid y_1 \oplus y_2 = \beta, y_0 \oplus y_3 = \beta'\}$. If there exit $(x_0^0, x_1^0, x_2^0, x_3^0) \in I$, $(x_0^r, x_1^r, x_2^r, x_3^r) \in O$ and independent round keys $(k_0, \cdots, k_{r-1})$, such that*

$$x_j^{i+1} = E_{i,k_i}(x_j^i)(0 \leq i \leq r-1, 0 \leq j \leq 4),$$

then $(x_0^0, x_1^0, x_2^0, x_3^0) \rightarrow \cdots \rightarrow (x_0^r, x_1^r, x_2^r, x_3^r)$ is called an r-round T_4 boomerang trail.

This definition enables us to construct IBD in another way.

Construction 6 (T_4-IBD). *Given an r-round block cipher E, for two input differences α, α' and two output differences β, β', if there is no r-round T_4 boomerang trail, then $((\alpha, \alpha'), (\beta, \beta'))$ is an IBD, called an r-round T_4-IBD.*

To consider the relationship of round keys in the single-key setting according to the key schedule, we further present the following definition.

Definition 15. *Let $E = E_{r-1,KS_{r-1}(k)} \circ \cdots \circ E_{0,KS_0(k)}(x)$ be an r-round block cipher under the key schedule KS. Given $\alpha, \alpha', \beta, \beta'$ four differences, let $I = \{(x_0, x_1, x_2, x_3) \mid x_0 \oplus x_1 = \alpha, x_2 \oplus x_3 = \alpha'\}$ and $O = \{(y_0, y_1, y_2, y_3) \mid y_1 \oplus y_2 = \beta, y_0 \oplus y_3 = \beta'\}$. If there exit $(x_0^0, x_1^0, x_2^0, x_3^0) \in I$, $(x_0^r, x_1^r, x_2^r, x_3^r) \in O$ and an master k such that*

$$x_j^{i+1} = E_{i,KS_i(k)}(x_j^i)(0 \leq i \leq r-1, 0 \leq j \leq 4),$$

then $(x_0^0, x_1^0, x_2^0, x_3^0) \rightarrow \cdots \rightarrow (x_0^r, x_1^r, x_2^r, x_3^r)$ is called an r-round T_5 boomerang trail.

This definition for the first time allows us to take into account the validity involving round keys' compact when constructing IBDs.

Construction 7 (T_5-IBD). *Given an r-round block cipher E, for two input differences α, α' and two output differences β, β', if there is no r-round T_5 boomerang trail, then $((\alpha, \alpha'), (\beta, \beta'))$ is an IBD, called an r-round T_5-IBD.*

The inclusion relationship between T_4-IBD and T_5-IBD is direct for their definitions.

Theorem 6. *An r-round T_4-IBD $((\alpha, \alpha'), (\beta, \beta'))$ is an r-round T_5-IBD.*

Construction 7 enables us to construct IBDs considering both the details of the operations and the key schedule in the single-key setting. Next, we prove that Construction 7 is equivalent to the original definition of IBDs given in Definition 5. That is, any approach for constructing IBDs will not be superior to Construction 7.

Theorem 7. *Construction 7 is the tightest method for constructing IBDs.*

3.3 The relationship between IBDs constructed from the perspective of differential propagation and state propagation

We discuss the relationship between the four IBDs constructed from the perspective of differential propagation and the two IBDs constructed from the perspective of state propagation. At first, we prove that the definition of T_3-IBD is equivalent with that of T_4-IBD within SPN-network block ciphers. A schematic diagram is shown in Figure 2.

Theorem 8. *Given an SPN-network block cipher, $((\alpha, \alpha'), (\beta, \beta'))$ is an r-round T_3-IBD if and only if it is an r-round T_4-IBD.*

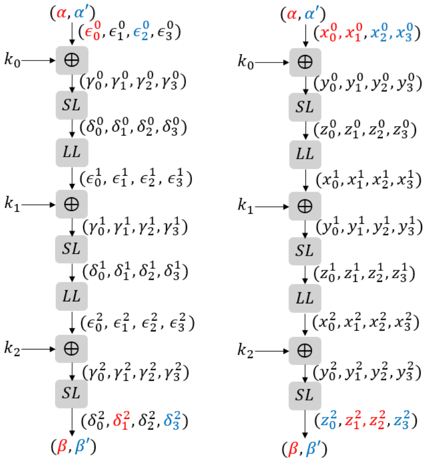
(α,α')
$(\epsilon_0^0,\epsilon_1^0,\epsilon_2^0,\epsilon_3^0)$
k_0
$(\gamma_0^0,\gamma_1^0,\gamma_2^0,\gamma_3^0)$
SL
$(\delta_0^0,\delta_1^0,\delta_2^0,\delta_3^0)$
LL
$(\epsilon_0^1,\epsilon_1^1,\epsilon_2^1,\epsilon_3^1)$
k_1
$(\gamma_0^1,\gamma_1^1,\gamma_2^1,\gamma_3^1)$
SL
$(\delta_0^1,\delta_1^1,\delta_2^1,\delta_3^1)$
LL
$(\epsilon_0^2,\epsilon_1^2,\epsilon_2^2,\epsilon_3^2)$
k_2
$(\gamma_0^2,\gamma_1^2,\gamma_2^2,\gamma_3^2)$
SL
$(\delta_0^2,\delta_1^2,\delta_2^2,\delta_3^2)$
(β,β')
(α,α')
$(x_0^0,x_1^0,x_2^0,x_3^0)$
k_0
$(y_0^0,y_1^0,y_2^0,y_3^0)$
SL
$(z_0^0,z_1^0,z_2^0,z_3^0)$
LL
$(x_0^1,x_1^1,x_2^1,x_3^1)$
k_1
$(y_0^1,y_1^1,y_2^1,y_3^1)$
SL
$(z_0^1,z_1^1,z_2^1,z_3^1)$
LL
$(x_0^2,x_1^2,x_2^2,x_3^2)$
k_2
$(y_0^2,y_1^2,y_2^2,y_3^2)$
SL
$(z_0^2,z_1^2,z_2^2,z_3^2)$
(β,β')

Fig. 2. The equivalence between T_3-IBD and T_4-IBD

Relations of all IBDs. Let S_{T_i} donates the set containing all T_i-IBDs. Based on the above theorems, it holds that

$$S_{T_0} \subseteq S_{T_1} \subseteq S_{T_2} \subseteq S_{T_3} = S_{T_4} \subseteq S_{T_5}.$$

4 New Automatic Search Methods for IBDs

In this section, we introduce our method for searching IBDs. Firstly, we propose the approach to determine IBDs from the perspective of the differential propagation. Then, we study the method for identifying IBDs from the aspect of state propagation. After that, we propose the composite model method for determining IBDs. Finally, we deliberate on the search space and propose the mixed model method to effectively search for IBDs within this space.

4.1 Determining the IBDs from the perspective of the differential propagation

In this part, we propose our method for determining T_0-IBDs and T_1-IBDs from the perspective of the differential propagation. First, we propose the SAT-based method for modeling the differential propagation through each operations. Those operations include **Xor**, **Copy**, **KeyAdd**, **Matrix multiplication** and **S-box**. In particular, we also propose the method for modeling the propagation of differences through S-box in the arbitrary model.

Model 1. *The method for modeling the differential propagation through the operations* **Xor***,* **Copy***,* **KeyAdd** *is presented in Table 1.*

Table 1. Modeling the Differential Propagation Through Xor, Copy and KeyAdd

Operation	Input Diff	Output Diff	Modeling Method
Copy	$\alpha \in \mathbb{F}_2$	$\beta_0, \beta_1 \in \mathbb{F}_2$	$\beta_0 = \alpha, \beta_1 = \alpha$
Xor	$\alpha_0, \alpha_1 \in \mathbb{F}_2$	$\beta \in \mathbb{F}_2$	$\beta = \alpha_0 \oplus \alpha_1$
KeyAdd	$\alpha \in \mathbb{F}_2$	$\beta \in \mathbb{F}_2$	$\beta = \alpha$

Model 2. *For the operation* **Matrix multiplication** $M = (m_{i,j})_{u\times v}$*, let* $\alpha_i (0 \le i \le v-1)$ *and* $\beta_i (0 \le i \le u-1)$ *be the input and output differences of* M*, it holds* $\beta_i = \oplus_{j=0}^{v-1} m_{i,j}\alpha_j$*, thus the differential propagation through* **Matrix multiplication** *can be expressed according to* **Xor***.*

Model 3. *For the operation* **S-box** S*, let* $\alpha_i (0 \le i \le v-1)$ *and* $\beta_i (0 \le i \le u-1)$ *be the input and output differences of* S*, the possible values of* $\alpha_i (0 \le i \le v-1)$ *and* $\beta_i (0 \le i \le u-1)$ *is restricted by the DDT of* S*, the differential propagation through* S *can be expressed with the help of the third party tool* ***Logic Friday***[6]*.*

Model 4. *For the operation* **S-box** S*, let* $\alpha_i (0 \le i \le v-1)$ *and* $\beta_i (0 \le i \le v-1)$ *be the input and output differences of* S*, in the arbitrary mode, only the following transitions are impossible:*

$$(\mathbf{0}, \mathbf{1}), \cdots, (\mathbf{0}, \mathbf{2^v} - \mathbf{1}), (\mathbf{1}, \mathbf{0}), \cdots, (\mathbf{2^v} - \mathbf{1}, \mathbf{0}).$$

Those impossible points can be removed by the following boolean expressions:

$$\alpha_{v-1}||\cdots||\alpha_0||\neg\beta_0 = 1, \neg\alpha_0||\beta_{v-1}||\cdots||\beta_0 = 1,$$
$$\vdots$$
$$\alpha_{v-1}||\cdots||\alpha_0||\neg\beta_{v-1} = 1, \neg\alpha_{v-1}||\beta_{v-1}\cdots||\alpha_0 = 1.$$

Example 1. For the 4-bit S-box, the following boolean expressions can be used to model the differential propagation through the S-box in the arbitrary mode.

$$\alpha_3||\alpha_2||\alpha_1||\alpha_0||\neg\beta_0 = 1, \neg\alpha_0||\beta_3||\beta_2||\beta_1||\beta_0 = 1,$$
$$\alpha_3||\alpha_2||\alpha_1||\alpha_0||\neg\beta_1 = 1, \neg\alpha_1||\beta_3||\beta_2||\beta_1||\beta_0 = 1,$$
$$\alpha_3||\alpha_2||\alpha_1||\alpha_0||\neg\beta_2 = 1, \neg\alpha_2||\beta_3||\beta_2||\beta_1||\beta_0 = 1,$$
$$\alpha_3||\alpha_2||\alpha_1||\alpha_0||\neg\beta_3 = 1, \neg\alpha_3||\beta_3||\beta_2||\beta_1||\beta_0 = 1.$$

Based on the model method for the differential propagation, we propose Algorithm 1 to determine whether a given $\mathcal{D} = (\alpha, \alpha', \beta, \beta')$ is an r-round T_0-IBD or T_1-IBD or not. *BuildDiffPropagation*(T, r', a, b) is a function that models the propagation of the input difference a to the output difference b through r'-round E. In the case of $T = T_0$, the differential propagation through the S-box is modeled according to Model 4, In the case of $T = T_1$, the differential

[6] The Logic Friday (http://sontrak.com/) is a third party tool, it can be used to derive the minimum (or as small as possible in a reasonable time) product-of-sum representation of a given Boolean function from its truth table. Such product-of-sum representation can be transform to a set of logic expressions equivalently. See [36] for detailed usage.

propagation via the S-box is modeled according to Model 3. Finally, this function returns two output variables, where $C_{a,b}$ represents the constraints for modeling the differential propagation, and $V_{a,b}$ represents the auxiliary variables.

Algorithm 1: Model for Determining the r-round T_0-IBD and T_1-IBD

```
Input: a tuple $(\alpha, \alpha', \beta, \beta')$, number of rounds $r$, type of IBD $T$
Output: variables of model $V$, constraints of model $C$
$V = [\ ], C = [\ ]$
$r_0 = \lceil r/2 \rceil$, $r_1 = r - r_0$
$\gamma, \gamma', \delta, \delta' = \mathit{VarDeclare}()$
$V.\mathit{AddList}(\gamma, \gamma', \delta, \delta')$
foreach $(r', a, b)$ in $[(r_0, \alpha, \gamma), (r_0, \alpha', \gamma'), (r_1, \delta, \beta), (r_1, \delta', \beta')]$ do
    $C_{a,b}, V_{a,b} = \mathit{BuildDiffPropagation}(T, r', a, b)$
    $V.\mathit{AddList}(V_{a,b})$
    $C.\mathit{AddList}(C_{a,b})$
end
$C.\mathit{AddList}(\gamma \oplus \gamma' \oplus \delta \oplus \delta' = 0)$
return $V, C$
```

With the return values of Algorithm 1, we declare the variables in V in accordance with the grammar of the SAT solver STP, and add the constraints in C to form a constraint set. Subsequently, we invoke STP to determine whether such a constraint set has a solution or not. In case there is no solution existing, then $\mathcal{D} = (\alpha, \alpha', \beta, \beta')$ is an r-round T-IBD.

4.2 Determining the IBDs from the perspective of the state propagation

In this part, we propose our method for determining T_4-IBDs and T_5-IBDs from the perspective of the state propagation. First, we recall the method for modeling the state propagation via each operation [21].

Model 5. *The method for modeling the state propagation through the operations* **Xor**, **Copy**, **Matrix multiplication** *is identical to that of Model 1 and Model 2. Regarding the operation* **KeyAdd**, *the method for modeling the state propagation is the same as that of* **Xor**.

Model 6. *For the operation* **S-box** S, *let* $\alpha_i (0 \le i \le v-1)$ *and* $\beta_i (0 \le i \le u-1)$ *be the input and output states of* S, *the possible values of* $\alpha_i (0 \le i \le v-1)$ *and* $\beta_i (0 \le i \le u-1)$ *is restricted by the truth table of* S, *the state propagation through* S *can also be expressed with the help of the third party tool Logic Friday.*

Based on the model method for the state propagation, we propose Algorithm 2 to determine whether a given $\mathcal{D} = (\alpha, \alpha', \beta, \beta')$ is an r-round T_4-IBD or T_5-IBD or not. The algorithm is briefly introduced as follows.

Algorithm 2: Model for Determining the r-round T_4-IBD and T_5-IBD

```
   Input: a tuple $(\alpha, \alpha', \beta, \beta')$, number of rounds $r$, type of IBD $T$
   Output: variables of model $V$, constraints of model $C$
1  $V = [\ ]$, $C = [\ ]$
2  $x_0, x_1, x_2, x_3, y_0, y_1, y_2, y_3 = \mathit{VarDeclare}()$
3  $V.\mathit{AddList}(x_0, x_1, x_2, x_3, y_0, y_1, y_2, y_3)$
4  foreach $i$ in $[0, 1, 2, 3]$ do
5      $C_i, V_i, RK = \mathit{BuildStatePropagation}(r, x_i, y_i)$
6      $V.\mathit{AddList}(V_i)$, $C.\mathit{AddList}(C_i)$
7  end
8  if $T == T_5$ then
9      $MK = \mathit{VarDeclare}()$
10     $V.\mathit{AddList}(MK)$
11     $C_K = \mathit{BuildKeyRelation}(MK, RK)$
12     $C.\mathit{AddList}(C_K)$
13 end
14 $C.\mathit{AddList}(x_0 \oplus x_1 = \alpha, x_2 \oplus x_3 = \alpha')$
15 $C.\mathit{AddList}(y_1 \oplus y_2 = \beta, y_0 \oplus y_3 = \beta')$
16 return $V, C$
```

- Line 4-6: The function $\mathit{BuildStatePropagation}(r, x_i, y_i)$ is a function that models the propagation of the input state x_i to the output state y_i via r-round E. This function returns three output variables, among which C_i represents the constraints for modeling the state propagation, V_i represents the auxiliary variables, and RK represents the round keys.
- Line 7-11: In the case of determining the T_5-IBDs, we declare the master key MK and establish the relationship between MK and RK according to the key schedule.

4.3 Determining the IBDs with the composite model method

In this part, we propose the composite model method to determine T_4-IBDs and T_5-IBDs efficiently. The efficiency advantage stems from our experimental observation of the two algorithms. One of them is AES [31]. On average, we can determine a 4-round IBD by the composite model method in 2123 seconds, while we cannot obtain a result within 3600 seconds in accordance with Algorithm 2. The other one is PRESENT [33]. On average, we can determine a 6-round IBD in 4.7 seconds by the composite model method, while it requires 13.4 seconds according to Algorithm 2, which indicates that the composite model method is around three times as fast as the pure state propagation. The core idea of the composite model method is to build the model by the propagation of both states and differences. For the convenience of description, we introduce the concept of a composite unit.

Definition 16. *For an integer t, let $x_{0,i}, x_{1,i}, x_{2,i}, x_{3,i} \in \mathbb{F}_2 (0 \le i \le t-1)$ be $4t$ states and $\alpha_{0,i}, \alpha_{1,i} \in \mathbb{F}_2 (0 \le i \le t-1)$ be $2t$ differences, the tuple*

$$(\underbrace{x_{0,0},\dots,x_{0,t-1}}_{t},\dots,\underbrace{x_{3,0},\dots,x_{3,t-1}}_{t},\underbrace{\alpha_{0,0},\dots,\alpha_{0,t-1}}_{\alpha_{0,i}=x_{0,i}\oplus x_{1,i}},\underbrace{\alpha_{1,0},\dots,\alpha_{1,t-1}}_{\alpha_{1,i}=x_{2,i}\oplus x_{3,i}})$$

*is called the **upper composite unit**, and the tuple*

$$(\underbrace{x_{0,0},\dots,x_{0,t-1}}_{t},\dots,\underbrace{x_{3,0},\dots,x_{3,t-1}}_{t},\underbrace{\alpha_{0,0},\dots,\alpha_{0,t-1}}_{\alpha_{0,i}=x_{1,i}\oplus x_{2,i}},\underbrace{\alpha_{1,0},\dots,\alpha_{1,t-1}}_{\alpha_{1,i}=x_{0,i}\oplus x_{3,i}})$$

*is called the **lower composite unit**.*

Subsequently, we illustrate the method for depicting the propagation of the upper composite unit via the operation **Xor**, **Copy**, **KeyAdd**, **Matrix multiplication** and **S-box**, and the method for depicting the propagation of the lower composite unit can be inferred in a similar manner.

Model 7. *For the operation* **Xor***, let $x = (x_0, x_1, x_2, x_3, \alpha_0, \alpha_1)$ and $y = (y_0, y_1, y_2, y_3, \beta_0, \beta_1)$ be the two input upper composite units, and $z = (z_0, z_1, z_2, z_3, \gamma_0, \gamma_1)$ be the output upper composite unit, then the following boolean expressions are capable of modeling the propagation of the upper composite unit.*

$$z_i = x_i \oplus y_i (0 \le i \le 3), \gamma_i = \alpha_i \oplus \beta_i (0 \le i \le 1).$$

Model 8. *For the operation* **Copy***, let $x = (x_0, x_1, x_2, x_3, \alpha_0, \alpha_1)$ be the input upper composite unit, $y = (y_0, y_1, y_2, y_3, \beta_0, \beta_1)$ and $z = (z_0, z_1, z_2, z_3, \gamma_0, \gamma_1)$ are the two output upper composite units, then the following boolean expressions are capable of modeling the propagation of the upper composite unit.*

$$y_i = x_i, z_i = x_i (0 \le i \le 3), \beta_i = \alpha_i, \gamma_i = \alpha_i (0 \le i \le 1).$$

Model 9. *For the operation* **KeyAdd***, let $x = (x_0, x_1, x_2, x_3, \alpha_0, \alpha_1)$ be the input upper composite unit, $y = (y_0, y_1, y_2, y_3, \beta_0, \beta_1)$ be the output upper composite unit, and k be the key, then the following boolean expressions are capable of modeling the propagation of the upper composite unit.*

$$y_i = x_i \oplus k (0 \le i \le 3), \beta_i = \alpha_i (0 \le i \le 1).$$

Model 10. *For the operation* **Matrix multiplication** *$M = (m_{i,j})_{u \times v}$, let*

$$(\underbrace{x_{0,0},\dots,x_{0,v-1}}_{v},\dots,\underbrace{x_{3,0},\dots,x_{3,v-1}}_{v},\underbrace{\alpha_{0,0},\dots,\alpha_{0,v-1}}_{\alpha_{0,i}=x_{0,i}\oplus x_{1,i}},\underbrace{\alpha_{1,0},\dots,\alpha_{1,v-1}}_{\alpha_{1,i}=x_{2,i}\oplus x_{3,i}})$$

be the input upper composite unit, and

$$(\underbrace{y_{0,0},\dots,y_{0,u-1}}_{u},\dots,\underbrace{y_{3,0},\dots,y_{3,u-1}}_{u},\underbrace{\beta_{0,0},\dots,\beta_{0,u-1}}_{\beta_{0,i}=y_{0,i}\oplus y_{1,i}},\underbrace{\beta_{1,0},\dots,\beta_{1,u-1}}_{\beta_{1,i}=y_{2,i}\oplus y_{3,i}})$$

be the output upper composite unit, then the following boolean expressions are capable of modeling the propagation of the upper composite unit.

$$y_{i,j} = \oplus_{k=0}^{v-1} m_{j,k} x_{i,k} (0 \le i \le 3, 0 \le j \le u-1),$$
$$\beta_{i,j} = \oplus_{k=0}^{v-1} m_{j,k} \alpha_{i,k} (0 \le i \le 1, 0 \le j \le u-1).$$

Model 11. *For the operation* **S-box** *S, let*

$$(\underbrace{x_{0,0},\ldots,x_{0,v-1}}_{v},\ldots,\underbrace{x_{3,0},\ldots,x_{3,v-1}}_{v},\underbrace{\alpha_{0,0},\ldots,\alpha_{0,v-1}}_{\alpha_{0,i}=x_{0,i}\oplus x_{1,i}},\underbrace{\alpha_{1,0},\ldots,\alpha_{1,v-1}}_{\alpha_{1,i}=x_{2,i}\oplus x_{3,i}})$$

be the input upper composite unit, and

$$(\underbrace{y_{0,0},\ldots,y_{0,u-1}}_{u},\ldots,\underbrace{y_{3,0},\ldots,y_{3,u-1}}_{u},\underbrace{\beta_{0,0},\ldots,\beta_{0,u-1}}_{\beta_{0,i}=y_{0,i}\oplus y_{1,i}},\underbrace{\beta_{1,0},\ldots,\beta_{1,u-1}}_{\beta_{1,i}=y_{2,i}\oplus y_{3,i}})$$

be the output upper composite unit. Suppose that Con_0 represents the constraints for modeling the arbitrary mode by Model 3, Con_1 represents the constraints for modeling the differential propagation by Model 4, Con_2 represents the constraints for modeling the state propagation by Model 6, then the following boolean expressions are capable of modeling the propagation of the upper composite unit.

$$\begin{aligned}
&Con_0((\alpha_{i,v-1},\ldots,\alpha_{i,0}),(\beta_{i,u-1},\ldots,\beta_{i,0}))(0\le i\le 1),\\
&Con_1((\alpha_{i,v-1},\ldots,\alpha_{i,0}),(\beta_{i,u-1},\ldots,\beta_{i,0}))(0\le i\le 1),\\
&Con_2((x_{i,v-1},\ldots,x_{i,0}),(y_{i,u-1},\ldots,y_{i,0}))(0\le i\le 3).
\end{aligned}$$

Algorithm 3: The Composite Model Method for Determining the r-round T_4-IBD and T_5-IBD

```
Input: a tuple $(\alpha, \alpha', \beta, \beta')$, number of rounds $r$, type of IBD $T$
Output: variables of model $V$, constraints of model $C$
$V = [\ ], C = [\ ]$
$r_0 = \lceil r/2 \rceil,\ r_1 = r - r_0$
$x_0, x_1, x_2, x_3, \alpha_0, \alpha_1, y_0, y_1, y_2, y_3, \beta_0, \beta_1 = VarDeclare()$
$z_0, z_1, z_2, z_3, \gamma_0, \gamma_1, u_0, u_1, u_2, u_3, \delta_0, \delta_1 = VarDeclare()$
$V.AddList(x_0, x_1, x_2, x_3, \alpha_0, \alpha_1, y_0, y_1, y_2, y_3, \beta_0, \beta_1)$
$V.AddList(z_0, z_1, z_2, z_3, \gamma_0, \gamma_1, u_0, u_1, u_2, u_3, \delta_0, \delta_1)$
$C_u, V_u, RK_u = BuildUpperR(r_0, (x_0, x_1, x_2, x_3, \alpha_0, \alpha_1), (z_0, z_1, z_2, z_3, \gamma_0, \gamma_1))$
$C_l, V_l, RK_l = BuildLowerR(r_1, (u_0, u_1, u_2, u_3, \delta_0, \delta_1), (y_0, y_1, y_2, y_3, \beta_0, \beta_1))$
$V.AddList(V_u, V_l)$, $C.AddList(C_u, C_l)$
$C.AddList(u_0 = z_0, u_1 = z_1, u_2 = z_2, u_3 = z_3, \gamma_0 \oplus \gamma_1 \oplus \delta_0 \oplus \delta_1 = 0)$
if $T == T_5$ then
    $MK = VarDeclare()$
    $V.AddList(MK)$
    $C_K = BuildKeyULRelation(MK, RK_u, RK_l)$
    $C.AddList(C_K)$
end
$C.AddList(x_0 \oplus x_1 = \alpha, x_2 \oplus x_3 = \alpha')$
$C.AddList(y_1 \oplus y_2 = \beta, y_0 \oplus y_3 = \beta')$
return $V, C$
```

Based on the model method for the propagation of composite unit, we propose Algorithm 3 to determine whether a given $\mathcal{D} = (\alpha, \alpha', \beta, \beta')$ is an r-round T_4-IBD or T_5-IBD or not. The algorithm is briefly introduced as follows.

- Line 7-8: The function *BuildUpperR* is a function that models the propagation of the input upper composite unit $(x_0, x_1, x_2, x_3, \alpha_0, \alpha_1)$ to the output upper composite unit $(z_0, z_1, z_2, z_3, \gamma_0, \gamma_1)$ via r_0-round E. This function returns three output variables, among which C_u represents the constraints for modeling the propagation of upper composite unit, V_u represents the auxiliary variables, and RK_u represents the round keys. The function *BuildLowerR* is similar to function *BuildUpperR*.
- Line 11-16: In the case of determining the T_5-IBDs, we declare the master key MK and establish the relationship between MK, RK_u and RK_l according to the key schedule.

4.4 Searching the IBDs with mixed model method

Since it is possible to determine whether a given $\mathcal{D} = (\alpha, \alpha', \beta, \beta')$ is an r-round IBD or not, by confining $\mathcal{D}$ within a specific set, we can conduct a search for the IBDs. The goal of our search for IBDs is to find the IBDs that cover the largest possible number of rounds, and such distinguishers are conducive to key recovery. Thus, we focus on the following two search space and their subset. Here, if not specially pointed out, we always suppose the block size is n, the number of S-boxes per round is t, and the size of each S-box is c.

Searching for 1 active word IBDs. For S-box based block ciphers, we restrict the active S-box of α and α' to be the same, and the active S-box of β and β' to be the same. When the number of active S-boxes is 1, the size of the search space for searching $(1, 1, 1, 1)$-active word IBDs is $(t \times 2^c \times 2^c)^2 = t^2 2^{4c}$, such type of IBDs is called 1 active word IBDs.

Searching for 1 active word truncated IBDs. The size of the search space for searching $(1, 1, 1, 1)$-active word truncated IBDs is $(t \times t)^2 = t^4$, such type of IBDs is called 1 active word truncated IBDs.

As discussed above, the entire search space is rather large. Therefore, we propose the mixed model method to further accelerate our search. The core idea is detailed as follows. Based on the experimental results, it is maintained that $C_{T_0} < C_{T_4} < C_{T_5}$, where C_{T_i} represent the cost time for determining T_i-IBD. Particularly, it is rather fast to determine whether $((\alpha, \alpha'), (\beta, \beta'))$ is an r-round T_0-IBD or not. Meanwhile, based on the theoretical results, it is held that $S_{T_0} \subseteq S_{T_4} \subseteq S_{T_5}$. Thus, prior to determining whether $((\alpha, \alpha'), (\beta, \beta'))$ is an r-round T_5-IBD or not, we can determine whether $((\alpha, \alpha'), (\beta, \beta'))$ is an r-round T_0-IBD or T_4-IBD or not, In this way, it enables us to search for the IBDs effectively.

The overview of our algorithm is shown as Algorithm 4, it is briefly introduced as follows.

Algorithm 4: The Mixed Model Method for Searching the r-round T_5-IBD

```
   Input: the search space $\mathcal{S}$, number of rounds $r$
   Output: a set $\mathcal{G}$ that contains all $r$-round $T_5$-IBD in $\mathcal{S}$
1  $\mathcal{G} = [\,]$
2  $flag\_direct = DetermineT_5^t(r, \mathcal{S}_t, pre\_time)$
3  foreach $\mathcal{D}$ in $\mathcal{S}$ do
4      $flag = DetermineT_0(r, \mathcal{D})$
5      if not $flag$ then
6          if $flag\_direct$ then
7              $flag = DetermineT_5(r, \mathcal{D})$
8          else
9              $flag\_direct = false$
10             foreach $i$ in $\{0, \ldots, pre\_number - 1\}$ do
11                 $flag, result = DetermineT_4(r, \mathcal{D})$
12                 if $flag$ then
13                     $flag\_determine = true$
14                     break
15                 end
16                 $flag = DetermineT_{5'}(r, \mathcal{D}, InferDiff(result))$
17                 if not $flag$ then
18                     $flag\_determine = true$
19                     break
20                 end
21             end
22             if not $flag\_determine$ then
23                 $flag = DetermineT_5(r, \mathcal{D})$
24             end
25         end
26     end
27     if $flag$ then
28         $\mathcal{G}.append(\mathcal{D})$
29     end
30 end
31 return $\mathcal{G}$
```

- Line 2: The function $DetermineT_5^t(r, \mathcal{S}_t, pre\_time)$ which determines whether all $\mathcal{D} \in \mathcal{S}_t$ is an r-round T_5-IBD or not within the time pre_time. Here, $\mathcal{S}_t$ is a small subset of $\mathcal{S}$ and pre_time is a pre-defined time.
- Line 4: The function $DetermineT_0(r, \mathcal{D})$ is function that determines whether $\mathcal{D}$ is an r-round T_0-IBD or not according to Algorithm 1.
- Line 6-7: If we are able to determine the T_5-IBD within a reasonable time, then we will determine those IBDs directly.
- Line 11: The function $DetermineT_4(r, \mathcal{D})$ is function that determines whether $\mathcal{D}$ is an r-round T_4-IBD or not according to Algorithm 3.
- Line 15: In the case of $\mathcal{D}$ is not an r-round T_4-IBD, then we can obtain an T_4 boomerang trail. The function $DetermineT_{5'}(r, \mathcal{D}, InferDiff(result))$ is

function that determines whether $\mathcal{D}$ is an r-round T_5-IBD or not by simply modifying Algorithm 3. The *InferDiff*$(result)$ infers the difference values from T_4 boomerang trail, and those values are used to restrict the states values for determining the T_5-IBD. It should be noted that when this function returns true, it does not necessarily imply that $\mathcal{D}$ is an r-round T_5-IBD, and when it returns false, it means that $\mathcal{D}$ is not an r-round T_5-IBD.
- Line 19-20: If we fail to obtain the determined result from Line 10-18, then we will determine the T_5-IBD directly.

5 Applications to Search the Impossible Boomerang Distinguishers

In this section, we utilize our method on various block ciphers, including the large S-boxes based block AES-128 [31], the lightweight block cipher Midori64 [32] which adopts the almost MDS matrix, and the lightweight block cipher PRESENT-80 [33] which employs the bit permutation. Only brief descriptions of those block ciphers are provided here. For more details, please refer to their corresponding references. All the experiments in this paper are conducted on this platform: AMD(R) @2.60GHz, 80.00G RAM, 64-bit Ubuntu18.04 system.

5.1 AES-128

AES [31] is a 128-bit block cipher which supports key sizes of 128, 192, and 256 bits. There is undoubtedly that AES is one of the most renowned block ciphers across the world. Its design philosophy has exerted a profound influence on block ciphers. AES-128 is one version of AES, the specifications of AES-128 is detailed in Appendix B.1.

Configurations. By adopting the composite model method and the mixed model method, we utilize our tool to search for the 1 active word truncated IBDs. The size of the search space is $16^4 = 65536$, in order to get the result rapidly, we enable Cryptominisat to utilize 1 thread, and divide the search space into 16 parts and make a parallel call to Cryptominisat.

Results. After approximately 149.04 hours, we obtain 59584 4-round 1 active word truncated IBDs, this is the first result that getting all such IBDs. Those IBDs can be divided into two two types.

- Type-I: the IBDs that are constructed by the pure differential propagation, they can be detected in the AS mode, and the total number of such IBDs is 59392.
- Type-II: the IBDs that are constructed via the state propagation, they cannot be cannot be detected by the $\mathcal{UB}$-method, and the total number of such IBDs is 192.

Example of IBD and Manual Verification. We pick one of the type-I 4-round 1 active word truncated IBDs, and one of the type-II 4-round 1 active word truncated IBDs, and verify those two IBDs manually. For the type-I 4-round 1 active word truncated IBD, we get the following theorem.

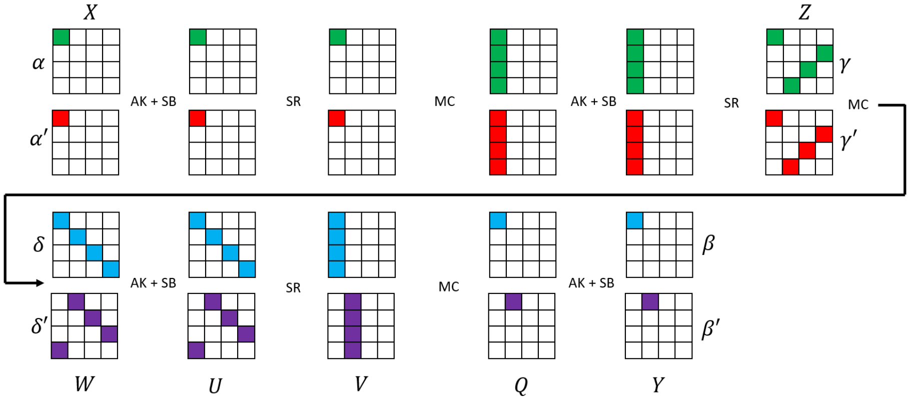

Fig. 3. The type-I 4-round impossible boomerang distinguisher for AES-128

Theorem 9. *The input differences* $(\alpha, \alpha') \in \{(0xa00000000000000, 0xa'000000 00000000)|a, a' \in \mathbb{F}_2^8/\{0\}\}$ *cannot propagate to the output differences* $(\beta, \beta') \in \{(0xb00000000000000b, 0x0b'00000000000000)|b, b' \in \mathbb{F}_2^8/\{0\}\}$ *after 4 rounds of AES-128 without the last SR and MC layer.*

Proof (proof by contradiction). Assume (α, α') can propagate to (β, β'), as shown in Figure 3, for $X_0, X_1 = X_0 \oplus \alpha, X_2, X_3 = X_2 \oplus \alpha'$, and $Y_0, Y_1, Y_2 = Y_1 \oplus \beta, Y_3 = Y_0 \oplus \beta'$, let Z_i be the value obtained by encrypting X_i after 2 rounds without the last MC layer, and W_i be the value obtained by decrypting Y_i after 2 rounds. Then $Z_0 \oplus Z_1 = \gamma$, $Z_2 \oplus Z_3 = \gamma'$, $W_1 \oplus W_2 = \delta$, and $W_0 \oplus W_3 = \delta'$.

One the one hand, since

$$W_{1,0} \oplus W_{2,0} = \delta_0 \neq 0, W_{0,0} \oplus W_{3,0} = 0,$$
$$W_{1,1} \oplus W_{2,1} = 0, W_{0,1} \oplus W_{3,1} = 0.$$

Therefore, $W_{0,0} \oplus W_{1,0} \oplus W_{2,0} \oplus W_{3,0} = \delta_0 \neq 0$ and $W_{0,1} \oplus W_{1,1} \oplus W_{2,1} \oplus W_{3,1} = 0$.

One the other hand, since

$$\begin{pmatrix} W_{0,0} \oplus W_{1,0} \oplus W_{2,0} \oplus W_{3,0} \\ W_{0,1} \oplus W_{1,1} \oplus W_{2,1} \oplus W_{3,1} \\ W_{0,2} \oplus W_{1,2} \oplus W_{2,2} \oplus W_{3,2} \\ W_{0,3} \oplus W_{1,3} \oplus W_{2,3} \oplus W_{3,3} \end{pmatrix} = M \cdot \begin{pmatrix} Z_{0,0} \oplus Z_{1,0} \oplus Z_{2,0} \oplus Z_{3,0} \\ Z_{0,1} \oplus Z_{1,1} \oplus Z_{2,1} \oplus Z_{3,1} \\ Z_{0,2} \oplus Z_{1,2} \oplus Z_{2,2} \oplus Z_{3,2} \\ Z_{0,3} \oplus Z_{1,3} \oplus Z_{2,3} \oplus Z_{3,3} \end{pmatrix}$$
$$= M \cdot \begin{pmatrix} \gamma_0 \oplus \gamma_0' \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Thus, $W_{0,0} \oplus W_{1,0} \oplus W_{2,0} \oplus W_{3,0} = 0$ and $W_{0,3} \oplus W_{1,3} \oplus W_{2,3} \oplus W_{3,3} = 0$, or $W_{0,0} \oplus W_{1,0} \oplus W_{2,0} \oplus W_{3,0} \neq 0$ and $W_{0,3} \oplus W_{1,3} \oplus W_{2,3} \oplus W_{3,3} \neq 0$, which is a contradiction. □

For the type-II 4-round 1 active word truncated IBD, we get the following theorem.

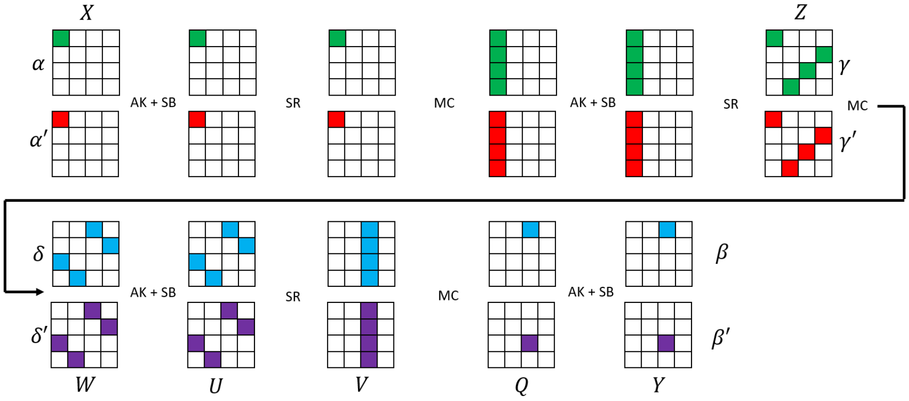

Fig. 4. The type-II 4-round impossible boomerang distinguisher for AES-128

Theorem 10. *The input differences* $(\alpha, \alpha') \in \{(0xa00000000000000, 0xa'00000000000000)|a, a' \in \mathbb{F}_2^8/\{0\}\}$ *cannot propagate to the output differences* $(\beta, \beta') \in \{(0x00b00000000000000, 0x000000000b'00000)|b, b' \in \mathbb{F}_2^8/\{0\}\}$ *after* 4 *rounds of AES-128 without the last SR and MC layer.*

Proof (proof by contradiction). Assume (α, α') can propagate to (β, β'), as shown in Figure 4, for $X_0, X_1 = X_0 \oplus \alpha, X_2, X_3 = X_2 \oplus \alpha'$, and $Y_0, Y_1, Y_2 = Y_1 \oplus \beta, Y_3 = Y_0 \oplus \beta'$, let Z_i be the value obtained by encrypting X_i after 2 rounds without the last MC layer and W_i be the value obtained by decrypting Y_i after 2 rounds. Then $Z_0 \oplus Z_1 = \gamma$, $Z_2 \oplus Z_3 = \gamma'$, $W_1 \oplus W_2 = \delta$, and $W_0 \oplus W_3 = \delta'$.

One the one hand, since

$$Q_{1,2} \oplus Q_{2,2} = \eta_2 \neq 0, Q_{0,2} \oplus Q_{3,2} = 0,$$
$$Q_{1,10} \oplus Q_{2,10} = 0, Q_{0,10} \oplus Q_{3,10} = \eta_{10} \neq 0.$$

then,

$$\begin{pmatrix} V_{0,2} \oplus V_{1,2} \oplus V_{2,2} \oplus V_{3,2} \\ V_{0,6} \oplus V_{1,6} \oplus V_{2,6} \oplus V_{3,6} \\ V_{0,10} \oplus V_{1,10} \oplus V_{2,10} \oplus V_{3,10} \\ V_{0,14} \oplus V_{1,14} \oplus V_{2,14} \oplus V_{3,14} \end{pmatrix} = (M^{-1}) \cdot \begin{pmatrix} Q_{0,2} \oplus Q_{1,2} \oplus Q_{2,2} \oplus Q_{3,2} \\ Q_{0,6} \oplus Q_{1,6} \oplus Q_{2,6} \oplus Q_{3,6} \\ Q_{0,10} \oplus Q_{1,10} \oplus Q_{2,10} \oplus Q_{3,10} \\ Q_{0,14} \oplus Q_{1,14} \oplus Q_{2,14} \oplus Q_{3,14} \end{pmatrix}$$
$$= (M^{-1}) \cdot \begin{pmatrix} \eta_2 \\ 0 \\ \eta_{10} \\ 0 \end{pmatrix}.$$

Thus, the values of $V_{0,10} \oplus V_{1,10} \oplus V_{2,10} \oplus V_{3,10}$ and $Q_{0,14} \oplus Q_{1,14} \oplus Q_{2,14} \oplus Q_{3,14}$ cannot be equal to 0 simultaneously. Assume $V_{0,10} \oplus V_{1,10} \oplus V_{2,10} \oplus V_{3,10} \neq 0$, thus $\delta_2 \oplus \delta_2' \neq 0$ and $\delta_3 \oplus \delta_3' = 0$.

One the other hand, similar to the proof in Theorem 9, it is the case that either $\delta_2 \oplus \delta_2' = 0$ and $\delta_3 \oplus \delta_3' = 0$, or $\delta_2 \oplus \delta_2' \neq 0$ and $\delta_3 \oplus \delta_3' \neq 0$, which is a contradiction. □

5.2 Midori64

Midori [32] is a lightweight block cipher which was designed by Banik et al. at AISACRYPT 2015. There exist two versions of Midori with state sizes of 64-bit and 128-bit, denoted as Midori64 and Midori128 respectively. The specifications of Midori64 is detailed in Appendix B.2.

Configurations. By adopting the composite model method and the mixed model method, we utilize our tool to search for the 1 active word truncated IBDs. The size of the search space is $16^4 = 65536$, in order to get the result rapidly, we enable Cryptominisat to utilize 1 thread, and divide the search space into 32 parts and make a parallel call to Cryptominisat.

Results. After around 65.57 hours, we demonstrate that there do not exist 7-round 1 active word truncated IBDs. Hence, we turn to search for the 6-round 1 active word truncated IBDs. After around 15.21 hours, we obtain 7296 6-round 1 active word truncated IBDs. In the case where the active S-boxes of α and α' are the same, or the the active S-boxes of β and β' are the same, the number of 6-round 1 active word truncated IBDs is 3456. Note that, the $\mathcal{UB}$-method also cannot take into account the details of the linear layers, all above IBDs cannot be detected by this method. Besides, Hu et al. showed that there no exist 6-round 1 active word truncated IDs [21], our result reveals that IBDs has an edge over IDs from the perspective of the number of rounds of the distinguishers.

Example of IBD and Manual Verification. We pick one of the 7296 6-round 1 active word truncated IBDs and verify this IBD manually, which leads to the following theorem.

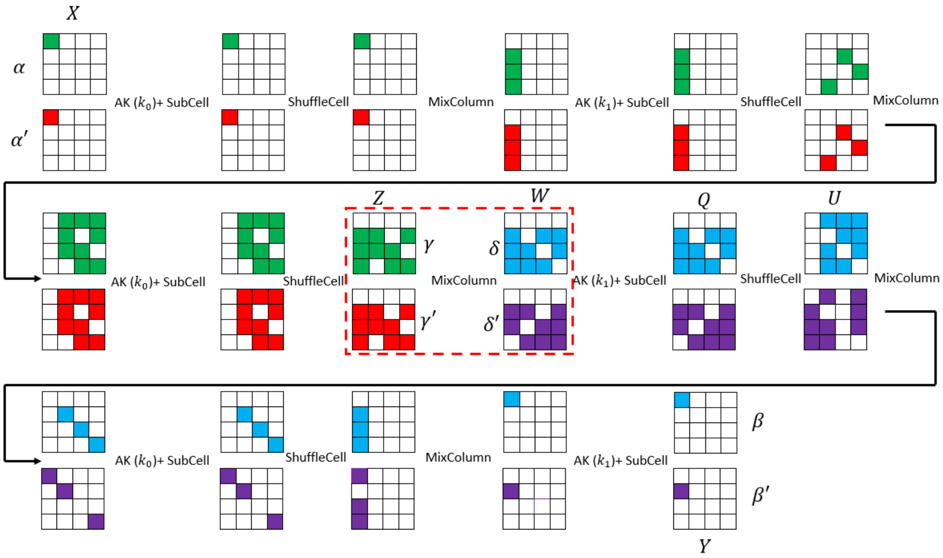

Fig. 5. The 6-round impossible boomerang distinguisher for Midori64

Theorem 11. *The input differences* $(\alpha, \alpha') \in \{(0xa000000000000000, 0xa'00000 0000000000)|a, a' \in \mathbb{F}_2^4/\{0\}\}$ *cannot propagate to the output differences* $(\beta, \beta') \in \{(0xb000000000000000, 0x0b'00000000000000)|b, b' \in \mathbb{F}_2^4/\{0\}\}$ *after* 6 *rounds of Midori64 without the last ShuffleCell and MixColumn layer.*

Proof (proof by contradiction). Assume (α, α') can propagate to (β, β'), as shown in Figure 5, for $X_0, X_1 = X_0 \oplus \alpha, X_2, X_3 = X_2 \oplus \alpha'$, and $Y_0, Y_1, Y_2 = Y_1 \oplus \beta, Y_3 = Y_0 \oplus \beta'$, let Z_i be the value obtained by encrypting X_i after 3 rounds without the last MixColumn layer, and W_i be the value obtained by decrypting Y_i after 3 rounds. Then $Z_0 \oplus Z_1 = \gamma$, $Z_2 \oplus Z_3 = \gamma'$, $W_1 \oplus W_2 = \delta$, and $W_0 \oplus W_3 = \delta'$, where $\gamma_8 \neq 0$, $\gamma_9 \neq 0$, $\gamma'_8 = 0$, and $\gamma_9 = 0$.

One the one hand, since

$$\begin{aligned}
\begin{pmatrix} W_{0,8} \oplus W_{1,8} \oplus W_{2,8} \oplus W_{3,8} \\ W_{0,9} \oplus W_{1,9} \oplus W_{2,9} \oplus W_{3,9} \\ W_{0,10} \oplus W_{1,10} \oplus W_{2,10} \oplus W_{3,10} \\ W_{0,11} \oplus W_{1,11} \oplus W_{2,11} \oplus W_{3,11} \end{pmatrix} &= M \cdot \begin{pmatrix} Z_{0,8} \oplus Z_{1,8} \oplus Z_{2,8} \oplus Z_{3,8} \\ Z_{0,9} \oplus Z_{1,9} \oplus Z_{2,9} \oplus Z_{3,9} \\ Z_{0,10} \oplus Z_{1,10} \oplus Z_{2,10} \oplus Z_{3,10} \\ Z_{0,11} \oplus Z_{1,11} \oplus Z_{2,11} \oplus Z_{3,11} \end{pmatrix} \\
&= M \cdot \begin{pmatrix} 0 \\ 0 \\ \gamma_{10} \oplus \gamma'_{10} \\ \gamma_{11} \oplus \gamma'_{11} \end{pmatrix} \\
&= \begin{pmatrix} \gamma_{10} \oplus \gamma'_{10} \oplus \gamma_{11} \oplus \gamma'_{11} \\ \gamma_{10} \oplus \gamma'_{10} \oplus \gamma_{11} \oplus \gamma'_{11} \\ \gamma_{11} \oplus \gamma'_{11} \\ \gamma_{10} \oplus \gamma'_{10} \end{pmatrix},
\end{aligned}$$

then $W_{0,8} \oplus W_{1,8} \oplus W_{2,8} \oplus W_{3,8} = W_{0,9} \oplus W_{1,9} \oplus W_{2,9} \oplus W_{3,9}$.

One the one hand, since $W_{1,8} \oplus W_{2,8} = \delta_8$, $W_{1,9} \oplus W_{2,9} = \delta_9$, $W_{1,8} \oplus W_{2,8} = \delta'_8$, and $W_{1,9} \oplus W_{2,9} = \delta'_9$, then $W_{0,8} \oplus W_{1,8} \oplus W_{2,8} \oplus W_{3,8} = \delta_8 \oplus \delta'_8 = 0$, and $W_{0,9} \oplus W_{1,9} \oplus W_{2,9} \oplus W_{3,9} = \delta_9 \oplus \delta'_9 \neq 0$. This is a contradiction. □

5.3 PRESENT-80

The PRESENT block cipher was designed by Bogdanov et al. in 2007 [33]. PRESENT-80 is one version of PRESENT, the specifications of PRESENT-80 is detailed in Appendix B.3.

Configurations. By adopting the composite model method, we search for the 1 active word IBDs, and restrict only the 0-th S-box of the input two differences to be active and the 0-th S-box of the output two differences to be active. The size of the search space is $15^4 = 50625$, in order to get the result rapidly, we enable Cryptominisat to utilize 1 thread, and divide the search space into 32 parts and make a parallel call to Cryptominisat.

Results. After around 24.52 hours, we demonstrate that there do not exist 7-round IBDs in our search space. Hence, we turn to search for the 6-round IBDs. After around 7.13 hours, we obtain 58 6-round 1 active word IBDs, all those IBDs cannot be detected by $\mathcal{UB}$-method.

Example of IBD and Manual Verification. We pick one of the 58 6-round 1 active word IBDs and verify this IBD manually. Our verification makes use

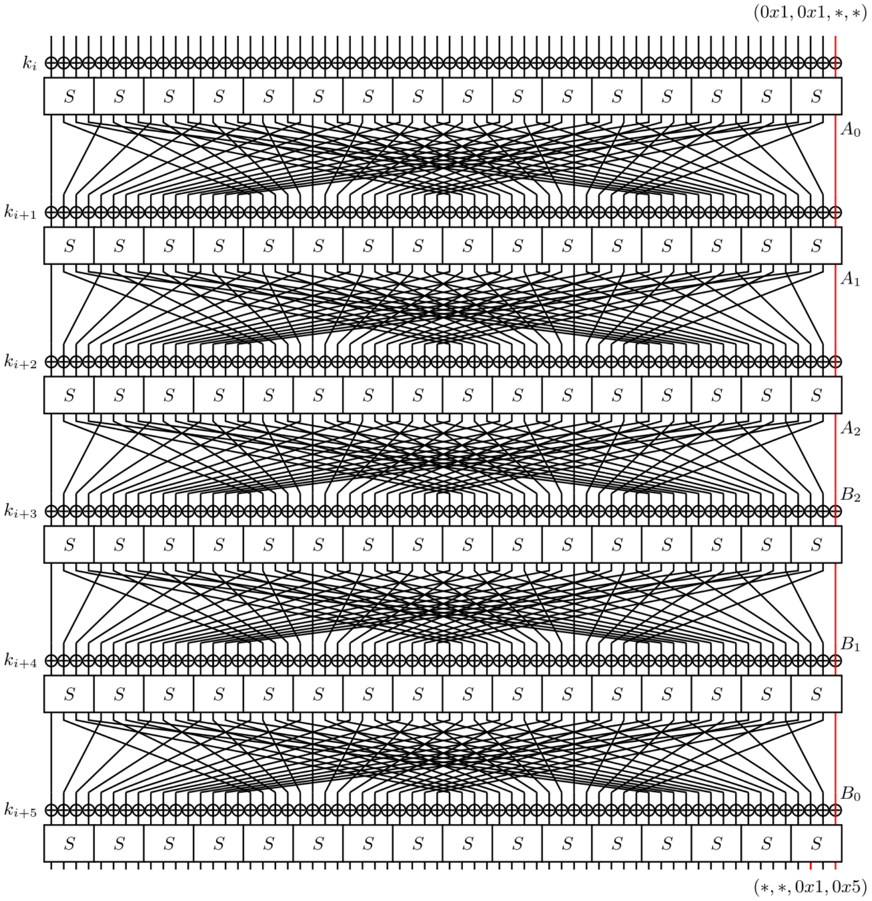

Fig. 6. The 6-round impossible boomerang distinguisher for PRESENT-80

of the definition of GEBCT, thus, we demonstrate some basic properties of the S-box of PRESENT-80 in the view of such table. The analysis reveals some new properties of the S-box of PRESENT.

Property 1. For $\rho, \rho', \theta, \theta', \varphi, \varphi' \in \mathbb{F}_2^4$, let $\mathcal{T}$ be the GEBCT of S, and $\theta_0, \theta_0', \varphi_0, \varphi_0'$ be the least bit of $\theta, \theta', \varphi, \varphi'$ respectively, the set $\{(\theta_0, \theta_0', \varphi_0, \varphi_0') | \mathcal{T}(1, 1, \rho, \rho', \theta, \theta', \varphi, \varphi') \neq 0\} = \{(1, 1, 0, 0), (1, 1, 1, 1)\}$.

Property 2. For $\mu, \mu', \rho, \rho', \theta, \theta', \varphi, \varphi' \in \mathbb{F}_2^4$, let $\mathcal{T}_{inv}$ be the GEBCT of invertible S-box S, and $\mu_0, \mu_0', \rho_0, \rho_0'$ be the least bit of μ, μ', ρ, ρ' respectively, the set $\{(\mu_0, \mu_0', \rho_0, \rho_0') | \mathcal{T}_{inv}(\mu, \mu', \rho, \rho', \theta, \theta', 1, 5)\} = \{(1, 0, 1, 0), (0, 1, 1, 0)\}$, and the set $\{(\mu_0, \mu_0', \rho_0, \rho_0') | \mathcal{T}_{inv}(\mu, \mu', \rho, \rho', \theta, \theta', \varphi, \varphi'), \theta, \theta', \varphi, \varphi' \in \{(1, 0, 1, 0), (0, 1, 1, 0)\}\} = \{(1, 0, 1, 0), (0, 1, 1, 0)\}$.

Theorem 12. *The input differences* $(0x0000000000000001, 0x0000000000000001)$ *cannot propagate to the output differences* $(0x0000000000000001, 0x0000000000000005)$ *after* 6 *rounds of PRESENT-80 without the last bit permutation layer.*

Proof. We use proof by contradiction to prove this theorem. Assume the input differences $(0x0000000000000001, 0x00000000000000\ 01)$ can propagate to the output differences $(0x0000000000000001, 0x00000000\ 00000005)$. As shown in Figure 6, let $(\theta_{i,0}, \theta'_{i,0}, \varphi_{i,0}, \varphi'_{i,0})$ be the least bit of the output of the S-boxes layer in the i-th round ($i = 0, 1, 2$), according to Property 1, it is holds that $A_0 = A_1 = A_2 = \{(1, 1, 0, 0), (1, 1, 1, 1)\}$, and $(\theta_{2,0}, \theta'_{2,0}, \varphi_{2,0}, \varphi'_{2,0}) \in A_2$.

Analogously, let $(\mu_{i,0}, \mu'_{i,0}, \rho_{i,0}, \rho'_{i,0})$ be the least bit of the input of the S-boxes layer in the i-th round ($i = 3, 4, 5$), according to Property 2, it is holds that $B_0 = B_1 = B_2 = \{(1, 0, 1, 0), (0, 1, 1, 0)\}$, and $(\mu_{3,0}, \mu'_{3,0}, \rho_{3,0}, \rho'_{3,0}) \in B_2$.

Since $(\theta_{2,0}, \theta'_{2,0}, \varphi_{2,0}, \varphi'_{2,0}) = (\mu_{3,0}, \mu'_{3,0}, \rho_{3,0}, \rho'_{3,0})$, $\theta_{2,0} = 1, \theta'_{2,0} = 1$, and one value of $\mu_{3,0}$ and $\mu'_{3,0}$ is 0. This is a contradiction. □

6 Conclusion and Future Work

In this paper, we explore the construction theory and automatic search approaches of the impossible boomerang distinguishers. Based on the novel technique for establishing ID and BD, we propose a series of IBDs in line with different construction methods, and examine the relationship among these IBD constructions. Finally, we develop a SAT-based tool for automatically searching the IBDs, and propose the composite model method and the mixed model method so as to achieve efficient and rapid search. This is the first automatic search tool which not only can take into account the details of each operation but also considers the impact of the key schedule in a single-key setting.

In our work, we only apply our method to the SPN structure block ciphers. Although our method is also capable of being applied to the Feistel structure block ciphers and the ARX structure block ciphers as well, whether there exists a better searching method for such block ciphers still requires further exploration. Additionally, our work merely focuses on the single-key setting, how to search for the IBDs in the related-key setting still awaits to be studied. All these aspects are considered as the future tasks.

References

1. Eli Biham and Adi Shamir. Differential cryptanalysis of des-like cryptosystems. In Alfred Menezes and Scott A. Vanstone, editors, *Advances in Cryptology - CRYPTO '90, 10th Annual International Cryptology Conference, Santa Barbara, California, USA, August 11-15, 1990, Proceedings*, volume 537 of *Lecture Notes in Computer Science*, pages 2–21. Springer, 1990.
2. Eli Biham, Alex Biryukov, and Adi Shamir. Cryptanalysis of skipjack reduced to 31 rounds using impossible differentials. In Jacques Stern, editor, *Advances in Cryptology - EUROCRYPT '99, International Conference on the Theory and Application of Cryptographic Techniques, Prague, Czech Republic, May 2-6, 1999, Proceeding*, volume 1592 of *Lecture Notes in Computer Science*, pages 12–23. Springer, 1999.

3. Lars R. Knudsen. Deal - a 128-bit block cipher. *Complexity*, 1998.
4. David A. Wagner. The boomerang attack. In Lars R. Knudsen, editor, *Fast Software Encryption, 6th International Workshop, FSE '99, Rome, Italy, March 24-26, 1999, Proceedings*, volume 1636 of *Lecture Notes in Computer Science*, pages 156–170. Springer, 1999.
5. Patrick Derbez and Pierre-Alain Fouque. Automatic search of meet-in-the-middle and impossible differential attacks. In Matthew Robshaw and Jonathan Katz, editors, *Advances in Cryptology - CRYPTO 2016 - 36th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 14-18, 2016, Proceedings, Part II*, volume 9815 of *Lecture Notes in Computer Science*, pages 157–184. Springer, 2016.
6. Hosein Hadipour, Sadegh Sadeghi, and Maria Eichlseder. Finding the impossible: Automated search for full impossible-differential, zero-correlation, and integral attacks. In Carmit Hazay and Martijn Stam, editors, *Advances in Cryptology - EUROCRYPT 2023 - 42nd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Lyon, France, April 23-27, 2023, Proceedings, Part IV*, volume 14007 of *Lecture Notes in Computer Science*, pages 128–157. Springer, 2023.
7. Alex Biryukov and Dmitry Khovratovich. Related-key cryptanalysis of the full AES-192 and AES-256. In Mitsuru Matsui, editor, *Advances in Cryptology - ASIACRYPT 2009, 15th International Conference on the Theory and Application of Cryptology and Information Security, Tokyo, Japan, December 6-10, 2009. Proceedings*, volume 5912 of *Lecture Notes in Computer Science*, pages 1–18. Springer, 2009.
8. Orr Dunkelman, Nathan Keller, and Adi Shamir. A practical-time related-key attack on the KASUMI cryptosystem used in GSM and 3g telephony. In Tal Rabin, editor, *Advances in Cryptology - CRYPTO 2010, 30th Annual Cryptology Conference, Santa Barbara, CA, USA, August 15-19, 2010. Proceedings*, volume 6223 of *Lecture Notes in Computer Science*, pages 393–410. Springer, 2010.
9. Jongsung Kim, Seokhie Hong, Bart Preneel, Eli Biham, Orr Dunkelman, and Nathan Keller. Related-key boomerang and rectangle attacks: Theory and experimental analysis. *IEEE Trans. Inf. Theory*, 58(7):4948–4966, 2012.
10. Jiqiang Lu. Cryptanalysis of block ciphers. *mat.uniroma3.it*.
11. Jiqiang Lu. The (related-key) impossible boomerang attack and its application to the AES block cipher. *Des. Codes Cryptogr.*, 60(2):123–143, 2011.
12. Nicolas T. Courtois and Gregory V. Bard. Algebraic cryptanalysis of the data encryption standard. In Steven D. Galbraith, editor, *Cryptography and Coding, 11th IMA International Conference, Cirencester, UK, December 18-20, 2007, Proceedings*, volume 4887 of *Lecture Notes in Computer Science*, pages 152–169. Springer, 2007.
13. Abdel Alim Kamal and Amr M. Youssef. Applications of SAT solvers to AES key recovery from decayed key schedule images. In Reijo Savola, Masaru Takesue, Rainer Falk, and Manuela Popescu, editors, *Fourth International Conference on Emerging Security Information Systems and Technologies, SECURWARE 2010, Venice, Italy, July 18-25, 2010*, pages 216–220. IEEE Computer Society, 2010.
14. Nicky Mouha and Bart Preneel. Towards finding optimal differential characteristics for arx: Application to salsa20. Cryptology ePrint Archive, Paper 2013/328, 2013. https://eprint.iacr.org/2013/328.
15. Nicky Mouha, Qingju Wang, Dawu Gu, and Bart Preneel. Differential and linear cryptanalysis using mixed-integer linear programming. In Chuankun Wu, Moti

Yung, and Dongdai Lin, editors, *Information Security and Cryptology - 7th International Conference, Inscrypt 2011, Beijing, China, November 30 - December 3, 2011. Revised Selected Papers*, volume 7537 of *Lecture Notes in Computer Science*, pages 57–76. Springer, 2011.

16. Siwei Sun, Lei Hu, Peng Wang, Kexin Qiao, Xiaoshuang Ma, and Ling Song. Automatic security evaluation and (related-key) differential characteristic search: Application to simon, present, lblock, DES(L) and other bit-oriented block ciphers. In Palash Sarkar and Tetsu Iwata, editors, *Advances in Cryptology - ASIACRYPT 2014 - 20th International Conference on the Theory and Application of Cryptology and Information Security, Kaoshiung, Taiwan, R.O.C., December 7-11, 2014. Proceedings, Part I*, volume 8873 of *Lecture Notes in Computer Science*, pages 158–178. Springer, 2014.
17. David Gérault, Marine Minier, and Christine Solnon. Constraint programming models for chosen key differential cryptanalysis. In Michel Rueher, editor, *Principles and Practice of Constraint Programming - 22nd International Conference, CP 2016, Toulouse, France, September 5-9, 2016, Proceedings*, volume 9892 of *Lecture Notes in Computer Science*, pages 584–601. Springer, 2016.
18. Siwei Sun, David Gérault, Pascal Lafourcade, Qianqian Yang, Yosuke Todo, Kexin Qiao, and Lei Hu. Analysis of aes, skinny, and others with constraint programming. *IACR Trans. Symmetric Cryptol.*, 2017(1):281–306, 2017.
19. Tingting Cui, Keting Jia, Kai Fu, Shiyao Chen, and Meiqin Wang. New automatic search tool for impossible differentials and zero-correlation linear approximations. *IACR Cryptol. ePrint Arch.*, page 689, 2016.
20. Yu Sasaki and Yosuke Todo. New impossible differential search tool from design and cryptanalysist aspects - revealing structural properties of several ciphers. In Jean-Sébastien Coron and Jesper Buus Nielsen, editors, *Advances in Cryptology - EUROCRYPT 2017 - 36th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Paris, France, April 30 - May 4, 2017, Proceedings, Part III*, volume 10212 of *Lecture Notes in Computer Science*, pages 185–215, 2017.
21. Xichao Hu, Yongqiang Li, Lin Jiao, Shizhu Tian, and Mingsheng Wang. Mind the propagation of states - new automatic search tool for impossible differentials and impossible polytopic transitions. In Shiho Moriai and Huaxiong Wang, editors, *Advances in Cryptology - ASIACRYPT 2020 - 26th International Conference on the Theory and Application of Cryptology and Information Security, Daejeon, South Korea, December 7-11, 2020, Proceedings, Part I*, volume 12491 of *Lecture Notes in Computer Science*, pages 415–445. Springer, 2020.
22. Sean Murphy. The return of the cryptographic boomerang. *IEEE Trans. Inf. Theory*, 57(4):2517–2521, 2011.
23. Alex Biryukov, Christophe De Cannière, and Gustaf Dellkrantz. Cryptanalysis of SAFER++. In Dan Boneh, editor, *Advances in Cryptology - CRYPTO 2003, 23rd Annual International Cryptology Conference, Santa Barbara, California, USA, August 17-21, 2003, Proceedings*, volume 2729 of *Lecture Notes in Computer Science*, pages 195–211. Springer, 2003.
24. Orr Dunkelman, Nathan Keller, and Adi Shamir. A practical-time related-key attack on the KASUMI cryptosystem used in GSM and 3g telephony. *J. Cryptol.*, 27(4):824–849, 2014.
25. Carlos Cid, Tao Huang, Thomas Peyrin, Yu Sasaki, and Ling Song. Boomerang connectivity table: A new cryptanalysis tool. In Jesper Buus Nielsen and Vincent Rijmen, editors, *Advances in Cryptology - EUROCRYPT 2018 - 37th Annual*

International Conference on the Theory and Applications of Cryptographic Techniques, Tel Aviv, Israel, April 29 - May 3, 2018 Proceedings, Part II, volume 10821 of *Lecture Notes in Computer Science*, pages 683–714. Springer, 2018.

26. Haoyang Wang and Thomas Peyrin. Boomerang switch in multiple rounds. application to AES variants and deoxys. *IACR Trans. Symmetric Cryptol.*, 2019(1):142–169, 2019.
27. Hamid Boukerrou, Paul Huynh, Virginie Lallemand, Bimal Mandal, and Marine Minier. On the feistel counterpart of the boomerang connectivity table introduction and analysis of the FBCT. *IACR Trans. Symmetric Cryptol.*, 2020(1):331–362, 2020.
28. Stéphanie Delaune, Patrick Derbez, and Mathieu Vavrille. Catching the fastest boomerangs application to SKINNY. *IACR Trans. Symmetric Cryptol.*, 2020(4):104–129, 2020.
29. Dachao Wang, Baocang Wang, and Siwei Sun. Sat-aided automatic search of boomerang distinguishers for ARX ciphers. *IACR Trans. Symmetric Cryptol.*, 2023(1):152–191, 2023.
30. Jiali Choy and Huihui Yap. Impossible boomerang attack for block cipher structures. In Tsuyoshi Takagi and Masahiro Mambo, editors, *Advances in Information and Computer Security, 4th International Workshop on Security, IWSEC 2009, Toyama, Japan, October 28-30, 2009, Proceedings*, volume 5824 of *Lecture Notes in Computer Science*, pages 22–37. Springer, 2009.
31. Joan Daemen and Vincent Rijmen. *The Design of Rijndael: AES - The Advanced Encryption Standard*. Information Security and Cryptography. Springer, 2002.
32. Subhadeep Banik, Andrey Bogdanov, Takanori Isobe, Kyoji Shibutani, Harunaga Hiwatari, Toru Akishita, and Francesco Regazzoni. Midori: A block cipher for low energy. In Tetsu Iwata and Jung Hee Cheon, editors, *Advances in Cryptology - ASIACRYPT 2015 - 21st International Conference on the Theory and Application of Cryptology and Information Security, Auckland, New Zealand, November 29 - December 3, 2015, Proceedings, Part II*, volume 9453 of *Lecture Notes in Computer Science*, pages 411–436. Springer, 2015.
33. Andrey Bogdanov, Lars R. Knudsen, Gregor Leander, Christof Paar, Axel Poschmann, Matthew J. B. Robshaw, Yannick Seurin, and C. Vikkelsoe. PRESENT: an ultra-lightweight block cipher. In Pascal Paillier and Ingrid Verbauwhede, editors, *Cryptographic Hardware and Embedded Systems - CHES 2007, 9th International Workshop, Vienna, Austria, September 10-13, 2007, Proceedings*, volume 4727 of *Lecture Notes in Computer Science*, pages 450–466. Springer, 2007.
34. Stephen A. Cook. The complexity of theorem-proving procedures. In Michael A. Harrison, Ranan B. Banerji, and Jeffrey D. Ullman, editors, *Proceedings of the 3rd Annual ACM Symposium on Theory of Computing, May 3-5, 1971, Shaker Heights, Ohio, USA*, pages 151–158. ACM, 1971.
35. Ling Song, Xianrui Qin, and Lei Hu. Boomerang connectivity table revisited. application to SKINNY and AES. *IACR Trans. Symmetric Cryptol.*, 2019(1):118–141, 2019.
36. Ahmed Abdelkhalek, Yu Sasaki, Yosuke Todo, Mohamed Tolba, and Amr M. Youssef. MILP modeling for (large) s-boxes to optimize probability of differential characteristics. *IACR Trans. Symmetric Cryptol.*, 2017(4):99–129, 2017.

A Proofs

Theorem 2

Proof (proof by contradiction). If an r-round T_0-IBD $((\alpha, \alpha'), (\beta, \beta'))$ is not an r-round T_1-*IBD*, there must exist one r-round T_1 boomerang trail:

$$(\alpha, \alpha') \to \cdots \to \underbrace{(\gamma, \gamma')(\delta, \delta')}_{\gamma \oplus \gamma' \oplus \delta \oplus \delta' = 0} \to \cdots \to (\beta, \beta'),$$

which is an r-round T_0 boomerang trail. Thus, $((\alpha, \alpha'), (\beta, \beta'))$ is neither an r-round T_0-*IBD*. □

Theorem 3

Proof (proof by contradiction). If an r-round T_1-IBD $((\alpha, \alpha'), (\beta, \beta'))$ is not an r-round T_2-*IBD*, there must exist one r-round T_2 boomerang trail:

$$(\alpha, \alpha') \to \cdots \to \underbrace{(\gamma, \gamma') \to (\delta, \delta')}_{(\gamma, \gamma') \xrightarrow{GBCT} (\delta, \delta')} \to \cdots \to (\beta, \beta').$$

As depicted in Figure 1, for each parallel S-box in E^m, $GBCT(\mu, \mu', \varphi, \varphi') \neq 0$, i.e., there exist u_1, u_2 such that $S(u_1) \oplus S(u_2) = \varphi$ and $S(u_1 \oplus \mu) \oplus S(u_2 \oplus \mu') = \varphi'$. Let $\rho = u_1 \oplus u_2$ and $\rho' = u_0 \oplus u_3 = u_1 \oplus \mu \oplus u_2 \oplus \mu'$, it holds that $\mu \oplus \mu' \oplus \rho \oplus \rho' = 0$. Since other operations in E^m are linear, there exist ω, ω' such that $\gamma \oplus \gamma' \oplus \omega \oplus \omega' = 0$, $\omega \overset{E^m}{\to} \delta$ and $\omega' \overset{E^m}{\to} \delta'$. Hence,

$$(\alpha, \alpha') \to \cdots \to \underbrace{(\gamma, \gamma')(\omega, \omega')}_{\gamma \oplus \gamma' \oplus \omega \oplus \omega' = 0} \to (\delta, \delta') \to \cdots \to (\beta, \beta').$$

Thus, $((\alpha, \alpha'), (\beta, \beta'))$ is neither an r-round T_1-*IBD*. □

Theorem 4

Proof (proof by contradiction). If an r-round T_2-IBD $((\alpha, \alpha'), (\beta, \beta'))$ is not an r-round T_3-*IBD*, there must exist one r-round T_3 boomerang trail:$(\epsilon_0^0 = \alpha, \epsilon_1^0, \epsilon_2^0 = \alpha', \epsilon_3^0) \xrightarrow{GEBCT} \cdots \xrightarrow{GEBCT} (\epsilon_0^{r_0}, \epsilon_1^{r_0}, \epsilon_2^{r_0}, \epsilon_3^{r_0}) \xrightarrow{GEBCT} (\epsilon_0^{r_0+1}, \epsilon_1^{r_0+1}, \epsilon_2^{r_0+1}, \epsilon_3^{r_0+1}) \xrightarrow{GEBCT} \cdots \xrightarrow{GEBCT} (\epsilon_0^r, \epsilon_1^r = \beta, \epsilon_2^r, \epsilon_3^r) = \beta'$. According to the definitions of GEBCT, DDT, and GBCT for S-boxes, it holds that

$$\begin{aligned}
&\mathrm{GEBCT}(\mu, \mu', \rho, \rho', \theta, \theta', \varphi, \varphi') \subseteq \mathrm{DDT}(\mu, \theta) \times \mathrm{DDT}(\mu', \theta'),\\
&\mathrm{GEBCT}(\mu, \mu', \rho, \rho', \theta, \theta', \varphi, \varphi') \subseteq \mathrm{DDT}(\rho, \varphi) \times \mathrm{DDT}(\rho', \varphi'),\\
&\mathrm{GEBCT}(\mu, \mu', \rho, \rho', \theta, \theta', \varphi, \varphi') \subseteq \mathrm{GBCT}(\mu, \mu', \theta, \theta').
\end{aligned}$$

Hence,

$$(\alpha, \alpha') \to \cdots \to \underbrace{(\gamma, \gamma') \to (\delta, \delta')}_{(\gamma, \gamma') \xrightarrow{GBCT} (\delta, \delta')} \to \cdots \to (\beta, \beta')$$

is an r-round T_2 boomerang trail. Thus, $((\alpha, \alpha'), (\beta, \beta'))$ is neither an r-round T_2-*IBD*. □

Theorem 5

Proof (proof by contradiction). If an r-round T_P-IBD $((\alpha, \alpha'), (\beta, \beta'))$ is not an r-round T_3-*IBD*, there must exist at least one r-round T_3 boomerang trail: $(\epsilon_0^0, \epsilon_1^0, \epsilon_2^0, \epsilon_3^0) \xrightarrow{GEBCT} \cdots \xrightarrow{GEBCT} (\epsilon_0^{r_0}, \epsilon_1^{r_0}, \epsilon_2^{r_0}, \epsilon_3^{r_0}) \xrightarrow{GEBCT} (\epsilon_0^{r_0+1}, \epsilon_1^{r_0+1}, \epsilon_2^{r_0+1}, \epsilon_3^{r_0+1}) \xrightarrow{GEBCT} \cdots \xrightarrow{GEBCT} (\epsilon_0^r, \epsilon_1^r, \epsilon_2^r, \epsilon_3^r)$. According to the definitions of various tables, it holds that

$$\begin{aligned}
&\mathrm{GEBCT}(\mu, \mu', \rho, \rho', \theta, \theta', \varphi, \varphi') \subseteq \mathrm{DDT}^2_{upper}(\mu, \mu', \theta, \theta'),\\
&\mathrm{GEBCT}(\mu, \mu', \rho, \rho', \theta, \theta', \varphi, \varphi') \subseteq \mathrm{DDT}^2_{lower}(\rho, \rho', \varphi, \varphi'),\\
&\mathrm{GEBCT}(\mu, \mu', \rho, \rho', \theta, \theta', \varphi, \varphi') \subseteq \mathrm{GBCT}(\mu, \mu', \theta, \theta'),\\
&\mathrm{GEBCT}(\mu, \mu', \rho, \rho', \theta, \theta', \varphi, \varphi') \subseteq \mathrm{GUBCT}(\mu, \mu', \theta, \theta', \varphi, \varphi'),\\
&\mathrm{GEBCT}(\mu, \mu', \rho, \rho', \theta, \theta', \varphi, \varphi') \subseteq \mathrm{GLBCT}(\mu, \mu', \rho, \rho', \varphi, \varphi').
\end{aligned}$$

Hence, it is also an r-round T_P boomerang trail. Thus, $((\alpha, \alpha'), (\beta, \beta'))$ is neither an r-round T_P-*IBD*. □

Theorem 6

Proof (proof by contradiction). According to the definitions, an r-round T_5 boomerang trail is also an r-round T_4 boomerang trail. □

Theorem 7

Proof. (Definition 5⇒ Construction 7) Let $((\alpha, \alpha'), (\beta, \beta'))$ be an r-round IBD, then any pair of plaintexts (x_0, x_3) cannot simultaneously satisfy $E_k(x_0) \oplus E_k(x_3) = \beta$ and $E_k(x_0 \oplus \alpha) \oplus E_k(x_3 \oplus \alpha') = \beta'$. If $((\alpha, \alpha'), (\beta, \beta'))$ is not an r-round T_5-IBD. Let $x_0^0 = x_0$, $x_1^0 = x_0 \oplus \alpha$, $x_3^0 = x_3$, $x_2^0 = x_3 \oplus \alpha'$, there exist an r-round T_5 boomerang trail $(x_0^0, x_1^0, x_2^0, x_3^0) \to \cdots \to (x_0^r, x_1^r, x_2^r, x_3^r)$, where $x_1^r \oplus x_2^r = \beta$ and $x_0^r \oplus x_3^r = \beta'$. Thus $E_k(x_0) \oplus E_k(x_3) = \beta$ and $E_k(x_0 \oplus \alpha) \oplus E_k(x_3 \oplus \alpha') = \beta'$, which is a contradiction.

(Construction 7⇒ Definition 5) Let $((\alpha, \alpha'), (\beta, \beta'))$ be an r-round T_5-IBD. then there is not any r-round T_5 boomerang trail $(x_0^0, x_1^0, x_2^0, x_3^0) \to \cdots \to (x_0^r, x_1^r, x_2^r, x_3^r)$. Thus, any pair of (x_0^0, x_3^0) cannot simultaneously meet $E_k(x_0^0) \oplus E_k(x_3^0) = \beta$ and $E_k(x_0^0 \oplus \alpha) \oplus E_k(x_3^0 \oplus \alpha') = \beta'$, which is according with Definition 5. □

Theorem 8

Proof. This is equivalent to prove that $(\epsilon_0^0, \epsilon_1^0, \epsilon_2^0, \epsilon_3^0) \overset{\text{AddKey}}{\to} (\gamma_0^0, \gamma_1^0, \gamma_2^0, \gamma_3^0) \overset{\text{GEBCT}}{\to} (\delta_0^0, \delta_1^0, \delta_2^0, \delta_3^0) \overset{\text{LL}}{\to} (\epsilon_0^1, \epsilon_1^1, \epsilon_2^1, \epsilon_3^1) \overset{\text{AddKey}}{\to} \cdots \overset{\text{GEBCT}}{\to} (\delta_0^{r-1}, \delta_1^{r-1}, \delta_2^{r-1}, \delta_3^{r-1})$ is an r-round T_3 boomerang trail if and only if $(x_0^0, x_1^0, x_2^0, x_3^0) \overset{\text{AddKey}}{\to} (y_0^0, y_1^0, y_2^0, y_3^0) \overset{\text{SL}}{\to} (z_0^0, z_1^0, z_2^0, z_3^0) \overset{\text{LL}}{\to} (x_0^1, x_1^1, x_2^1, x_3^1) \overset{\text{AddKey}}{\to} \cdots \overset{\text{SL}}{\to} (z_0^{r-1}, z_1^{r-1}, z_2^{r-1}, z_3^{r-1})$ is an r-round T_4 boomerang trail, where $\alpha = \epsilon_0^0, \alpha' = \epsilon_0^2, \beta = \epsilon_1^{r-1}$, $\beta' = \epsilon_3^{r-1}$, and $\alpha = x_0^0 \oplus x_1^0, \alpha' = x_2^0 \oplus x_3^0$, $\beta = z_1^{r-1} \oplus z_2^{r-1}$ and $\beta' = z_0^{r-1} \oplus z_3^{r-1}$. In particular, we prove this in the case of $r = 3$. The other cases can be proved analogously. Suppose $(\epsilon_0^0, \epsilon_1^0, \epsilon_2^0, \epsilon_3^0) \overset{\text{AddKey}}{\to} (\gamma_0^0, \gamma_1^0, \gamma_2^0, \gamma_3^0) \overset{\text{GEBCT}}{\to} (\delta_0^0, \delta_1^0, \delta_2^0, \delta_3^0) \overset{\text{LL}}{\to} (\epsilon_0^1, \epsilon_1^1, \epsilon_2^1, \epsilon_3^1) \overset{\text{AddKey}}{\to} \cdots \overset{\text{GEBCT}}{\to} (\delta_0^2, \delta_1^2, \delta_2^2, \delta_3^2)$ is an 3-round T_3 boomerang trail. Since $(\gamma_0^i, \gamma_1^i, \gamma_2^i, \gamma_3^i) \overset{\text{SL, GEBCT}}{\to} (\delta_0^i, \delta_1^i, \delta_2^i, \delta_3^i)$, there exists $(y_0^i, y_1^i, y_2^i, y_3^i)$ and $(z_0^i, z_1^i, z_2^i, z_3^i)$, such that

$$\begin{aligned}
&y_0^i \oplus y_1^i = \gamma_0^i, y_1^i \oplus y_2^i = \gamma_1^i, y_2^i \oplus y_3^i = \gamma_2^i, y_0^i \oplus y_3^i = \gamma_3^i,\\
&z_0^i \oplus z_1^i = \delta_0^i, z_1^i \oplus z_2^i = \delta_1^i, z_2^i \oplus z_3^i = \delta_2^i, z_0^i \oplus z_3^i = \delta_3^i.
\end{aligned}$$

Let $x_i^0 = y_i^0 \oplus k_0$ $(0 \leq i \leq 3)$ and $k_j = LL(z_x^{j-1}) \oplus y_i^j$ $(0 \leq i \leq 3, j = 1, 2)$, then $(x_0^0, x_1^0, x_2^0, x_3^0) \overset{\text{AddKey}}{\to} (y_0^0, y_1^0, y_2^0, y_3^0) \overset{\text{SL}}{\to} (z_0^0, z_1^0, z_2^0, z_3^0) \overset{\text{LL}}{\to} (x_0^1, x_1^1, x_2^1, x_3^1) \overset{\text{AddKey}}{\to} \cdots \overset{\text{SL}}{\to} (z_0^{r-1}, z_1^{r-1}, z_2^{r-1}, z_3^{r-1})$ is an 3-round T_4 boomerang trail. The above process is invertible. □

B Specifications of Block Ciphers

B.1 Specifications of AES-128

AES-128 [31] is a 128-bit block cipher whose key sizes is 128 bits. There is undoubtedly that AES is one of the most renowned block ciphers across the world. Its design philosophy has exerted a profound influence on block ciphers. The internal state is regarded as a square array of bytes as follows where $s_i \in \mathbb{F}_2^8$ $(0 \leq i \leq 15)$.

$$S = \begin{pmatrix} s_0 & s_4 & s_8 & s_{12} \\ s_1 & s_5 & s_9 & s_{13} \\ s_2 & s_6 & s_{10} & s_{14} \\ s_3 & s_7 & s_{11} & s_{15} \end{pmatrix}.$$

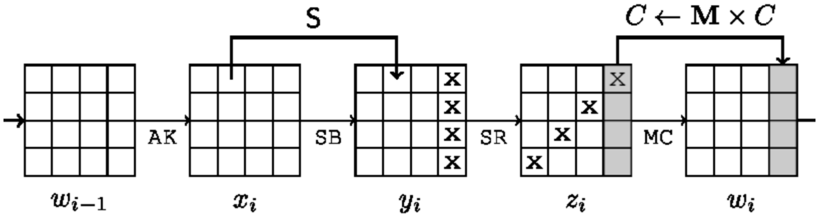

Fig. 7. One Round of Block Cipher AES

One encryption round of AES-128 is depicted in Figure 7, and it consists of the following four operations:

AddRoundKey(AK): The 128-bit round key which is derived from the key schedule is XORed with the state.

SubBytes(SB): Applying the -bit S-box to each byte in parallel to the cipher's internal state.

ShiftRows(SR): The i-th rows $(0 \leq i \leq 3)$ of the internal state is rotated by i bytes form right to left.

MixColumns(MC): Each column of the internal state is multiplied with the MDS matrix.

The key schedule of AES-128 is shown as Figure 8. The function g is a 32-bit to 32-bit function which consists of: 1) a right rotation of the input by 1 word; 2) processing all four bytes of this rotated input through the AES S-box; 3) the addition of a fixed round coefficient to the output of the first S-box.

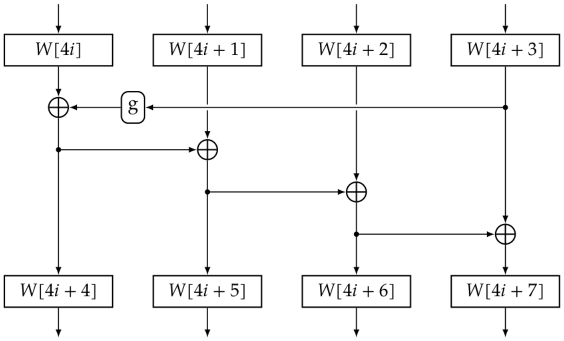

Fig. 8. The Key Schedule of AES-128

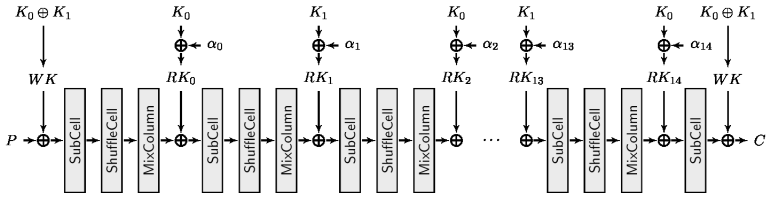

Fig. 9. The Overview of Block Cipher Midori64

B.2 Specifications of Midori64

Midori64 [32] is a lightweight block cipher which was designed by Banik et al. at AISACRYPT 2015. The overall process is as shown in the Figure 9. Midori64 has a 64-bit state size and its key size is 128-bit. It utilizes the following 4×4 array as a data expression:

$$S = \begin{pmatrix} s_0 & s_4 & s_8 & s_{12} \\ s_1 & s_5 & s_9 & s_{13} \\ s_2 & s_6 & s_{10} & s_{14} \\ s_3 & s_7 & s_{11} & s_{15} \end{pmatrix}$$

where the size of each cell is 4-bit.

Table 2. The S-box of Midori64

x	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$S(x)$	12	10	13	3	14	11	15	7	8	9	1	5	0	2	4	6

The round function of Midori64 is composed of the following 4 operations:

SubCell Apply the non-linear 4×4 S-box (as shown in Table 2) in parallel to each nibble of the state.

ShuffleCell Each nibble of the state is performed in the following way:

$$(s_0, s_1, \ldots, s_{15}) \leftarrow (s_0, s_{10}, s_5, s_4, s_{11}, s_1, s_9, s_3, s_{12}, s_6, s_7, s_{13}, s_2, s_8)\,.$$

MixColumn Midori64 utilizes an almost MDS matrix M, it is applied to every 4-nibble column of the state S:

$$M = \begin{pmatrix} 0\,1\,1\,1 \\ 1\,0\,1\,1 \\ 1\,1\,0\,1 \\ 1\,1\,1\,0 \end{pmatrix}.$$

KeyAdd The 64-bit round key rk_i is XORed to the state S.

The key schedule of Midori64 is rather simple. A 128-bit key K is represented as two 64-bit keys k_0 and k_1, that is $K = k_0 \| k_1$. The whitening key and the last sub-key are $rk_{-1} = rk_{R-1} = k_0 \oplus k_1$, and the sub-key for round i is $rk_i = k_{(imod2)} \oplus \alpha_i$, where $0 \leq i \leq R - 2$ and α_i is constant.

B.3 Specifications of PRESENT-80

Table 3. The S-box of PRESENT

x	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$S(x)$	12	5	6	11	9	0	10	13	3	14	15	8	4	7	1	2

Table 4. The bit-permutation of PRESENT

i	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$P(i)$	0	16	32	48	1	17	33	49	2	18	34	50	3	19	35	51
i	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
$P(i)$	4	20	36	52	5	21	37	53	6	22	38	54	7	23	39	55
i	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47
$P(i)$	8	24	40	56	9	25	41	57	10	26	42	58	11	27	43	59
i	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63
$P(i)$	12	28	44	60	13	29	45	61	14	30	46	62	15	31	47	63

The PRESENT-80 block cipher was designed by Bogdanov et al. in 2007 [33]. It employs a 64-bit state where the state can be viewed as a concatenation of 16 nibbles. The round function of it involves an XOR with the round key, the application of a 4-bit S-box (as shown in Table 3) in parallel to the state and a bit permutation (as shown in Table 4).

For the key schedule of PRESENT-80, the master key is stored in a register K and is represented as $\mathrm{k}_{79}\mathrm{k}_{78}\cdots\mathrm{k}_{0}$. At round i, the round key K_i consists of the 64 leftmost bits of the current content of the register K:

$$K_i = \mathrm{k}_{79}\mathrm{k}_{78}\ldots\mathrm{k}_{16}.$$

Once the round key has been extracted, the register K is updated as follows:

$$\begin{aligned}
&[\mathrm{k}_{79}\mathrm{k}_{78}\ldots\mathrm{k}_{1}\mathrm{k}_{0}] = [\mathrm{k}_{18}\mathrm{k}_{17}\ldots\mathrm{k}_{20}\mathrm{k}_{19}]\\
&[\mathrm{k}_{79}\mathrm{k}_{78}\mathrm{k}_{77}\mathrm{k}_{76}] = S\,[\mathrm{k}_{79}\mathrm{k}_{78}\mathrm{k}_{77}\mathrm{k}_{76}]\\
&[\mathrm{k}_{19}\mathrm{k}_{18}\mathrm{k}_{17}\mathrm{k}_{16}\mathrm{k}_{15}] = [\mathrm{k}_{19}\mathrm{k}_{18}\mathrm{k}_{17}\mathrm{k}_{16}\mathrm{k}_{15}] \oplus \mathrm{round\_counter}.
\end{aligned}$$