Improved (related-key) Attacks on Round-Reduced KATAN-32/48/64 Based on the Extended Boomerang Framework

Jiageng Chen[1(✉)], Je Sen Teh[2(✉)], Chunhua Su[3], Azman Samsudin[2], and Junbin Fang[4]

[1] Computer School, Central China Normal University, Wuhan 430079, China
jiageng.chen@mail.ccnu.edu.cn
[2] School of Computer Sciences, Universiti Sains Malaysia, George Town, Malaysia
jesen_teh@hotmail.com, Azman.samsudin@usm.my
[3] School of Information Science,
Japan Advanced Institute of Science and Technology,
1-1 Asahidai, Nomi, Ishikawa 923-1292, Japan
chsu@jaist.ac.jp
[4] Department of Optoelectronic Engineering,
Jinan University, Guangzhou 510632, China
junbinfang@gmail.com

Abstract. The boomerang attack is one of the many extensions of the original differential attack. It has been widely applied to successfully attack many existing ciphers. In this paper, we investigate an extended version of the boomerang attack and show that it is still a very powerful tool especially in the related-key setting. A new branch-and-bound searching strategy which involves the extended boomerang framework is then introduced. We provide an improved cryptanalysis on the KATAN family (a family of hardware-oriented block ciphers proposed in CHES 2009) based on the boomerang attack. In the related-key setting, we were able to greatly improve upon the previous results to achieve the best results, namely 150 and 133 rounds by far for KATAN48/64 respectively. For KATAN32 in the related-key setting and all KATAN variants in the single-key setting, our results are the best ones in the differential setting although inferior to the meet-in-the-middle attack.

Keywords: KATAN32/48/64 · Related-key attack · Boomerang attack · Differential attack

1 Introduction

The statistical attack is one of the most effective attacks against symmetric key cryptography. It includes many popular cryptanalysis techniques such as the linear attack, differential attack and so on. Among these methods, the differential attack is one of the most popular approaches due to its wide range of

J.K. Liu and R. Steinfeld (Eds.): ACISP 2016, Part II, LNCS 9723, pp. 333–346, 2016.
DOI: 10.1007/978-3-319-40367-0_21

applications to many ciphers including DES and AES. More importantly, it has many variations such as the impossible differential attack [5], multi-differential attack [7] and others which make differential attacks more flexible compared to linear attacks. Among these variations, the boomerang attack [22] proposed by Wagner back in 1999 provides an interesting approach to differential cryptanalysis. By considering quartets of differences instead of pairs, the attack separates traditional cipher distinguishers into two parts. This way, the burden of finding good differential characteristics can be greatly eased, leading to better distinguishers. The amplified boomerang attack [14] and rectangle attack [3] were later proposed to improve the efficiency of the boomerang attack. Unlike the original version which requires adaptive chosen plaintext and ciphertext queries, the modified boomerang attacks only require chosen plaintext queries which is a more practical attack assumption. The power of this attack has been demonstrated when it was used to break the full-round AES-192/256 [6] in the related-key setting. Since the boomerang attack falls under the differential attack framework, one natural question is which of these two methods will lead to better results. Although there are a lot of recent research work focusing on exploiting the relationship between statistical attacks such as the differential and linear attacks [8] as well as the zero correlation linear and integral attacks [21], the relationship between the boomerang and differential attack has not been fully investigated. However, the boomerang attack often outperforms the differential attack which suggests that under the condition of limited computing resources, the boomerang attack is a feasible option.

The design of lightweight block ciphers and cryptanalysis of these ciphers have recently attracted a lot of research attention. The KATAN family proposed in CHES 2009 [9] is one example. Although the cipher KTANTAN [9] proposed by the same authors was broken with a meet-in-the-middle attack [23], the KATAN family still remains secure after many years of cryptanalysis. There have been several attacks on the KATAN family in both single-key and related-key settings. In the single-key setting, a conditional differential attack [15] was able to break 78, 70 and 68 rounds of KATAN32/48/64 respectively. In [2], the authors took advantage of the full differential distribution to improve the attack on KATAN32, breaking 115 rounds. However, this approach cannot be applied on KATAN48/64 since the full differential distribution cannot be easily computed. Later on, more research work put focus on meet-in-the-middle attacks (MITM) [10,12,13,24]. In particular, [20] was recently published on e-print claiming to break 206 rounds of KATAN32 using MITM. The cube attack was also applied to KATAN32 in the single key model [1] with better results than the differential attack.

In the related-key setting, [16] introduced 120, 103 and 90-round attacks on KATAN32/48/64 respectively. By taking advantage of the key scheduling, [11] further improved the result to 174, 145 and 130 rounds respectively using the boomerang attack. In this paper, we investigate the extended boomerang technique to improve upon the previously found boomerang differential characteristics. As a result, we can achieve the best records in attacking KATAN48/64 in

the related-key setting. In all the other cases, while the results are inferior to the MITM attacks, we are able to deliver the best differential attack results so far. Particularly in the single key setting, our approach is able to outperform the attack on KATAN32 [2] which uses the full differential distribution. Although their distinguisher is better than ours, we point out that using the full distribution will result in an inefficient key recovery attack. Their methods are also not applicable to larger block sizes. From this point of view, our approach is more realistic in practice. We summarize our results along with previous related results in Table 1.

We organize the paper as follows: In Sect. 2, the boomerang attack and its extended version are described. In Sect. 3, we demonstrate the boomerang distinguisher search and key recovery attack on the KATAN family in both single-key and related-key settings. Finally, we conclude our paper with a summary of findings in Sect. 4.

2 The Framework of the Boomerang Cryptanalysis

Ever since its proposal, differential cryptanalysis [5] quickly became one of the main cryptanalytic methods used today. Based on its original form, researchers have later derived many extended variants such as truncated differential cryptanalysis, multi-differential cryptanalysis and so on. The boomerang attack can also be viewed as an extension of differential attack, but it is more unique because it modifies the original attack in a structural manner. Let m be the block size of a block cipher E, and we assume E to be a cascade cipher consisting of three concatenated parts $E_K = E_2 \circ E_1 \circ E_0$ influenced by a secret key K. Here E is a n-bit to n-bit keyed permutation $E : \{0,1\}^n \times \{0,1\}^k \rightarrow \{0,1\}^n$. E_2 is the final rounds where the subkey bits are the primary target whereas $E_1 \circ E_0$ is the distinguisher.

Boomerang Attack. The motivation behind the boomerang attack is that finding two short efficient differential distinguishers is easier than finding a long one. The original version of the boomerang attack is a combination of a chosen plaintext and ciphertext attack, which is a cryptanalysis model that makes very strong assumptions with regard to the capabilities of an attacker. Furthermore, the original boomerang attack is not efficient when performing the last round attack due to its "boomerang" property. Later, the amplified boomerang attack was proposed to solve these problems. Given that the rectangle attack is an extension of the original boomerang attack, we will refer to the amplified or rectangle attack as a boomerang throughout the paper.

In the chosen plaintext setting, an attacker chooses plaintext pairs with differences (α, α), and expect the differences between C_1, C_3 and C_2, C_4 to be (δ, δ). Randomly, $P_R((\alpha, \alpha) \rightarrow (\delta, \delta)) = 2^{-2m}$, thus the boomerang distinguisher should have probability greater than 2^{-2m}. For E_0, the attacker searches for high probability differential paths $\alpha \rightarrow \beta_i$, where $0 \leq i \leq 2^m - 1$. For any differential path

Table 1. Comparison of attacks against KATAN family

Cipher	Attacking Technique Technique	# Attacking Rounds	Time Complexity	Data Complexity	Memory Complexity	Reference
KATAN32	Differential (Single Key)	78	2^{76}	2^{16} CP	Not given	[15]
	MitM (Single Key)	110	2^{77}	138 KP	$2^{75.1}$	[12]
	Differential (Single Key)	115	2^{79}	138 KP	$2^{75.1}$	[2]
	Boomerang (Single Key)	**117**	$\mathbf{2^{79.3}}$	$\mathbf{2^{27.3}}$ **CP**	$\mathbf{2^{29.9}}$	**Ours**
	MitM (Single Key)	119	$2^{79.1}$	144 CP	$2^{79.1}$	[13]
	MitM (Single Key)	153	$2^{78.5}$	2^5 CP	2^{76}	[10]
	Cube (Single Key)	155	$2^{78.3}$	2^{32} CP	$2^{33.5}$	[1]
	MitM (Single Key)	175	$2^{79.3}$	3 KP	$2^{79.58}$	[24]
	MitM (Single Key)	206	2^{79}	3 KP	$2^{78.1}$	[20]
	Differential (Related Key)	120	2^{31}	Practical (CP)	Practical	[16]
	Boomerang (Related Key)	174	$2^{78.8}$	$2^{27.6}$ CP	$2^{26.6}$	[11]
	Boomerang (Related Key)	**187**	$\mathbf{2^{78.4}}$	$\mathbf{2^{31.8}}$ **CP**	$\mathbf{2^{33.9}}$	**Ours**
KATAN48	Differential (Single Key)	70	2^{78}	2^{31} CP	Not given	[15]
	Boomerang (Single Key)	**87**	$\mathbf{2^{78}}$	$\mathbf{2^{36.7}}$ **CP**	$\mathbf{2^{39.3}}$	**Ours**
	MitM (Single Key)	100	2^{78}	128 KP	2^{78}	[12]
	MitM (Single Key)	105	$2^{79.1}$	144 KP	$2^{79.1}$	[13]
	MitM (Single Key)	129	$2^{78.5}$	2^5 CP	2^{76}	[10]
	MitM (Single Key)	130	$2^{79.45}$	2 KP	2^{79}	[24]
	MitM (Single Key)	148	2^{79}	2 KP	2^{77}	[20]
	Differential (Related Key)	103	2^{25}	Practical (CP)	Practical	[16]
	Boomerang (Related Key)	145	$2^{78.5}$	$2^{38.4}$ CP	$2^{37.4}$	[11]
	Boomerang (Related Key)	**150**	$\mathbf{2^{77.6}}$	$\mathbf{2^{47.2}}$ **CP**	$\mathbf{2^{49.8}}$	**Ours**
KATAN64	Differential (Single Key)	68	2^{78}	2^{32} CP	Not given	[15]
	Boomerang (Single Key)	**72**	$\mathbf{2^{78}}$	$\mathbf{2^{55.1}}$ **CP**	$\mathbf{2^{58.1}}$	**Ours**
	MitM (Single Key)	94	$2^{77.68}$	116 KP	$2^{77.68}$	[12]
	MitM (Single Key)	99	$2^{79.1}$	142 KP	$2^{79.1}$	[13]
	MitM (Single Key)	112	$2^{79.45}$	2 KP	2^{79}	[24]
	MitM (Single Key)	119	$2^{78.5}$	2^5 CP	2^{74}	[10]
	MitM (Single Key)	129	2^{79}	2 KP	2^{77}	[20]
	Differential (Related Key)	90	2^{27}	Practical (CP)	Practical	[16]
	Boomerang (Related Key)	130	$2^{78.1}$	$2^{53.1}$ CP	$2^{52.1}$	[11]
	Boomerang (Related Key)	**133**	$\mathbf{2^{78.5}}$	$\mathbf{2^{58.4}}$ **CP**	$\mathbf{2^{61.4}}$	**Ours**

KP: Known Plaintext, CP: Chosen Plaintext

$\alpha \to \beta_i$ starting from a message pair P_1, P_2, the attacker expects that the differential path starting from the message pair P_3, P_4 should have the same form. Thus after E_0, the probability cost is $\sum_{i=0}^{r-1} p_i^2$ where $r < 2^m$ and $p_i = P(\alpha \to \beta_i)$.

Two edges in the middle quartet have the difference value β_i. Therefore if we assume the third edge to have a difference γ_j with a random probability of 2^{-m}, then the last edge will have difference value γ_j with probability 1 since the XOR sum of the quartet edges should be 0. Here again we can choose as many γ_j as possible where j is also bounded by the block size 2^m. For E_1 due to the middle quartet shift, we start from two γ_j differences and hope to reach the output difference δ. Denote $q_j = P(\gamma_j \rightarrow \delta)$, then the probability can be similarly computed as $\sum_{j=0}^{t-1} q_j^2, t < 2^m$. The total probability can be computed as:

$$P_{bmg}((\alpha, \alpha) \rightarrow (\delta, \delta)) = \sum_{i=0}^{r-1} p_i^2 \cdot \sum_{j=0}^{t-1} q_j^2 \cdot 2^{-m}$$

Since $P_{bmg}((\alpha, \alpha) \rightarrow (\delta, \delta)) > 2^{-2m}$, thus we need:

$$P_{bmg-dist} = \sum_{i=0}^{r-1} p_i^2 \cdot \sum_{j=0}^{t-1} q_j^2 > 2^{-m} \tag{1}$$

Here $P_{bmg-dist}$ denotes the distinguisher probability, which is consistent with previous work such as in [11]. Please refer to Fig. 1 for the boomerang model.

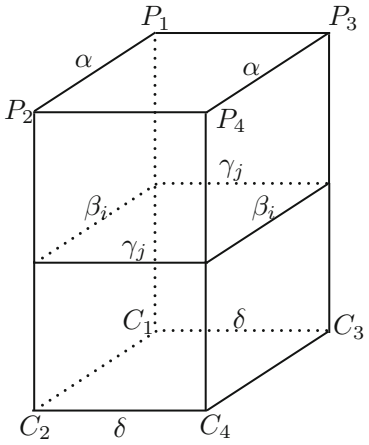

Fig. 1. The model of Boomerang attack

The framework of the boomerang can be further improved by considering various differential quartets in the middle. The idea was first introduced in [22] and was later mentioned in [4]. We refer to the concept as an extended boomerang in this paper. In the boomerang setting, we are assuming that E_0 has two differential paths $\alpha \rightarrow \beta_i$ that has to appear at the same time so that the middle quartet has the format such as $(\beta_i, \beta_i, \gamma_j, \gamma_j)$. However, the two differential paths

in E_0 need not be the same, thus we actually missed a lot of combinations in the middle. For example, let us consider the following scenario:

$$E_0 : p_i = P(\alpha \to \beta_i), p_j = P(\alpha \to \beta_j)$$

$$E_1 : q_s = P(\gamma_s \to \delta), q_t = P(\gamma_t \to \delta)$$

$$Quartet : (\beta_i, \beta_j, \gamma_s, \gamma_t) \text{ satisfying } \beta_i \oplus \beta_j \oplus \gamma_s \oplus \gamma_t = 0$$

Now we have all combinations in the middle quartet that can still lead to the output difference δ. This will potentially increase the total probability when all these cases are taken into consideration. Let u and v be the size of the differential set for $\alpha \to \beta_i$ and $\gamma_s \to \delta$ respectively. This leads us to the new calculation formula:

$$P_{exBmg} = \sum_{j=0}^{u-1} \sum_{i=0}^{u-1} p_{\beta_i} \cdot p_{\beta_j} \times \sum_{s} \sum_{t} q_{\gamma_s} \cdot q_{\gamma_t} \times 2^{-m}$$

Once $\beta_i, \beta_j, \gamma_s$ is decided in the middle quartet, γ_t is determined with probability one, namely, $\gamma_t = \beta_i \oplus \beta_j \oplus \gamma_s$, thus we have:

$$P_{exBmg} = \sum_{j=0}^{u-1} \sum_{i=0}^{u-1} p_{\beta_i} \cdot p_{\beta_j} \times \sum_{i=0}^{u-1} \sum_{j=0}^{u-1} \sum_{s=0}^{v-1} q_{\gamma_s} \cdot q_{\beta_i \oplus \beta_j \oplus \gamma_s} \times 2^{-m} \tag{2}$$

To be consistent with the previous boomerang distinguisher for ease of comparison, we denote the first part of probability term to be $\hat{p}^2$, and second part to be $\hat{q}^2$. We then define the probability for the extended boomerang distinguisher to be:

$$P_{exBmg-dist} = \hat{p}^2 \times \hat{q}^2 > 2^{-m}$$

which should be greater than the random case 2^{-m}.

3 KATAN Family

The KATAN block cipher family comprises of three lightweight block ciphers KATAN32, KATAN48 and KATAN64 whose block sizes are 32 bits, 48 bits and 64 bits respectively. It was proposed in CHES 2009 [9] and it is a well known cipher in the area. The design is based on the linear feedback shift register (LFSR) and supports an 80-bit key.

The key scheduling function expands an 80-bit user-provided key k_i $(0 \leq i < 80)$ into a 508-bit subkey sk_i $(0 \leq i < 508)$ by the following linear operations,

$$sk_i = \begin{cases} k_i \ (0 \leq i < 80), \\ k_{i-80} \oplus k_{i-61} \oplus k_{i-50} \oplus k_{i-13} \ (80 \leq i < 508). \end{cases}$$

These operations are expressed as an 80-bit LFSR whose polynomial is $x_{80} + x_{61} + x_{50} + x_{13} + 1$ as shown in Fig. 2.

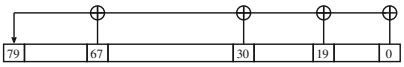

Fig. 2. Key scheduling function of KATAN32/48/64

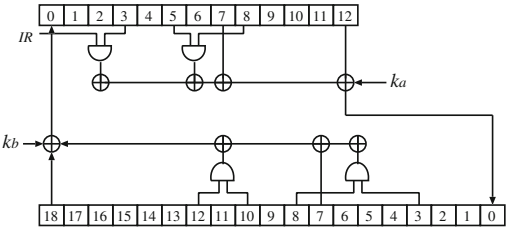

Fig. 3. Round function of KATAN32

In the round function, each bit of a plaintext is loaded into registers L_1 and L_2. Then, these are updated as follows:

$$
\begin{aligned}
f_a(L_1) &= L_1[x_1] \oplus L_1[x_2] \oplus (L_1[x_3] \cdot L_1[x_4]) \oplus (L_1[x_5] \cdot IR) \oplus k_a, \\
f_b(L_2) &= L_2[y_1] \oplus L_2[y_2] \oplus (L_2[y_3] \cdot L_2[y_4]) \oplus (L_2[y_5] \cdot L_2[y_6]) \oplus k_b, \\
L_1[i] &= L_1[i-1] \ (1 \le i < |L_1|), \quad L_1[0] = f_b(L_2), \\
L_2[i] &= L_2[i-1] \ (1 \le i < |L_2|), \quad L_2[0] = f_a(L_1),
\end{aligned}
$$

where $\oplus$ and $\cdot$ are bitwise XOR and AND operations, respectively, and $L[x]$ denotes the x-th bit of L, IR is the round constant value defined in the specification, and k_a and k_b are two subkey bits. Table 2 shows the detailed parameters of KATAN32/48/64. For round i, k_a and k_b correspond to $sk_{2(i-1)}$ and $sk_{2(i-1)+1}$, respectively. After 254 rounds (1-254 round), values of registers are output as a ciphertext. Fig. 3 illustrates the round function of KATAN32.

Table 2. Parameters of KATAN family

Algorithm	$\|L_1\|$	$\|L_2\|$	x_1	x_2	x_3	x_4	x_5	y_1	y_2	y_3	y_4	y_5	y_6
KATAN32	13	19	12	7	8	5	3	18	7	12	10	8	3
KATAN48	19	29	18	12	15	7	6	28	19	21	13	15	6
KATAN64	25	39	24	15	20	11	9	38	25	33	21	14	9

4 Improved Attack on KATAN Family

4.1 Novel Searching Strategy

The basic searching strategy used to find differentials is a branch-and-bound algorithm divided into two parts. The first part (*single trail search*) is based on the branch-and-bound algorithm proposed by Matsui [19]. It performs a search for individual differential paths that have the best probabilities. These paths are then used in the second part of the algorithm (*cluster search*) which expands

the search to find other paths that start from the same input difference and lead to the same output difference. Any paths found by the *cluster search* improves the differential probability of the paths found by *single trail search*.

As an exhaustive search using this algorithm would take a long time, several bounds are imposed onto the search. The first bound, θ is used in the *single trail search*. When $\theta = 1$, only paths with the best probabilities will be stored for the *cluster search* whereas $\theta = 0$ will store every path exhaustively. When the θ bound is loosened, the paths found can range from high probability paths to extremely low probability paths. To filter out paths with low probability, a second bound λ is used. As an example, if $\lambda = 2^{-16}$, only paths with probabilities larger than 2^{-16} will be stored for the *cluster search*. The *cluster search* itself has a bound μ which ranges from [0,1] (similar to θ).

For ciphers with block size less than 32-bit, it is possible to derive all the differential paths, so that the size of the differential set u or v could reach $2^m - 1$. However, for larger size greater than say 48 bits, we are still bounded to searching a subset of all differential paths with relatively high probabilities. Based on the extended boomerang framework, we derive the following advanced algorithm which can be used to improve the cryptanalytic capability of the boomerang attack:

Extended Boomerang Characteristic Searching Algorithm.

1. For E_0 precompute the good differential paths $(\alpha \rightarrow \beta_i)$ using branch-and-bound algorithm where $i \leq u$. Store all the β_i in a set Φ.
2. Proceed similarly for E_1 to find paths $(\gamma_j \rightarrow \delta)$, $j \leq v$, and save the output differences in a set Ω.
3. For all the $\beta_i, \beta_j \in \Phi$ and $\gamma_s \in \Omega$, compute $\gamma_t = \beta_i \oplus \beta_j \oplus \gamma_s$. If $\gamma_t \in \Omega$, then $(\beta_i, \beta_j, \gamma_s, \gamma_t)$ is a valid quartet, and we can add the corresponding paths' probability to the total boomerang probability.

4.2 Related Key Boomerang Distinguisher Search

To perform a basic boomerang search, the *single trail search* and *cluster search* algorithms are performed separately for E_0 and E_1. As the clustering effect for E_1 starts from one output difference δ to multiple intermediary differences γ, the branch-and-bound algorithm has to be applied in reverse (decryption) starting from δ to find multiple γ. The search is performed for various combinations of E_0 and E_1 rounds to find the optimal middle point for the boomerang attack. In the related key setting, the search algorithm also involves key differences. As a starting point, we build upon the findings of Isobe et al. [11] who found a 140-round distinguisher with a probability of $2^{-27.2}$. In their paper, they identified sets of plaintext/key differences that lead to *blank steps* that have no differences in registers and subkeys. We use these sets as the inputs of our E_0 search and also use them to find ciphertext/key sets for the reverse E_1 search.

To find starting points for the E_1 search (ciphertext and key differences), we first perform the E_0 search starting from a designated intermediate round.

E.g. if the number of rounds for E_0 is 70, we start our search from round 71 onwards. By using the same sets from [11] as a guide, we obtain the output and key differences which are used as inputs to the E_1 search. The best results for the basic related key boomerang search is shown in Table 3 with the following settings: $(\theta = 0, \mu = 0.5, \lambda = 2^{-20})$. It can be seen that the branch-and-bound algorithm is able to improve Isobe's 140-round distinguisher probability $(2^{-27.2} \rightarrow 2^{-26.58214})$. We are also able to push a valid distinguisher for 2 more rounds to obtain a 142-round distinguisher with probability of $2^{-30.58214}$.

Next, the extended search algorithm is applied where β_i and γ_s from the basic boomerang search are stored in sets Φ and Ω respectively. We found that for certain values of α and δ, the extended search is unable to find additional quartets, therefore did not improve the existing distinguisher probability. However, there also exist other α and δ values that lead to a large amount of additional quartets. The following settings were used for the branch-and-bound search: $(\theta = 0, \mu = 0.5)$ whereas the λ bound varies based on input. We provide only the best result found in Table 4 where large improvements to the overall distinguisher probability are obtained. We are able to improve upon the previously found 142-round distinguisher by 12 rounds, obtaining a 154-round distinguisher with a probability of $2^{-29.7209}$ after applying the extended boomerang search.

The conditional difference is another technique which has been used in previous research work such as [11,16]. For KATAN, the only non-linear part is the AND logic gate. According to the AND table, if we fix one of the two inputs to the AND gate to be 0, then any difference in the other input will be canceled out and the final output difference will be 0. Based on this observation, we can improve the probability of E_0 by fixing some of the plaintext bits. The downside of using this technique is that the message space will be reduced, thus we have to determine if the probability gain will surpass the decrease of the message space. Fortunately, the extended boomerang technique can potentially amplify the effect of the conditional difference approach due to the extra quartets we can collect in the middle. For KATAN32, we set $L_2[1] = L_2[4] = L_2[8] = 0$ for the input difference $\alpha = 10020040$ and key differences located at k_6, k_{25}. As a result, we can improve the distinguisher probability to $2^{-23.7209}$. The results of the distinguisher for KATAN32/48/64 are located in Table 4. The application of the extended boomerang in the single key setting follows the same steps, but with the exclusion of key differences. Please refer to the appendix for the distinguishers in the single-key setting. The overall related key extended boomerang search algorithm is summarized below:

1. Identify an input set that leads to *blank rounds* for the E_0 search. For this input, determine the fixed bits for the conditional difference technique.
2. Perform *single trail search* and *cluster search* for $\#E_0$ rounds. Store all intermediary differences, β_i in a set Φ along with their probabilities (which have been improved using the sufficient condition technique).
3. Identify an input set that leads to *blank rounds* for the E_1 search. Using this input set, start from $(\#E_0 + 1)$ rounds and perform the *single trail search* for $\#E_1$ rounds to obtain the corresponding output difference, δ and output key difference.

4. Using δ and output key difference as a starting point, the *single trail search* and *cluster search* is performed in reverse (decryption) starting from round-$(\#E_0+\#E_1)$ for $\#E_1$ rounds. Store all intermediary differences, γ_i in a set Ω along with their probabilities.
5. For all the $\beta_i, \beta_j \in \Phi$ and $\gamma_s \in \Omega$, compute $\gamma_t = \beta_i \oplus \beta_j \oplus \gamma_s$. If $\gamma_t \in \Omega$, then $(\beta_i, \beta_j, \gamma_s, \gamma_t)$ is a valid quartet, and we can add the corresponding paths' probability to the total boomerang probability.

Table 3. Related Key Boomerang Distinguisher on KATAN32 (before extended search)

α	$\#E_0$	Prob p (log_2)	δ	$\#E_1$	Prob q (log_2)	Total Rounds	Final Prob (log_2)
10020040	70	−6.79	280184	70	−6.5	140	−26.58
10020040	70	−6.79	280184	71	−7.5	141	−28.58
10020040	70	−6.79	280184	72	−8.5	142	−30.58

4.3 Key Recovery Attacks

Finally, we demonstrate the concrete key recovery attack for the KATAN family in both related-key and single-key setting. [11] has already provided an optimized key recovery framework. Because each round is rather cheap for the KATAN family and we want to add many rounds in E_2, the differential pattern will be lost. This makes sieving techniques impossible. In other words, the key recovery technique in [11] is not related to the exact output difference values, thus it is easy to seamlessly apply here for a fair comparison. The principle of the attack and some facts of KATAN family used in the attack are listed below:

1. Use meet-in-the-middle approach to recover the key. This is achieved by storing all the ciphertexts pairs in a table, guessing the subkey bits for decryption then checking for matches in the table.
2. The differential state is known after $\#E_2$ rounds by only guessing $(\#E_2 - 4)$-round subkeys.

Table 4. Boomerang distinguisher for KATAN32/48/64 in the related-key setting

Ciphers	Total rounds	$\#E_0$	$\#E_1$	α	δ	Before ES$(log2)$	After ES$(log2)$	After SC	$\lambda_{E_0}/\lambda_{E_1}$ difference	key conditions	plaintext
KATAN32	154	70	84	1002 0040	280 184	-42.43	-29.72	-23.72	$2^{-24}/2^{-22}$	k_6, k_{25}	$L_2[1] = L_2[4] = L_2[8] = 0$
KATAN48	126	63	63	c00018 0c0600	800000 001051	-58.15	-46.40	-32.40	$2^{-22}/2^{-22}$	k_0, k_{19}	$L_2[0] = L_2[1] = L_2[2] = L_2[11] = L_2[17] = 0$, $L_2[10] \neq L_2[18]$
KATAN64	116	56	60	1c00e 00000	100000 703800	-62.43	-50.84	-42.84	$2^{-18}/2^{-26}$	k_{11}	$L_2[9] = L_2[10] = L_2[11] = L_2[33] = 0$

3. A trade-off trick can be achieved by using the partial matching method which involves matching only part of the differential state instead of the whole. This technique is also known as the "early abort" mentioned in paper [17] and [18]. Denote P_r as the probability that a subkey candidate is the correct key, which is supposed to be $N^2 \times 2^{-2m}$, where N is the number plaintext pairs. Let r denote the number of rounds that we do not guess subkeys (except for the first skipped round, we guess 1 bit). By using the partial matching technique, we can improve the probability as follows:
   (a) KATAN32: $P_r = N^2 \times 2^{-86+4r}$, $r \geq 6$, known difference bits when matching is $S_{matching} = 43 - 2r$.
   (b) KATAN48: $P_r = N^2 \times 2^{-120+8r}$, $r \geq 4$, known difference bits when matching is $S_{matching} = 59 - 4r$.
   (c) KATAN64: $P_r = N^2 \times 2^{-152+12r}$, $r \geq 3$, known difference bits when matching is $S_{matching} = 74 - 6r$.

$\#E_2$ denotes the number of rounds for the key recovery phase, then the subkey bits we need to guess is denoted by $2(\#E_2 - r) + 1$. Since N is the number plaintext pairs required, then we can generate N^2 quartets. To assure that the right quartet will appear, we set $N = 2^{\frac{m}{2}} \times P_r^{-1/2}$. Since we adapt the meet-in-the-middle approach, two pairs of plaintexts and ciphertexts need to be processed independently, thus the data complexity D is $2^{\frac{m}{2}+1} \times P_r^{-1/2}$. The key recovery steps are as follows:

1. Choose N plaintext pairs (P_1, P_2) and (P_3, P_4) such that $P_1 \oplus P_2 = P_3 \oplus P_4 = \alpha$, ask for ciphertexts C_1, C_2, C_3 and C_4 under secret key K_1, K_2, K_3 and K_4.
2. For each guess of $2(\#E_2 - r) + 1$ bits of subkey for K_i (Guess one K_i and others are determined), do the following:
   (a) For both (C_1, C_2), derive $S_{matching}$ bits of known differences by decrypting t rounds. XOR with δ and store in the big table.
   (b) For each pair (C_3, C_4), do
      i. Decrypt $\#E_2$ rounds and compute the $S_{matching}$ bits of the known difference.
      ii. Check if the value matches the ones stored in the table. If it exists, proceed the following step.
      iii. Brute force search the rest of the $80 - (2(\#E_2 - r) + 1) = 79 - 2(\#E_2 - r)$ unknown bits. Verify with fresh plaintext and ciphertext pairs, output the correct key if passed.

Step 2(a) and 2(b)-i requires to compute $(\frac{2^{2(\#E_2-r-1)} \times N \times 2 \times \#E_2}{\#E_0+\#E_1+\#E_2}\ \#E_0+\#E_1+\#E_2)$ rounds of KATAN32/48/64. Then after filtering, we have $2^{2(\#E_2-r)+1} \times P_r$ key candidates remaining. To brute force search the rest key bits, step(b)-iii takes $2^{2(\#E_2-r)+1} \times P_r \times 2^{79-2(\#E_2-r)} = 2^{80} \times P_r$. As a result, the total time complexity can be denoted as

$$T = 2 \times \frac{2^{2(\#E_2-r-1)} \times N \times 2 \times \#E_2}{\#E_0 + \#E_1 + \#E_2} + 2^{80} \times P_r$$

The memory complexity depends on Step2(a) where $2 \times N$ state values need to be stored.

Now based on the derived distinguishers for both single-key and related-key settings, we test all the possible variables for $\#E_2$ and r to derive the optimal results shown in Tables 5 and 6 respectively.

Table 5. Cryptanalysis results for KATAN family in the single-key setting

Ciphers	Total rounds	$\#E_0$	$\#E_1$	$\#E_2$	r	Dist Prob(log_2)	$T(log_2)$	$D(log_2)$	$MEM(bytes)$
KATAN32	117	35	48	34	7	-21.78	79.25	27.89	29.89
KATAN48	87	35	25	27	5	-23.36	78.00	36.68	39.26
KATAN64	72	30	26	16	3	-44.26	77.99	55.13	58.13

Table 6. Cryptanalysis results for KATAN family in the Related-key setting

Ciphers	Total rounds	$\#E_0$	$\#E_1$	$\#E_2$	r	Dist Prob(log_2)	$T(log_2)$	$D(log_2)$	$MEM(bytes)$
KATAN32	187	70	84	33	7	-23.72	78.39	31.86	33.86
KATAN48	150	63	63	24	4	-32.40	77.60	47.20	49.79
KATAN64	133	56	60	17	3	-42.84	78.46	58.42	61.42

5 Conclusion

In this paper, we investigated the extended boomerang attacks. Our study showed that by considering the extended version of the original boomerang attack, the efficiency of distinguishers can be greatly improved. For situations where the full differential distribution is not available or computing resources are limited, our results have shown that the extended boomerang attack can lead to strong results in practical cryptanalysis. Furthermore, we observed that the extended boomerang framework is able to amplify the effect of the conditional difference technique due to the large number of differential paths involved in the computation. As a result, we are able to derive the best cryptanalysis results by far on KATAN48/64 in the related-key setting. For all the other versions of the family, the best differential attacks are derived.

Acknowledgment. This work has been partly supported by the research funds of CCNU from colleges' basic research and operation of MOE under grand No. CCNU16A05040, and Fundamental Research Grant Scheme (FRGS - 203/PKOMP/6711427) funded by the Ministry of Higher Education of Malaysia (MOHE). The authors would like to thank anonymous reviewers for their comments. A special mention is needed for Jiqiang Lu for all the help and suggestions to improve this paper.

Appendix - Distinguisher Results in the Single-key Setting

By applying the same searching methodology, we derive the distinguishers for KATAN32/48/64 in the single-key setting as follows (Table 7).

Table 7. Boomerang distinguisher for KATAN32/48/64 in the single-key setting

Ciphers	Total rounds	$\#E_0$	$\#E_1$	α	δ	Before ES	After ES	$\lambda_{E_0}/\lambda_{E_1}$
KATAN32	83	35	48	8010	801081	-38.58	-21.78	$-17/-24$
KATAN48	60	35	25	904000	402000000	-36.60	-23.36	$-22/-18$
KATAN64	56	30	26	4002001	20110080000000	-52.52	-44.26	$-22/-22$

References

1. Ahmadian, Z., Rasoolzadeh, S., Salmasizadeh, M., Aref, M.R.: Automated Dynamic Cube Attack on Block Ciphers: Cryptanalysis of SIMON and KATAN. IACR Cryptology ePrint Archive 2015 (2015)
2. Albrecht, M.R., Leander, G.: An all-in-one approach to differential cryptanalysis for small block ciphers. In: Knudsen, L.R., Wu, H. (eds.) SAC 2012. LNCS, vol. 7707, pp. 1–15. Springer, Heidelberg (2013)
3. Biham, E., Dunkelman, O., Keller, N.: New results on boomerang and rectangle attacks. In: Daemen, J., Rijmen, V. (eds.) FSE 2002. LNCS, vol. 2365, pp. 1–16. Springer, Heidelberg (2002)
4. Biham, E., Dunkelman, O., Keller, N.: The rectangle attack - rectangling the serpent. In: Pfitzmann, B. (ed.) EUROCRYPT 2001. LNCS, vol. 2045, pp. 340–357. Springer, Heidelberg (2001)
5. Biham, E., Shamir, A.: Differential cryptanalysis of DES-like cryptosystems. In: Menezes, A., Vanstone, S.A. (eds.) CRYPTO 1990. LNCS, vol. 537, pp. 2–21. Springer, Heidelberg (1991)
6. Biryukov, A., Khovratovich, D.: Related-key cryptanalysis of the full AES-192 and AES-256. In: Matsui, M. (ed.) ASIACRYPT 2009. LNCS, vol. 5912, pp. 1–18. Springer, Heidelberg (2009)
7. Blondeau, C., Gérard, B.: Multiple differential cryptanalysis: theory and practice. In: Joux, A. (ed.) FSE 2011. LNCS, vol. 6733, pp. 35–54. Springer, Heidelberg (2011)
8. Blondeau, C., Nyberg, K.: New links between differential and linear cryptanalysis. In: Johansson, T., Nguyen, P.Q. (eds.) EUROCRYPT 2013. LNCS, vol. 7881, pp. 388–404. Springer, Heidelberg (2013)
9. De Cannière, C., Dunkelman, O., Knežević, M.: KATAN and KTANTAN — a family of small and efficient hardware-oriented block ciphers. In: Clavier, C., Gaj, K. (eds.) CHES 2009. LNCS, vol. 5747, pp. 272–288. Springer, Heidelberg (2009)
10. Fuhr, T., Minaud, B.: Match box meet-in-the-middle attack against KATAN. In: Cid, C., Rechberger, C. (eds.) FSE 2014. LNCS, vol. 8540, pp. 61–81. Springer, Heidelberg (2015)
11. Isobe, T., Sasaki, Y., Chen, J.: Related-key boomerang attacks on KATAN32/48/64. In: Boyd, C., Simpson, L. (eds.) ACISP. LNCS, vol. 7959, pp. 268–285. Springer, Heidelberg (2013)

12. Isobe, T., Shibutani, K.: All subkeys recovery attack on block ciphers: extending meet-in-the-middle approach. In: Knudsen, L.R., Wu, H. (eds.) SAC 2012. LNCS, vol. 7707, pp. 202–221. Springer, Heidelberg (2013)
13. Isobe, T., Shibutani, K.: Improved all-subkeys recovery attacks on FOX, KATAN and SHACAL-2 block ciphers. In: Cid, C., Rechberger, C. (eds.) FSE 2014. LNCS, vol. 8540, pp. 104–126. Springer, Heidelberg (2015)
14. Kelsey, J., Kohno, T., Schneier, B.: Amplified boomerang attacks against reduced-round MARS and serpent. In: Schneier, B. (ed.) FSE 2000. LNCS, vol. 1978, pp. 75–93. Springer, Heidelberg (2001)
15. Knellwolf, S., Meier, W., Naya-Plasencia, M.: Conditional differential cryptanalysis of NLFSR-based cryptosystems. In: Abe, M. (ed.) ASIACRYPT 2010. LNCS, vol. 6477, pp. 130–145. Springer, Heidelberg (2010)
16. Knellwolf, S., Meier, W., Naya-Plasencia, M.: Conditional differential cryptanalysis of trivium and KATAN. In: Miri, A., Vaudenay, S. (eds.) SAC 2011. LNCS, vol. 7118, pp. 200–212. Springer, Heidelberg (2012)
17. Lu, J., Kim, J.-S., Keller, N., Dunkelman, O.: Differential and rectangle attacks on reduced-round SHACAL-1. In: Barua, R., Lange, T. (eds.) INDOCRYPT 2006. LNCS, vol. 4329, pp. 17–31. Springer, Heidelberg (2006)
18. Lu, J., Kim, J.-S., Keller, N., Dunkelman, O.: Related-key rectangle attack on 42-round SHACAL-2. In: Katsikas, S.K., López, J., Backes, M., Gritzalis, S., Preneel, B. (eds.) ISC 2006. LNCS, vol. 4176, pp. 85–100. Springer, Heidelberg (2006)
19. Matsui, M.: On correlation between the order of S-Boxes and the strength of DES. In: De Santis, A. (ed.) EUROCRYPT 1994. LNCS, vol. 950, pp. 366–375. Springer, Heidelberg (1995)
20. Rasoolzadeh, S., Raddum, H.: Improved Multi-Dimensional Meet-in-the-Middle Cryptanalysis of KATAN. IACR Cryptology ePrint Archive 2016 (2016)
21. Sun, B., Liu, Z., Rijmen, V., Li, R., Cheng, L., Wang, Q., Alkhzaimi, H., Li, C.: Links among impossible differential, integral and zero correlation linear cryptanalysis. In: Gennaro, R., Robshaw, M. (eds.) Advances in Cryptology-CRYPTO 2015. LNCS, vol. 9215, pp. 95–115. Springer, Heidelberg (2015)
22. Wagner, D.: The boomerang attack. In: Knudsen, L.R. (ed.) FSE 1999. LNCS, vol. 1636, pp. 156–170. Springer, Heidelberg (1999)
23. Wei, L., Rechberger, C., Guo, J., Wu, H., Wang, H., Ling, S.: Improved meet-in-the-middle cryptanalysis of KTANTAN (Poster). In: Parampalli, U., Hawkes, P. (eds.) ACISP 2011. LNCS, vol. 6812, pp. 433–438. Springer, Heidelberg (2011)
24. Zhu, B., Gong, G.: Multidimensional meet-in-the-middle attack and its applications to KATAN32/48/64. Crypt. Commun. **6**, 313–333 (2014)