Rectangle Attacks on 49-Round SHACAL-1*

Eli Biham,[1] Orr Dunkelman,[1] Nathan Keller[2]

[1] Computer Science Department, Technion.
Haifa 32000, Israel
{biham,orrd}@cs.technion.ac.il
[2] Mathematics Department, Technion.
Haifa 32000, Israel
nkeller@tx.technion.ac.il

Abstract. SHACAL-1 is a 160-bit block cipher with variable key length with up to 512-bit key based on the hash function SHA-1. It was submitted to the NESSIE project and was accepted as a finalist for the 2nd phase of the evaluation. In this paper we present rectangle attacks on 49 rounds out of the 80 rounds of SHACAL-1. The attacks require $2^{151.9}$ chosen plaintexts or ciphertexts and have time complexity of $2^{508.5}$ 49-round SHACAL-1 encryptions. These are the best known attacks against SHACAL-1. In this paper we also identify and fix some flaws in previous attacks on SHACAL-1.

1 Introduction

In 1993 NIST has issued a standard hash function called Secure Hash Algorithm [13]. Later this version was named SHA-0, as in 1995 NIST published a small tweak to this standard called SHA-1. Both SHA-0 and SHA-1 are based on padding the message and dividing it to blocks of 512 bits. Then iteratively compressing those blocks into a 160-bit digest. Recently, NIST has published (besides SHA-1) 3 more standard hash functions as part of FIPS-180: SHA-256, SHA-384 and SHA-512. Each of the new hash function has a digest size corresponding to its number, i.e., SHA-256 has 256-bit digest, etc.

Both SHA-0 and SHA-1 were subjected to a great deal of analysis. In [4] an attack producing pseudo-collisions in SHA-0 was suggested. A pseudo-collision is not a true collision, as it assumes that the attacker can control the input to the compression function, which is a constant value both in SHA-0 and in SHA-1. The attack requires 2^{61} computations of SHA-0 to produce a pseudo-collision. This result does not apply to SHA-1.

As hash functions can also be attacked using differential cryptanalysis [1], there is a continuous search for differentials in SHA-1. In [5] several of these differentials were presented.

Recently, it was shown how to generate slid pairs in SHA-1 with about 2^{32} computations of SHA-1, under some conditions [9].

* The work described in this paper has been supported by the European Commission through the IST Programme under Contract IST-1999-12324.

As SHA-1 was throughly examined and analyzed, it was suggested to use its compression function as a block cipher [5]. Later this suggestion was named SHACAL-1 [6]. It is a 160-bit block cipher with variable key length (0–512 bits) and 80 rounds based on the compression function of SHA-1. The cipher, which was submitted as a candidate to the NESSIE project [10], was selected as a NESSIE finalist, but was not selected for the NESSIE portfolio [12].

As mentioned before, it is possible to use the results of differential cryptanalysis obtained on SHA-1, and apply them to SHACAL-1. In [5, 6] a 5-round differential characteristic with probability 1 is presented. Few 10-round differentials are also presented in [5, 6]. Recently, differentials with higher probabilities than claimed in [5, 6] were presented in [14, 8]. In [8] these new differentials were combined to mount an amplified boomerang attack on 47-round SHACAL-1. The attack requires $2^{158.5}$ chosen plaintexts and has time complexity equivalent to $2^{508.4}$ 47-round SHACAL-1 encryptions.

In this paper we present some flaws in the analysis done in [8], and show how to correct them without affecting the data complexity nor the time complexity of the original attack.

We further improve the corrected results of [8] by applying the successor of the amplified boomerang attack – the rectangle attack to SHACAL-1. This allows us to reduce the data complexity to $2^{151.9}$ chosen plaintexts, and the time complexity of attacking 47 rounds (rounds 0–46) of SHACAL-1 from $2^{508.4}$ to $2^{482.6}$ encryptions.

By moving the rectangle attack to other rounds (rounds 22–70 or rounds 29-77) we also get an attack on 49-round SHACAL-1. The attack requires $2^{151.9}$ chosen plaintexts or chosen ciphertexts (depending on the rounds which we attack) and $2^{508.5}$ 49-round SHACAL-1 encryptions. This is the best known attack against SHACAL-1.

This paper is organized as follows: In Section 2 we describe the block cipher SHACAL-1. In Section 3 we present the previously best known results on SHACAL-1 (we also describe the flaws, and present a modification to the attack to correct those flaws). In Section 4 we give a short description of the amplified boomerang and the rectangle attacks, and we present a rectangle attack on 47-round SHACAL-1. In Section 5 we add two more rounds to the attack to present a rectangle attack on 49-round SHACAL-1. Finally, Section 6 summarizes the paper.

2 Description of SHACAL-1

SHACAL-1 [6] is a 160-bit block cipher supporting a variable key length (0–512 bits). It is based on the compression function of the hash function SHA-1. The cipher has 80 rounds (also referred as steps) grouped into 4 types of 20 rounds each[1].

[1] To avoid confusion, we adopt the standard and common notations for rounds. In [6] the notation step means round, where round is used for a group of 20 steps.

The 160-bit plaintext is divided into five 32-bit words – A, B, C, D and E. We denote by X_i the value of word X before the ith round, i.e., the plaintext P is divided into A_0, B_0, C_0, D_0 and E_0, and the ciphertext is composed of $A_{80}, B_{80}, C_{80}, D_{80}$ and E_{80}.

In each round the words are updated according to the following rule:

$$\begin{aligned} A_{i+1} &= W_i + ROTL_5(A_i) + f_i(B_i, C_i, D_i) + E_i + K_i \\ B_{i+1} &= A_i \\ C_{i+1} &= ROTL_{30}(B_i) \\ D_{i+1} &= C_i \\ E_{i+1} &= D_i \end{aligned}$$

where + denotes addition modulo 2^{32}, $ROTL_j(X)$ represents rotation to the left by j bits, W_i is the round subkey, and K_i is the round constant[2]. There are three different functions f_i, selected according to the round number:

$$\begin{array}{lll} f_i(X,Y,Z) = f_{if} = & (X\&Y)|(\neg X\&Z) & 0 \le i \le 19 \\ f_i(X,Y,Z) = f_{xor} = & (X \oplus Y \oplus Z) & 20 \le i \le 39, 60 \le i \le 79 \\ f_i(X,Y,Z) = f_{maj} = & ((X\&Y)|(X\&Z)|(Y\&Z)) & 40 \le i \le 59 \end{array}$$

In [6] it is strongly advised to use keys of at least 128 bits, even though shorter keys are supported. The first step in the key schedule algorithm is to pad the supplied key into a 512-bit key. Then, the 512-bit key is expanded into eighty 32-bit subkeys (or a total of 2560 bits of subkey material). The expansion is done in a linear manner using a a linear feedback shift register (over $GF(2^{32})$).

We omit the way the round subkeys are computed and the values of the round constants as these details do not affect our analysis. However, we do use the fact that the subkeys are linearly dependent in the key.

3 Previous Results

SHACAL-1 is based on SHA-1 and it is widely presumed that any attack on one of them would lead to an attack on the other (as demonstrated lately in [9]). Moreover, great deal of the analysis done to SHA-1 can be applied to SHACAL-1 as well.

In [5] the properties of the compression function of SHA-1 as a block cipher were studied. Differential and linear properties of SHACAL-1 were studied in [5, 6]: There is a 4-round linear approximation with bias 1/2 (the maximal bias), and for rounds with f_{xor} this approximation can be extended into a 7-round approximation with the same bias. There is also a 5-round differential with probability 1. These papers also contain results on 10-round linear approximations and differentials. We summarize these results in Table 1.

[2] This time we adopt the notations of [6], and alert the reader of the somewhat confusing notations.

Type	f Type in Use	Rounds	Probability/Bias
Linear [6]	any	4	1/2
Linear [6]	f_{xor}	7	1/2
Linear [6]	f_{xor}	10	2^{-6}
Linear [6]	f_{if}	10	$2^{-7.2}$
Linear [6]	f_{maj}	10	$2^{-6.4}$
Differential [6]	any	5	1
Differential [6]	f_{if}, f_{maj}	10	2^{-13}
Differential [6]	f_{xor}	10	2^{-26}
Differential [14]	f_{if}	10	2^{-12}
Differential [14]	f_{xor}	10	2^{-12}
Differential [14]	f_{if}, f_{maj}	20	2^{-41}
Differential [8]	20 f_{if} then f_{xor}	21	2^{-45}
Differential [8]	20 f_{if} then f_{xor}	28	2^{-107}
Differential [8]	20 f_{if} then f_{xor}	30	2^{-130}
Differential [8]	f_{xor}	15	2^{-31}

Table 1. Previously Known Differentials and Linear Approximations of SHACAL-1.

In [14] the differentials presented in [5,6] were improved, and 20-round differentials with probability 2^{-41} are presented. In [8] another set of differentials of SHACAL-1 is presented. We summarize these results in Table 1.

In [9] an algorithm for identifying whether two SHACAL-1 encryptions use related keys is presented. The attack is based on finding slid pairs, once a slid pair is encountered, the attacker can determine whether the two encryptions have related keys. The attack requires about 2^{96} encryptions under each of the two keys to find a slid pair.

In [8] the 21-round differential for rounds 0–20 and the 15-round differential for rounds 21–35 were combined to build an amplified boomerang [7] distinguisher for 36-round SHACAL-1. This distinguisher is used to attack a 39-round SHACAL-1 using $2^{158.5}$ chosen plaintexts and about $2^{250.8}$ 39-round SHACAL-1 encryptions. The attack is based on guessing (or trying) the subkeys of the 3 additional rounds, and then checking whether the distinguisher succeeds. This approach is further extended to attack 47-round SHACAL-1 before exhaustive key search becomes faster than this attack. Another attack presented in [8] is a differential attack on 41-round SHACAL-1. The attack uses the 28-round differential characteristics with probability 2^{-107} for 128-bit keys, and the 30-round differential characteristic with probability 2^{-138} for longer key lengths. We summarize the data and time complexities of the attacks presented in [8] in Table 2.

3.1 Problems in the 47-Round Amplified Boomerang Attack and How to Fix Them

As mentioned before, in [8] an amplified boomerang attack on 47-round SHACAL-1 is presented. The attack is based on a 36-round amplified boomerang distinguisher, and guessing the remaining 11 rounds subkeys. The basic idea, is to try

Key Size	Type of Attack	Number of	Complexity	
		Rounds	Data	Time
128-bit	Amplified Boomerang	28	$2^{127.5}$CP	$2^{127.2}$
	Differential	30	2^{110}CP	$2^{75.1}$
160-bit	Amplified Boomerang	37	$2^{158.5}$CP	$2^{87.8}$
	Differential	32	2^{141}CP	2^{105}
256-bit	Amplified Boomerang	39	$2^{158.5}$CP	$2^{250.8}$
	Differential	34	2^{141}CP	2^{234}
512-bit	Amplified Boomerang	47	$2^{158.5}$CP	$2^{508.4}$
	Differential	41	2^{141}CP	2^{491}

CP - Chosen Plaintexts

Table 2. Complexities of Previous Attacks on SHACAL-1 ([8]).

each and every subkey for those 11 rounds, partially decrypt all the ciphertexts, and run the distinguisher. If the distinguisher succeeds (i.e., succeeds to distinguish the remaining 36 rounds from a random permutation), then the subkey for those 11 rounds is considered to be the right subkey.

We shall concentrate on the 39-round attack, which is based on guessing the subkeys of the last 3 rounds. The 39-round attack is actually considered as a procedure in the 47-round attack, hence, if this attack fails, so does the 47-round attack presented in [8].

The 36-round distinguisher needs $2^{158.5}$ chosen plaintexts that compose $2^{157.5}$ pairs. According to the analysis in [8] it is expected that for the right subkey the number of right amplified boomerang quartets[3] is 8. However, the number of right quartets has a Poisson distribution. Hence, if the expected number of quartets is 8, and denoting by a random variable X the number of quartets, we get that $X \sim Poi(8)$. This means that $\Pr[X \geq 8] = 0.505$ which is much lower than expected by the authors of [8].

Usually this confusion between the expected value and the true distribution of the value has no implication on the correctness of the attack. However, the authors of [8] claim that

> "But, for a wrong subkey, the expected value of counter is equal to 0 or 1, since the expected number of quartets passed through Step 4 is $2^{-5}(= 2^{187} \cdot (2^{-96})^2)$."

We agree that for most subkeys, the value of the counter (how many quartets suggest this subkey) is 0 or 1. The values of these counters also behave like Poisson random variables. Let us examine a Poisson random variable $Y \sim Poi(1/32)$, the probability $\Pr[Y \geq 8] \approx e^{-1/32} \cdot (1/32)^8/8! = 2^{-55.3}$ is truly very small. If we take into consideration the fact that we have 2^{96} such variables (each corresponds to a wrong subkey guess), each with probability of $2^{-55.3}$ to pass the

[3] We present in Section 4 a detailed description of the amplified boomerang attack. Meanwhile, we note that for a random permutation with 160-bit block, the probability that $2^{157.5}$ pairs create an amplified boomerang quartet is 2^{-5}.

filtering, then about $2^{40.7}$ (out of the possible 2^{96} subkeys) also have at least 8 "right" quartets. Therefore the 39-round attack fails, as there are $2^{40.7}$ subkeys suggested by the attack, and only in half of the cases the right subkey is among them!

At first it appears that these observations indicate that the attack is incorrect. We correct the attack by exploiting the fact that all subkeys and the user key are all linearly dependent. Given a subkey with 8 (or more) candidate quartets, we can check all keys which generate this subkey. Enlisting those keys is easy and can be done efficiently, as the subkeys are linearly dependent in the key. Thus, we can reduce the time complexity of exhaustive key search by a factor of $2^{55.3}$ and find the right key with probability 50.5%. By reducing the threshold (examining subkeys with 7 or more quartets) we increase the success rate of the attack to 69%. In exchange, we examine 1 out of every $2^{47.3}$ keys, i.e., 255 times more keys than when 8 quartets are the requirement.

We can further correct and improve the above results. Right quartets are composed of two pairs of right pairs with respect to the first differential used in the amplified boomerang distinguisher[4], thus, if a subkey gets 7 quartets or more, we have at least 14 pairs with respect to the first differential. We use those pairs to mount a regular differential attack. Our analysis shows that 14 pairs are sufficient to determine the first round subkey uniquely. Moreover, for a wrong subkey guess, the probability that these 14 pairs suggest a subkey for the first round is about 2^{-24}, hence, can be used for eliminating a wrong subkey guess.

This variant of the attack retrieves 128-bit subkey material, 32 bits in the first round (these bits are key material), and the remaining 96 bits that are linearly derived from the key. We can further retrieve subkey material, continuing the differential attack and using auxiliary techniques.

Combining these corrections and improvements we get a valid attack with the same time and data complexity as the one mentioned in [8]. The most time consuming part of the attack is still the basic 39-round attack, while the remaining steps has a negligible time complexity.

4 Rectangling the Attack – Attacking 47-Round SHACAL-1

In this section, we improve the corrected results of [8] to attack 47-round SHACAL-1 more efficiently. Our improvements are based on transforming the amplified boomerang attack into a rectangle attack. This allows us to reduce the data complexity of the attack to $2^{151.9}$ chosen plaintexts, and the time complexity to $2^{482.6}$ 47-round SHACAL-1 encryptions.

We first upgrade the distinguisher from an amplified boomerang distinguisher [7] into a rectangle distinguisher [2]. These attacks are closely related to the boomerang attack [15]. Both the amplified boomerang and the rectangle attacks

[4] Again, we present a detailed description of the amplified boomerang attack in the next section.

Probability (p)	2^{-45}	2^{-46}	2^{-47}	2^{-48}	2^{-49}	2^{-50}	2^{-51}
Number of Differentials (l)	1	7	24	73	182	351	677
Contribution to $\hat{p}^2$ ($= lp^2$)	2^{-90}	$2^{-89.2}$	$2^{-89.4}$	$2^{-89.8}$	$2^{-90.5}$	$2^{-91.5}$	$2^{-92.6}$

Table 3. Number of Differentials for Rounds 0–20 of SHACAL-1.

are based on treating the distinguished part of the cipher as composed of two sub-ciphers. Formally, we treat the block cipher E as a cascade of 4 sub-ciphers: $E = E_f \circ E_1 \circ E_0 \circ E_b$, where E_b are the attacked rounds before the distinguisher, E_f are the attacked rounds after the distinguisher, and $E' = E_1 \circ E_0$ is the distinguished part.

In the amplified boomerang attack, we take a differential $\alpha \rightarrow \beta$ through E_0 with probability p and a differential $\gamma \rightarrow \delta$ through E_1 with probability q. If we take N pairs of plaintexts with input difference α, about Np of them has a difference β after E_0. Those Np pairs can be combined in $(Np)^2/2$ quartets. Denoting such a quartet by $((P_1, P_2), (P_3, P_4))$ we know that $P_1 \oplus P_2 = P_3 \oplus P_4 = \alpha$ and that $E_0(P_1) \oplus E_0(P_2) = E_0(P_3) \oplus E_0(P_4) = \beta$. Assuming that $\{E_0(P_i)\}$ are distributed uniformally, then with probability 2^{-128} it is true that $E_0(P_1) \oplus E_0(P_3) = E_0(P_2) \oplus E_0(P_4) = \gamma$. When this happens, with probability q^2 we get that $E_1(E_0(P_1)) \oplus E_1(E_0(P_3)) = E_1(E_0(P_2)) \oplus E_1(E_0(P_4)) = \delta$. Hence, starting with N plaintext pairs with input difference α we expect about $N^2 \cdot (pq)^2 \cdot 2^{-129}$ quartets which satisfy the condition $C_1 \oplus C_3 = C_2 \oplus C_4 = \delta$, where C_i is the corresponding ciphertext to the plaintext P_i.

The rectangle attack is based on the same basic idea. However, the attack allows using any differential $\alpha \rightarrow \beta'$ in E_0 and any differential $\gamma' \rightarrow \delta$ in E_1 (as long as $\beta' \neq \gamma'$). Besides these improvements, the attack gains a factor of 2 in the number of expected quartets by checking the two different quartets $((P_1, P_2), (P_3, P_4))$ and $((P_1, P_2), (P_4, P_3))$. Hence, starting with N plaintext pairs with input difference α, we expect to get $N^2 \cdot (\hat{p}\hat{q})^2 \cdot 2^{-128}$ right quartets, where:

$$\hat{p} = \sqrt{\sum_{\beta} \Pr{}^2[\alpha \rightarrow \beta]}, \qquad \hat{q} = \sqrt{\sum_{\gamma} \Pr{}^2[\gamma \rightarrow \delta]}.$$

Moreover, for a rectangle distinguisher, there is a better key recovery algorithm presented in [3]. Therefore, using the rectangle attack instead of the amplified boomerang attack is much more efficient and requires less data.

In order to compute $\hat{p}$ we need to summarize the square of the probability of all the differentials with input difference α through E_0. This task is computationally infeasible, and thus, we try to count over as many differentials as we can. We settle for counting over all differentials which have the same first 20 or 19 rounds as the 21-round differential used in the original amplified boomerang attack. In Table 3 we gathered the number of counted differentials according to their probabilities, and in Appendix A we present some of these differentials. Given these results we are able to compute a lower bound for $\hat{p} = 2^{-43.64}$.

Probability (p)	2^{-31}	2^{-32}	2^{-33}	2^{-34}	2^{-35}	2^{-36}	2^{-37}
Number of Differentials (l)	1	3	8	18	32	48	56
Contribution to $\hat{p}^2$ ($= lp^2$)	2^{-62}	$2^{-62.4}$	2^{-63}	$2^{-63.8}$	2^{-65}	$2^{-66.4}$	$2^{-68.2}$

Table 4. Number of Differentials for Rounds 21–35 of SHACAL-1.

Due to the same reasons, computing the exact value of $\hat{q}$ is computationally infeasible. Hence, we count only differentials which have the same 13 or 14 last rounds as the 15-round differential used in the original amplified boomerang attack. In Table 4 we list the number of differentials with their respective probability, and in Appendix A we present some of these differentials. We take all these differentials into account and get that $\hat{q} = 2^{-30.28}$.

These improvements reduces the data complexity of the attack from $2^{158.5}$ chosen plaintexts to $2^{155.9}$ chosen plaintexts. Due to the nature of the amplified boomerang and the rectangle attacks, this reduces the time complexity by a factor of $2^{5.2} \approx 35.5$.

Our third improvement to the data and time complexity of the attack is based on reducing the number of rounds in the distinguisher itself.

In the attack presented in [8], there are 0 rounds in E_b, 21 rounds in E_0, 15 rounds in E_1, and between 3 to 11 rounds in E_f. We can move one round from the E_0 to the rounds before the distinguisher E_b. This changes the division to sub-ciphers a little bit into 1 round in E_b and 20 in E_0. This increases the probability of the differentials of E_0 by 2^{-4} in exchange for a more complex attack algorithm.

In these new settings we use the results of [3], where a generic key recovery algorithm based on the rectangle distinguisher is presented. We shortly describe the notations used in that paper. Let $E = E_f \circ E_1 \circ E_0 \circ E_b$ be an n-bit block cipher. Assume that for E_0 we have an input difference α and a related $\hat{p}$, and an output difference δ and $\hat{q}$ for E_1. We denote by r_b the number of bits which are active or can be active in the plaintext given that there is an α difference after E_b (and before E_0). We denote by 2^{t_b} the number of of possibilities for these bits. For example, if an α difference after E_b requires that there is some plaintext bit that always differs in the pair, then this bit is counted by r_b but has no affect on t_b. We denote by m_b the number of subkey bits in E_b that we attack (i.e., the number of subkey bits in E_b that affect the α difference after E_b). For E_f, we denote by r_f the number of ciphertext bits whose difference is unknown after E_f if the input difference of E_f was δ. Similarly, 2^{t_f} is the number of possible differences in these r_f bits. We denote by m_f the number of subkey bits in E_f that affect the δ difference. See [3] for more details of these notations and the detailed attack.

The figures for this decomposition of 39-round SHACAL-1 are as follows: $n = 160$ (as SHACAL-1 is a 160-bit block cipher). If we truncate the 21-round characteristics by the first round (which becomes E_b), there are are 25 bits that have or might have difference in the plaintext: The 22 most significant bits of register E, bit 5 of register A, bit 20 of register C and bit 15 of register D, thus

$r_b = 25$. If we look at these 25 bits we observe that not all the 2^{25} differences are possible: 3 out of these 25 bits always differ, i.e., for a pair with difference α after E_b these bits are always active, and for some other bits not all the differences are possible. For example, bits 10–14 of register E can have only one of the following five patterns – $\{01_x, 03_x, 07_x, 0F_x, 1F_x\}$. Similarly, our analysis reveals that the 22 bits of register E can have at most $5 \cdot 2 \cdot 4 \cdot 2 \cdot 2 \cdot 12 = 960 = 2^{9.9}$ differences before round 0, and therefore $t_b = 9.9$.

If the output difference of a pair after E_1 is δ, after E_f we get that out of the 160 ciphertext bits, 68 bits have no difference or always have a difference (bits 0,1 of register B, bits 30,31 of register C and the entire D and E registers). Therefore, the number of bits which may differ in a right pair is $r_f = 160 - 68 = 92$. Still, for a right pair, not all the 2^{92} possible differences in these 92 bits can be achieved. Our analysis reveals that at most $7 \cdot 2^{22} \cdot 7 \cdot 2^{22} \cdot 2^{32} = 2^{81.6}$ of these differences can be achieved if the input difference to E_f is δ, thus, $t_b = 81.6$.

Finally, for this attack $m_b = 32$ (as we attack one round in E_b) and $m_f = 96$ (as we attack 3 rounds in E_f).

Assigning these figures to the complexity analysis from [3] we obtain that the data complexity of the attack is $N = 2^{n/2+2}/\hat{p}\hat{q} = 2^{151.9}$ chosen plaintexts, and the time complexity of the attack is $N^2(2^{r_f-n-1} + 2^{t_f-n} + 2^{2t_f+2r_b-2n-2} + 2^{m_b+2t_f+t_b-2n-1} + 2^{m_f+2t_b+t_f-2n-1}) + N$ memory accesses, which are $2^{235.4}$ memory accesses. These $2^{235.4}$ memory accesses are equivalent to $2^{227.5}$ 39-round SHACAL-1 encryptions[5].

We can extend this attack to a 47-round SHACAL-1 using the following algorithm:

1. Try all 2^{256} possible values for the eight 32-bit subkeys value (of rounds 39–46). For each guess partially decrypt all ciphertexts these 8 rounds, and:
   (a) Apply the 39-round rectangle attack.
   (b) In case the 39-round attack suggests a subkey with 3 or more quartets, check all the 2^{128} keys that generate this 128-bit subkey value (the one suggested by the 39-round attack) and the 256-bit subkey values for rounds 39–46.

The loop is repeated 2^{256} times, and it is expected that no more than 2^{66} subkeys have 3 or more quartets for each 256-bit subkey guess. Thus, the time complexity of Step (b) is $2^{256} \cdot 2^{66} \cdot 2^{128} = 2^{450}$ 47-round SHACAL-1 encryptions. The time complexity of Step 1 is $2^{256} \cdot 2^{151.9} \cdot \frac{8}{47} = 2^{405.3}$ 47-round SHACAL-1 encryptions, and the time complexity of Step (a) is $2^{490.8}$ memory accesses (which are equivalent to $2^{482.6}$ 47-round SHACAL-1 encryptions). Thus, the total time complexity of the attack is $2^{482.6}$ 47-round SHACAL-1 encryptions.

5 Attacking 49-Round SHACAL-1

In this section we present a rectangle attack on 49-round SHACAL-1. The attack is quite similar to the one presented in the previous section. In order to improve

[5] The conversion was done according to the best performance figures presented in [11].

the attack, we remove one round of the basic 39-round attack (obtaining a 38-round rectangle attack), and perform a chosen ciphertext attack on rounds 29–78 (or on rounds 22–70 using a chosen plaintext attack) of SHACAL-1. This results in an attack on these 49 rounds that requires $2^{508.5}$ encryptions, or an attack on 47 rounds that requires $2^{444.5}$ encryptions.

The first observation is that the inclusion of third round of E_f (the rounds after the distinguisher) in the 39-round attack does not contribute to the attack. After adding this third round to E_f, the number of subkey bits checked (m_f) is increased by 32. The number of possible differences that a δ difference causes after E_f is increased by 2^{32}, hence, t_f and r_f are increased by 32 each. Therefore, adding this third round is equivalent for guessing this last round subkey, and decrypting all ciphertexts another round.

This observation points out that we can use in the 47-round attack, a 38-round attack which contains one round before the rectangle (E_b), 20 rounds in E_0, 15 in E_1 and 2 rounds after the rectangle (E_f), with the following parameters: $r_b = 25$, $t_b = 9.9$, $m_b = 32$, $r_f = 60$, $t_f = 49.6$, $m_f = 64$ and $\hat{p} = 2^{-39.87}, \hat{q} = 2^{-30.32}$. Using these figures in the rectangle attack presented in [3] yields an attack which requires $2^{151.9}$ chosen plaintexts and $2^{203.4}$ memory accesses.

Using this observation about the distinguisher, we change the attack algorithm from the previous section accordingly to:

1. Try all 2^{288} possible values for the nine 32-bit subkeys value (of rounds 38–46). For each guess partially decrypt all ciphertexts these 9 rounds and:
   (a) Apply the 38-round rectangle attack.
   (b) In case the 38-round attack suggests a subkey with 3 or more quartets, check all the 2^{128} keys that generate this 96-bit subkey value (the one suggested by the 38-round attack) and the 288-bit subkey values for rounds 38–46.

This new attack algorithm has the same data complexity as before ($2^{151.9}$ chosen plaintexts) with time complexity of $2^{151.9} \cdot 2^{288} \cdot 9/47 = 2^{437.5}$ encryptions for step 1, $2^{490.8}$ memory accesses for step (a), and 2^{450} encryptions for Step (b).

As the reader surely observed, this change does not improve the attack. First, we transform the 38-round attack into a chosen ciphertext attack. This might pose a problem, as the distinguisher starts with the first round, and we cannot decrypt the round before it. To solve this problem, we use our second observation that f_{if} and f_{maj} behave almost in the same manner with respect to differential cryptanalysis. Especially, the differentials that we use can be transferred to rounds 41–60 without affecting their probability. Therefore, we can move the 38-round attack to rounds 40–77. The attack is a chosen ciphertext attack, thus we treat the cipher in a reversed order. The cipher which we attack starts just after round 77 and ends just before round 40: it has two rounds in the new E_b (rounds 76–77), 35 rounds in the distinguisher itself (rounds 41–75), and one round afterwards in the new E_f (round 40). For these settings the following values for the rectangle attack are: $m_b = 64$, $m_f = 32$, $t_b = 49.6$, $t_f = 9.9$, $r_b = 68$ and $r_f = 22$. As the same differentials are used, The values of $\hat{p}$ and $\hat{q}$

remain the same. Thus, our attack requires $2^{151.9}$ chosen ciphertexts but only $2^{165.4}$ memory accesses.

We can now attack 11 more rounds after the distinguisher (actually before the distinguisher), thus attacking rounds 29–77 using the following algorithm:

1. Try all 2^{352} possible values for the eleven 32-bit subkeys value (of rounds 29–39). For each guess partially encrypt all plaintexts these 11 rounds and:
   (a) Apply the 38-round rectangle attack on rounds 40–77.
   (b) In case the 38-round attack suggests a subkey with 3 or more quartets, check all the 2^{64} keys which generate this 96-bit subkey value (the one suggested by the 38-round attack) and the 352-bit subkey values for rounds 29–39.

The time complexity of this 49-round attack is: $2^{151.9} \cdot 2^{352} \cdot 11/47 = 2^{501.8}$ 49-round SHACAL-1 encryptions for step 1, $2^{516.8}$ memory accesses for step (a), and 2^{478} encryptions for Step (b). Translating the time complexity to units of 49-round SHACAL-1 encryptions we get that the total time complexity of the attack is $2^{508.5}$ 49-round SHACAL-1 encryptions.

We can use a chosen plaintext attack on rounds 22–70 using the same attack algorithm with chosen plaintexts and guessing the subkeys of the 11 rounds after the end (i.e., rounds 60–70), with the same data and time complexity.

We can also apply this attack to 47-round SHACAL-1, with the same data complexity, but with time complexity of only $2^{444.5}$ 47-round SHACAL-1 encryptions.

6 Summary and Conclusions

In this paper we improved the cryptanalysis results on SHACAL-1. The improvements allow us to attack 47-round SHACAL-1 (rounds 0–46) using the rectangle attack with data complexity of $2^{151.9}$ chosen plaintexts and time complexity of $2^{482.6}$ encryptions. We can also attack rounds 22-68 or rounds 31–77 with time complexity of $2^{444.5}$ encryptions. Another attack presented in this paper is on 49 rounds that requires the same amount of data, and $2^{508.5}$ encryptions. This is the best currently known results on SHACAL-1.

We summarize our results and compare them with the previously known results in Table 5.

References

1. Eli Biham, Adi Shamir, *Differential Cryptanalysis of the Data Encryption Standard*, Springer-Verlag, 1993.
2. Eli Biham, Orr Dunkelman, Nathan Keller, *The Rectangle Attack – Rectangling the Serpent*, Advances in Cryptology, proceeding of EUROCRYPT 2001, Lecture Notes in Computer Science 2045, pp. 340–357, Springer-Verlag, 2001.
3. Eli Biham, Orr Dunkelman, Nathan Keller, *New Results on Boomerang and Rectangle Attacks*, proceeding of Fast Software Encryption 9, Lecture Notes in Computer Science 2365, pp. 1–16, Springer-Verlag, 2002.

Attack	Number of	Rounds	Complexity	
	Rounds		Data	Time
Differential [8]	41	0–40	2^{141}CP	2^{491}
Amplified Boomerang [8]	47	0–46	$2^{158.5}$CP	$2^{508.4}$
Rectangle – this paper	47	0–46	$2^{151.9}$CP	$2^{482.6}$
Rectangle – this paper	47	22–68	$2^{151.9}$CP	$2^{444.5}$
Rectangle – this paper	47	31–77	$2^{151.9}$CC	$2^{444.5}$
Rectangle – this paper	49	22–70	$2^{151.9}$CP	$2^{508.5}$
Rectangle – this paper	49	29–77	$2^{151.9}$CC	$2^{508.5}$

Complexity is measured in encryption units.
CP - Chosen Plaintexts, CC - Chosen Ciphertexts

Table 5. Summary of Our Results and Previously Known Results.

4. Florent Chabaud, Antoine Joux, *Differential Collisions in SHA-0*, Advances in Cryptology, proceeding of CRYPTO 1998, Lecture Notes in Computer Science 1462, pp. 56–71, Springer-Verlag, 1998.
5. Helena Handschuh, Lars R. Knudsen, Matthew J. Robshaw *Analysis of SHA-1 in Encryption Mode*, proceedings of CT-RSA 2001, Springer-Verlag Lecture Notes in Computer Science, vol. 2020, pp. 70–83, 2001.
6. Helena Handschuh, David Naccache, *SHACAL*, preproceedings of NESSIE first workshop, Leuven, 2000.
7. John Kelsey, Tadayoshi Kohno, Bruce Schneier, *Amplified Boomerang Attacks Against Reduced-Round MARS and Serpent*, proceedings of Fast Software Encryption 7, Lecture Notes in Computer Science 1978, pp. 75–93, Springer-Verlag, 2000.
8. Jonsung Kim, Dukjae Moon, Wonil Lee, Seokhie Hong, Sangjin Lee, Seokwon Jung, *Amplified Boomerang Attack against Reduced-Round SHACAL*, Advances in Cryptology, proceedings of ASIACRYPT 2002, To appear.
9. Markku-Juhani O. Saarinen, *Cryptanalysis of Block Ciphers Based on SHA-1 and MD5*, these proceedings.
10. NESSIE – New European Schemes for Signatures, Integrity and Encryption. *http://www.nessie.eu.org/nessie*
11. NESSIE, *Performance of Optimized Implementations of the NESSIE Primitives*, NES/DOC/TEC/WP6/D21/2.
12. NESSIE, *Portfolio of recommended cryptographic primitives.*
13. US National Bureau of Standards, *Secure Hash Standard*, Federal Information Processing Standards Publications No. 180-2, 2002.
14. Eteinee Van Den Bogeart, Vincent Rijmen, *Differential Analysis of SHACAL*, NESSIE internal report NES/DOC/KUL/WP3/009/a, 2001.
15. David Wagner, *The Boomerang Attack*, proceedings of Fast Software Encryption 6, Lecture Notes in Computer Science 1636, pp. 156–170, 1999.

A Differentials of SHACAL-1

In this appendix we describe the differentials used for the rectangle attacks on SHACAL-1. The basic differentials were previously presented in [8].

The first differential (before truncating the first round) is for rounds 0–20 or for rounds 40–60, and is presented in Table 6. We use the notation e_i to present

Round (i)	ΔA_i	ΔB_i	ΔC_i	ΔD_i	ΔE_i	Probability
Input ($i = 0$)	0	e_{22}	e_{15}	e_{10}	e_5	2^{-4}
1	e_5	0	e_{20}	e_{15}	e_{10}	2^{-3}
2	0	e_5	0	e_{20}	e_{15}	2^{-3}
3	e_{15}	0	e_3	0	e_{20}	2^{-2}
4	0	e_{15}	0	e_3	0	2^{-2}
5	0	0	e_{13}	0	e_3	2^{-2}
6	e_3	0	0	e_{13}	0	2^{-2}
7	e_8	e_3	0	0	e_{13}	2^{-2}
8	0	e_8	e_1	0	0	2^{-2}
9	0	0	e_6	e_1	0	2^{-2}
10	0	0	0	e_6	e_1	2^{-2}
11	e_1	0	0	0	e_6	2^{-1}
12	0	e_1	0	0	0	2^{-1}
13	0	0	e_{31}	0	0	2^{-1}
14	0	0	0	e_{31}	0	2^{-1}
15	0	0	0	0	e_{31}	1
16	e_{31}	0	0	0	0	2^{-1}
17	e_4	e_{31}	0	0	0	2^{-2}
18	e_9	e_4	e_{29}	0	0	2^{-3}
19	e_{14}	e_9	e_2	e_{29}	0	2^{-4}
20	e_{19}	e_{14}	e_7	e_2	e_{29}	2^{-5}
Output ($i = 21$)	$e_{2,7,14,24,29}$	e_{19}	e_{12}	e_7	e_2	

Differences are presented before the round, i.e., ΔA_0 is the input difference.

Table 6. Differential Characteristic for Rounds 0–20 (or 40–60) of SHACAL-1.

the 32-bit word composed of 31 0's and 1 in the ith place. We use $e_{i,j}$ to denote $e_i \oplus e_j$ and $e_{i,j,k} = e_{i,j} \oplus e_k$, etc. Recall that for the rectangle attacks presented in this paper, we changed this differential by removing its first round.

Due to the nature of the rectangle attack, we count over several differentials. We have counted over differentials which have the same 19 rounds as the one presented in [8] after it was truncated to a 20-round differential. In Table 7 we list the differentials with the highest probabilities which has the same 19 rounds as the original one. As in the last round the only affected register is A, the list contains only the various differences in register A (the remaining registers have the same differences as in the original one of Table 6).

We can also alter the round before last of this differential, gaining more characteristics with probabilities 2^{-46} at most. These changes are based on activating one bit in the output of the non-linear function f_{if} (or f_{maj}). We added their numbers to the table presented in Section 4, as they affect the probability of the distinguisher, but we omit their description here.

The second differential for rounds 21–35 (or rounds 61–75) is presented in Table 8. It was also previously appeared in [8].

Again, due to the nature of the rectangle attack, we counted probabilities of several differentials. We counted over various similar characteristics, by changing

ΔA_{21}	Prob.	ΔA_{21}	Prob.	ΔA_{21}	Prob.	ΔA_{21}	Prob.
$e_{2,7,14,24,29}$	2^{-45}	$e_{2,3,7,14,24,29}$	2^{-46}	$e_{2,7,8,14,24,29}$	2^{-46}	$e_{2,7,14,15,24,29}$	2^{-46}
$e_{2,7,14,24,25,29}$	2^{-46}	$e_{2,3,7,14,24,29,30}$	2^{-46}	$e_{2,7,8,14,24,29,30,31}$	2^{-46}	$e_{2,3,4,7,14,15,24,29}$	2^{-47}
$e_{2,3,7,8,14,24,29}$	2^{-47}	$e_{2,3,7,14,15,24,29}$	2^{-47}	$e_{2,3,7,14,24,25,29}$	2^{-47}	$e_{2,3,7,14,15,24,29,30}$	2^{-47}
$e_{2,3,7,8,14,24,29,30,31}$	2^{-47}	$e_{2,7,8,9,14,24,29}$	2^{-47}	$e_{2,7,8,14,15,24,29}$	2^{-47}	$e_{2,7,8,14,24,25,29}$	2^{-47}
$e_{2,7,8,14,24,29,30}$	2^{-47}	$e_{2,7,8,14,24,29,30,31}$	2^{-47}	$e_{2,7,14,15,16,24,29}$	2^{-47}	$e_{2,7,14,15,24,25,29}$	2^{-47}
$e_{2,7,14,15,24,29,30}$	2^{-47}	$e_{2,7,14,15,24,29,30,31}$	2^{-47}	$e_{2,7,14,24,25,26,29}$	2^{-47}	$e_{2,7,14,24,25,29,30}$	2^{-47}
$e_{2,7,14,24,25,29,30,31}$	2^{-47}						

Table 7. Possible ΔA_{21} Values for the First Characteristic with the Respective Probabilities.

Round (i)	ΔA_i	ΔB_i	ΔC_i	ΔD_i	ΔE_i	Probability
Input ($i = 21$)	$e_{1,5,8}$	$e_{1,3,5}$	$e_{3,13}$	$e_{1,5,13,31}$	$e_{6,10,13,31}$	2^{-3}
22	0	$e_{1,5,8}$	$e_{1,3,31}$	$e_{3,13}$	$e_{1,5,13,31}$	2^{-4}
23	$e_{1,8}$	0	$e_{3,6,31}$	$e_{1,3,31}$	$e_{3,13}$	2^{-4}
24	$e_{1,3}$	$e_{1,8}$	0	$e_{3,6,31}$	$e_{1,3,31}$	2^{-4}
25	0	$e_{1,3}$	$e_{6,31}$	0	$e_{3,6,31}$	2^{-3}
26	e_1	0	$e_{1,31}$	$e_{6,31}$	0	2^{-2}
27	e_1	e_1	0	$e_{1,31}$	$e_{6,31}$	2^{-1}
28	0	e_1	e_{31}	0	$e_{1,31}$	2^{-1}
29	0	0	e_{31}	e_{31}	0	1
30	0	0	0	e_{31}	e_{31}	1
31	0	0	0	0	e_{31}	1
32	e_{31}	0	0	0	0	2^{-1}
33	e_4	e_{31}	0	0	0	2^{-1}
34	$e_{9,31}$	e_4	e_{29}	0	0	2^{-3}
35	$e_{14,29}$	$e_{9,31}$	e_2	e_{29}	0	2^{-4}
Output	$e_{9,19,29,31}$	$e_{14,29}$	$e_{7,29}$	e_2	e_{29}	

Differences are presented before the round, i.e., ΔA_{21} is the input difference.

Table 8. Differential Characteristic for Rounds 21–35 (or 61–75) of SHACAL-1.

the first one or two rounds of this differential. In Table 9 we present some of the differentials which agree with the original characteristic but the first round (round 21). As in this round the only change is in the difference of register E, hence, only the various differences in that register are presented.

ΔE_{21}	Prob.	ΔE_{21}	Prob.	ΔE_{21}	Prob.	ΔE_{21}	Prob.
$e_{6,10,13,31}$	2^{-31}	$e_{6,7,10,13,31}$	2^{-32}	$e_{6,10,11,13,31}$	2^{-32}	$e_{6,10,13,14,31}$	2^{-32}
$e_{6,7,8,10,13,31}$	2^{-33}	$e_{6,7,10,11,13,31}$	2^{-33}	$e_{6,7,10,13,14,31}$	2^{-33}	$e_{6,10,11,12,31}$	2^{-33}
$e_{6,10,11,12,13,31}$	2^{-33}	$e_{6,10,11,13,14,31}$	2^{-33}	$e_{6,10,13,14,15,31}$	2^{-33}	$e_{6,7,8,9,13,31}$	2^{-34}
$e_{6,7,10,11,12,31}$	2^{-34}	$e_{6,7,8,9,10,13,31}$	2^{-34}	$e_{6,7,8,10,11,13,14,31}$	2^{-34}	$e_{6,7,8,10,11,13,31}$	2^{-34}
$e_{6,7,10,11,12,13,31}$	2^{-34}	$e_{6,7,10,11,13,14,31}$	2^{-34}	$e_{6,7,10,13,14,15,31}$	2^{-34}	$e_{6,10,11,12,13,14,31}$	2^{-34}
$e_{6,10,11,12,13,14,15,31}$	2^{-34}	$e_{6,10,11,13,14,15,31}$	2^{-34}	$e_{6,10,13,14,15,16,31}$	2^{-34}		

Table 9. Possible ΔE_{21} Values for the Second Characteristic with the Respective Probabilities.