Related-Key Differential Cryptanalysis of 192-bit Key AES Variants*

Goce Jakimoski and Yvo Desmedt

Computer Science Department, Florida State University
Tallahassee, Florida FL 32306-4530, USA
{jakimosk,desmedt}@cs.fsu.edu

Abstract. A related-key differential cryptanalysis is applied to the 192-bit key variant of AES. Although any 4-round differential trail has at least 25 active bytes, one can construct 5-round related-key differential trail that has only 15 active bytes and break six rounds with 2^{106} plaintext/ciphertext pairs and complexity 2^{112}. The attack can be improved using truncated differentials. In this case, the number of required plaintext/ciphertext pairs is 2^{81} and the complexity is about 2^{86}. Using impossible related-key differentials we can break seven rounds with 2^{111} plaintext/ciphertext pairs and computational complexity 2^{116}. The attack on eight rounds requires 2^{88} plaintext/ciphertext pairs and its complexity is about 2^{183} encryptions. In the case of differential cryptanalysis, if the iterated cipher is Markov cipher and the round keys are independent, then the sequence of differences at each round output forms a Markov chain and the cipher becomes resistant to differential cryptanalysis after sufficiently many rounds, but this is not true in the case of related-key differentials. It can be shown that if in addition the Markov cipher has $\mathcal{K} - f$ round function and the hypothesis of stochastic equivalence for related keys holds, then the iterated cipher is resistant to related-key differential attacks after sufficiently many rounds.

Keywords: Differential cryptanalysis, related keys, Markov ciphers, Advanced Encryption Standard

1 Introduction

On October 2, 2000, after a long and complex evaluation process, NIST announced that it has selected Rijndael [9] to propose for the Advanced Encryption Standard. A draft standard was published for public review and comment, and in 2001, FIPS-197 was approved as a Federal Information Processing Standard for the AES [1]. The AES algorithm is a symmetric block cipher that can process data blocks of 128 bits. It may be used with three different key lengths (128, 192, and 256 bits), and these different "flavors" are referred to as "AES-128", "AES-192", and "AES-256". It is expected that the algorithm "will be used by the U.S.

* A part of this research was funded by NSF CCR-0109425.

M. Matsui and R. Zuccherato (Eds.): SAC 2003, LNCS 3006, pp. 208–221, 2004.

Government, and on a voluntary basis, by the private sector" [8]. Therefore, it is very important to constantly reevaluate the security of AES.

Differential cryptanalysis [2] is a chosen-plaintext attack that can be applied to a large class of encryption algorithms. It analyzes the effect of the difference of a pair of plaintexts on the difference of succeeding round outputs in an iterated cipher. Because of its generality, differential cryptanalysis is extensively exploited tool for cryptanalysis of encryption algorithms [3] and for defining new attacks [15, 14]. Resistance to differential cryptanalysis became one of the basic criteria in the evaluation of the security of block encryption algorithms. Rijndael [9] is designed according to the wide trail strategy [10], and one of the goals of this strategy is resistance to differential cryptanalysis.

Related-key attacks [4, 13] allow the cryptanalyst to obtain p/c pairs by using different, but related keys. The relation between the keys can be chosen by the attacker. Ferguson et al [11] noticed that "compared to the cipher itself, the Rijndael key schedule appears to be more of an ad hoc design. It has much slower diffusion structure than the cipher, and contains relatively few non-linear elements." They exploited the slow diffusion of the Rijndael key schedule to mount a related-key attack on 9 rounds of 256-bit key Rijndael. The time complexity of the attack is 2^{224} and it requires 2^{85} plaintext/ciphertext pairs obtained under 256 related keys.

In this paper, we apply combinations of related-key and differential attacks to AES-192. The attacks exploit the effect of the difference of a pair of plaintexts on the difference of succeeding round outputs when the plaintexts are encrypted using distinct keys. The chosen relation between the keys is the key difference, which can be possibly obtained by fault insertion [17]. The complexity of the attack depends on the ability of the attacker to predict the propagation of the key difference during the key schedule. If we view the expanded key as a sequence of words, then the key schedule of AES-192 applies non-linear transformations to every sixth word, whereas the key schedules of AES-128 and AES-256 apply non-linear transformations to every fourth word. Therefore, we believe that AES-192 is more susceptible to related-key differential cryptanalysis than the other AES variants. We were able to break eight rounds of AES-192 using 2^{88} plaintext/ciphertext pairs. The time complexity of the attack is about 2^{183} encryptions and it suggests that the problem of designing block ciphers being resistant to related-key differential cryptanalysis is worth investigating.

The techniques used by related-key differential cryptanalysis are similar to the techniques used by differential cryptanalysis (e.g. finding highly probable differential trail). Markov cipher is a concept that was introduced to analyze the resistance of iterated ciphers to differential cryptanalysis. Namely, in [16], the authors showed that the sequence of differences at each round output forms a Markov chain if the iterated cipher is Markov and its round keys are independent. Assuming that the hypothesis of stochastic equivalence holds, then the cipher is secure against differential cryptanalysis after sufficiently many rounds. Revisiting the concept of Markov ciphers can give us some insight about the properties of a cipher that make it resistant to related-key differential cryptanalysis.

Hence, in this paper, we also address the question of sufficient conditions such that the sequence of differences forms a Markov chain in the case of related-keys.

Here is the outline of the paper. In Section 2, we give a definition of related-key differentials and describe a general related-key differential attack. In Section 3, we examine the resistance of the AES to related-key differential cryptanalysis. Some results about the properties that make iterated ciphers resistant to related-key differential cryptanalysis are derived in Section 4. The paper ends with the concluding remarks.

2 Related-Key Differential Attack

Differential cryptanalysis exploits the propagation of the differences when a pair of distinct plaintexts is submitted for encryption under the same key. Related-key differential cryptanalysis exploits the properties of the difference propagation when the plaintexts x_1 and x_2, which can be equal, are submitted for encryption under distinct keys k_1 and k_2 correspondingly.

Formally, an r-round related-key differential is a triple (α, β, δ), where α is the difference of the inputs at the input of the block encryption algorithm, β is the difference of the outputs of the r-th round and δ is the difference of the keys. The probability of r-round related-key differential is the probability that the difference of the outputs of the r-th round will be β, when the input difference is α, the key difference is δ, and the plaintext x_1 and the key k_1 are selected uniformly at random.

A possible related-key differential attack works as follows:

- Find highly probable $R-1$-round related-key differential, where R is the number of rounds of the block cipher.
- Select randomly x_1 and submit it for encryption under key k_1 to obtain ciphertext y_1. Compute $x_2 = x_1 + \alpha$ and submit it for encryption under key $k_2 = k_1 + \delta$ to obtain the ciphertext y_2.
- Find all possible last round key pairs (k_1^R, k_2^R) such that the difference between $d_{k_1^R}(y_1)$ and $d_{k_2^R}(y_2)$ is β, where $d_k(y)$ is the output of the first round of the decryption algorithm for input y and round key k. Add one to each counter that corresponds to one of the previously computed key pairs.
- Repeat previous two steps until one or more last round key pairs are counted significantly more than the others. Check these keys if they are the right keys.

Let K be the number of possible last round key pairs (k_1^R, k_2^R) and let l be the average number of suggested key pairs in the third step. Furthermore, let $p_{max} >> 2^{-m}$, where m is the block length and p_{max} is the probability of the related-key differential found in step 1 and let $N = c/p_{max}$ be the number of repetitions of steps two and three. Then the wrong key pairs will be counted lN/K times on average and the right key pair will be counted about c times on average. If $l \times c < K \times p_{max}$, then the wrong key pairs will be counted less then once on average. The order of the number of required plaintext/ciphertext pairs is $1/p_{max}$.

The related-key differential attack does not have to follow the pattern described above. In general, we will refer to any attack that exploits a related-key differentials as related-key differential attack.

3 Related-Key Differential Attacks on AES-192

In this section, we describe some attacks on reduced 192-bit key variants of AES. Description of the algorithm can be found in [1]. We will use the following notation: $x_i^I, x_i^S, x_i^P, x_i^M$ and x_i^O denote the input of the round i, the output after SubBytes, the output after ShiftRows, the output after MixColumns and the output after AddRoundKey transformation, correspondingly; k_i denotes the round i key and we will use $a_{i,j}, i, j \in \{0, 1, 2, 3\}$ to denote the byte j of the 32-bit word (column) i of a. For example, $x^P_{3,0,2}$ denotes the third byte of the first column of the output after the ShiftRows transformation in the round 3 (the initial key addition of the algorithm is considered as round 0). The encryption round can be transformed into equivalent one that uses key addition before the MixColumns. The round keys used in the equivalent encryption round will be denoted by z_i. Finally, we assume that the last round of the analyzed variants is a FinalRound.

3.1 Basic Attack

Differential cryptanalysis attacks are based on difference propagations over all but few rounds that have large enough prop ratio. For Rijndael, it is proven that any 4-round differential trail has at least 25 active bytes, that is there are no 4-round differential trails with predicted prop ratio above 2^{-150} [9, 10]. The idea of the attack described here is to use the round key differences in order to cancel the differences that exist before the key addition and reduce the number of active bytes.

The propagation of the key difference (0000)(0000)(0Δ00)(0Δ00)(0000)(0000) is depicted in Table 1. If we submit two plaintexts x and x' for encryption under the keys k and $k' = k \oplus \Delta k$ correspondingly, such that $\Delta x =$ (0000)(0000)(0Δ00)(0Δ00), then a possible propagation of the difference is the one shown in Table 2.

Table 1. Propagation of the key difference

j	Δk_j
0	(0000)(0000)(0Δ00)(0Δ00)
1	(0000)(0000)(0000)(0000)
2	(0Δ00)(0000)(0000)(0000)
3	(0000)(0000)(0Δ00)(0Δ00)
4	(0Δ00)(0Δ00)($\Delta_1$000)($\Delta_1$000)
5	($\Delta_1\Delta$00)($\Delta_1$000)($\Delta_1\Delta$00)($\Delta_1$000)
6	($\Delta_1$00Δ_2)(000Δ_2)($\Delta_1\Delta$0Δ_2)(0Δ0Δ_2)

Table 2. Possible propagation of the plaintext difference

j	Δx_j^I
0	$(0000)(0000)(0\Delta 00)(0\Delta 00)$
1	$(0000)(0000)(0000)(0000)$
2	$(0000)(0000)(0000)(0000)$
3	$(0\Delta 00)(0000)(0000)(0000)$
4	$(0000)(0000)(0\Delta 00)('03' \cdot \Delta' \| \| 0\Delta'\Delta')$
	$\Delta x_4^S = (0000)(0000)(0\Delta 00)(\Delta'' 0\Delta\Delta)$
5	$(\Delta 0 \| \| '03' \cdot \Delta \| \| '02' \cdot \Delta)('02' \cdot \Delta \| \| 0 \| \| '03' \cdot \Delta \| \| 0)$
	$(\Delta_1 000)('02' \cdot \Delta'' \oplus \Delta_1 \| \| \Delta''\Delta'' \| \| '03' \cdot \Delta'')$

The difference Δ' is selected to satisfy the relation $'02' \cdot \Delta' = \Delta$. In addition, the differences Δ, Δ' and Δ'' should be selected so that the probabilities $P_S(\Delta \to \Delta), P_S(\Delta \to \Delta'), P_S(\Delta' \to \Delta)$ and $P_S('03' \cdot \Delta' \to \Delta'')$ are greater than zero, where $P_S(a \to b)$ is the probability that the output difference of the S-box will be b when the input difference is a. When the previous conditions are satisfied, the 5-round related-key differential trail from Table 2 has 15 active bytes and its probability is $2^{-7\times 15} = 2^{-105}$ in the worst case. If we use 2^{106} plaintext/ciphertext pairs in a related-key differential attack on the six round variant, then the right key will be counted at least twice on average. The time complexity of the attack is about 2^{112} encryptions.

3.2 Improving the Basic Attack: Truncated Differentials

Let us consider the same plaintext and key differences as in the previous subsection. The probability that $\Delta x_{5,2}^I = \Delta_1 000$ is the same as the probability $P_S(\Delta \to \Delta')$ and it is 2^{-7} instead of 2^{-32}, which is the probability when the difference $\Delta x_{5,2}^I$ is uniformly distributed. This highly probable truncated differential [14] is exploited in the attack on six round version of AES. The attack is described below.

We assign counter to every 10-tuple

$$(\Delta_1, \Delta_2, k_{6,0,0}, k_{6,0,1}, k_{6,0,2}, k_{6,0,3}, k_{6,1,1}, k_{6,2,0}, k_{6,3,6}, z_{5,2,0})$$

that is possible for a given Δ. The attack is as follows:

- Select randomly a plaintext pair (x, x') such that the plaintext difference is $(0000)(0000)(0\Delta 00)(0\Delta 00)$ and the ciphertext difference is $\Delta y = (****)(0*0\Delta_2)(*\Delta 0\Delta_2)(0\Delta 0*)$, where '*' means any value, $y = E_k(x), y' = E_{k\oplus\delta}(x')$ and $\delta = (0000)(0000)(0\Delta 00)(0\Delta 00)(0000)(0000)$.
- For every possible 10-tuple, check whether $\Delta x_{5,2}^I = \Delta_1 000$. The value of $\Delta x_{5,2,0}^I$ can be determined from the ciphertext pair using the key differences and the five bytes of the key $k_{6,0,2}, k_{6,1,1}, k_{6,2,0}, k_{6,3,3}$ and $z_{5,2,0}$. Further, the difference $\Delta x_{6,i,j}^P$ is zero for all i and j such that $k_{6,i,j}$ is unknown. Hence, one can compute the differences Δx_5^O, and therefore,the value Δx_5^S can also

be computed due to the linearity of the MixColumns transformation. Now, it is easy to check whether the particular difference before the SubBytes transformation of the round 5 is zero. If $\Delta x^I_{5,2} = \Delta_1 000$, then add one to the counter that corresponds to the particular 10-tuple.

– Repeat the previous two steps until one or more of the 10-tuples are counted significantly more than the others. Take this values as possible values of the specified bytes of the key.

If we repeat the first two steps 2^8 times, then the right 10-tuple will be counted twice on average, while the wrong 10-tuples will be counted 2^{-24} times assuming that when we use wrong key values the probability distribution of $\Delta x^I_{5,2}$ is uniform. Further, the probability that the output difference Δy will be $(***)(0*0\Delta_2)(*\Delta 0\Delta_2)(0\Delta 0*)$, when the plaintext difference is $\Delta x = (0000)(0000)(0\Delta 00)\ (0\Delta 00)$, is $2^{-9\times 8} = 2^{-72}$. Therefore, the number of plaintext pairs required for the attack is about $2^{72} \times 2^8 = 2^{80}$. There are at most 2^{14} possible values of (Δ_1, Δ_2) for a given Δ. Hence, the complexity of the attack is about $2^8 \times 2^{14} \times 2^{8\times 8} = 2^{86}$ encryptions. The previously described attack is used to determine eight bytes of the key. It is not difficult to find the rest of the key using similar methods.

3.3 Impossible Related-Key Differentials Attack

Impossible differential attack against Rijndael reduced to five rounds was proposed by Biham and Keller [6]. Later, this attack was extended to six rounds [7]. In this section, we describe related-key impossible differentials attacks on 192-bit key variant reduced to seven and eight rounds.

The attack exploits a similar weakness in the key schedule as the previous attacks. Namely, if the key difference is $(0000)(0000)(\Delta 000)(\Delta 000)(0000)(0000)$, then this difference is propagated during the key generation as depicted in Table 3. We can see that the round 1 key difference is zero and the round 2 keys differ in only one byte. If we submit two plaintexts x and x' for encryption, such that $\Delta x = (0000)(0000)(\Delta 000)(\Delta 000)$, then Δx^I_1 is zero, and so is $\Delta x^O_1 = \Delta x^I_2$. Because of the round 2 key difference, the inputs of the third

Table 3. Propagation of the key difference

j	Δk_j
0	$(0000)(0000)(\Delta 000)(\Delta 000)$
1	$(0000)(0000)(0000)(0000)$
2	$(\Delta 000)(0000)(0000)(0000)$
3	$(0000)(0000)(\Delta 000)(\Delta 000)$
4	$(\Delta 000)(\Delta 000)(000\Delta_1)(000\Delta_1)$
5	$(\Delta 00\Delta_1)(000\Delta_1)(\Delta 00\Delta_1)(000\Delta_1)$
6	$(00\Delta_2\Delta_1)(00\Delta_20)(\Delta 0\Delta_2\Delta_1)(\Delta 0\Delta_20)$
7	$(00\Delta_2\Delta_1)(00\Delta_20)(0\Delta_3\Delta_2\Delta_1)(0\Delta_30\Delta_1)$

round will differ in only one byte $x^I_{3,0,0}$. Due to the MixColumn transformation and the round 3 key difference, the inputs of the round 4 will differ in six bytes $x^I_{4,0,1}, x^I_{4,0,2}, x^I_{4,0,3}, x^I_{4,1,1}, x^I_{4,1,2}$, and $x^I_{4,1,3}$. Hence, $\Delta x^M_{5,3} \neq 0000$ and $\Delta x^O_{5,3} \neq 000\Delta_1$.

The aforementioned fact can be used to find values of seven bytes of the last round key. Given Δ, for every possible 10-tuple

$$(\Delta_1, \Delta_2, k_{7,0,0}, k_{7,1,0}, k_{7,1,1}, k_{7,1,2}, k_{7,1,3}, k_{7,2,2}, k_{7,3,1}, z_{6,0,3})$$

do the following:

- Compute Δ_3 using $k_{7,1,2}$ and Δ_2.
- For a plaintext pair (x, x') such that $\Delta x = (0000)(0000)(\Delta 000)(\Delta 000)$ and the ciphertext difference is $\Delta y = (*0\Delta_2\Delta_1)(****)(0\Delta_3 * \Delta_1)(0 * 0\Delta_1)$, where $y = E_k(x), y' = E_{k\oplus\delta}(x')$ and $\delta = (0000)(0000)(\Delta 000)(\Delta 000)(0000)\ (0000)$, check whether $\Delta x^O_{5,3} = 000\Delta_1$. The value of $\Delta x^O_{5,3,3}$ can be determined from the ciphertext pair using the key differences and the five bytes of the key $k_{7,0,0}, k_{7,1,3}, k_{7,2,2}, k_{7,3,1}$ and $z_{6,0,3}$. Further, the difference $\Delta x^P_{7,i,j}$ is zero for all i and j such that $k_{7,i,j}$ is unknown. Hence, one can compute the differences Δx^O_6 and therefore Δx^S_6 due to the linearity of the MixColumns transformation. Once the value of Δx^S_6 is determined, it is not difficult to check whether $\Delta x^O_{5,3} = 000\Delta_1$. If $\Delta x^O_{5,3} = 000\Delta_1$, then mark the current 10-tuple as wrong.
- Repeat the previous step until the 10-tuple is marked as wrong or the maximum of 2^{38} tried plaintext pairs is reached.

The probability that the ciphertext difference will be $\Delta y = (*0\Delta_2\Delta_1)(* * **)(0\Delta_3 * \Delta_1)(0 * 0\Delta_1)$, when the plaintext x is randomly selected, is $2^{-9\times 8} = 2^{-72}$. Hence, the number of plaintext pairs required to obtain 2^{38} plaintext pairs with the desired property is about 2^{110}. Given Δ, the number of possible values of (Δ_1, Δ_2) is less than 2^{14}. Thus, the complexity of finding the possible 10-tuples is of order $2^{38} \times 2^{14} \times 2^{8\times 8} = 2^{116}$ encryptions. The probability that particular wrong 10-tuple will be marked as wrong using only one pair of plaintexts is 2^{-32}. The number of wrong 10-tuples that are not marked as wrong after applying the procedure 2^{38} times is on average $2^{14} \times 2^{64} \times (1 - 2^{-32})^{2^{38}} \approx 2^{78} \times e^{-2^6} \approx 2^{-14}$ i.e. most of the time there will be no wrong keys that are not marked as wrong. The previous procedure is used to find eight bytes of the key. The rest of the key can be determined using similar techniques with complexity which is negligible compared to the complexity of the overall attack.

The attack can be extended to eight rounds. We will use the same plaintext and key differences, but we will use the fact that $\Delta x^M_{5,1} \neq 0000$ and $\Delta x^O_{5,1} \neq 000\Delta_1$, which can be proved to be true by similar arguments as in the previous case.

Given Δ, for every possible 3-tuple

$$(k_8, z_{7,2,0}, z_{7,3,3})$$

do the following:

- Compute: $k_{7,0}, k_{7,1}, k_{6,3}, k_{5,2}, k_{5,3}, k_{4,1}$ and $z_{8,0,3}$ using k_8; Δ_1 using $k_{4,1}$ and Δ; Δ_2 using $k_{5,3}$ and Δ_1; Δ_3 using $k_{7,1}$ and Δ_2; and finally, $z_{5,2,3}$ using $z_{7,3,3}$ and $z_{8,0,3}$.
- For a plaintext pair (x, x') such that $\Delta x = (0000)(0000)(\Delta 000)(\Delta 000)$ and the difference $\Delta_7^{a,P}$ is $(***)(***)(*000)(000*)$, check whether $\Delta x_{5,1}^O = 000\Delta_1$. The difference $\Delta_7^{a,P}$ is the difference after the ShiftRows transformation of the round 7 computed using the assumed value k_8 from the ciphertext pair obtained when x is submitted for encryption under k and x' is submitted for encryption under $k \oplus \delta$. The value of $\Delta x_{5,1,3}^O$ can be determined from the ciphertext pair using the key differences, $k_8, k_{7,0}, k_{7,1}, z_{7,2,0}, z_{7,3,3}$ and $z_{5,2,3}$.Further, the difference $\Delta x_{7,i,j}^P$ is zero for all i and j such that $z_{7,i,j}$ can not be computed. Hence, one can compute the differences Δx_6^O, and therefore Δx_6^S also due to the linearity of the MixColumns transformation. Now, it is easy to check whether the particular difference before the SubBytes transformation of the round 6 is zero. If $\Delta x_{5,1}^O = 000\Delta_1$, then mark the current 3-tuple as wrong.
- Repeat the previous step until the 3-tuple is marked as wrong or the maximum of 2^{39} tried plaintext pairs is reached.

The probability that the difference $\Delta_7^{a,P}$ will be$(***)(***)(*000)(000*)$, when the plaintext x is randomly selected, is $2^{-6\times 8} = 2^{-48}$. The number of plaintext pairs required to obtain 2^{39} plaintext pairs with the desired property is about 2^{87}. There are $2^{128} \times 2^{2\times 8} = 2^{144}$ values of $(k_8, z_{7,2,0}, z_{7,3,3})$. Thus, the complexity of finding the "right" 3-tuples is of order $2^{39} \times 2^{144} = 2^{183}$ encryptions. The probability that particular wrong 3-tuple will be marked as wrong using only one pair of plaintexts is 2^{-32}. The number of wrong 3-tuples that are not marked as wrong after applying the procedure 2^{39} times is on average $2^{144} \times (1 - 2^{-32})^{2^{39}} \approx 2^{-40}$ i.e. the probability that only the right key will not be marked as wrong is very large. Once the right 3-tuple is determined, it is easy to determine the rest of the key using exhaustive search. One naive way to select the set of 2^{39} plaintext pairs with desired property from the set of 2^{87} available plaintext pairs is to check whether each pair leads to the required difference $\Delta x_7^{a,P}$ for the particular key k_8. In that case, the complexity will be $2^{87+144} = 2^{231}$. The differences $\Delta x_{7,2}^{a,P}$ and $\Delta x_{7,3}^{a,P}$ depend only on eight bytes of the key k_8 and the key differences Δ_1, Δ_2 and Δ_3. Hence, a better way to select the set is to assume first the values of these eight bytes and then compute the set for every possible value of Δ_1, Δ_2 and Δ_3. Then, we can assume the rest of the key, compute the real values of Δ_1, Δ_2 and Δ_3, and select the set that corresponds to the real values of the key differences. Selection can be made by selecting those pairs such that $\Delta x_{7,3}^{a,P} = (000*)$, and then selecting the pairs that satisfy $\Delta x_{7,2}^{a,P} = (*000)$ from the previously selected pairs. The complexity in this case is about $2^{4\times 8} \times 2^{3\times 7} \times 2^{87} = 2^{140}$. Table 4[1] compares the complexities of impossible related-key differential attacks to the complexities of the partial

[1] RK-CP stands for related-key chosen plaintext, and CP stands for chosen plaintext.

Table 4. Comparison of the impossible related-key differential attacks to partial sums attacks on AES-192

# of rounds	p/c pairs	Time	Attack
7	2^{111} RK-CP	2^{116}	impossible related-key differential
8	2^{88} RK-CP	2^{183}	impossible related-key differential
7	19×2^{32} CP	2^{155}	partial sums
7	$2^{128} - 2^{119}$ CP	2^{120}	partial sums
8	$2^{128} - 2^{119}$ CP	2^{188}	partial sums

sums attacks proposed in [11], which is the best attack on the 192-bit key variant known to the authors of this paper.

4 Markov Cipher Property Is Not Enough

The concept of Markov ciphers was introduced in order to analyze the security of iterated block ciphers against differential cryptanalysis. We give the following definition taken from [16]:

Definition 1. *An iterated cipher with round function* $y = f(x, k)$ *is* a Markov cipher *if there is a group operation for defining differences such that, for all choices of* α, $\alpha \neq e$ *and* β, $\beta \neq e$

$$P_o(\Delta y = \beta | \Delta x = \alpha, x = \gamma)$$

is independent of γ *when the subkey* k *is uniformly random, where* $P_o(\Delta y = \beta | \Delta x = \alpha, x = \gamma)$ *is the probability when the same round key is used to encrypt* γ *and* $\gamma + \alpha$, *and* e *is the identity element.*

One can easily notice that *if an iterated cipher is Markov cipher, then the previous property holds even when* $\alpha = e$ *or* $\beta = e$. The differences in the previous definition are computed when the ciphertexts are obtained using the same key. It is shown that, if an iterated cipher is Markov and its round keys are independent, then the sequence of differences at each round output forms a Markov chain. Furthermore, if the Markov chain of differences has a steady state probability distribution, then this steady state distribution must be the uniform distribution. If we additionally assume that the hypothesis of stochastic equivalence holds for the Markov cipher, then, for almost all subkeys, this cipher is secure against a differential cryptanalysis attack after sufficiently many rounds (see [16] for more details).

The differences in the previous discussion are computed when the ciphertexts are obtained using the same key. In general, we can consider differences in the case when the ciphertexts are obtained using different keys. When the round keys are independent, it is obvious that we can construct highly probable related-key differentials by encrypting the same plaintext using keys that differ in one round key (the key of one of the last rounds). This is demonstrated by the following example.

Magenta [12] is 128-bit block encryption algorithm submitted for AES by Deutsche Telekom AG. It supports 128-bit, 192-bit and 256-bit key sizes. We will consider the 128-bit key variant, which consist of of six Feistel rounds. The key is divided into two 64-bit halves K_1 and K_2. The first part K_1 is used in rounds 1,2,5 and 6, and the second part K_2 is used in the remaining rounds 3 and 4. The algorithm is given by

$$E_K(M) = F_{K_1}(F_{K_1}(F_{K_2}(F_{K_2}(F_{K_1}(F_{K_1}(M)))))),$$

where

$$F_y(x) = ((x_8, \ldots, x_{15}), (x_0, \ldots, x_7) \oplus E^{(3)}(x_8, \ldots, x_{15}, y_0, \ldots, y_7)).$$

Let Δy and ΔE be two differences such that $P(\Delta E^{(3)} = \Delta E | \Delta y, \Delta x = 0)$ is significantly greater[2] than 2^{-64}. If we submit the same plaintext for encryption under the keys (K_1, K_2) and $(K_1, K_2 \oplus \Delta y)$, then the difference between the left halves at the input of the fourth round will be ΔE with probability significantly higher than 2^{-64}. We must note that, although the attack that exploits such related-key differential is more efficient than exhaustive search, the complexity of the attack is large compared to the attack proposed in [5]. It is obvious that we must take the subkey differences into account if we want to analyze the resistance of iterated ciphers to related-key differential cryptanalysis.

Definition 2. *We say that* the round function $y = f(x, k)$ is $\mathcal{K} - f$ *if for every* α *,* β *and* δ *one can find* α_1 *such that*

$$P(\Delta y = \beta | \Delta x = \alpha, \Delta k = \delta, x = \gamma) = P_o(\Delta y = \beta | \Delta x = \alpha_1, x = \gamma)$$

for any γ *and uniform distribution of the subkey* k*.*

Often, the round function is composed of key addition using bitwise XOR and bijective transformation (e.g. AES). In this case, the difference α_1 can be simply computed[3] as $\alpha_1 = \alpha \oplus \delta$. The definition of $\mathcal{K} - f$ round functions enforces relation between the probability distributions of the round output differences in the cases of zero and nonzero key differences. This is formally stated by the following theorem.

Theorem 1. *If the round function is* $\mathcal{K} - f$ *and the input* x *is independent of the input difference* Δx *and round key difference* Δk*, then for every* α *,* β *and* δ *one can find* α_1 *such that*

$$P(\Delta y = \beta | \Delta x = \alpha, \Delta k = \delta) = P_o(\Delta y = \beta | \Delta x = \alpha_1).$$

[2] Authors mention 2^{-40} as an upper bound for transition probabilities of $E^{(3)}$.
[3] This is the reason why we use a somewhat strange notation $\mathcal{K} - f$.

Proof.

$$\begin{aligned}
&P(\Delta y = \beta | \Delta x = \alpha, \Delta k = \delta) = \\
&= \sum_{\gamma} P(\Delta y = \beta | \Delta k = \delta, \Delta x = \alpha, x = \gamma) \times P(x = \gamma) = \\
&= \sum_{\gamma} P_o(\Delta y = \beta | \Delta x = \alpha_1, x = \gamma) \times P(x = \gamma) = \\
&= P_o(\Delta y = \beta | \Delta x = \alpha_1). \square
\end{aligned}$$

The Markov cipher property (round output difference to depend only on the the round input difference and not on the particular round inputs) is crucial in proving that the sequence of the round output differences forms a homogenous Markov chain. Therefore, it is convenient to define a similar property in the case of related-key differentials.

Definition 3. *An iterated cipher with round function $y = f(x, k)$ possesses* a Markov cipher property for related keys *if there is a group operation for defining differences such that, for all choices of α and β*

$$P(\Delta y = \beta | \Delta x = \alpha, x = \gamma) = P(\Delta y = \beta | \Delta x = \alpha)$$

for any probability distribution of the round key differences and uniformly distributed round key k.

The $\mathcal{K} - f$ property of the round function enables us to analyze the propagation of the related-key differences by observing the propagation of the differences when we use the same key for encryption of the pair of plaintexts. Therefore, it is not surprising that Markov ciphers with $\mathcal{K} - f$ round function possess a Markov cipher property for related keys.

Theorem 2. *If an iterated cipher is a Markov cipher with $\mathcal{K} - f$ round function, the round key is uniformly distributed, and the round key difference is independent of the input and the input difference, then the cipher possesses a Markov cipher property for related keys.*

Proof.

$$\begin{aligned}
&P(\Delta y = \beta | \Delta x = \alpha, x = \gamma) = \\
&= \sum_{\delta} P(\Delta y = \beta, \Delta k = \delta | \Delta x = \alpha, x = \gamma) = \\
&= \sum_{\delta} P(\Delta y = \beta | \Delta x = \alpha, \Delta k = \delta, x = \gamma) \times P(\Delta k = \delta | \Delta x = \alpha, x = \gamma) = \\
&= \sum_{\delta} P(\Delta k = \delta) \times P_o(\Delta y = \beta | \Delta x = \alpha_1, x = \gamma) \\
&= \sum_{\delta} P(\Delta k = \delta) \times P_o(\Delta y = \beta | \Delta x = \alpha_1)
\end{aligned}$$

$$= \sum_{\delta} P(\Delta k = \delta) \times P(\Delta y = \beta | \Delta x = \alpha, \Delta k = \delta)$$
$$= \sum_{\delta} P(\Delta k = \delta, \Delta y = \beta | \Delta x = \alpha)$$
$$= P(\Delta y = \beta | \Delta x = \alpha). \square$$

The previous results provide intuition for proving the following theorem.

Theorem 3. *If* (i) *an r-round iterated cipher is a Markov cipher with $\mathcal{K} - f$ round function,* (ii) *the round keys k_i are independent and uniformly random and* (iii) *the round key differences are independent random variables, then the sequence of differences $\Delta x = \Delta y(0), \Delta y(1), \ldots \Delta y(r)$ is a Markov chain. If additionally* (iv) *the probability distributions $p(\Delta k_i)$ are identical, then the Markov chain is homogeneous.*

Proof.

$$\begin{aligned}
&P(\Delta y(i) = \beta_i | \Delta y(i-1) = \beta_{i-1}, \ldots, \Delta x = \alpha) = \\
&= \sum_{\gamma} P(\Delta y(i) = \beta_i, y(i-1) = \gamma | \Delta y(i-1) = \beta_{i-1}, \ldots, \Delta x = \alpha) = \\
&= \sum_{\gamma} P(\Delta y(i) = \beta_i | y(i-1) = \gamma, \Delta y(i-1) = \beta_{i-1}, \ldots, \Delta x = \alpha) \times \\
&\quad P(y(i-1) = \gamma | \Delta y(i-1) = \beta_{i-1}, \ldots, \Delta x = \alpha) = \\
&= \sum_{\gamma} P(\Delta y(i) = \beta_i | y(i-1) = \gamma, \Delta y(i-1) = \beta_{i-1}) \times \\
&\quad P(y(i-1) = \gamma | \Delta y(i-1) = \beta_{i-1}, \ldots, \Delta x = \alpha) = \\
&= \sum_{\gamma} P(\Delta y(i) = \beta_i | \Delta y(i-1) = \beta_{i-1}) \times \\
&\quad P(y(i-1) = \gamma | \Delta y(i-1) = \beta_{i-1}, \ldots, \Delta x = \alpha) = \\
&= P(\Delta y(i) = \beta_i | \Delta y(i-1) = \beta_{i-1})
\end{aligned}$$

If the probability distributions $P(\Delta k_i)$ are identical, then

$$\begin{aligned}
&P(\Delta y(i) = \beta | \Delta y(i-1) = \alpha) = \\
&= \sum_{\delta} P(\Delta y(i) = \beta, \Delta k_i = \delta | \Delta y(i-1) = \alpha) = \\
&= \sum_{\delta} P(\Delta y(i) = \beta | \Delta k_i = \delta, \Delta y(i-1) = \alpha) \times P(\Delta k_i = \delta) = \\
&= \sum_{\delta} P_o(\Delta y = \beta | \Delta x = \alpha_1) \times P(\Delta k_i = \delta) = \\
&= \sum_{\delta} P_o(\Delta y = \beta | \Delta x = \alpha_1) \times P(\Delta k_{i-1} = \delta) = \\
&= P(\Delta y(i-1) = \beta | \Delta y(i-2) = \alpha). \square
\end{aligned}$$

Now, suppose that the round keys and round key differences are independent and uniformly distributed. If the round input difference is uniformly distributed, then the round output difference is also uniformly distributed. Hence, if the Markov chain formed by the round output differences has steady-state probability distribution, then this steady-state distribution must be the uniform distribution. Usually, the round keys are derived using some key generation algorithm, and, given the key and the key difference, the round keys and the round key differences are uniquely determined. If we assume that the probability of the related-key differentials when the round keys and round key differences are fixed, and the probability of the related-key differentials when the round keys and round key differences are independent and uniformly distributed are approximately equal, then the previously discussed Markov ciphers are secure against related-key differential attack after sufficiently many rounds. We will refer to this assumption as *hypothesis of stochastic equivalence for related keys*.

The previous discussion suggests that one way of dealing with related-key differential cryptanalysis is to use key scheduling algorithms whose output is "close" to random. We already mentioned that the success of a related-key attack depends on the attacker's ability to find highly probable (or impossible) related-key differential trails. Unpredictable key differences make the task of constructing such related-key differential trails very difficult.

5 Conclusion

We applied the related-key differential cryptanalysis to the 192-bit key variant of AES. The related-key differential attack on six rounds requires 2^{106} plaintext/ciphertext pairs and its complexity is 2^{112}. Using truncated differentials, we can improve the attack on six rounds. In this case, the number of required plaintext/ciphertext pairs is 2^{81} and the complexity is about 2^{86}. Impossible related-key differential cryptanalysis gave best results. The complexity of the attack on seven rounds is 2^{116} and requires 2^{111} plaintext/ciphertext pairs, and the complexity of the attack on eight rounds is about 2^{183} encryptions and requires 2^{88} plaintext/ciphertext pairs.

We also examined the additional constraints that should be satisfied so that the sequence of round output differences forms a Markov chain in the case of related keys. If the Markov cipher has $\mathcal{K} - f$ round function and the round key differences are independent variables with identical probability distribution, then the sequence forms a homogenous Markov chain. Assuming that the hypothesis of stochastic equivalence for related keys holds and steady-state probability distribution exists, then the steady-state probability distribution is the uniform distribution and the cipher is secure against related-key differential cryptanalysis after sufficiently many rounds.

References

[1] Advanced Encryption Standard (AES), FIPS Publication 197, November 26, 2001, available at http://csrc.nist.gov/encryption/aes. 208, 211

[2] E.Biham and A.Shamir, *"Differential cryptanalysis of DES-like cryptosystems,"* Journal of Cryptology, 4(1):3-72, 1991. 209

[3] E. Biham and A. Shamir, *"Differential Cryptanalysis of Snefru, Khafre, REDOC II, LOKI, and Lucifer,"* Advances in Cryptology, CRYPTO '91 Proceedings, Springer- Verlag, 1992, pp. 156-171. 209

[4] E. Biham, *"New Types of Cryptanalytic Attacks Using Related Keys,"* Journal of Cryptology, v. 7, n. 4, 1994, pp. 229-246. 209

[5] E. Biham, A. Biryukov, N. Ferguson, L. Knudsen, B. Schneier, A. Shamir, *"Cryptanalysis of MAGENTA",* http:// csrc.nist.gov/ encryption/ aes/ round1/ conf2/ aes2conf.htm 217

[6] E.Biham and N.Keller, *"Cryptanalysis of Reduced Variants of Rijndael,"* http:// csrc.nist.gov/ encryption/ aes/ round2/ conf3/ aes3papers.html 213

[7] J. Cheon, M. Kim, K. Kim, J. Lee, and S. Kang, *"Improved Impossible Differential Cryptanalysis of Rijndael and Crypton,"* Information Security and Cryptology - ICISC 2001 4th International Conference Seoul, Korea, December 6-7, 2001, Proceedings, LNCS 2288,p. 39 ff. 213

[8] http:// csrc.nist.gov/ CryptoToolkit/ aes/ round2/ r2report.pdf 209

[9] J.Daemen and V.Rijmen, *"AES Proposal: Rijndael,"* http:// csrc.nist.gov/ encryption/ aes. 208, 209, 211

[10] J.Daemen, *"Cipher and hash function design strategies based on linear and differential cryptanalysis,"* Doctoral Dissertation, March 1995, K. U.Leuven. 209, 211

[11] N. Ferguson, J. Kelsey, B. Schneier, M. Stay, D. Wagner, D. Whiting, *"Improved Cryptanalysis of Rijndael",* 7th International Workshop, FSE 2000, New York, NY, USA, April 2000, Proceedings, LNCS 1978, p. 213 ff. 209, 216

[12] M. J.Jacobson,Jr and K.Huber, *"The MAGENTA Block Cipher Algorithm,"* AES candidate, http:// csrc.nist.gov/ encryption/ aes. 217

[13] J.Kelsey, B.Schneier and D.Wagner, *"Key-schedule cryptanalysis of IDEA, GDES, GOST, SAFER, and Triple-DES,"* Advances in Cryptology, Proceedings Crypto'96, LNCS 1109, pp.237-252. 209

[14] L. R. Knudsen, *"Truncated and Higher Order Differentials,"* Fast Software Encryption, 2nd International Workshop Proceedings, Springer-Verlag, 1995, pp. 196-211. 209, 212

[15] X. Lai, *"Higher Order Derivations and Differential Cryptanalysis,"* Communications and Cryptography: Two Sides of One Tapestry, Kluwer Academic Publishers, 1994, pp. 227-233. 209

[16] X. Lai, J. Massey, and S. Murphy, *"Markov Ciphers and Differential Cryptanalysis,"* Advances in Cryptology, CRYPTO '91 Proceedings, Springer-Verlag, 1991, pp. 17-38. 209, 216

[17] Jean-Jacques Quisquater and David Samyde, *"Eddy current for Magnetic Analysis with Active Sensor",* Proceedings of Esmart 2002 3rd edition, Nice, France, September 2002. 209