THE PUBLISHING HOUSE
OF THE ROMANIAN ACADEMY

PROCEEDINGS OF THE ROMANIAN ACADEMY, Series A,
Volume 13, Number 4/2012, pp. 395–400

ON THE RELATED-KEY ATTACKS AGAINST AES*

Joan DAEMEN[1], Vincent RIJMEN[2]

[1] STMicroElectronics, Belgium
[2] KU Leuven & IBBT (Belgium), Graz University of Technology, Austria
E-mail: Vincent.Rijmen@esat.kuleuven.be

Alex Biryukov and Dmitry Khovratovich presented related-key attacks on AES and reduced-round versions of AES. The most impressive of these were presented at Asiacrypt 2009: related-key attacks against the full AES-256 and AES-192. We discuss the applicability of these attacks and related-key attacks in general. We model the access of the attacker to the key in the form of key access schemes. Related-key attacks should only be considered with respect to sound key access schemes. We show that defining a sound key access scheme in which the related-key attacks against AES-256 and AES-192 can be conducted, is possible, but contrived.

Key words: Advanced Encryption Standard, AES, security, related-key attacks.

1. INTRODUCTION

Since the start of the process to select the Advanced Encryption Standard (AES), the block cipher Rijndael, which later became the AES, has been scrutinized extensively for security weaknesses. The initial cryptanalytic results can be grouped into three categories. The first category contains attacks variants that were weakened by reducing the number of rounds [0]. The second category contains observations on mathematical properties of sub-components of the AES, which don't lead to a cryptanalytic attack [0]. The third category consists of side-channel attacks, which target deficiencies in hardware or software implementations [0].

In 2009, the first cryptanalytic results on the full AES were published [0, 0]. These results have been obtained in the related-key security model. Although related-key attacks on block ciphers have been described in the literature [0], the security model isn't stable yet. A scenario where the adversary can query the block cipher under related keys, or even multiple keys, inevitably leads to security erosion [0]. A good view on the best possible security that can be achieved in such a scenario, is still lacking. In this paper, we recall some basic facts about related key attacks, give a concise description of the related-round-key attack described in [0], and describe the problems we see with this attack model.

Note that in 2011 a paper on biclique cryptanalysis of AES was presented, which doesn't require a related-key scenario [0]. We don't consider biclique cryptanalysis in this paper.

2. NOTATIONS AND DEFINITIONS

2.1. Block cipher security

An n-bit block cipher B is a family of permutations on the set of n-bit strings. We denote by B_k the permutation of the family B indexed by the key k.

The security of a block cipher, or any other cryptographic primitive, can be defined in different ways. One frequently used approach is to study how well the block cipher B behaves compared to some well-

* This material was in part presented at the international conference "Romanian Cryptology Days, RCD-2011."

defined idealized abstraction. An example in this category is *pseudo-random-permutation security*, where the abstraction is the set of all permutations on the set of n-bit strings [0], here denoted by Π. We denote by β a permutation randomly selected from Π. Let a be a randomly selected value in $\{0,1\}$. A *distinguisher* for B is an algorithm that can submit two types of *queries*. Firstly, it can submit any n-bit input x and receives in return the values $aB_k(x)+(1-a)\beta(x)$ and $(1-a)B_k(x)+a\beta(x)$. Secondly, it can submit any n-bit output y and receives in return the values $aB_k^{-1}(y)+(1-a)\beta^{-1}(y)$ and $(1-a)B_k^{-1}(y)+a\beta^{-1}(y)$. Note that the distinguisher doesn't get the values a,k, nor does it have any information on β other than what it obtains through the queries. The advantage of the distinguisher is the expected value (over all choices of a,k,β) of the probability that it outputs the correct value of a, minus one half. Sometimes, the minimum number of queries required by any distinguisher in order to have a significant advantage, is used as a measure of the security of the cipher. Sometimes, the number of computations to be performed by the distinguisher is taken into account as well. A shortcoming of this definition for security is that it doesn't take into account related-key attacks, which are currently a topic of active research.

The security of a cryptographic primitive F that uses a block cipher as a building element, can be proven in the *ideal-cipher model* [0]. In such a security proof, one measures how easy it is to distinguish between two idealised abstractions. On the one hand, we have Φ, the idealized abstraction of F. On the other hand, we have F_Π, the hybrid construction that is formed by taking the real construction F but replacing the block cipher building block by Π, the idealised abstraction of the block cipher. The ideal-cipher security of F against generic attacks is proven by showing that any distinguisher for F_Π leads to a distinguisher for B. One can argue that any property that is present in a larger (or smaller) fraction of the permutations of the block cipher B than in the permutations of the abstraction Π, has to be considered as a weakness, because it may be used to construct an algorithm that distinguishes F_Π from F.

Both security notions have the problem that their reference points exist only as mathematical abstractions. All the block ciphers that are currently in use, share the property that they have a compact description, because this is necessary to be implementable in hardware or software. For realistic values of n, only a negligible fraction of the permutations in Π (or primitives in F_Π) have a compact description. In the case of ideal-cipher security, this property can be used to define constructions that are secure when Π is used, but insecure when Π is replaced by any block cipher [0,0].

2.2. Related-key security

In related-key security, the adversary can submit queries, which contain as before an input x or an output y for the block cipher or the permutation. Secondly, the adversary can specify for each query a function G, which needs to be applied to the secret key k before it is used to select a permutation from the block cipher. In the oldest examples of related-key attacks, the functions G are affine functions: xor-ing the secret key with a constant chosen by the adversary [0].

The *key access scheme* (KAS) of a related-key attack defines the relations between the keys under which the adversary can query the block cipher. A KAS consists of a set of functions Γ and a set of domains α. The adversary is allowed the query the block cipher under all keys l for which there is a function G_i in Γ and a constant a_j in a domain A_j in α such that $l=G_i(k,a_j)$. Following [0], in the ideal system each choice of G_i and a_j randomly selects a permutation from Π.

Some key access schemes lead to properties in block ciphers, which may at first sight appear to be weaknesses in the examined block cipher, but on closer inspection don't lead to a distinguisher with a nonzero advantage. For example, consider the KAS containing two sets of functions $G_1(k,a_j)=k\oplus a_j$ and $G_2(k,a_j)=k+a_j$, where a_j can be any valid key. It can easily be verified that for any block cipher B, if the least significant bit of k equals 0 then for all x:

$$B_{G_1(k,1)}(x) = B_{G_2(k,1)}(x). \tag{1}$$

On the other hand, if the least significant bit of k equals 1, then for most block ciphers Eq. (1) will hold for a very low fraction of the inputs x. By repeating this set of queries for $a_j = 2, 4, 8, \ldots$, the adversary can recover the secret key k. However, this key recovery "attack" doesn't lead to a distinguisher with nonzero advantage, because the property also holds when the block cipher is replaced by Π.

A more subtle characteristic of related-key security follows from the following time/data-trade-off. If an adversary can obtain for an arbitrary input x the values $B_{k_1}(x), B_{k_2}(x), \ldots, B_{k_d}(x),$ then an exhaustive key search attack can be accelerated with a factor d. By setting $k_j = G\left(k, a_j\right)$ for $j = 1, 2, \ldots, d$, we can convert this result to the related-key scenario [0]. Hence all KAS result in an erosion of security proportional to the number of functions and constants.

2.3. The AES key schedule

For a full specification of the AES, we refer to [0, 0]. We recall here only the key schedule of AES-256. The AES-256 encryption operation uses 15 128-bit round keys, here denoted by $\mathbf{K}^i$, where i ranges from 0 to 14. The first two round keys are simple copies from the two halves of the key: $\left[\mathbf{K}^0\mathbf{K}^1\right] = \mathbf{K}$. The 13 remaining round keys are derived from the 256-bit key by the repeated application of a transformation that we denote here by φ. The transformation φ takes as input two round keys $\mathbf{K}^{2i}, \mathbf{K}^{2i+1}$ and outputs the two next round keys $\mathbf{K}^{2i+2}, \mathbf{K}^{2i+3}$. It is defined as follows:

$$\phi\left(\left[a_{ij}\right]\right) = \left[b_{ij}\right] \Leftrightarrow \begin{cases} b_{j1} = \mathrm{S}\left[b_{j+1,8}\right] \oplus a_{j1} \\ b_{j2} = b_{j1} \oplus a_{j2} \\ b_{j3} = b_{j2} \oplus a_{j3} \\ b_{j4} = b_{j3} \oplus a_{j4} \\ b_{j5} = \mathrm{S}\left[b_{j4}\right] \oplus a_{j5} \\ b_{j6} = b_{j5} \oplus a_{j6} \\ b_{j7} = b_{j6} \oplus a_{j7} \\ b_{j8} = b_{j7} \oplus a_{j8} \end{cases}, j = 1, 2, 3, 4,$$

where S denotes the nonlinear byte substitution used in SubBytes.

We denote by φ^i the map constructed by i times iterating φ. Finally, we denote by $\varphi^{0.5}$ the map that outputs the matrix constructed by taking the leftmost four columns of the output of φ and copying the rightmost four columns of the input, i.e.:

$$\varphi\left(\left[\mathbf{K}^{2i}\mathbf{K}^{2i+1}\right]\right) = \left[\mathbf{K}^{2i+2}\mathbf{K}^{2i+3}\right] \Rightarrow \phi^{0.5}\left(\left[\mathbf{K}^{2i}\mathbf{K}^{2i+1}\right]\right) = \left[\mathbf{K}^{2i+2}\mathbf{K}^{2i+1}\right].$$

3. THE BIRYUKOV-KHOVRATOVICH ATTACK

Biryukov and Khovratovich present attacks on AES-192 and on AES-256 in [0]. We briefly repeat here some key elements of the AES-256 attack, which has the lowest data complexity of the two attacks. The attack is based on an advanced variant of differential cryptanalysis [0]. In differential cryptanalysis, the adversary can query the block cipher with a number of plaintext pairs which have a difference chosen by the adversary. With a certain probability, encryption of a pair of plaintexts with a given input difference, leads to a pre-defined difference between the corresponding ciphertexts. When this happens, the adversary can determine or partially determine the secret key that was used.

The attack by Biryukov and Khovratovich uses several extensions on the basic differential attack. The first extension is the use of quartets rather than pairs. This extension is called the Boomerang attack [0]. The second extension is that the adversary can query the block cipher under different keys, which are related in a way specified by the adversary. This extension is called a related-key attack [0]. Biryukov and Khovratovich use a quartet of keys, where the relations between the keys are defined in terms of the round keys. Hence, we could call their attack a *related-round-key attack*. Let the quartet of round keys used in round i be denoted by $\left(\mathbf{K}_A^i, \mathbf{K}_B^i, \mathbf{K}_C^i, \mathbf{K}_D^i\right)$. The quartets are defined by four states $\mathbf{M}_1, \mathbf{M}_2, \mathbf{M}_3, \mathbf{M}_4,$, which are used in the following four relations between round keys:

$$\mathbf{K}_A^2 \oplus \mathbf{K}_B^2 = \mathbf{K}_C^2 \oplus \mathbf{K}_D^2 = \mathbf{M}_1 = \begin{bmatrix} 00 & 00 & 00 & 00 \\ 00 & 01 & 00 & 01 \\ 00 & 00 & 00 & 00 \\ 00 & 00 & 00 & 00 \end{bmatrix},$$

$$\mathbf{K}_A^3 \oplus \mathbf{K}_B^3 = \mathbf{K}_C^3 \oplus \mathbf{K}_D^3 = \mathbf{M}_2 = \begin{bmatrix} 3e & 00 & 3e & 00 \\ 21 & 00 & 21 & 00 \\ 1f & 00 & 1f & 00 \\ 1f & 00 & 1f & 00 \end{bmatrix},$$

$$\mathbf{K}_A^7 \oplus \mathbf{K}_C^7 = \mathbf{K}_B^7 \oplus \mathbf{K}_D^7 = \mathbf{M}_3 = \begin{bmatrix} 3e & 3e & 3e & 3e \\ 21 & 21 & 21 & 21 \\ 1f & 1f & 1f & 1f \\ 1f & 1f & 1f & 1f \end{bmatrix},$$

$$\mathbf{K}_A^8 \oplus \mathbf{K}_C^8 = \mathbf{K}_B^8 \oplus \mathbf{K}_D^8 = \mathbf{M}_4 = \begin{bmatrix} 01 & 00 & 01 & 00 \\ 00 & 00 & 00 & 00 \\ 00 & 00 & 00 & 00 \\ 00 & 00 & 00 & 00 \end{bmatrix}.$$

Using the notation φ^i, $\varphi^{0.5}$ defined in the previous section, we can describe the relations between the four keys of a quartet more compactly as follows:

$$\mathbf{K}_B = \varphi^{-1}\left(\varphi\left(\mathbf{K}_A\right) \oplus \left[\mathbf{M}_1\mathbf{M}_2\right]\right), \mathbf{K}_D = \varphi^{-1}\left(\phi\left(\mathbf{K}_C\right) \oplus \left[\mathbf{M}_1\mathbf{M}_2\right]\right),$$

$$\mathbf{K}_C = \varphi^{-3.5}\left(\varphi^{3.5}\left(\mathbf{K}_A\right) \oplus \left[\mathbf{M}_3\mathbf{M}_4\right]\right), \mathbf{K}_D = \varphi^{-3.5}\left(\varphi^{3.5}\left(\mathbf{K}_B\right) \oplus \left[\mathbf{M}_3\mathbf{M}_4\right]\right).$$

A related-key differential attack on AES can perform better than a standard differential attack on AES because the differences introduced in the key will propagate to the round keys and cancel out differences in the intermediate results of the encryption process (with a significant success probability). In the related-round-key attack by Biryukov and Khovratovich increases the success probability of a related-key attack by directly imposing the desired differences on the round keys, thereby making a part of the attack deterministic instead of probabilistic. The attack requires the en- or decryption of $2^{99.5}$ chosen plain- or ciphertexts, has an additional computational complexity of $2^{99.5}$ and requires a memory of 2^{77} 128-bit blocks.

4. DEFINING THE KEY ACCESS SCHEME

In order to determine whether the Biryukov-Khovratovich attack reduces the security of AES-256 under our definition of security, we have to check the soundness of the key access scheme. Since Biryukov and Khovratovich didn't specify a key access scheme, we will try to define here a sound key access scheme for their attack.

4.1. First attempt

As a first attempt to define a sound key access scheme, we propose the following function:

$$G(\mathbf{K}, \mathbf{A}_1, \mathbf{A}_2) = \varphi^{-1}\left(\varphi^{-2.5}\left(\varphi^{3.5}\left(\mathbf{K}\right) \oplus \mathbf{A}_2\right) \oplus \mathbf{A}_1\right),$$

where the adversary can choose freely the states $\mathbf{A}_1, \mathbf{A}_2$. Indeed, in order to construct a quartet of keys as needed for the Biryukov-Khovratovich attack, one can set $\mathbf{K}_A = \mathbf{K}$ and compute the 3 remaining keys as follows:

$$\mathbf{K}_B = G\left(\mathbf{K}_A, \left[\mathbf{M}_1\mathbf{M}_2\right], 0\right), \mathbf{K}_C = G\left(\mathbf{K}_A, 0, \left[\mathbf{M}_3\mathbf{M}_4\right]\right), \mathbf{K}_D = G\left(\mathbf{K}_A, \left[\mathbf{M}_1\mathbf{M}_2\right], \left[\mathbf{M}_3\mathbf{M}_4\right]\right)$$

4.2. Problems with the first KAS

We will now show that the KAS defined in the first attempt, is not sound. We do this by showing that with this KAS, one can recover the key of any block cipher. Let a, b be two values in GF(256) such that $\Pr\left(S\left(x \oplus a\right) = S\left(x\right) \oplus b\right) \geq 2^{-7}$. In other words, let $\left(a, b\right)$ be a differential with differential probability at least 2^{-7} over the AES S-box. Such differentials are known [0]. Define 4 states $\mathbf{D_1}, \mathbf{D}_2, \mathbf{D}_3, \mathbf{D}_4$, as follows:

$$\mathbf{D}_3 = \begin{bmatrix} 00 & 00 & 00 & 00 \\ 00 & 00 & 00 & 00 \\ 00 & 00 & 00 & 00 \\ 00 & 00 & 00 & 00 \end{bmatrix}, \mathbf{D}_2 = \begin{bmatrix} 00 & 00 & 00 & 00 \\ 00 & a & a & 00 \\ 00 & 00 & 00 & 00 \\ 00 & 00 & 00 & 00 \end{bmatrix}, \mathbf{D}_3 = \begin{bmatrix} b & b & b & b \\ 00 & 00 & 00 & 00 \\ 00 & 00 & 00 & 00 \\ 00 & 00 & 00 & 00 \end{bmatrix}, \mathbf{D}_4 = \begin{bmatrix} 00 & 00 & 00 & 00 \\ 00 & a & a & a \\ 00 & 00 & 00 & 00 \\ 00 & 00 & 00 & 00 \end{bmatrix},$$

Straightforward computation shows that $\left(\left[\mathbf{D}_1\mathbf{D}_2\right], \left[\mathbf{D}_3\mathbf{D}_4\right]\right)$ is a differential over $\varphi^{2.5}$ with differential probability at least 2^{-7}, since only one S-box gets a non-zero input difference. It follows that for all keys $\mathbf{K}$, the relation

$$G\left(\mathbf{K}, \mathbf{X}, \mathbf{Y}\right) = G\left(\mathbf{K}, \mathbf{X} \oplus \left[\mathbf{D}_1\mathbf{D}_2\right], \mathbf{Y} \oplus \left[\mathbf{D}_3\mathbf{D}_4\right]\right)$$

holds for a fraction of 2^{-7} of the states $\mathbf{X}, \mathbf{Y}$. Exploiting the fact that there are many good choices for $\left(a, b\right)$, 25 applications of the key access scheme suffice to create with high probability a collision for the key access scheme. As shown before, a collision in the key access scheme is trivially detectable by the adversary, allowing it to recover one byte of the round key. The attack can be repeated 19 times with slightly different choices for the states $\mathbf{D_1}, \mathbf{D}_2, \mathbf{D}_3, \mathbf{D}_4$, allowing the adversary to recover in total 160 bits of key material using only 500 queries.

Observe that the attack doesn't depend on the underlying encryption block cipher. It only depends on the properties of the key access scheme. The adversary described in this section, using the key access scheme defined in the previous section, can recover the key of *any* block cipher. Hence, this key access scheme is not sound.

4.3. Second attempt

In order to make the key access scheme sound, we have to modify it as follows. We forbid the use of values $\mathbf{A}_1, \mathbf{A}_2$ resulting in a differential over $\varphi^{2.5}$ with nonzero probability, while at the same time not excluding the values that are needed for the Biryukov-Khovratovich attack. Although this restriction restores the soundness of the KAS, it is specially designed for this attack to work and can be called contrived.

5. CONCLUSION

Due to the complex relations required between the related keys, the Biryukov-Khovratovich attacks is generally considered to be of academic rather than practical interest. In this paper, we show that even the academic interest is very limited due to the fact that any sound key access scheme for which it works is contrived.

The actual relevance of related-key attacks could well be that they can be a step up to conventional cryptanalysis because they provide insight in the limitations of key schedules and their interaction with the data path. As such, it is very likely that the insights of the Biryukov-Khovratovich attack have contributed to the more recent biclique attacks [0].

As mentioned in the paper, the usage of key access schemes in protocols leads to a degradation of security and avoiding them does not necessarily lead to less efficient protocols. One may wonder whether security against related-key attacks should be a design requirement. One may design block ciphers targeting PRP security only. This relaxation of requirements may lead to lighter key schedule and higher key agility.

REFERENCES

1. MIHIR BELLARE, TADAYOSHI KOHNO, *A Theoretical Treatment of Related-Key Attacks: RKA-PRPs, RKA-PRFs, and Applications*, Lecture Notes in Computer Science, **2656,** pp. 491–506 2003.
2. GUIDO BERTONI, VITTORIO ZACCARIA, LUCA BREVEGLIERI, MATTEO MONCHIERO, GIANLUCA PALERMO, *AES Power Attack Based on Induced Cache Miss and Countermeasure*, Proceedings ITCC'05, 2005, pp. 586–591.
3. ELI BIHAM, *How to Forge DES-Encrypted Messages in 2^{28} Steps*, Technion – Computer Science Department – Technical Report CS0884, 1996.
4. ELI BIHAM, ADI SHAMIR, *Differential cryptanalysis of DES-like cryptosystems,* Journal of Cryptology, **4**, *1*, pp. 3–72, 1991.
5. JOHN BLACK., *The ideal-cipher model, revisited: an uninstantiable blockcipher-based hash function,* Lecture Notes in Computer Science, **4047**, pp. 328–340, 2006.
6. ANDREY BOGDANOV, DMITRY KHOVRATOVICH, CHRISTIAN RECHBERGER, *Biclique cryptanalysis of the full AES,* Lecture Notes in Computer Science, **7073**, pp. 344–371, 2011.
7. ALEX BIRYUKOV, DMITRY KHOVRATOVICH, *Related-key cryptanalysis of the full AES-192 and AES-256,* Lecture Notes in Computer Science, **6477**, pp. 1–19, 2009
8. ALEX BIRYUKOV, DMITRY KHOVRATOVICH, IVICA NIKOLIC, *Distinguisher and related-key attack on the full AES-256,* Lecture Notes in Computer Science, **5677**, pp. 231–249, 2009.
9. RAN CANETTI, ODED GOLDREICH, SHAI HALEVI, *The random oracle methodology revisited,* Proceedings of the 30th ACM symposium on the theory of computation, 1998, pp. 209–218.
10. JOAN DAEMEN, VINCENT RIJMEN., *The design of Rijndael: AES – the Advanced Encryption Standard,* Springer-Verlag, 2002.
11. *FIPS 197: Specification for the Advanced Encryption Standard (AES),* NIST, 2001.
12. JOANNE FULLER, WILLIAM MILLAN*, Linear Redundancy in S-Boxes*, Lecture Notes in Computer Science, **2887**, pp. 74–86, 2003.
13. JOHN KELSEY, BRUCE SCHNEIER, DAVID WAGNER, *Key-Schedule Cryptoanalysis of IDEA, G-DES, GOST, SAFER and Triple-DES*, Lecture Notes in Computer Science, **1109**, pp. 237–251, 1996.
14. JIQIANG LU, ORR DUNKELMAN, NATHAN KELLER, JONGSUNG KIM, *New Impossible Differential Attacks on AES*, Lecture Notes in Computer Science, **5365**, pp. 279–293, 2008.
15. MICHAEL LUBY, CHARLES RACKOFF, *How to Construct Pseudorandom Permutations from Pseudorandom Functions*, SIAM Journal of Computing, **17**, pp. 373–386, 1988.
16. CLAUDE SHANNON, *Communication theory of secrecy systems*, Bell Systems Technical Journal, **4**, pp. 656–715, 1949.
17. DAVID WAGNER, *The Boomerang attack*, Lecture Notes in Computer Science**, 1636**, pp. 156–170, 1999.

Received June 26, 2012