Key Recovery Attacks of Practical Complexity on AES Variants

Alex Biryukov, Orr Dunkelman, Nathan Keller,
Dmitry Khovratovich, Adi Shamir

Département d'Informatique
École Normale Supérieure

France Telecom Chaire

17 September 2009

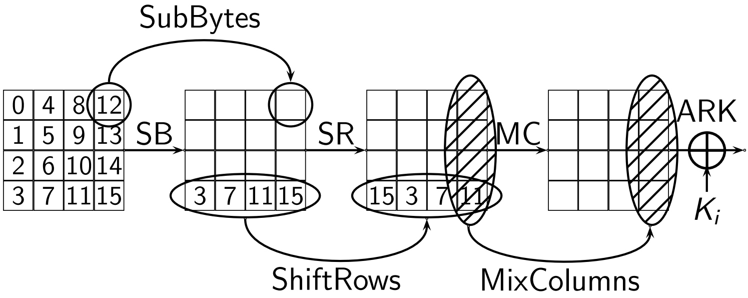
Outline

Outline

The Advanced Encryption Standard

- Designed by Vincent Rijmen and Joan Daemen, under the name Rijndael and submitted to NIST's competition in 1998.
- Selected after a three year competition as the new standard.
- The cipher has an SP network structure.
- Block size — 128 bits, Key size — 128, 192, or 256 bits.
- Number of rounds depends on the key length (10/12/14, respectively).

The Advanced Encryption Standard



AES' Key Schedule Algorithm

AES has three key schedules. One for each key size.

- AES-128 ($Nk = 4$) and AES-192 ($Nk = 6$):
  1. Initialize $W[0, \ldots, Nk - 1]$ with the user supplied key.
  2. For $i = Nk, \ldots, 43/51$ do
     - If $i \equiv 0 \bmod Nk$ then
       $W[i] = W[i - Nk] \oplus SB(W[i-1] \lll 8) \oplus RCON[i/Nk]$,
     - Otherwise $W[i] = W[i - 1] \oplus W[i - Nk]$,
- AES-256 ($Nk = 8$):
  1. Initialize $W[0, \ldots, 7]$ with the user supplied key.
  2. For $i = 8, \ldots, 59$ do
     - If $i \equiv 0 \bmod Nk$ then
       $W[i] = W[i - Nk] \oplus SB(W[i-1] \lll 8) \oplus RCON[i/Nk]$,
     - Else if $i \equiv 4 \bmod Nk$ then
       $W[i] = W[i - 8] \oplus SB(W[i - 1])$,
     - Otherwise $W[i] = W[i - 1] \oplus W[i - Nk]$,

Security Properties

- The S-boxes are based on inversion over $GF(2^8)$.
- The MixColumns operation is an MDS matrix, which along with the ShiftRows operation ensures full diffusion after two rounds.
- The “wide trail strategy” assures that the number of active S-boxes in any differential characteristic is at least five for two rounds, nine for three rounds, and 21 for four rounds.
- There structure offers some 4-round impossible differentials, and several sets of 4-round Square properties.

Differential/Linear Cryptanalysis

- The security against these attacks is derived from the fact that there are no good differentials (linear hulls) of high probability.

Differential/Linear Cryptanalysis

- The security against these attacks is derived from the fact that there are no good differentials (linear hulls) of high probability.
- In a series of papers, the maximal expected differential and linear probabilities for two and four rounds were computed.
- The results are that 4-round AES have no differentials or linear hulls with high enough probability for attacks.

Outline

History of Cryptanalytic Papers

- Early 80's — some experiments of flipping bits in the plaintext/key. Statistical tests (on small data sets). Attacks are verified.

History of Cryptanalytic Papers

- Early 80's — some experiments of flipping bits in the plaintext/key. Statistical tests (on small data sets). Attacks are verified.
- Mid 80's — attacks which uses up to 1000 known plaintexts. Time needs to be less than exhaustive search.

History of Cryptanalytic Papers

- Early 80's — some experiments of flipping bits in the plaintext/key. Statistical tests (on small data sets). Attacks are verified.
- Mid 80's — attacks which uses up to 1000 known plaintexts. Time needs to be less than exhaustive search.
- 1990/1 — differential cryptanalysis. Chosen plaintexts, and $2^{47.2}$ data and time!

History of Cryptanalytic Papers

- Early 80's — some experiments of flipping bits in the plaintext/key. Statistical tests (on small data sets). Attacks are verified.
- Mid 80's — attacks which uses up to 1000 known plaintexts. Time needs to be less than exhaustive search.
- 1990/1 — differential cryptanalysis. Chosen plaintexts, and $2^{47.2}$ data and time!
- 1992/3 — related-key attacks (known/chosen key relations).

History of Cryptanalytic Papers

- Early 80's — some experiments of flipping bits in the plaintext/key. Statistical tests (on small data sets). Attacks are verified.
- Mid 80's — attacks which uses up to 1000 known plaintexts. Time needs to be less than exhaustive search.
- 1990/1 — differential cryptanalysis. Chosen plaintexts, and $2^{47.2}$ data and time!
- 1992/3 — related-key attacks (known/chosen key relations).
- 1997 — AES competition. One strike and your out!

History of Cryptanalytic Papers

- Early 80's — some experiments of flipping bits in the plaintext/key. Statistical tests (on small data sets). Attacks are verified.
- Mid 80's — attacks which uses up to 1000 known plaintexts. Time needs to be less than exhaustive search.
- 1990/1 — differential cryptanalysis. Chosen plaintexts, and $2^{47.2}$ data and time!
- 1992/3 — related-key attacks (known/chosen key relations).
- 1997 — AES competition. One strike and your out!
- 1999 — Adaptive chosen plaintext and ciphertext attacks (boomerang attacks).

Current State of Affairs in Cryptanalysis

Time complexity of a related-key attack:

"Thus, the total time complexity of Step 2(b) is about $2^{256} \cdot 2^{167.0} = 2^{423.0}$ *SHACAL-1 encryptions."*

- Most cryptanalytic papers discuss certificational attacks:
  - Data complexity — just slightly less than the entire code book.
  - Time complexity — just slightly less than exhaustive search.
  - Memory — store more information than there are particles in the universe

Current State of Affairs in Cryptanalysis (cont.)

- These certificational attacks are of great importance:
  1. Why to use a primitive whose less secure than optimal?

Current State of Affairs in Cryptanalysis (cont.)

- These certificational attacks are of great importance:
  1. Why to use a primitive whose less secure than optimal?
  2. By publishing the first step of analysis, others may be able to improve the attacks.
  3. Attacks only get better!

Current State of Affairs in Cryptanalysis (cont.)

- These certificational attacks are of great importance:
  1. Why to use a primitive whose less secure than optimal?
  2. By publishing the first step of analysis, others may be able to improve the attacks.
  3. Attacks only get better!
- But they do not help answering questions by users:
  1. Does this attack affect my system?
  2. Should I still use AES-256 for encryption?
  3. MD5 is still OK for certificates, right?

What a Break is?

- There is an ongoing debate what a broken scheme is.

What a Break is?

- There is an ongoing debate what a broken scheme is. Even from the theoretical point of view.
- The extreme approach: max(Time, Data, Memory) less than Exhaustive search' time.
- Another approach: (Time, Data, Memory) better then generic attacks (time-memory-data tradeoff attacks).
- Time $\times$ Memory $<$ Exhaustive search.
- Money for finding a key in a given time $<$ for a generic attack.

What is a Practical Attack?

- We upper-bound the complexities of the attack.

What is a Practical Attack?

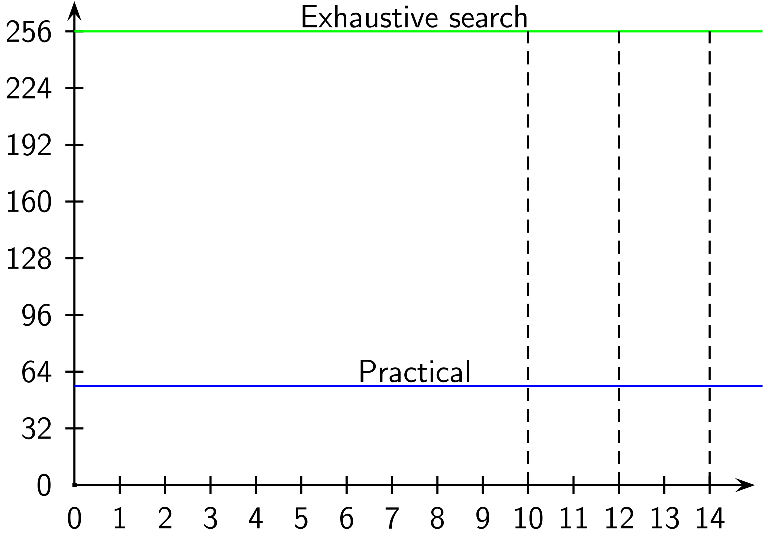
- We upper-bound the complexities of the attack.
- 2^{55} DES encryptions are feasible ...
- 2^{61} SHA-1 evaluations did not complete ...

What is a Practical Attack?

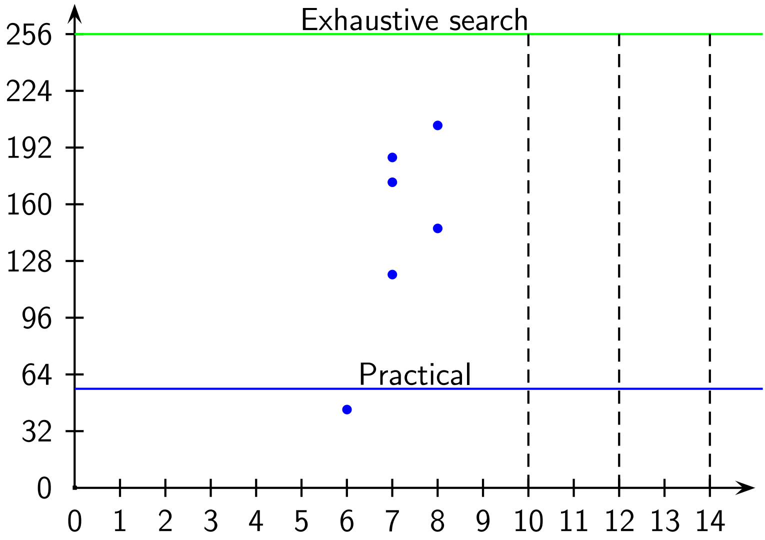
- We upper-bound the complexities of the attack.
- 2^{55} DES encryptions are feasible ...
- 2^{61} SHA-1 evaluations did not complete ...
- So, let's take 2^{64} cycles
    - which are about 2^{56} AES encryptions.
- This is also a restriction on the data complexity.

Outline

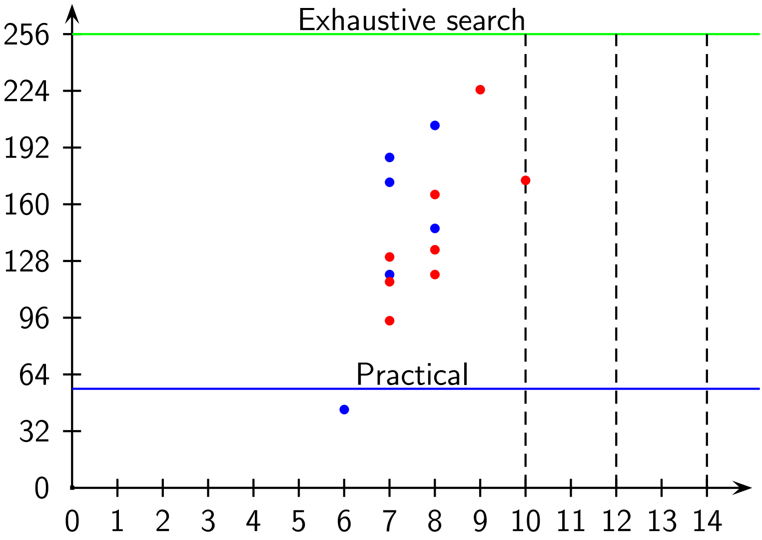
Time Complexity of Attacks on AES-256



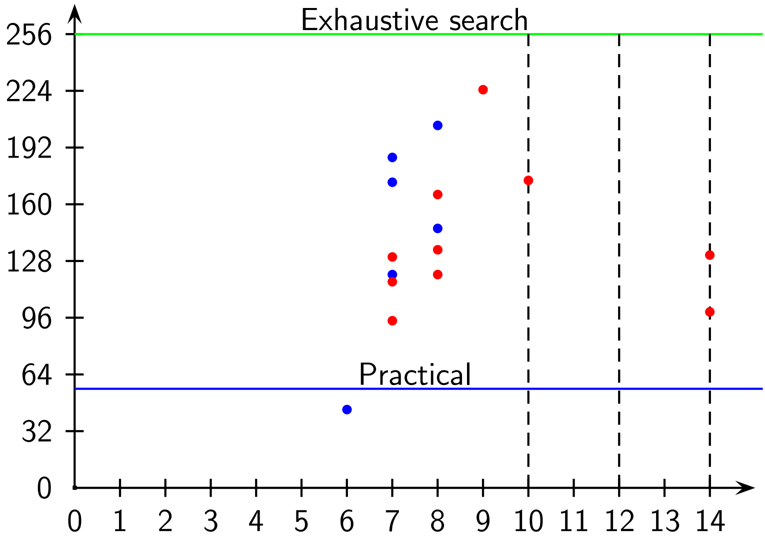
Time Complexity of Attacks on AES-256



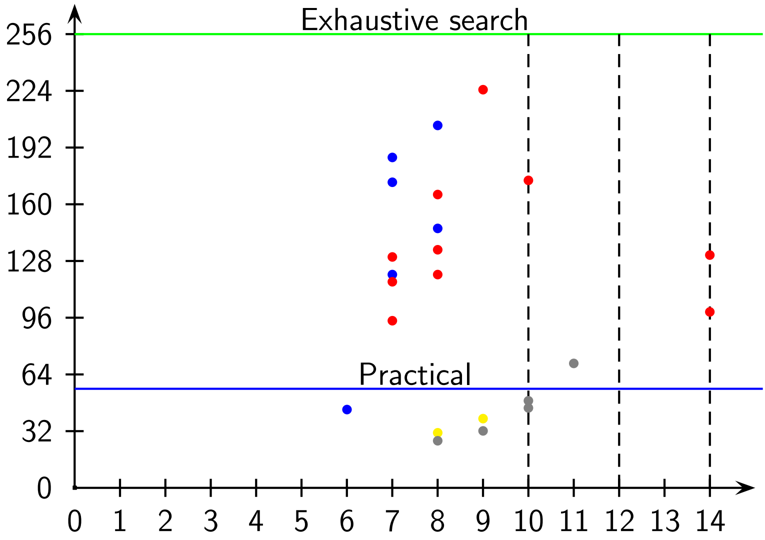
Time Complexity of Attacks on AES-256



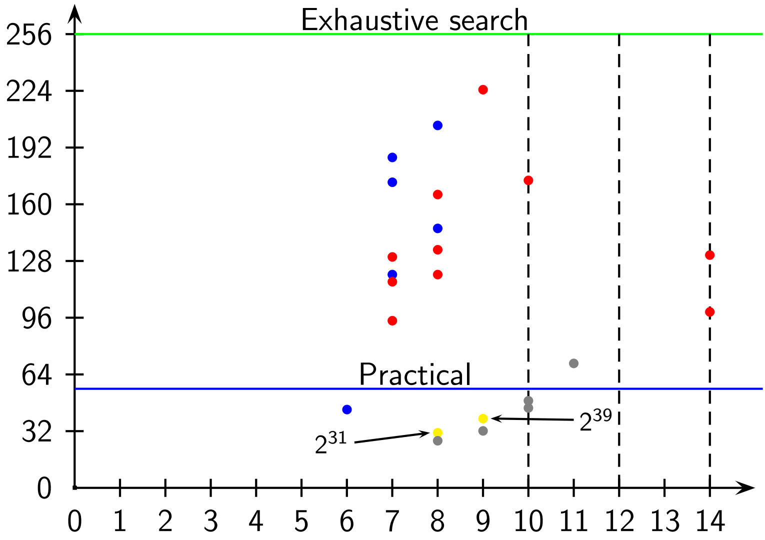
Time Complexity of Attacks on AES-256



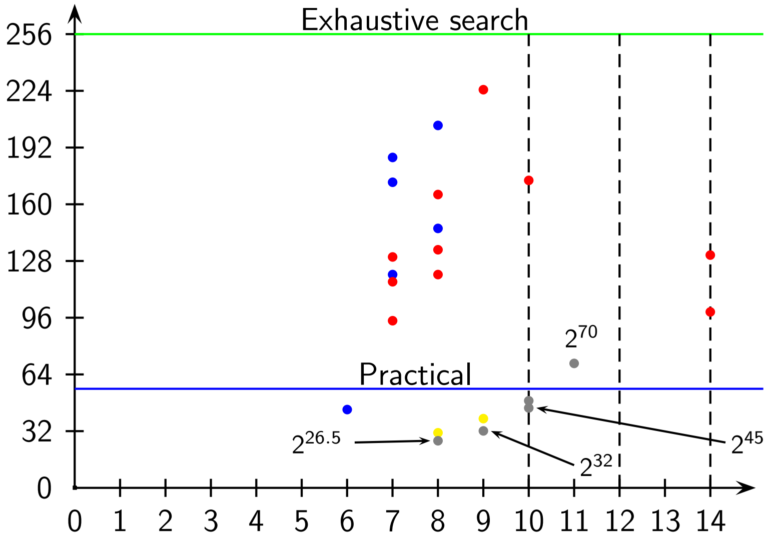
Time Complexity of Attacks on AES-256



Time Complexity of Attacks on AES-256



Time Complexity of Attacks on AES-256



The Related-Key Model

- First introduced by Knudsen and Biham, independently.
- The adversary is assumed to have some knowledge on the relation.
- In 1996/7, the concept of related-key differentials was introduced, along with it, the concept where the adversary is allowed to chose the key relation.
- There are "good relations" (XORs, rotations, or additions), and "bad relations" (AND, ORs, XORs + additions together).

The Related-Key Model

- First introduced by Knudsen and Biham, independently.
- The adversary is assumed to have some knowledge on the relation.
- In 1996/7, the concept of related-key differentials was introduced, along with it, the concept where the adversary is allowed to chose the key relation.
- There are "good relations" (XORs, rotations, or additions), and "bad relations" (AND, ORs, XORs + additions together).
- At the end, the main issue is applicability — does the attack scenario allows this relation or not.

Example: Related-Key Differentials

- The probability of a regular differential is:

$$\Pr_{P,K}[E_K(P) \oplus E_K(P \oplus \Delta P) = \Delta C]$$

- The probability of a related-key differential is:

$$\Pr_{P,K}[E_K(P) \oplus E_{K \oplus \Delta K}(P \oplus \Delta P) = \Delta C]$$

- The key difference leads to subkey differences, that may be used to cancel the differences in the input to the round function.

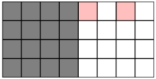
The Related-Subkey Model

- This new model was recently introduced in [BK09].
- In related-key attacks, a simple relation R is used for the keys K_1, K_2.
- In related-subkey attacks, R is a simple relation between two subkeys, RK_1, RK_2.
- The two subkeys are then handled by the key schedule algorithm to obtain the actual keys.
- This slightly less intuitive approach (and less practical one) can be "covered" by the theoretical treatment by just expanding the set of "good relations".

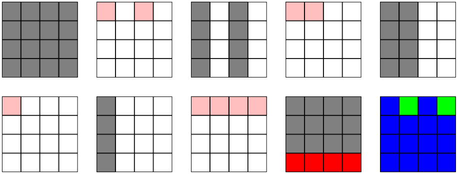
Outline

An Interesting Property of the Key Schedule Algorithm of AES-256

Our results are based on the fact that key difference

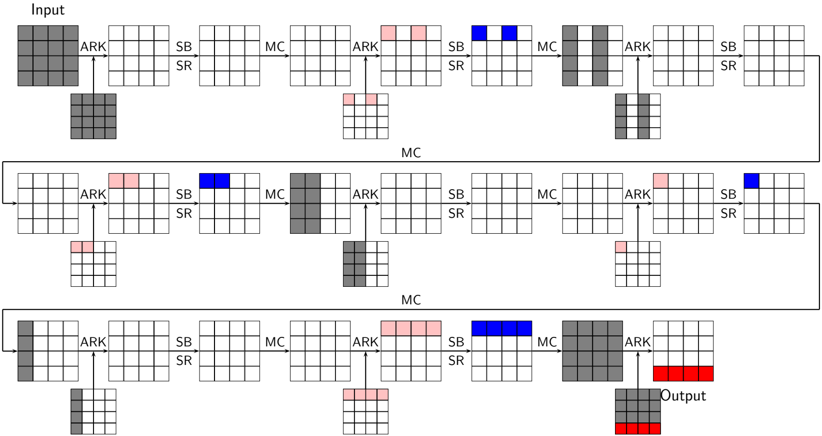


leads to the 10 subkey differences



With probability 1!

An 8-Round Related-Key Differential



The probability is 2^{-56}. It can be transformed into a truncated one predicting 24 bits of difference with probability 2^{-36}.

Verification of the Differential

- We have verified experimentally the correctness of the 7-round related-key differential derived from the 8-round one (it has probability 2^{-30}).
- We preformed 100 experiments, each with a random key and 2^{32} random plaintext pairs.

Pairs	0	1	2	3	4	5	6
Theory	1.8	7.3	14.7	19.5	19.5	15.6	10.4
Experiment	0	10	18	10	28	18	6

Pairs	7	8	9	10	11	12
Theory	6.0	3.0	1.3	0.5	0.2	0.06
Experiment	8	1	0	0	0	1

A 10-Round Related-Subkey Differential

- In the related-subkey model, it is possible to pick two keys which satisfy the difference in a slightly different manner.
- The related-subkey allows for shifting the differential by one round.
- This allows an extension of the differential in the backwards direction (despite having a highly active state).
- Which in turn, allows for attacks of practicaly complexity of up to 10 rounds, and semi-practical of up to 11 rounds.

Other Attack Scenarios

- The attacks work when the plaintexts are generated not randomly as well.
- For example, when counter mode is used. The encryption system is initialized to two initial states and are allowed to generate data sequentially. This simplifies the attack model.
- The attacks are applicable when the plaintexts are ASCII characters (as some key differences are suitable).
- Or even when they are ASCII characters representing only numeric values.
- The minimal hamming weight of the key difference is 24.

Outline

Summary of the Attacks

Rounds	Scenario	Time	Data	Memory	Result
8	Key Diff. – CP	2^{31}	2^{31}	2	Distinguisher
8	Subkey Diff. – CC	$2^{26.5}$	$2^{26.5}$	$2^{26.5}$	35 subkey bits
9	Key Diff. – CP	2^{39}	2^{38}	2^{32}	Full key
9	Subkey Diff. – CC	2^{32}	2^{32}	2^{32}	56 key bits
10	Subkey Diff. – CP	2^{49}	2^{48}	2^{33}	Distinguisher
10	Subkey Diff. – CC	2^{45}	2^{44}	2^{33}	35 subkey bits
11	Subkey Diff. – CP	2^{70}	2^{70}	2^{33}	50 subkey bits

Security Implications

- Extending AES-128 key to 256 bits actually reduces security!
- The security margins of AES-256 are smaller than expected.
- The related-subkey model — many new results awaiting.

Security Implications

- Extending AES-128 key to 256 bits actually reduces security!
- The security margins of AES-256 are smaller than expected.
- The related-subkey model — many new results awaiting.
- This is a good time to check that Serpent-support...

Conclusions

- Did we break the full AES with practical complexity?

Conclusions

- Did we break the full AES with practical complexity?

Conclusions

- Did we break the full AES with practical complexity?
- Should users be worried?

Conclusions

- Did we break the full AES with practical complexity?
- Should users be worried?

Questions?

Thank you for your attention!

The paper is available on ePrint (2009/374)