Structural Evaluation of AES and Chosen-Key Distinguisher of 9-round AES-128

Pierre-Alain Fouque\textsuperscript{1} \hspace{1cm} Jérémy Jean\textsuperscript{2} \hspace{1cm} Thomas Peyrin\textsuperscript{3}

\textsuperscript{1}Université de Rennes 1, France
\textsuperscript{2}École Normale Supérieure, France
\textsuperscript{3}Nanyang Technological University, Singapore

CRYPTO’2013 – August 19, 2013
Block Ciphers

Iterated SPN Block Ciphers

- **Internal Permutation**: $f$
- **Number of Iterations**: $r$
- **SPN**: $f = P \circ S$ applies Substitution (S) and Permutation (P) layers.
- **Secret Key**: $k$
- **Key Scheduling Algorithm**: $k \rightarrow (k_0, \ldots, k_r)$
- **Ex**: AES, PRESENT, SQUARE, Serpent, etc.

![Key Scheduling Algorithm Diagram]
Differentials and Differential Characteristics

Differential Characteristics

- Used in differential cryptanalysis
- Sequence of differences at each round for an iterated primitive
- The success probability of a differential attack depends on the differential with maximal differential probability $p$.

Example: 4-round AES

- 4-round characteristic with 25 active S-Boxes (minimal).
- AES S-Box: $p_{\text{max}} = 2^{-6}$.
- Differential probability: $p \leq 2^{-6 \times 25} = 2^{-150}$. 
Design of the AES

- AES Permutation: structurally bounded diffusion for any rounds
- Provably resistant to non-RK differential attacks
- Ad-hoc key schedule

\[ \implies \text{RK Attacks} \ [BKN-C09], \ [BK-A09], \ [BN-E10]. \]

Minimal Number of Active S-Boxes for AES

<table>
<thead>
<tr>
<th>Rounds</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>min</td>
<td>1</td>
<td>5</td>
<td>9</td>
<td>25</td>
<td>26</td>
<td>30</td>
<td>34</td>
<td>50</td>
<td>51</td>
<td>55</td>
</tr>
</tbody>
</table>

Question: Similar numbers for AES structure in the RK model?
Our Contributions

- We propose an algorithm finding all the “smallest” RK characteristics.
- It improves previous works: runs in time linear in the number of rounds.
- We focus on AES-128.
- We provide a distinguisher for 9-round AES-128.
Existing Algorithms (1/2)

Matsui’s Algorithm (e.g., for DES)

- **Works by induction**: derive best \( n \)-round char. from best chars. on \( 1, \ldots, n - 1 \) rounds
- Compute best char. for 1R
- Traverse a **tree** of depth 2 for 2R
- Pruning possible (A* optim.)

Tree Example

\[ p^j_i \overset{\text{def}}{=} P(\Delta_i \rightarrow \Delta_j) \]

\[ \Delta_1 \]
Existing Algorithms (1/2)

Matsui’s Algorithm (e.g., for DES)

- Works by **induction**: derive best $n$-round char. from best chars. on $1, \ldots, n-1$ rounds
- Compute best char. for 1R
- Traverse a **tree** of depth 2 for 2R
- Pruning possible ($A^*$ optim.)

**Tree Example**

$P_i \overset{\text{def}}{=} \mathbb{P}(\Delta_i \rightarrow \Delta_j)$
Existing Algorithms (1/2)

Matsui’s Algorithm (e.g., for DES)

- Works by **induction**: derive best $n$-round char. from best chars. on $1, \ldots, n-1$ rounds
- Compute best char. for 1R
- Traverse a **tree** of depth 2 for 2R
- Pruning possible ($A^*$ optim.)

Tree Example

\[
p^j_i \overset{\text{def}}{=} \mathbb{P}(\Delta_i \rightarrow \Delta_j)
\]
Existing Algorithms (1/2)

Matsui’s Algorithm (e.g., for DES)

- Works by **induction**: derive best $n$-round char. from best chars. on $1, \ldots, n-1$ rounds
- Compute best char. for 1R
- Traverse a **tree** of depth 2 for 2R
- Pruning possible ($A^*$ optim.)

Tree Example

\[
p_i^j \overset{\text{def}}{=} \mathbb{P}(\Delta_i \rightarrow \Delta_j)
\]

\[
\begin{align*}
\Delta_1 & \quad p_1^1 \rightarrow \Delta_3 \quad p_3^3 \rightarrow \Delta_7 \\
\Delta_2 & \quad p_2^2 \rightarrow \Delta_4 \\
\Delta_3 & \quad p_3^4 \rightarrow \Delta_1 \\
\Delta_4 & \\
\Delta_6 & \\
\Delta_7 &
\end{align*}
\]
Existing Algorithms (1/2)

Matsui’s Algorithm (e.g., for DES)

- Works by **induction**: derive best \( n \)-round char. from best chars. on \( 1, \ldots, n-1 \) rounds
- Compute best char. for 1R
- Traverse a **tree** of depth 2 for 2R
- Pruning possible (A* optim.)

Tree Example

\[ p_i^j \overset{\text{def}}{=} \mathbb{P}(\Delta_i \rightarrow \Delta_j) \]
Existing Algorithms (1/2)

Matsui’s Algorithm (e.g., for DES)

- Works by **induction**: derive best \( n \)-round char. from best chars. on 1, \ldots, \( n-1 \) rounds
- Compute best char. for 1R
- Traverse a tree of depth 2 for 2R
- Pruning possible (A* optim.)

Tree Example

\[ p_i^j \overset{\text{def}}{=} \Pr(\Delta_i \rightarrow \Delta_j) \]
Existing Algorithms (1/2)

**Matsui’s Algorithm (e.g., for DES)**
- Works by *induction*: derive best $n$-round char. from best chars. on $1, \ldots, n-1$ rounds
- Compute best char. for 1R
- Traverse a tree of depth 2 for 2R
- Pruning possible ($A^*$ optim.)

**Pros**
- Very efficient on DES

**Drawbacks**
- Rely on non-equivalent differential probabilities
- Need for dominant characteristic(s)
- Poor performances for AES
- Differences visited several times

---

Tree Example

\[ p_i \stackrel{\text{def}}{=} \mathbb{P}(\Delta_i \rightarrow \Delta_j) \]

![Tree Diagram](image-url)
Existing Algorithms (2/2)

Biryukov-Nikolic [BN-E10]
- Adapt Matsui’s algorithm
- Different algos for several KS

Pros
- No need for a predominant char.
- Switch to truncated differences \(\Rightarrow\) less edges
- Representation of trunc. differences \(\Rightarrow\) handle branching in the KS
- Work on AES

Cons
- Differences visited several times
- Nodes visited exponential in the number of rounds

Tree Example
\[
p^i_j \overset{\text{def}}{=} P(\Delta_i \rightarrow \Delta_j)
\]

\[
\begin{align*}
\Delta_1 & \quad p^4_1 \quad \Delta_4 \\
\Delta_2 & \quad p^6_2 \quad \Delta_6 \\
\Delta_3 & \quad p^7_3 \quad \Delta_7 \\
\Delta_4 & \quad p^4_4 \\
\Delta_5 & \quad p^9_5 \\
\Delta_6 & \quad p^1_6 \\
\Delta_7 & \quad p^1_7 \\
\Delta_8 & \quad p^4_8 \\
\Delta_9 & \quad p^1_9 \\
\end{align*}
\]
Our Algorithm

- Switch to a graph representation

Algorithm

Graph Example

\[ \Delta_1 \rightarrow \Delta_2 \rightarrow \Delta_4^* \rightarrow \Delta_6 \rightarrow \Delta_3 \rightarrow \Delta_4 \rightarrow \Delta_7 \rightarrow \Delta_8 \rightarrow \Delta_5 \rightarrow \Delta_9 \]
Our Algorithm

Algorithm

- Switch to a graph representation
- Merge equal diff. of the same round

Graph Example

\[ \Delta_1 \rightarrow \Delta_2 \rightarrow \Delta_4 \star \rightarrow \Delta_6 \rightarrow \Delta_3 \rightarrow \Delta_7 \rightarrow \Delta_4 \star \rightarrow \Delta_8 \rightarrow \Delta_9 \]
Our Algorithm

- Switch to a graph representation
- Merge equal diff. of the same round
- Graph traversal similar as Dijkstra
- Dynamic programming approach

Graph Example

\[ \Delta_1 \rightarrow \Delta_2 \rightarrow \Delta_4^* \]
\[ \Delta_3 \rightarrow \Delta_4^* \]
\[ \Delta_4 \rightarrow \Delta_6 \rightarrow \Delta_1^* \rightarrow \Delta_7 \rightarrow \Delta_8 \rightarrow \Delta_5 \rightarrow \Delta_9 \]
Our Algorithm

Algorithm

- Switch to a graph representation
- Merge equal diff. of the same round
- Graph traversal similar as Dijkstra
- Dynamic programming approach

Pros

- Path search seen as Markov process
- Each difference in each round is visited only once
- Numbers of nodes and edges are linear in the number of rounds
- $A^*$ optimization still applies

Notes

- Only partial information propagated
- Need to adapt the Markov process
Different Levels of Analysis

Truncated Differences

- Basic Markov process
- Apply to any SPN cipher: we focus on AES-like ciphers
- Provide a structural evaluation of the cipher in regard to RK attacks
- For AES, similar results as the seminal work [DR-02] (for non-RK)

Actual Differences

- Enhanced Markov process:
  - More complete representation of differences
  - Add information for local system resolutions
- Need to be adapted to a particular cipher
- For AES, recover all the truncated results from [BN-E10]
- Full instantiation of characteristics while maximizing its probability
- Running time linear in the number of rounds

In reality: Mixing the two concepts
Application to the Structure of \textbf{AES–128}

### Structural Analysis

- We ignore the semantic definition of the S-Box and the MDS matrix
- We count the number of active S-Boxes (truncated differences)
- Do not apply to \texttt{AES–128} with the instantiated S and P
- Give an estimation of the structural quality of the \texttt{AES} family

### Related-Key Model (XOR difference of the keys)

<table>
<thead>
<tr>
<th>Rounds</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>min</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>9</td>
<td>11</td>
<td>13</td>
<td>15</td>
<td>21</td>
<td>23</td>
<td>25</td>
</tr>
</tbody>
</table>
There exists a characteristic on 10 rounds with only 25 active S-Boxes \( \implies \) best RK differential attack in \( p_{\text{max}}^{-25} \) computations.

**Result 1**

It is impossible to prove the security of the full AES-128 against related-key differential attacks without considering the differential property of the S-Box.

**Notes**

- With a random S-Box, \( p_{\text{max}}^{-25} \) might be smaller than \( 2^{128} \) \( \implies \) when \( p_{\text{max}} \geq 2^{-5} \)
- AES structure on its own not enough for RK security
- For a specified S-Box with bounded \( p_{\text{max}} \leq 2^{-6} \) \( \implies \) security against RK attacks
Impossibility Results for the Structure of AES-128 (2/2)

There exists a characteristic on 8 rounds with only 21 active S-Boxes
\[ \implies \text{best RK differential attack in } p_{\text{max}}^{-21} \text{ computations.} \]

Result 2

It is impossible to prove the security of 8-round AES-128 against related-key differential attacks without considering both the differential property of the S-Box and the P layer.

Notes

- With a random S-Box, same reason as before
- For a specified S-Box with bounded \( p_{\text{max}} \leq 2^{-6} \):
  - Best attack might be \( 2^{6\times 21} = 2^{126} \leq 2^{128} \)
  - For AES, we have exhausted all the possible attacks, no valid one
  - P layer and KS introduce linear dependencies in the characteristic
  - P can be chosen such that there is/isn’t solutions
Related-Key attacks on AES-128

RK attacks against AES-128

- After 6 rounds, there is no RK characteristic for AES-128 with a probability greater than $2^{-128}$.
- For 1, ..., 5 rounds, our algorithm has found the best characteristics.
- Same truncated characteristics as [BN-E10].
- Best instantiations of differences: maximal probabilities.

Best RK attacks on AES-128

<table>
<thead>
<tr>
<th>Rounds</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>#S-Boxes</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>13</td>
<td>17</td>
</tr>
<tr>
<td>[BN-E10]</td>
<td>0</td>
<td>-6</td>
<td>-30</td>
<td>-78</td>
<td>-102</td>
</tr>
<tr>
<td>max $\log_2(p)$</td>
<td>0</td>
<td>-6</td>
<td>-31</td>
<td>-81</td>
<td>-105</td>
</tr>
</tbody>
</table>
Distinguishing model [KR-A07, BKN-C09]

Solve Open-Problem

We can use the best 5-round characteristic to construct a chosen-key distinguisher for 9-round AES-128.

Let $E_k$ be the 9-round AES-128 block cipher using key $k$.

Limited Birthday Problem [GP-FSE10]

Given

- a fully instantiated difference $\delta$ in the key,
- a partially instantiated difference $\Delta_{IN}$ in the plaintext,
- a partially instantiated difference $\Delta_{OUT}$ in the ciphertext,

find

- a key $k$,
- a pair of messages $\langle m, m' \rangle$,

such that:

\[
m \oplus m' \in \Delta_{IN}
\]

and:

\[
E_k(m) \oplus E_{k\oplus\delta}(m') \in \Delta_{OUT}.
\]
Construction of the characteristic

Take the best 5-round characteristic for AES-128 we have found.
9-Round characteristic for AES-128

Construction of the characteristic
Prepend three rounds to be controlled by the SuperSBox technique.

Controlled by SuperSBox
9-Round characteristic for AES-128

Construction of the characteristic
Prepend one other round, as inactive as possible.

Controlled by SuperSBox

\[ \delta \]

\[ \Delta_{IN} \]

\[ AK0 \]
\[ KS \]
\[ SB \]
\[ SR \]
\[ MC \]

\[ AK1 \]
\[ KS \]
\[ SB \]
\[ SR \]
\[ MC \]

\[ AK2 \]
\[ KS \]
\[ SB \]
\[ SR \]
\[ MC \]

\[ AK3 \]
\[ KS \]
\[ SB \]
\[ SR \]
\[ MC \]

\[ AK4 \]
\[ KS \]
\[ SB \]
\[ SR \]
\[ MC \]

\[ AK5 \]
\[ KS \]
\[ SB \]
\[ SR \]
\[ MC \]

\[ AK6 \]
\[ KS \]
\[ SB \]
\[ SR \]
\[ MC \]

\[ AK7 \]
\[ KS \]
\[ SB \]
\[ SR \]
\[ MC \]

\[ AK8 \]
\[ KS \]
\[ SB \]
\[ SR \]
\[ MC \]

\[ AK9 \]
\[ KS \]
\[ SB \]
\[ SR \]
\[ MC \]

\[ S_{start} \]
\[ S'_{start} \]

\[ S_{end} \]

\[ \Delta_{OUT} \]
9-Round CK Distinguisher for AES-128

Distinguishing algorithm

- Generate a valid pair of keys (about $2^{27}$ of them, since $P_{KS} = 2^{-101}$)
  - Store the $i$th SuperSBox from $S'_{start}$ to $S_{end}$ in $T_i$
  - For all 5 differences at $S_{start}$, check the tables and:
    - Check backward direction: $p = 2^{-7}$ (a single S-Box)
    - Check forward direction: $p = 2^{-6 \times 8} = 2^{-48}$ (6 S-Boxes)
**Time complexity**

### Complexity of the distinguishing algorithm

- **Check probability**: $2^{-7-48} = 2^{-55}$
- **Time complexity**:
  
  $$2^{15} \times (2^{32} + 2^{40}) \approx 2^{55} \text{ computations}$$

- **For** $2^{15}$ different pairs of keys:
  - Construct the SuperSBoxes in $2^{32}$ operations
  - Try all values for the 5 byte-differences in $2^{40}$ operations

### Generic time complexity

- **Limited-Birthday Problem** [GP-FSE10]
- **Input space** ($\Delta_{IN}$) of size $4 \times 8 + 7 = 39$ bits
- **Output space** ($\Delta_{OUT}$) of size $3 \times 7 = 21$ bits
- **Time complexity**: $2^{68}$ encryptions
Conclusion

- New algorithm for SPN ciphers
  - Graph-based approach: Dijkstra and A* optimization
  - Search the best truncated differential characteristics
  - Instantiation $\Rightarrow$ best differential characteristics
  - Time complexity linear in the number of rounds considered

- Applications to the structure of AES-128:
  - Impossibility results for related-key attacks
  - Impossibility results for the hash function setting

- Chosen-key distinguisher for 9-rounds AES-128
  - Solve open problem
  - Time Complexity: $2^{55}$ encryptions
  - Generic Complexity: $2^{68}$ encryptions

- More details in the paper and its extended version (ePrint/2013/366)
Conclusion

- New algorithm for SPN ciphers
  - Graph-based approach: Dijkstra and A* optimization
  - Search the best truncated differential characteristics
  - Instantiation $\Rightarrow$ best differential characteristics
  - Time complexity linear in the number of rounds considered

- Applications to the structure of AES-128:
  - Impossibility results for related-key attacks
  - Impossibility results for the hash function setting

- Chosen-key distinguisher for 9-rounds AES-128
  - Solve open problem
  - Time Complexity: $2^{55}$ encryptions
  - Generic Complexity: $2^{68}$ encryptions

- More details in the paper and its extended version (ePrint/2013/366)

Thank you!

Thanks to the organizing committee and sponsors for waiving my registration fee.