Proving the Security of AES Substitution-Permutation Network

T. Baignères S. Vaudenay

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

SAC 2005

Outline

Introduction

What does Cryptanalysis mean?

- Breaking a cryptographic algorithm? Not only!
- Proving the security of a construction/algorithm

As breaking $\neq$ proving security

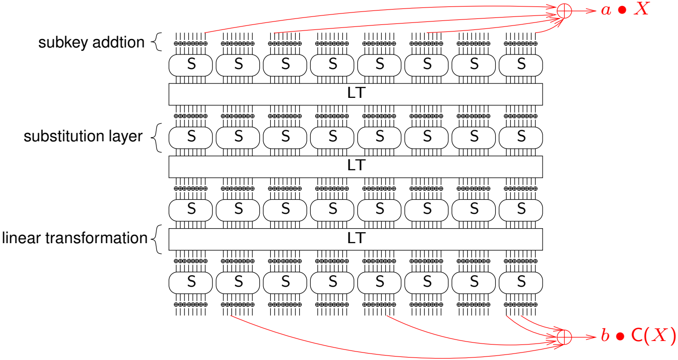
$\Downarrow$

different techniques must be applied.

$\leadsto$ An example: Linear Cryptanalysis

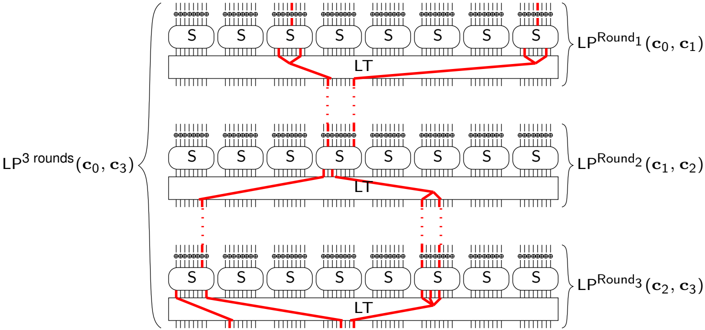
Example: Linear Cryptanalysis (LC)

Efficency of LC on a cipher C is measured by the Linear Probability: $\mathrm{LP}^C(a,b) = (2\Pr_X[a \bullet X = b \bullet C(X)] - 1)^2$



Example: Linear Cryptanalysis (LC)

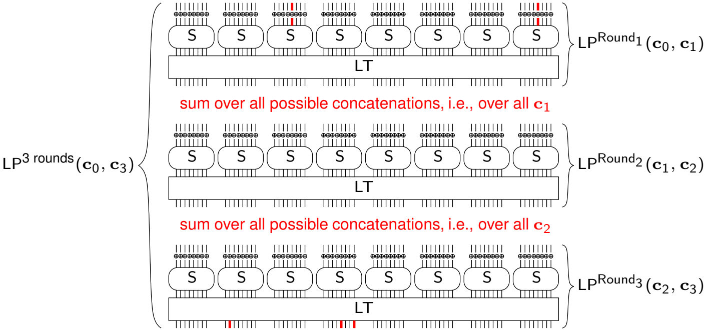
Computing the exact LP of a SPN is usually not practical.
$\rightsquigarrow$ concatenate round-LP's and apply the Piling-up Lemma



$$\mathsf{LP}^{3\text{ rounds}}(\mathbf{c}_0, \mathbf{c}_3) \approx \prod_{i=1,2,3} \mathsf{LP}^{\mathsf{Round}_i}(\mathbf{c}_{i-1}, \mathbf{c}_i)$$

Example: Linear Cryptanalysis (LC)

Following [Nyberg94], the approximation corresponds to considering only one characteristic among a linear hull.



$$\mathsf{LP}^{\text{3 rounds}}(\mathbf{c}_0,\mathbf{c}_3)=\sum_{\mathbf{c}_1,\mathbf{c}_2}\prod_{i=1,2,3}\mathsf{LP}^{\mathsf{Round}_i}(\mathbf{c}_{i-1},\mathbf{c}_i)$$

Example: Linear Cryptanalysis (LC)

How accurate is the approximation?

- It is ok when one characteristic is overwhelming (ex: DES)
- It is ok when it leads to an efficient attack
- This is not always the case (ex: AES)

It actually underestimates the LP!

$\rightsquigarrow$ an attack can only work better than expected...

$\rightsquigarrow$ ...a security proof becomes meaningless

Example: Linear Cryptanalysis (LC)

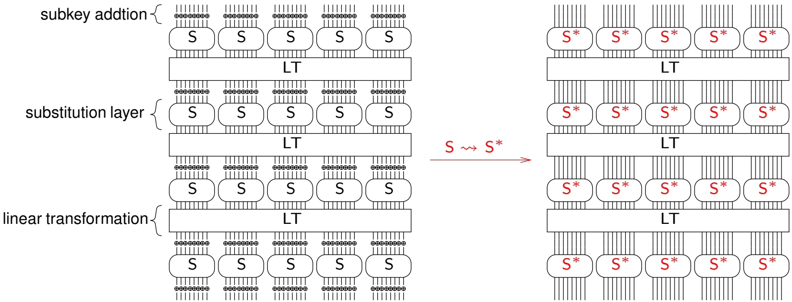
Conclusion

For security proofs, the LP cannot be approximated by the LP of one characteristic $\rightsquigarrow$ linear hull must be taken into account.

For AES, two (rigorous) alternatives have been studied:

- Upperbound the LP (e.g., [Keliher-Meijer-Tavares01], [Park-Sung-Chee-Yoon-Lim02], and [Keliher04])
- Adopt a Luby-Rackoff-like approach (e.g., [Moriai-Vaudenay00] and [Keliher-Meijer-Tavares03])

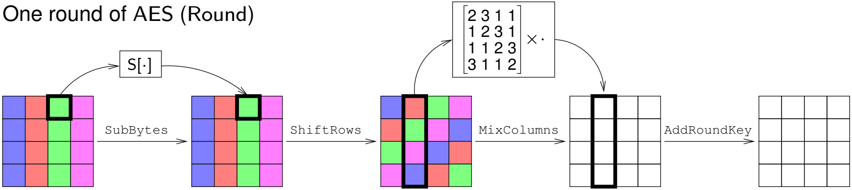
A Luby-Rackoff-like approach in a SPN



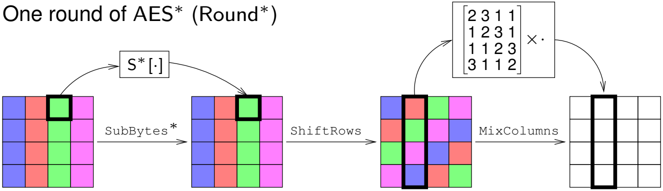
- S^* is a random permutation, uniformly distributed
- all random S-boxes are independent from each-other
- the subkey addition is included in S^*

AES vs. AES*

One round of AES (Round)



One round of AES* (Round*)



Results on AES*

- AES* is made of all identical rounds, except for the last one which excludes both linear transformations
- The LP on AES* is taken on average over all the random S-boxes

Summary of our results

- AES* is protected against linear and differential cryptanalysis after 4 inner rounds
- AES* is protected against iterated attacks of order one after 10 inner rounds
- LP^{AES^*} tends towards the LP of the perfect cipher as the number of rounds increases

On the Complexity of the Exact LP Computation

Given input/output masks $\mathbf{c}_0$ and $\mathbf{c}_r$,

$$\mathsf{LP}^{\mathsf{AES}^*}(\mathbf{c}_0, \mathbf{c}_r) = \sum_{\mathbf{c}_1,\ldots,\mathbf{c}_{r-1}} \prod_{i=1,\ldots,r} \mathsf{LP}^{\mathsf{Round}_i^*}(\mathbf{c}_{i-1}, \mathbf{c}_i)$$

Needs about $(2^{128})^3 \log r$ field operations $\rightsquigarrow$ Prohibitive!

First reduction: summing over intermediate supports instead of intermediate masks

Masks and Supports

The support of a mask $\mathbf{c}$ is the 4×4 array γ indicating which entries of $\mathbf{c}$ are zero and which are not:

mask $\mathbf{c}$

0x45	0x91	0xAB	0x00
0xF3	0x23	0x2C	0x37
0xFE	0x11	0x00	0xD1
0xB1	0x63	0x0A	0x00

support γ

1	1	1	0
1	1	1	1
1	1	0	1
1	1	1	0

Hamming weight of γ is denoted $|\gamma|$ (in this example, $|\gamma| = 13$)

Supports are useful to compute the LP on one round of AES*...

Average LP on $\texttt{SubBytes}^*$

For any non-zero input/output masks a, b on S^*

$$\mathsf{E}_{\mathsf{S}^*}[\mathsf{LP}^{\mathsf{S}^*}(a,b)] = \frac{1}{2^8 - 1} = \sigma^{-1}$$

Lemma

For any non-zero masks $\mathbf{a}, \mathbf{b} \in \mathsf{GF}(2^8)^{16}$ of respective supports α and β

$$\mathsf{E}[\mathsf{LP}^{\texttt{SubBytes}^*}(\mathbf{a},\mathbf{b})] = \begin{cases} \sigma^{-|\alpha|} & \text{if } \alpha = \beta \\ 0 & \text{otherwise.} \end{cases}$$

LP on LT $= \texttt{MixColumns} \circ \texttt{ShiftRows}$

- LT denotes $\texttt{MixColumns} \circ \texttt{ShiftRows}$
- For any state $\mathbf{x}$ and masks $\mathbf{a}, \mathbf{b}$

$$\mathbf{a} \bullet \mathbf{x} = \mathbf{b} \bullet (\mathsf{LT} \times \mathbf{x}) \quad \Leftrightarrow \quad \mathbf{a} = \mathsf{LT}^T \times \mathbf{b}$$

  We say that $\mathbf{a}$ and $\mathbf{b}$ are connected through LT
- N$[\alpha, \beta]$ denotes the number of possible connections through LT, given the input/output supports α and β.

Average LP on AES*

Theorem

For any non-zero masks $\mathbf{c}_0, \mathbf{c}_r \in \mathrm{GF}(2^8)^{16}$ of respective supports γ_0 and γ_r

$$\mathrm{E}[\mathrm{LP}^{\mathrm{AES}^*}(\mathbf{c}_0, \mathbf{c}_r)] = \sigma^{-|\gamma_r|} \times (\mathcal{M}^{r-1})_{\gamma_0,\gamma_r}$$

where $\mathcal{M}$ is a $2^{16} \times 2^{16}$ matrix indexed by pairs of masks (γ_{i-1}, γ_i) such that

$$\mathcal{M}_{\gamma_{i-1},\gamma_i} = \sigma^{|\gamma_{i-1}|}\mathrm{N}[\gamma_{i-1}, \gamma_i]$$

The computation roughly needs $(2^{16})^3 \log r$ field operations $\rightsquigarrow$ almost feasible!

Exploiting MDS properties of LT

- In order to further reduce the complexity, we used properties inherent to any MDS matrix (not only the one in LT) which induce symmetries in the table N[·].
- After some (frightening) computations…
- …using rather (horrible) notations…

Final Expression for the LP

(Simplified) Theorem

For any non-zero masks $\mathbf{c}_0, \mathbf{c}_r \in \mathrm{GF}(2^8)^{16}$ of respective supports γ_0 and γ_r

$$\mathrm{E}[\mathrm{LP}^{\mathrm{AES}^*}(\mathbf{c}_0, \mathbf{c}_r)] = \mathcal{U}^T \times \mathcal{L}^{r-2} \times \mathcal{V}$$

where

- $\mathcal{U}$ only depends on the diagonal weights of $\mathbf{c}_0$
- $\mathcal{V}$ only depends on the column weights of $\mathbf{c}_r$
- $\mathcal{L}$ is a matrix 1001×1001 matrix

Computing all the LP for AES* can be done on a laptop.

Experimental Results

- Maximum value of $\mathrm{E}[\mathrm{LP}^{\mathrm{AES}^*}(\mathbf{a},\mathbf{b})]$ for various number of rounds:

2	3	4	5	6	7	8	9
$2^{-33.9774}$	$2^{-55.9605}$	$2^{-127.9096}$	$2^{-127.9096}$	$2^{-127.9999}$	$2^{-127.9999}$	$2^{-128.0}$	$2^{-128.0}$

- Conclusion: AES* is protected against linear cryptanalysis after 4 rounds
- These results can be extended to differential cryptanalysis and to various S-box sizes

Properties of the matrix $\mathcal{M}$

In the previous Theorem

$$\mathsf{E}[\mathsf{LP}^{\mathsf{AES}^*}(\mathbf{c}_0, \mathbf{c}_r)] = \sigma^{-|\gamma_r|} \times (\mathcal{M}^{r-1})_{\gamma_0,\gamma_r}$$

The $2^{16} \times 2^{16}$ matrix $\mathcal{M}$ actually looks like

$$\left(\begin{array}{c|c} 1 & \mathbf{0} \\ \hline \mathbf{0} & \mathcal{M}' \end{array}\right)$$

where $\mathcal{M}'$ is a $(2^{16} - 1) \times (2^{16} - 1)$ indexed by non-zero supports.

Properties of the matrix $\mathcal{M}'$

Property

$\mathcal{M}'$ is the transition matrix of a Markov chain, i.e., $\mathcal{M}'_{\gamma,\gamma'}$ is the transition probability from a non-zero support γ to a non-zero support γ'.

From the study of supports propagation (based on the MDS criterion) $\rightsquigarrow$ the Markov chain is irreducible and aperiodic.

$\Rightarrow$ there exists a stationary distribution π, which can be determined. Then

$$(\mathcal{M}'^r)_{\gamma,\gamma'} \xrightarrow[r\to\infty]{} \pi_{\gamma'}$$

Towards the LP of the True Random Cipher

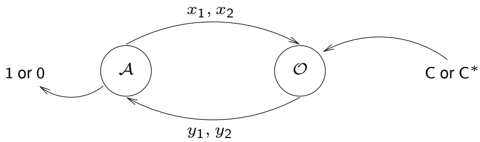
Theorem

For any non-zero input/output masks $\mathbf{a}, \mathbf{b}$,

$$\lim_{r \to \infty} \mathsf{E}[\mathsf{LP}^{\mathsf{AES}^*}(\mathbf{a}, \mathbf{b})] = \frac{1}{2^{128} - 1}$$

Iterated Attacks of Order 1

Consider an adversary $\mathcal{A}$ in the Luby-Rackoff model: unlimited computational power, limited access to an oracle $\mathcal{O}$ implementing either AES* or the perfect cipher C*. $\mathcal{A}$ must guess which is the case.



$\mathcal{A}$ can adapt x_2 depending on y_1

$\rightsquigarrow$ 2-limited adaptative distinguisher of advantage $\mathrm{Adv}_{\text{2-limited}}$

Iterated Attacks of Order 1

- Iterated attacks of order 1 are similar to linear cryptanalysis, except that the bit of information is not necessarily derived in a linear way (and that can make a huge difference, see [Baignères-Junod-Vaudenay04])
- Resistance against 2-limited adaptative distinguishers is sufficient to resist iterated attacks of order 1 (result from Decorrelation theory)

(Simplified) Theorem

Let $\epsilon = \max_{\mathbf{a}\neq\mathbf{0},\mathbf{b}} \mathsf{E}[\mathsf{DP}^{\mathsf{AES}^*}(\mathbf{a},\mathbf{b})] - \frac{1}{2^{128}-1}$, then

$$\mathsf{Adv}_{\text{2-limited}} \leq 2^{128}\epsilon$$

Iterated Attacks of Order 1: practical results

- Experimental values of ϵ depending on the number of rounds r:

2	3	4	5	6	7	8	9	10
$2^{-33.98}$	$2^{-55.96}$	$2^{-131.95}$	$2^{-131.95}$	$2^{-152.17}$	$2^{-174.74}$	$2^{-200.39}$	$2^{-223.93}$	$2^{-270.82}$

- Conclusion: provable security achieved for 10 rounds of AES*

Conclusion

- Study of the SPN of AES using a Luby-Rackoff-like approach $\rightsquigarrow$ AES*
- AES* is protected against linear and differential cryptanalysis after 4 inner rounds
- $\mathrm{LP}^{\mathrm{AES}^*}$ tends towards the LP of the perfect cipher as the number of rounds increases
- AES* is protected against iterated attacks of order one after 10 inner rounds

Thank you for your attention!