Block Cipher Cryptanalysis - part 2

Alessandro Barenghi

Department of Electronics, Information and Bioengineering (DEIB)
Politecnico di Milano

barenghi - at - elet.polimi.it

Recap

Cipher flaws

- Last time we tackled block cipher cryptanalysis from two points of view:
  - **Algebraic**: Consider the cipher as a system of equations with binary coefficients and try to solve them
  - **Statistical**: Exploit the fact that the ciphertext is not statistically independent from the plaintext
- We were able to:
  - **Algebraic**: Solve completely a fully linear cipher, obtaining the key from a single ptx/ctx pair
  - **Statistical**: Exploit a direct relation between ptx and ctx in terms of noting that they occour either too often or too seldom
- This time we exploit the fact that a difference in the ptx may imply a difference in the ctx with nontrivial probability

Differential Cryptanalysis

Same crate, different crowbar

- In a perfect n bit block cipher, flipping an input bit causes all the output bits to flip with probability $\frac{1}{2}$
- Given a difference in the input d_{in}, all the output differences d_{out} have the same probability $\Pr(d_{out}|d_{in}) = \frac{1}{2^n}$ of happening
- We will exploit the imperfections of the real ciphers in the way they distribute the differences over the outputs to retrieve the key
- Goal: find a key independent difference pair $\Delta p = a, \Delta y = b$ between ptx and pre-last-key addition states, with $\Pr(\Delta y = b|\Delta p = a) \neq \frac{1}{2^n}$

Differential Cryptanalysis

Notation

- Given two 4-bit states $X' = [X'_0, X'_1, X'_2, X'_3]$,$X'' = [X''_0, X''_1, X''_2, X''_3]$, the difference $\Delta X = [\Delta X_0, \Delta X_1, \Delta X_2, \Delta X_3]$ is their bitwise $\oplus$
- The difference ΔY among two different outputs is defined analogously
- A pair of values for $(\Delta X, \Delta Y)$ is commonly called a differential
- All the pairs $(\Delta X, \Delta Y)$ for which $\Pr((\Delta Y = b | \Delta X = a)) \neq \frac{1}{2^n}$ holds represent a viable vulnerability
- As we will be tackling garden variety differential cryptanalysis we will only consider the ones with $\Pr((\Delta Y = b | \Delta X = a)) > \frac{1}{2^n}$
- A different approach it to the ones where $\Pr((\Delta Y = b | \Delta X = a)) = 0$ this is commonly called impossible cryptanalysis

Differential Cryptanalysis

Linear Elements

- A difference in the input of a linear element of a cipher always yields the same output difference: i.e. its differential probability is 1
- Bitwise permutations simply change which bits are involved in the differential, not the probability of the differential holding
- Duplicating bits (also known as bit expansion, see DES) simply involves more bits in the differential, without probability change
- Bottom line: all the linear elements of a cipher do not change the differential probabilities

Differential Cryptanalysis

Nonlinear Elements - SBoxes

- No general method to determine the differentials $\rightarrow$ exhaustive search
- Notation: we will denote ΔX and ΔY as the nibble in binary obtained as $X' \oplus X''$ and $Y' \oplus Y''$ respectively
- For an n-bit S-box, there are only 2^{2n} possible differentials to be checked (as they are pairs of n bit differences)
- The trivial differential $(0000, 0000)$ always holds (no input difference, always means no output difference)
- Note: this time the quantity we exploit is the probability of a differential holding throughout an S-box, not the difference from $\frac{1}{2}$

Differential Cryptanalysis

SBox Differential Probability computation

- For each possible differential $\Delta X, \Delta Y$ out of $(2^4)^2$=256 choices:
  $\langle(\Delta X_0, \Delta X_1, \Delta X_2, \Delta X_3), (\Delta Y_0, \Delta Y_1, \Delta Y_2, \Delta Y_3)\rangle \in \{\langle 0000, 0000\rangle, \ldots, \langle 1111, 1111\rangle\}$
  - Initialize a counter: $\texttt{ctr} \leftarrow 0$
  - For each of the 16 S-box inputs $I = (i_0, i_1, i_2, i_3) \in \{0000, 0001, \ldots\}$
    - Compute the output values for both $\texttt{sbox}(I)$ and $\texttt{sbox}(I \oplus \Delta X)$
    - Check if the difference of the output values is ΔY
    - If it is, increment $\texttt{ctr}$
  - Store the differential probability value $\frac{\texttt{ctr}}{16}$ as "$\texttt{ctr}$"
- The differential table entries which we will consider interesting are the ones $\neq 1$

S-Box Description

Small, yet troublesome

- We will analyze the same 4×4 bit S-Box, we employed for the linear cryptanalysis
- It is obtained considering the first S-Box of DES and fixing to 00 the two leftmost input bits
- The complete description is provided below (input nibble on the top row, output on the bottom one)

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
14	4	13	1	2	15	11	8	3	10	6	12	5	9	0	7

Differential Cryptanalysis

Table: Frequency table for the Simple cipher S-Box

$\Delta X \downarrow \ \Delta Y \rightarrow$	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	**16**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	**2**	0	0	0	**2**	0	**2**	**4**	0	**4**	**2**	0	0
2	0	0	0	**2**	0	**6**	**2**	**2**	0	**2**	0	0	0	0	**2**	0
3	0	0	**2**	0	**2**	0	0	0	0	**4**	**2**	0	**2**	0	0	**4**
4	0	0	0	**2**	0	0	**6**	0	0	**2**	0	**4**	**2**	0	0	0
5	0	**4**	0	0	0	**2**	**2**	0	0	0	**4**	0	**2**	0	0	**2**
6	0	0	0	**4**	0	**4**	0	0	0	0	0	0	**2**	**2**	**2**	**2**
7	0	0	**2**	**2**	**2**	0	**2**	0	0	**2**	**2**	0	0	0	0	**4**
8	0	0	0	0	0	0	**2**	**2**	0	0	0	**4**	0	**4**	**2**	**2**
9	0	**2**	0	0	**2**	0	0	**4**	**2**	0	**2**	**2**	**2**	0	0	0
A	0	**2**	**2**	0	0	0	0	0	**6**	0	0	**2**	0	0	**4**	0
B	0	0	**8**	0	0	**2**	0	**2**	0	0	0	0	0	**2**	0	**2**
C	0	**2**	0	0	**2**	**2**	**2**	0	0	0	0	**2**	0	**6**	0	0
D	0	**4**	0	0	0	0	0	**4**	**2**	0	**2**	0	**2**	0	**2**	0
E	0	0	**2**	**4**	**2**	0	0	0	**6**	0	0	0	0	0	**2**	0
F	0	**2**	0	0	**6**	0	0	0	0	**4**	0	**2**	0	0	**2**	0

Combining differential biases

- The differential bias is the probability of a differential holding for all possible inputs
- Given two differentials $(\Delta X = a, \Delta Y = b), (\Delta Y = b, \Delta Z = c)$, holding with probabilities p_1 and p_2, we know that $(\Delta X = a, \Delta Z = c)$ will hold with probability $p_1 p_2$
- This,again, assumes that the probability of a certain differential $(\Delta X, \Delta Y)$ holding is independent from all the others[a]
- We can thus obtain the differential bias for a "longer" relation in the cipher chaining the differentials and multiplying the probabilities

[a] once again, this works on real world ciphers

Target cipher

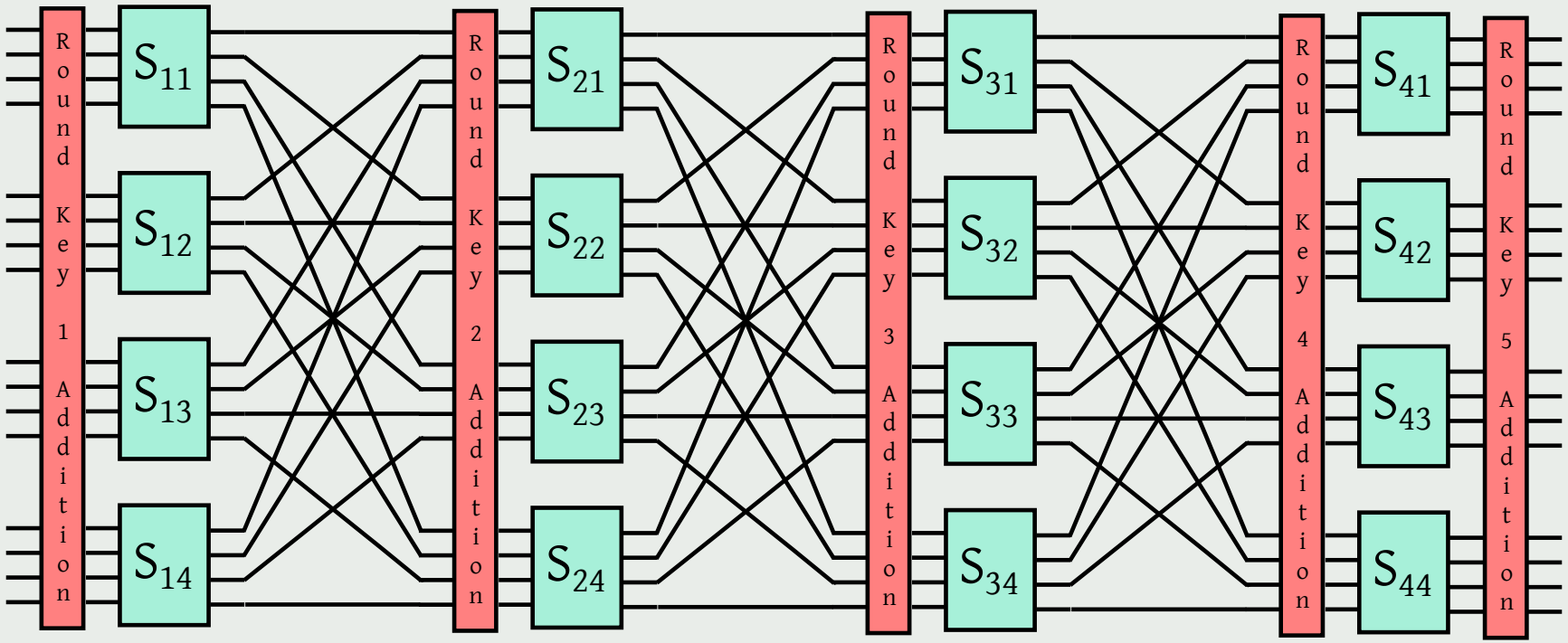
Simple

- Once again, 16 bit wide block, 4 rounds, 5x16 bit key (without any keyschedule: a 16-bit subkey per round)
- SPN design: [AddRoundKey, SBox, Perm]$\times$3, [AddRoundKey, SBox]$\times$1 , plus a final AddRoundKey
- Four 4-to-4 bit S-boxes, obtained as a part of the first DES Sbox[a]
- The *Simple* permutation layer: i-th bit from j-th box goes into the j-th bit of the i-th Sbox
- Note: the key is added through a simple bitwise `xor`

[a] Keeping the two highest bit set to 00

Target cipher

Simple



Differential Cryptanalysis

Getting rid of the keys

- Dealing with the key bits which get in our way to obtain a differential with a given probability is easier than with the linear cipher approximations
- Consider two one-bit plaintext X', X'', and a single bit key addition through $\oplus$
- The outputs can be expressed as $Y' = X' \oplus K$, $Y'' = X'' \oplus K$
- The output difference $\Delta Y = (X' \oplus K) \oplus (X'' \oplus K) = X' \oplus X'' = \Delta X$
- The output difference is already key independent: we can skip considering the key additions!

Differential Cryptanalysis

Breaking the cipher

- We have all the required notions to build a whole cipher differential
- The target is obtaining a differential involving only input and pre-last-round-key-addition bits
- After those are obtained, we need a statistically significant amount of ptx/ctx pairs to check that the differential bias holds
- Note: To check this, we need ptx-ctx pairs with the a specific difference between ptxs! (related ptx attack)
- By contrast, linear cryptanalysis works on a known plaintext assumption only, which may be easier to achieve in practice

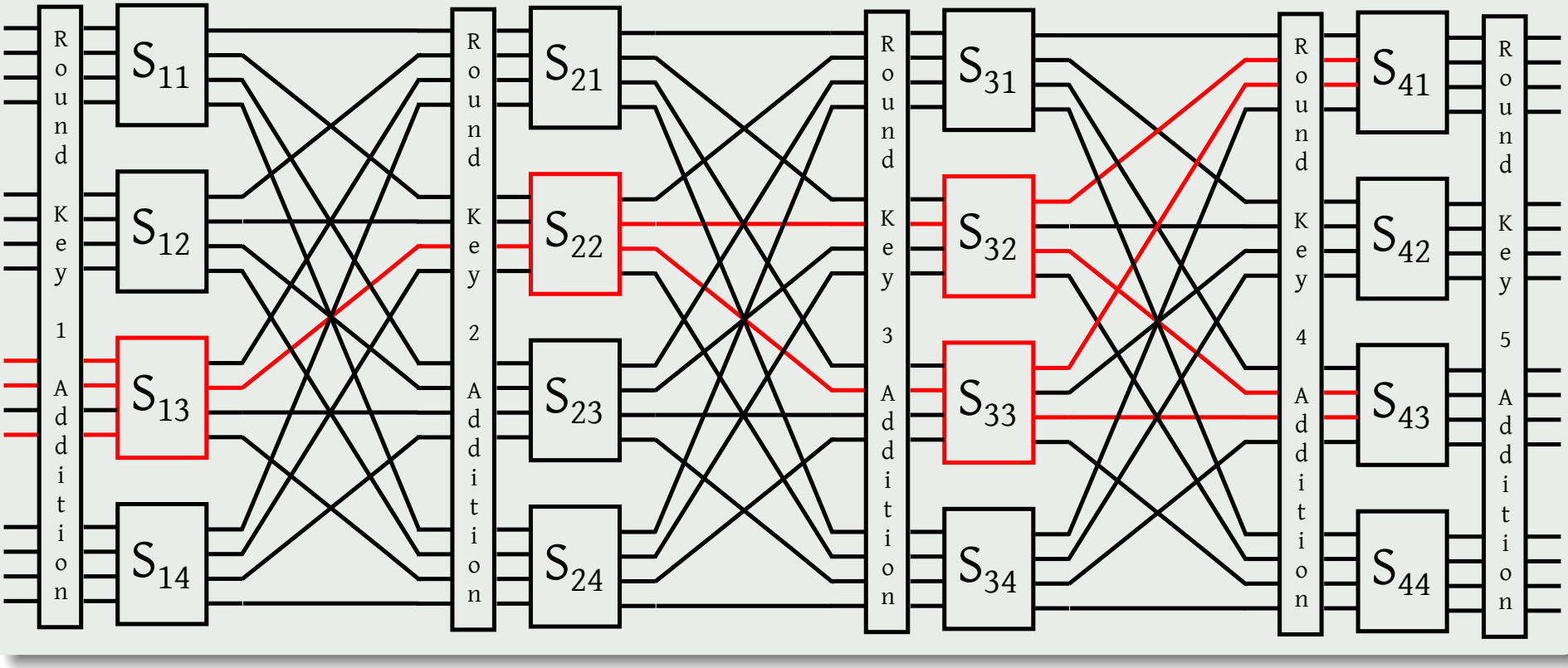
Differential Cryptanalysis

Choice of the relations and path building

- To build differential holding with high probability we need to follow the same guidelines employed in linear cryptanalysis
  - Minimize the number of involved bits (so to minimize the number of active S-boxes, which effectively reduce the probability)
  - Employ high probabilities so to obtain an approximation working as "often" as possible
- We employ the following differential relations to build the path
  - For S_{13} $(\Delta X = B, \Delta Y = 2)$ with bias $\frac{8}{16}$
  - For S_{23} $(\Delta X = 4, \Delta Y = 6)$ with bias $\frac{6}{16}$
  - For S_{32} $(\Delta X = 2, \Delta Y = 5)$ with bias $\frac{6}{16}$
  - For S_{33} $(\Delta X = 2, \Delta Y = 5)$ with bias $\frac{6}{16}$
- The probability for which the differential holds is simply computed as $\frac{8}{16} \times \frac{6}{16} \times \frac{6}{16} \times \frac{6}{16} = \frac{27}{1024}$

Differential Cryptanalysis

Complete differential path



Differential Cryptanalysis

Extracting the key bits for the last round of Simple

Key idea: Check for which value the bias holds out of all the possible partial subkeys added between the approximated path and the ciphertext:

1. Collect a significant number of ptx/ctx pairs with known differences among the ptxs
2. For each ptx/ctx pair
   - For each possible value of the partial subkeys
   - Initialize *count* to zero
     - Invert the effect of the partial subkeys on the affected ciphertext bits
     - Invert the last SBox and check if the differential approximation holds
     - If the approximation holds, increment *count*
3. After this, check which partial subkey makes the differential approximation hold exactly with the predicted bias
4. Lather, rinse, repeat for all the key bits you need to extract

Differential Cryptanalysis

Final Considerations

- Differential cryptanalysis needs at least $\frac{1}{p}$ ptx-ctx pairs, to confirm that a differential d holds with probability $\Pr(d) = p$
- Although this seems far stronger than linear cryptanalysis (which requires $\frac{1}{\epsilon^2}$ pairs), in this case the ptx/ctx pairs must be related, which may be difficult to achieve without choosing them
  - Obtaining them through the encryption of randomly distributed plaintexts (discarding the ones with wrong ptx difference) needs at least $\frac{1}{p}\frac{1}{2^{(n/2)}}$ ptx-ctx pairs, which may not be feasible (f.i. $n = 128$)
- Bottom line: differential cryptanalysis requires less ptx/ctx pairs than linear (assuming a flaw in the distribution of the ctx of the same entity), at the cost of a related ptx attacker

Cipher design criteria

Cryptanalysis summary

- The confusion principle is captured formally in statistical cryptanalysis by low linear biases/differential probabilities
- A robust nonlinear layer has negligible linear biases and differential probabilities close to $\frac{1}{2^n}$
- The diffusion principle is captured formally by the number of S-boxes into which the diffusion layer spreads the results of the nonlinear layer
- A robust linear diffusion layer is designed so that every bit of a nonlinear layer influences every bit of the next one
- These principles have been applied since linear/differential cryptanalyses have been known (first Turing paper in '50s, 60s NSA/IBM in DES design, 90s in AES design contest)

Cipher design criteria

A note on key addition

- The vast majority of block ciphers combine the key with the state through a simple $\oplus$ operation
- In light of differential cryptanalytic efforts, a valid alternative is to add the key via arithmetic addition modulo 2^n for some $n > 1$
- Note: the addition modulo 2^n is NOT linear modulo 2 for all of its input bits
  - Take the addition of a two bit key (k_1, k_0) to a two bit ptx (p_1, p_0) modulo 2^2: ctx=$(p_1 \oplus k_1 \oplus (p_0 \wedge k_0), p_0 \oplus k_0)$
- The assumption of the differentials being key independent must be adapted for any cipher performing such a key addition
- Note: the relation involving the least significant bit of the plaintext is still linear!

Real World Case studies

Case study, DES: broken

- The DES was the first cipher to be attacked with both linear and differential cryptanalyses
- Differential cryptanalysis was known during the design of DES (by NSA)
  - It turns out that the modifications suggested by NSA actually strengthened the cipher against differential cryptanalysis
- Current best differential cryptanalysis: 2^{47} (i.e. ptx/ctx to be chosen) for the whole cipher (16 rounds)
- Current best linear cryptanalysis 2^{43} (i.e. related ptx/ctx pairs) for the whole cipher (16 rounds)
- Significant, but not cathastrophic, exhaustive search is 2^{56}

Real World Case studies

Case study, AES: secure by design

- AES was designed with resistance against linear and differential cryptanalysis as a design criterion
- It has differential 4- uniform 8-to-8 bit S-Boxes (maximum probability for any differential $= \frac{4}{256}$)
- Since there are at least 25 SBoxes active every 4 rounds of AES, the maximum probability for any particular difference to happen is at most $(\frac{4}{256})^{25} \approx 2^{-150}$ after 4 rounds
- Exploiting such a statistical bias for the whole cipher (10 rounds) requires more than 2^{480} ptx-ctx pairs, way more than the ones available for a single key (only 2^{128})

How strong is a block cipher?

Summary

- Resistance against statistical cryptanalysis
  - Very low differential probabilities and linear biases for the S-box(es)
- Resistance against algebraic cryptanalysis
  - The algebraic relations mapping the key bits and ptx bits into the ctx bits are not trivial[a] and involve the whole key for all output bits
- Fast diffusion principle
  - The permutation layer should maximize the number of active S-Boxes
- Reduced round analyses provide a quantitative estimate of the confusion/diffusion properties of a single round
  - Maximum linear and differential biases w.r.t. number of rounds

[a] i.e. not solvable for key bits through any means but bruteforce/SAT