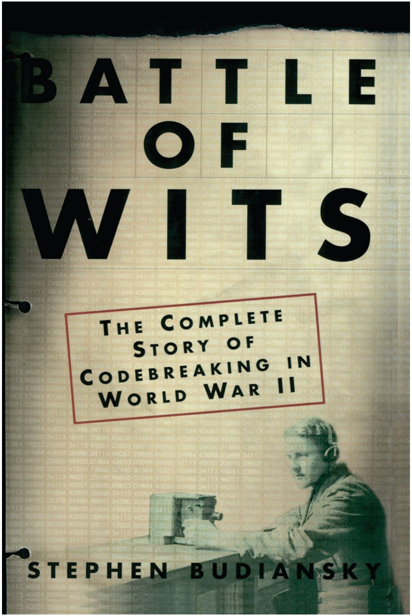
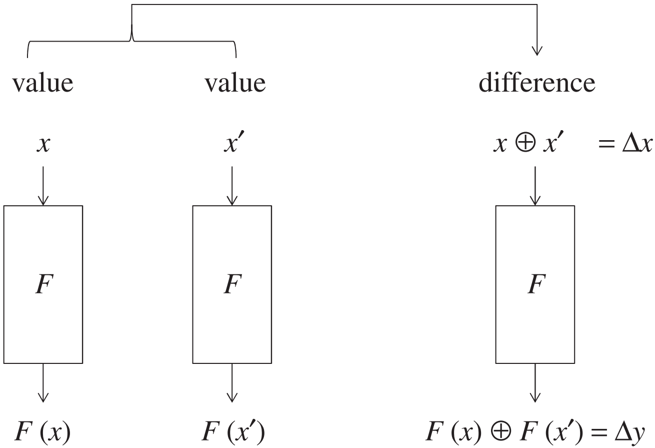


Image Source: Google

CS 553
CRYPTOGRAPHY

Lecture 6
Differential Cryptanalysis

Instructor
Dr. Dhiman Saha


value
value
difference
x
x'
$x \oplus x' \quad = \Delta x$
F
F
F
$F(x)$
$F(x')$
$F(x) \oplus F(x') = \Delta y$

Differential Cryptanalysis (DC)

- Differential? $\Delta_x \to \Delta_y$
- Notion of difference of inputs

Primary intuition

To study the propagation of differences through an SPN network focusing on the properties of the confusion (Sbox) and diffusion layers

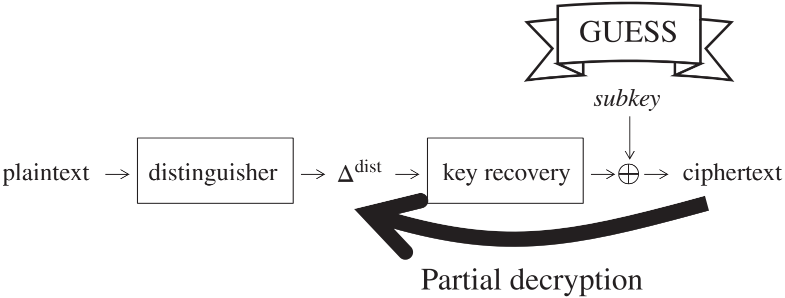
- Trace differences of pairs of plaintexts in the decryption process.
- Deduce information about the key

- Differential? $\Delta_x \to \Delta_y$
- Notion of difference of inputs Δ

Primary intuition

To study the propagation of differences through an SPN network focusing on the properties of the confusion (Sbox) and diffusion layers

- Trace differences of pairs of plaintexts in the decryption process.
- Deduce information about the key



****Will be clearer later*

Image Source: Security of Block Ciphers: From Algorithm Design to Hardware Implementation

DC

- The discovery is generally attributed to Eli Biham and Adi Shamir in the late 1980s.
- However, in 1994, IBM claimed that DC was known to IBM as early as 1974.
- Within IBM, it was known as the ”T-attack or ”Tickle attack.
- Invented to break DES, did not succeed though
- A chosen plaintext attack.
- Applicable to many iterated block ciphers.

Resistance against DC a prerequisite for present-day block cipher proposals.

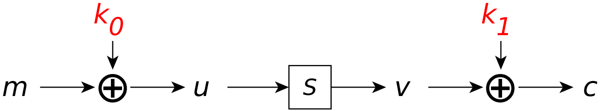
First Target

- Sypher001 encrypts 4 bits with two 4 bit keys

S-box

x	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
$S(x)$	6	4	c	5	0	7	2	e	1	f	3	d	8	a	9	b

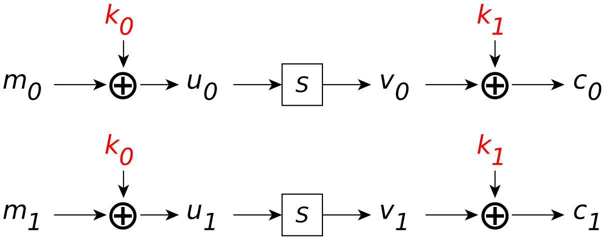
Encryption



Note

- No **diffusion** (permutation), only **confusion** (substitution)!

Assume we are given the encryptions of two messages m_0, m_1.



DC of Sypher001

$$\begin{tikzcd}
 & \color{red}{k_0} \arrow[d] & & & & \color{red}{k_1} \arrow[d] & \\
m_0 \arrow[r] & \oplus \arrow[r] & u_0 \arrow[r] & \boxed{S} \arrow[r] & v_0 \arrow[r] & \oplus \arrow[r] & c_0
\end{tikzcd}$$

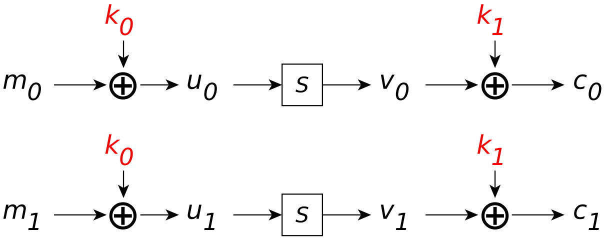
$$\begin{tikzcd}
 & \color{red}{k_0} \arrow[d] & & & & \color{red}{k_1} \arrow[d] & \\
m_1 \arrow[r] & \oplus \arrow[r] & u_1 \arrow[r] & \boxed{S} \arrow[r] & v_1 \arrow[r] & \oplus \arrow[r] & c_1
\end{tikzcd}$$

First Observation: Key Annihilation

We Know:

$$u_0 \oplus u_1 = (m_0 \oplus \color{red}{k_0}) \oplus (m_1 \oplus \color{red}{k_0}) = m_0 \oplus m_1$$

even though we do not know k_0

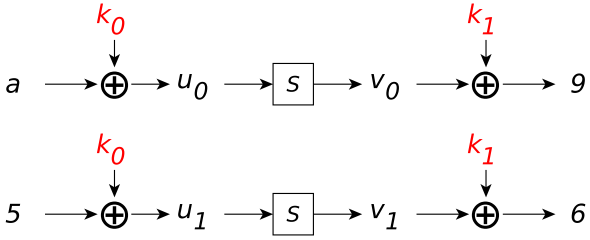


Strategy

- Guess k_1
- Compute v_0' and v_1'
- Compute $u_0' = S[v_0']^{-1}$ and $u_1' = S[v_1']^{-1}$
- Verify if $u_0 \oplus u_1 = u_0' \oplus u_1'$
- If not, then key guess was incorrect!

Example

- Given $m_0 = a, m_1 = 5$ and $c_0 = 9, c_1 = 6$



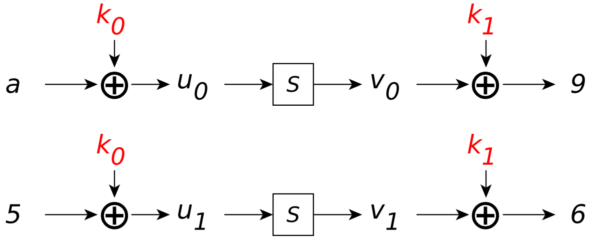
- Computer $u_0 \oplus u_1 = a \oplus 5 = f$
- Guess k_1

k_1	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
$u_0' \oplus u_1'$	e	b	e	e	d	8	d	**f**	**f**	d	8	d	e	e	b	e

- Compare $u_0 \oplus u_1$ and $u_0' \oplus u_1'$
- Only candidates for k_1 are $7, 8$

Example — Is this is approach **symmetrical**?

- Given $m_0 = a, m_1 = 5$ and $c_0 = 9, c_1 = 6$



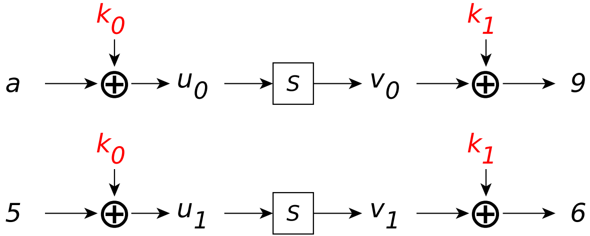
- Computer $u_0 \oplus u_1 = a \oplus 5 = f$
- Guess k_1

k_1	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
$u'_0 \oplus u'_1$	e	b	e	e	d	8	d	**f**	**f**	d	8	d	e	e	b	e

- Compare $u_0 \oplus u_1$ and $u'_0 \oplus u'_1$
- Only candidates for k_1 are $7, 8$

Exercise — **Repeat** DC from plaintext side

- Given $m_0 = a, m_1 = 5$ and $c_0 = 9, c_1 = 6$



- Computer $v_0 \oplus v_1 = 9 \oplus 6 =$ ___
- Guess k_0

k_0	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
$v_0' \oplus v_1'$																

- Compare $v_0 \oplus v_1$ and $v_0' \oplus v_1'$
- Only candidates for k_0 are ________

Take Away

- We know things about **differences** even though we do not know the individual values.
- We make a guess for the key and verify it by computing a bit backwards

Is it enough if we have a good guess for the difference?

Probably not, lets see out next cipher

Second Target

- Sypher002 encrypts 4 bits with **three** 4 bit keys

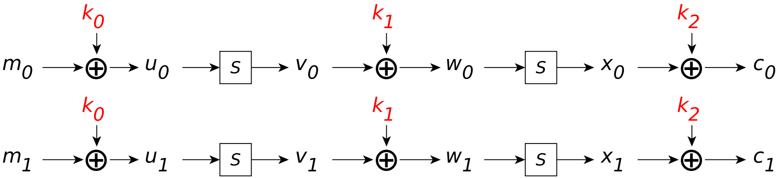
S-box

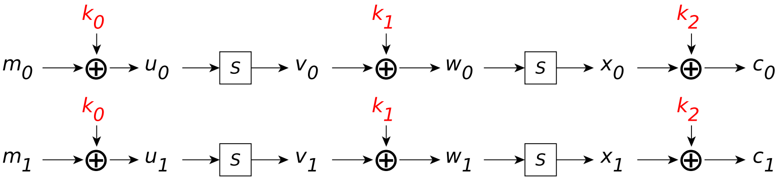
x	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
$S(x)$	6	4	c	5	0	7	2	e	1	f	3	d	8	a	9	b

Encryption

$$m \xrightarrow{\oplus\, k_0} u \xrightarrow{S} v \xrightarrow{\oplus\, k_1} w \xrightarrow{S} x \xrightarrow{\oplus\, k_2} c$$

Assume we are given the encryptions of two messages m_0, m_1.

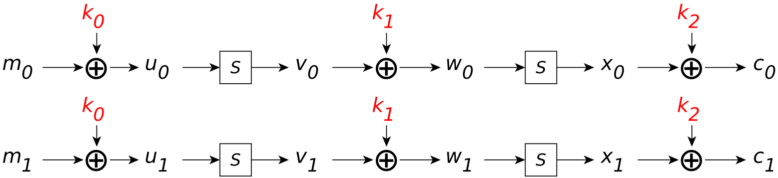




We can compute (after guessing k_2)

- $u_0 \oplus u_1$
- x'_0 and x'_1
- w'_0 and w'_1
- $v'_0 \oplus v'_1 = w'_0 \oplus w'_1$

But we still cannot check our guess for k_2



More Powerful Attacker

Now, we make it a **chosen plaintext attack**:

- Choose the starting difference

$$m_0 \oplus m_1 = u_0 \oplus u_1 = f$$

$$m_0 \rightarrow \overset{k_0}{\oplus} \rightarrow u_0 \rightarrow \boxed{S} \rightarrow v_0 \rightarrow \overset{k_1}{\oplus} \rightarrow w_0 \rightarrow \boxed{S} \rightarrow x_0 \rightarrow \overset{k_2}{\oplus} \rightarrow c_0$$

$$m_1 \rightarrow \overset{k_0}{\oplus} \rightarrow u_1 \rightarrow \boxed{S} \rightarrow v_1 \rightarrow \overset{k_1}{\oplus} \rightarrow w_1 \rightarrow \boxed{S} \rightarrow x_1 \rightarrow \overset{k_2}{\oplus} \rightarrow c_1$$

We can compute (after guessing k_2)

- $u_0 \oplus u_1 = f$
- $v_0' \oplus v_1'$

Question

Is there anything we can say about $v_0 \oplus v_1$ given that $u_0 \oplus u_1 = f$?

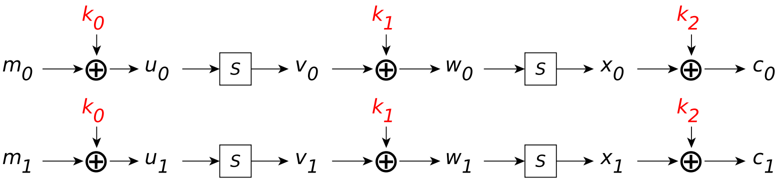
u_0	$u_1 = u_0 \oplus f$	$v_0 = S[u_0]$	$v_1 = S[u_1]$	$v_0 \oplus v_1$
0	f	6	b	d
1	e	4	9	d
2	d	c	a	6
3	c	5	8	d
4	b	0	d	d
5	a	7	3	4
6	9	2	f	d
7	8	e	1	f
8	7	1	e	f
9	6	f	2	d
a	5	3	7	4
b	4	d	0	d
c	3	8	5	d
d	2	a	c	6
e	1	9	4	d
f	0	b	6	d

Observations

- The difference is unevenly distributed.
- Not all values occur.
- The difference d occurs 10 out of 16 times.

Thus, we assume that $v_0 \oplus v_1 = d$ and this enables us to verify our guess for k_2 .

DC of Sypher002



We can compute (after guessing k_2)

- $u_0 \oplus u_1 = f$
- $v'_0 \oplus v'_1$

Thus, we assume that $v_0 \oplus v_1 = d$ and this enables us to verify our guess for k_2 .

$$v_0 \oplus v_1 = d = v'_0 \oplus v'_1$$

What if the assumption is right/wrong?

Right

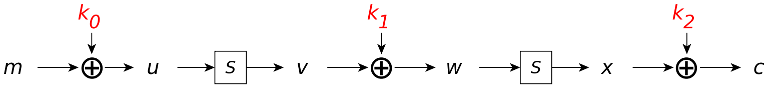
If the assumption is right, the right key is one of the possible candidates

Wrong

If the assumption is wrong, the right key might not be one of the possible candidates.

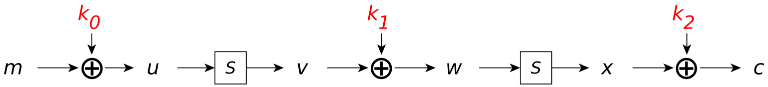
Still: If the assumption has a good probability (here : 10/16), the right key is a candidate more often than any wrong key.

Finally



- Initialize counters $T_i = 0$, one for each possible key k_2 .
- For each message/ciphertext pair do
  - For each guess i for k_2 do
    - Compute $v_0' \oplus v_1'$
    - If $v_0' \oplus v_1' = d$ increase counter T_i
- Assume that the right key k_2 corresponds to the highest counter.

Analysis



Assumption

Assume that a wrong guess for k_2 gives a random value for $v_0 \oplus v_1$

This implies that after processing t pairs we can expect

- The counter for the correct key is $\approx t \times \frac{10}{16}$
- The counter for the wrong key is $\approx t \times \frac{1}{16}$

Again: Influence of Sbox

Observation

The attack was possible because for the input difference f the output differences where highly unbalanced.

Question

What happens for other input differences?

in \out	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
0	16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1	-	-	6	-	-	-	-	2	-	2	-	-	2	-	4	-
2	-	6	6	-	-	-	-	-	-	2	2	-	-	-	-	-
3	-	-	-	6	-	2	-	-	2	-	-	-	4	-	2	-
4	-	-	-	2	-	2	4	-	-	2	2	2	-	-	2	-
5	-	2	2	-	4	-	-	4	2	-	-	2	-	-	-	-
6	-	-	2	-	4	-	-	2	2	-	2	2	2	-	-	-
7	-	-	-	-	-	4	4	-	2	2	2	2	-	-	-	-
8	-	-	-	-	-	2	-	2	4	-	-	4	-	2	-	2
9	-	2	-	-	-	2	2	2	-	4	2	-	-	-	-	2
a	-	-	-	-	2	2	-	-	-	4	4	-	2	2	-	-
b	-	-	-	2	2	-	2	2	2	-	-	4	-	-	2	-
c	-	4	-	2	-	2	-	-	2	-	-	-	-	-	6	-
d	-	-	-	-	-	-	2	2	-	-	-	-	6	2	-	4
e	-	2	-	4	2	-	-	-	-	-	2	-	-	-	-	6
f	-	-	-	-	2	-	2	-	-	-	-	-	-	10	-	2

How to interpret it?

Characteristics

Definition

Characteristic Given an Sbox S, a pair (α, β) is called a *characteristic* with probability p, if the probability that two inputs with difference α provide outputs with difference β equals p. This is denoted as

$$\alpha \xrightarrow{S} \beta$$

Examples for our Sbox

- $f \xrightarrow{S} d$ has probability $\frac{10}{16}$
- $d \xrightarrow{S} c$ has probability $\frac{6}{16}$
- $c \xrightarrow{S} a$ has probability 0: Impossible Characteristic