CS641

Modern Cryptology
Indian Institute of Technology, Kanpur

Group Name: Astralis
Kartik Rajiv Nirmal (180346), Sourabh Kulhari (180778), Sumit Jaiswal (180796)

Mid Semester Examination

Date of Submission:
March 10, 2021

Question 1

Consider a variant of DES algorithm in which the S-box S1 is changed as follows:

> For every six bit input α, the following property holds: $S1(\alpha) = S1(\alpha \oplus 001100) \oplus 1111$.

All other S-boxes and operations remain the same. Design an algorithm to break four rounds of this variant. In order to get any credit, your algorithm must make use of the changed behavior of S1.

Solution

As discussed in class, while mounting an attack on four round DES we have to predict the XOR of the output of S boxes in 2^{nd} round to get the XOR of the output of S boxes in 4^{th} round. Now in the given question we have variant of a four round DES but the strategy to break this variant will also be similar to what we discussed in lectures. We will first predict the XOR of output of S boxes in 2^{nd} round and further proceed. As we have a changed beahviour of S1 box in this variant, this changed behaviour leads to a theorem which helps us predicting the output of S boxes of 2^{nd} round in more accurate way, leading to analysis results which differ from what we had in normal variant of DES.

Theorem 1.1. Given two six bit inputs, α and α' to S1 box such that $\alpha \oplus \alpha' = 001100$, then

$$S1(\alpha) \oplus S1(\alpha') = 1111$$

Proof. Given :

$$\alpha \oplus \alpha' = 001100 \tag{1.2}$$

$$S1(\alpha) = S1(\alpha \oplus 001100) \oplus 1111 \tag{1.3}$$

On taking XOR with α on both sides in Equation 1.2, we get

$$\alpha' = \alpha \oplus 001100 \tag{1.4}$$

Substituting value of α' from Equation 1.4 in Equation 1.3, we get

$$S1(\alpha) = S1(\alpha') \oplus 1111 \tag{1.5}$$

Now take XOR with $S1(\alpha')$ on both sides in Equation 1.5, we get

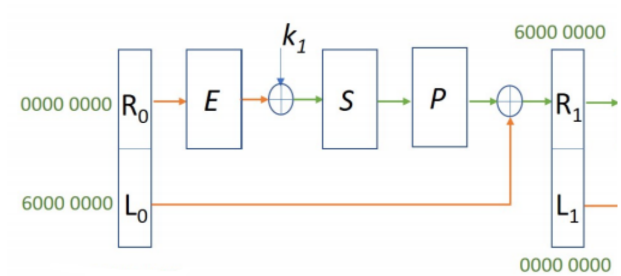
$$S1(\alpha) \oplus S1(\alpha') = 1111$$

Hence we have proved Theorem 1.1

□

Now to break the encryption we will mount a chosen plaintext attack with input plaintext L_0R_0 and another input plaintext $L_0^{'}R_0^{'}$ on this variant such that

$$L_0 \oplus L_0^{'} = 60000000 \text{ and } R_0 \oplus R_0^{'} = 00000000 \ \ \text{[hexadecimal notation]}$$



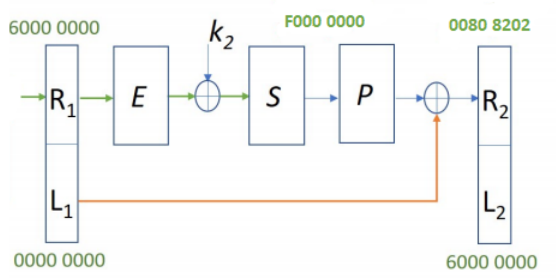
Now since the right halves of the plaintext inputs are same , all the output XOR values of expansion, S boxes and permutation will be 00000000 for 1^{st} round.

Since $L_i = R_{i-1}$ for any DES,

$$L_1 \oplus L_1^{'} = R_0 \oplus R_0^{'} = 00000000 \tag{1.6}$$

Also, since the permutation output XOR is 00000000,

$$R_1 \oplus R_1^{'} = L_0 \oplus L_0^{'} = 60000000 \tag{1.7}$$



For the 2^{nd} round, the output XOR of expansion can be found out using XOR values of R_1 and $R_1^{'}$

$$E[R_1] \oplus E[R_1^{'}] = 300000000000 \tag{1.8}$$

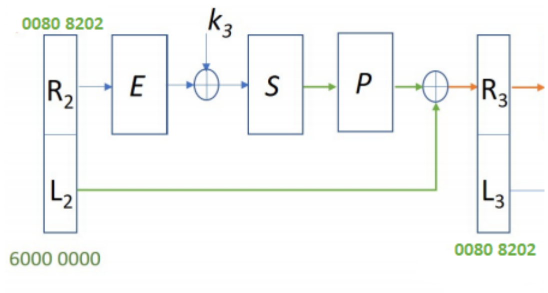
The input XOR to S boxes will be 300000000000. Now consider the input to S1 box to be α and α' in the two input cases ($|\alpha| = |\alpha^{'}| = 6$). Here we have $\alpha \oplus \alpha' = 001100$ and hence using Theorem 1.1 we can certainly say the output XOR will be $\mathrm{S1}(\alpha) \oplus \mathrm{S1}(\alpha') = 1111$. The difference in normal DES and the given variant of DES is in the probability with which we can predict the XOR of output of S1 box. In normal DES we had the case that with probability $p = \frac{14}{64}$ our output XOR will be 1110 but in the given variant with probability $p = 1$ our output XOR will be 1111. Output XOR of all other S boxes will be 0000.

Output XOR of S boxes in 2^{nd} round will be F0000000(hexadecimal representation).
The output XOR of permutation operation will be 00808202.

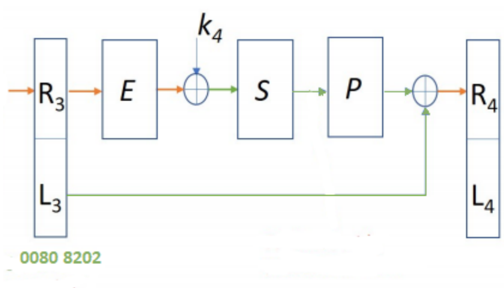
$$R_2 \oplus R_2^{'} = (L_1 \oplus L_1^{'}) \oplus 00808202 \tag{1.9}$$

using Equation 1.6

$$R_2 \oplus R_2^{'} = 00808202 \tag{1.10}$$



Since $L_3 = R_2$ and $L_3^{'} = R_2^{'}$
$L_3 \oplus L_3^{'} = R_2 \oplus R_2^{'} = 00808202$.



In 4^{th} round, as we know the XOR of L_3 and $L_3^{'}$ and the exact values of R_4 and $R_4^{'}$, we can get the XOR of output of permutation. As we know the XOR of output of permutation, we can get the XOR of output of S boxes. Also, as we know the exact values of L_4 and $L_4^{'}$, we know $L_4 \oplus L_4^{'}$.
Using the relation $R_3 = L_4$ and $R_3^{'} = L_4^{'}$,

$$R_3 \oplus R_3^{'} = L_4 \oplus L_4^{'}$$

Hence we know $R_3 \oplus R_3^{'}$.
Using $R_3 \oplus R_3^{'}$ we get, $E[R_3] \oplus E[R_3^{'}]$(E is the expansion operation)

Now we know the XOR of input to S boxes and XOR of output to S boxes. We have scenario similar to what we used to have for 3 round DES [1], using a similar approach we solve for k_4.

Let $E[R_3] = \alpha_1\alpha_2.....\alpha_8$ and $E[R_3^{'}] = \alpha_1^{'}\alpha_2^{'}....\alpha_8^{'}$ with $|\alpha_i| = |\alpha_i^{'}| = 6$.
Here, R_3 and $R_3^{'}$ are the right halves of output of third round on the plaintexts L_0R_0 and $L_0^{'}R_0^{'} = L_0^{'}R_0$.
Let k_4 be the key for the fourth round such that $k_4 = k_{4,1}k_{4,2}k_{4,3}....k_{4,8}$, $|k_{4,i}| = 6$
Let $\beta_i = \alpha_i \oplus k_{4,i}$ and $\beta_i^{'} = \alpha_i^{'} \oplus k_{4,i}$, $|\beta_i| = |\beta_i^{'}| = 6$.

Let $\gamma_i = S_i(\beta_i)$ and $\gamma_i^{'} = S_i(\beta_i^{'})$, $|\gamma_i| = |\gamma_i^{'}| = 4$.
We know, α_i, $\alpha_i^{'}$, $\beta_i \oplus \beta_i^{'} = \alpha_i \oplus \alpha_i^{'}$ and $\gamma_i \oplus \gamma_i^{'}$ for all $1 \leqslant i \leqslant 8$.
We define,

$$X_i = \{(\beta, \beta^{'}) | \beta \oplus \beta^{'} = \beta_i \oplus \beta_i^{'} \text{ and } S_i(\beta) \oplus S_i(\beta^{'}) = \gamma_i \oplus \gamma_i^{'}\} \tag{1.11}$$

Since the pair (β_i,$\beta_i^{'}$) satisfies all the properties of the set X_i, we can say (β_i,$\beta_i^{'}$) $\in X_i$.
Now we define,

$$K_i = \{k | k = \beta \oplus \alpha_i \text{ and } (\beta, \beta^{'}) \in X_i \text{ for some } \beta^{'}\} \tag{1.12}$$

Now since $(\beta_i,\beta_i^{'}) \in X_i$, we have $k_{4,i} \in X_i$.
We have $|K_i|$ = $|X_i|$ since α_i and $\beta \oplus \beta^{'}$ is fixed for $(\beta, \beta') \in X_i$.
As discussed in class,we have $|X_i| \leq 16$ for any choice of $\beta_i \oplus \beta_i^{'}$ and $\gamma_i \oplus \gamma_i^{'}$ and any i.
Therefore, $|K_i| \leq 16$ from the above analysis.
Doing the same for all S-boxes, we get at most $16^8 = 2^{32}$ possibilities for K_4.
By repeating the above attack for a few pairs of plain text that share the same right half, we can uniquely identify k_4.
Now as we know k_4, using decryption method for 1 round we can get L_3,R_3,$L_3^{'}$ and $R_3^{'}$.
Now we have a 3 round DES, with input L_0R_0, $L_0^{'}R_0^{'}$ and output L_3R_3 and $L_3^{'}R_3^{'}$ respectively. Now since the breaking technique for 3 round DES is independent of S box behaviour, we can solve the same using the approach discussed in lectures[1] to break 3 round DES .
This is a **chosen plain text** attack since plain text pairs with same right half are chosen for attack.

Question 2

The SUBSET-SUM problem is defined as follows:

> Given $(a_1, \ldots, a_n) \in \mathbb{Z}^n$ and $m \in \mathbb{Z}$, find $(b_1, \ldots, b_n) \in \{0,1\}^n$ such that $\sum_{i=1}^{n} a_i b_i = m$ if it exists.

This problem is believed to be a hard-to-solve problem in general. Consider a hypothetical scenario where Anubha and Braj have access to a fast method of solving SUBSET-SUM problem. They use the following method to exchange a secret key of AES:

> Anubha generates an $n = 128$ bit secret key k. She then chooses n positive integers $a_1, \ldots, a_n$ such that $a_i > \sum_{1 \le j < i} a_j$. She computes $m = \sum_{i=1}^{n} a_i k_i$ and sends $(a_1, a_2, \ldots, a_n, m)$ to Braj, where k_i is ith bit of k. Upon receiving numbers $(a_1, a_2, \ldots, a_n, m)$, Braj solves the SUBSET-SUM problem to extract the key k.

Show that an attacker Ela does not need to solve SUBSET-SUM problem to retrieve the key k from $(a_1, a_2, \ldots, a_n, m)$.

Solution

Given $(a_1, \ldots, a_n) \in \mathbb{Z}^n$ and $m \in \mathbb{Z}$, find $(k_1, \ldots, k_n) \in \{0,1\}^n$ such that $\sum_{i=1}^{n} a_i k_i = m$. Note here $n = 128$ (number of bits in k).

We can solve this problem by SUBSET-SUM problem, but SUBSET-SUM problem is hard-to-solve, hence, we need to find some other way to solve the problem.

We have to extract k from $(a_1, a_2, \ldots, a_n, m)$, without actually solving SUBSET-SUM problem, where $k = (k_1, k_2, k_3, \ldots, k_n)$

The most important information that we will be using is $a_i > \sum_{1 \le j < i} a_j$.

Theorem 2.1. For given $(a_1, a_2, \ldots, a_i, m)$ where $m = \sum_{j=1}^{i} a_j k_j, \quad (a_r > 0, k_r \in \{0,1\}) \ \forall \ r \in \{1, 2, \ldots, i\}$ and $a_i > \sum_{1 \le j < i} a_j$ then

$$k_i = \begin{cases} 1 & \text{if } m \ge a_i \\ 0 & \text{if } m < a_i \end{cases}$$

Proof. Consider following cases :

Case 1 : $m \geq a_i$

$$
\begin{aligned}
& m \geq a_i \\
\implies & \sum_{j=1}^{i} a_j k_j \geq a_i && [\because m = \sum_{j=1}^{i} a_j k_j \text{ and } m \geq a_i] \\
\implies & \sum_{j=1}^{i-1} a_j k_j + a_i k_i \geq a_i && [\text{taking } i_{th} \text{ term out of summation}] \\
\implies & a_i k_i \geq a_i - \sum_{j=1}^{i-1} a_j k_j && (2.2) \\
& \because a_i > \sum_{1 \leq j < i} a_j \\
\implies & a_i - \sum_{1 \leq j < i} a_j > 0 \\
& \text{Also note that } k \in \{0,1\}, \text{ hence } \sum_{1 \leq j < i} a_j \geq \sum_{1 \leq j < i} a_j k_j \\
\implies & a_i - \sum_{1 \leq j < i} a_j k_j > 0 && (2.3)
\end{aligned}
$$

From Equation 2.2 and Equation 2.3,

$$
\begin{aligned}
\implies & a_i k_i > 0 \\
\implies & k_i > 0 && [\because a_i > 0\,] \\
\implies & k_i = 1 && [\because k_i \in \{0,1\}\,]
\end{aligned}
$$

Case 2 : $m < a_i$

Let us assume $k_i = 1$, So

$$
\begin{aligned}
& m < a_i \\
\implies & \sum_{j=1}^{i} a_j k_j < a_i && [\because m = \sum_{j=1}^{i} a_j k_j \text{ and } m < a_i] \\
\implies & \sum_{j=1}^{i-1} a_j k_j + a_i k_i < a_i && [\text{taking } i_{th} \text{ term out of summation}]
\end{aligned}
$$

$$\implies \sum_{j=1}^{i-1} a_j k_j + a_i < a_i \qquad [\because \text{ according to our assumption } k_i = 1]$$

$$\implies \sum_{j=1}^{i-1} a_j k_j < 0$$

Above inequality leads to contradiction, because $a_j > 0$ and $k_j \in \{0,1\} \implies \sum_{j=1}^{i-1} a_j k_j \geq 0$
So, our assumption that $k_i = 1$ is wrong.
$\implies k_i = 0$

Hence,

$$k_i = \begin{cases} 1 & \text{if } m \geq a_i \\ 0 & \text{if } m < a_i \end{cases}$$

□

Conclusion of the above theorem : For given $(a_1, a_2, \dots, a_i, m)$ where $m = \sum_{j=1}^{i} a_j k_j$, $(a_r > 0, k_r \in \{0,1\})$ $\forall$ $r \in \{1, 2, \dots, i\}$ and $a_i > \sum_{1 \leq j < i} a_j$ then we can find k_i (which is last bit of $k = (k_1, k_2, \dots, k_i)$).

We can find $(k_1, \dots, k_n)$ by following the procedure mentioned below. We will start from $i = n$:-

1. Find k_i using Theorem 2.1.
2. $m = m - k_i * a_i$.
3. Now we know, $m = \sum_{j=1}^{i-1} a_j k_j$. Hence we apply Theorem 2.1 on $(a_1, a_2, \dots, a_{i-1}, m)$ and solve for k_{i-1}.
4. $i = i - 1$
5. Now we will repeat the same procedure from step 1 till we get the value of k_1 .

Hence, we can find k without actually solving SUBSET-SUM problem.

Question 3

Having falied to arrive at a secret key as above, Anubha and Braj try another method. Let G be the group of $n \times n$ invertible matrices over field F, $n = 128$. Let $a, b, g \in G$ such that $ab \neq ba$. The group G and the elements a, b, g are publicly known. Anubha and Braj wish to create a shared secret key as follows:

> Anubha chooses integers ℓ, m randomly with $1 < \ell, m \leq 2^n$, and sends $u = a^\ell g b^m$ to Braj. Braj chooses integers r, s randomly with $1 < r, s \leq 2^n$, and sends $v = a^r g b^s$ to Anubha. Anubha computes $k_a = a^\ell v b^m = a^{\ell+r} g b^{m+s}$. Braj computes $k_b = a^r u b^s = a^{\ell+r} g b^{m+s}$. The secret key is thus $k = k_a = k_b$.

Show that even this attempt fails as Ela can find k using u and v.

Hint: Show that Ela can

1. find elements x and y such that $xa = ax$, $yb = by$, and $u = xgy$,
2. use x, y, and v to compute k.

Solution

Task : To prove that we can find k (key), as mentioned in the question, without actually determining l, m, r, s.

Given : We know matrices a, b, g, u, v.
a, b, $g \in G$, where G is a group of $n \times n$ invertible matrices, where $ab \neq ba$.

Procedure :
Claim : $u, v \in G$.
Proof : It is sufficient to prove that u and v are $n \times n$ invertible matrices.

Since a, b, g are $n \times n$ dimentional matrices so the dimension of u, v will also be $n \times n$ as $u = a^\ell g b^m$ and $v = a^r g b^s$.

Since $u = a^\ell g b^m$
$\implies \det(u) = \det(a^\ell g b^m)$
$\implies \det(u) = (\det(a))^\ell \det(g)(\det(b))^m$ $\qquad$ [$\because a^l, g, b^m$ are all $n \times n$ matrices]

$\implies \det(u) \neq 0$ [Given that $a, b, g \in G \implies \det(a) \neq 0$, $\det(b) \neq 0$, $\det(g) \neq 0$]
$\implies u$ is Invertible

Similarly we can prove that v is also invertible

Part 1. We can find matrices x, y, such that $xa = ax$, $yb = by$, $u = xgy$.

Proof.
Claim 1 : x, y are invertible matrices.
Proof of claim 1 :
Since $u = xgy$
$\implies \det(u) = \det(xgy)$
$\implies \det(u) = \det(x)\det(g)\det(y)$ [$\because x, g, y$ are all $n \times n$ matrices.]
$\implies \det(x)\det(y) = \frac{\det(u)}{\det(g)}$
$\implies \det(x)\det(y) \neq 0$ [$\because$ g,u$\in G \implies \det(g) \neq 0$, $\det(u) \neq 0$]
$\implies \det(x) \neq 0$ and $\det(y) \neq 0$
$\implies x$ and y is Invertible

To find : x and y

$$xa = ax \tag{3.1}$$
$$yb = by \tag{3.2}$$
$$u = xgy \tag{3.3}$$

Equation 3.1 and Equation 3.2 are linear equations, but Equation 3.3 is not linear. We can convert Equation 3.3 into a linear equation.
$\because x$ is invertible, pre-multiply both sides of Equation 3.3 with x^{-1}.

$$x^{-1}u = gy \tag{3.4}$$

Let $z = x^{-1}$, z is a $n \times n$ matrix ,

$$zu = gy \tag{3.5}$$

Also pre-multiply and post-multiply both sides of Equation 3.1 by x^{-1}

We get,

$$\begin{aligned} ax^{-1} &= x^{-1}a \\ \implies az &= za \end{aligned} \tag{3.6}$$

Now we have three linear equations in y and z, namely

$$\begin{aligned} yb &= by && \textit{Equation } 3.2 \\ za &= az && \textit{Equation } 3.6 \\ zu &= gy && \textit{Equation } 3.5 \end{aligned}$$

We can solve these equations by substituion.
Post-multiply both sides of Equation 3.5 by u^{-1}, we get

$$z = gyu^{-1} \qquad [\because u \text{ is invertible}] \tag{3.7}$$

Substitute value of z into Equation 3.6.

$$\implies gyu^{-1}a = agyu^{-1} \tag{3.8}$$

Solving, Equation 3.8 and Equation 3.2 for y.
In this case we will have n^2 variables and $2n^2$ equations. We can find y using these equations.

Once we get y, we can find z using Equation 3.7.
We can get x, by simply inverting z. $[\because x^{-1} = z \implies z^{-1} = x]$

Hence we have found x and y using the provided information.

□

Part 2. If we have x and y, such that $ax = xa$, $by = yb$, $u = xgy$, then we can find key, $k = xvy$.

Proof. We have x and y, such that, $xa = ax$, $yb = by$ and $u = xgy$

Also, $u = a^l g b^m$ and $v = a^r g b^s$

$$v = a^r g b^s \tag{3.9}$$

Claim 2 : If $xa = ax$, then $xa^p = a^p x$, where p is a positive integer.
Proof for claim 2:
$xa^p = xaa^{p-1} = axa^{p-1} = axaa^{p-2} = a^2 xa^{p-2} = \ldots = a^{p-1}xa = a^p x \qquad \because xa = ax$
Hence proved.

Similarly we can prove that $b^p y = yb^p$, where p is positive integer.

Pre multiply by x and post multiply by y on both sides of Equation 3.9
$xvy = xa^r g b^s y$
$\implies xvy = a^r xgyb^s \qquad [\because xa^r = a^r x$ and $b^s y = yb^s$, as proved in the above claim.]
$\implies xvy = a^r u b^s \qquad [\because xgy = u]$
$\implies xvy = k, \qquad [\because k = xvy]$
here k is the key that is needed to be broken.

□

Hence, we can get the value of k without getting to know about l, m, r, s. Therefore, this method of sharing the key fails too.

References

[1] *Lecture-6 , Slide 6,7,8,9*.