Cryptography and Network Security Homework 1

Jacques Becker

September 23, 2018

Problem 1: For the simplified DES, consider Sbox S_0 and show how a DiffCrypto attack would work. Show your work for partial credit.

Begin by computing the differential distribution table for S_0. Since S_0 takes a 4-bit number as input, simply take every pair of 4-bit numbers x and $x*$, compute $S_0(x)$ and $S_0(x*)$, then compute $x' = x \oplus x*$ and $y' = S_0(x) \oplus S_0(x*)$. Use the table to keep track of the frequency with which x' and y' occur. I used a python script to automate this process. Here is the result:

	0	1	2	3
0	16	0	0	0
1	0	2	10	4
2	0	10	6	0
3	2	4	0	10
4	2	4	8	2
5	10	0	4	2
6	0	2	2	12
7	4	10	2	0
8	2	4	8	2
9	8	2	2	4
10	4	2	2	8
11	2	8	4	2
12	8	2	2	4
13	2	4	8	2
14	2	8	4	2
15	4	2	2	8

Consider the input $x' = 4$. Notice that $x' = 4$ produces $y' = 0$ only twice. This means that there exists a unique pair of numbers $(x, x*)$ such that $x \oplus x* \to y'$. This occurs twice in the table because $x \oplus x* \leftrightarrow x * \oplus x$. From the differential analysis table, we find that $9 \oplus 13 = 4$, and that $S_0(9) \oplus S_0(13) \Leftrightarrow 3 \oplus 3 = 0$. Now, we must determine the key. To do so, create a table listing the values that $x' = 4$ can produce from S_0, and the inputs x and $x*$ which can produce the result. Here is the table for $x' = 4$:

$4 \to 0$	$(9, 13)$
$4 \to 1$	$(8, 12), (10, 14)$
$4 \to 2$	$(0, 4), (1, 5), (2, 6), (3, 7)$
$4 \to 3$	$(11, 15)$

For convenience, I have grouped pairs of numbers whose XOR is 4. Because we know the inputs and outputs of the S-box, we can now start to reduce the number of possible keys. Suppose we know that the output of the S-box is $y' = 2$, and that the inputs are 10 and 14 ($10 \oplus 14 = 4$). We know that the

input to the S-box will be $S_{0i} = S_{0e} \oplus S_{0k}$, so it follows that $S_{0k} = S_{0i} \oplus S_{0e}$. From this, we get that any of the following values can be the key:

$$\begin{array}{ll} 10 \oplus 0 = 10 & 14 \oplus 0 = 14 \\ 10 \oplus 4 = 14 & 14 \oplus 4 = 10 \\ 10 \oplus 1 = 11 & 14 \oplus 1 = 15 \\ 10 \oplus 2 = 8 & 14 \oplus 2 = 12 \\ 10 \oplus 6 = 12 & 14 \oplus 6 = 8 \\ 10 \oplus 3 = 9 & 14 \oplus 3 = 13 \\ 10 \oplus 7 = 13 & 14 \oplus 7 = 9 \end{array}$$

So, we know the key is in the set $\{8, 10, 11, 12, 14, 15\}$. To reduce the number of possible keys, let's assume that we know the output of the S-box is $y' = 3$, and that the inputs are 0 and 4. Repeating the process again, we know that the key must be one of the following:

$$\begin{array}{l} 0 \oplus 11 = 11 \\ 0 \oplus 15 = 15 \\ 4 \oplus 11 = 15 \\ 4 \oplus 15 = 11 \end{array}$$

So, we know the key must be in the set $\{11, 15\}$. Since the key must be in both the sets we have deduced, it follows that the key must be either 11 or 15. At this point, we have reduced the number of possible keys enough to exhaustively search through each of them in a timely manner, so the Differential Cryptanalysis attack on S_0 is complete.

Problem 2: Consider the following cryptosystem and compute $H(K|C)$

- $P = \{a, b, c\}$ with $P_p(a) = 1/3, P_p(b) = 1/6, P_p(c) = 1/2$
- $K = (k_1, k_2, k_3)$ with $P_k(k_1) = 1/2, P_k(k_2) = 1/4, P_k(k_3) = 1/4$
- $C = \{1, 2, 3, 4\}$

$e_{k_1}(a) = 1 \quad e_{k_1}(b) = 2 \quad e_{k_1}(c) = 2$
$e_{k_2}(a) = 2 \quad e_{k_2}(b) = 3 \quad e_{k_2}(c) = 1$
$e_{k_e}(a) = 3 \quad e_{k_3}(b) = 4 \quad e_{k_3}(c) = 4$

We know that $H(K|C) = H(K)+H(P)-H(C)$, and computing $H(X)$ where X is a random variable is done with $H(X) = \sum_{i=1}^{n} p(X = x_i)log_2(p(X = x_i))$. The probability distribution for the random variable P is given in the problem, so we can easily compute $H(P)$:

$$H(P) = -\sum_{i=1}^{3} p(P = p_1)log_2(p(P = p_1))$$

$$H(P) = -\left[\frac{1}{3}log_2\left(\frac{1}{3}\right) + \frac{1}{6}log_2\left(\frac{1}{6}\right) + \frac{1}{2}log_2\left(\frac{1}{2}\right)\right]$$

$$H(P) \approx -[-0.5283 - 0.4308 - 0.5]$$

$$H(P) \approx 1.4591$$

The probability distribution for K is also given in the problem, so we can compute $P(K)$ in the same manner:

$$H(K) = -\sum_{i=1}^{3} p(K = k_1)log_2(p(K = k_1))$$

$$H(K) = -\left[\frac{1}{2}log_2\left(\frac{1}{2}\right) + 2\frac{1}{4}log_2\left(\frac{1}{4}\right)\right]$$

$$H(K) = -[-0.5 - 1]$$

$$H(K) = 1.5$$

Before we can compute $H(C)$, we must first compute the probability distribution $P(C)$. We must compute $P_c(y) = \sum_{\{k:y\in C(k)\}} P_k(k) \cdot P_p(d_k(y))$ for every y in C.

$$P(1) = P_k(k_1) \cdot P_p(a) + P_k(k_2) \cdot P_p(c) = \frac{1}{6} + \frac{1}{8} = \frac{7}{24}$$

$$P(2) = P_k(k_1) \cdot P_p(b) + P_k(k_1) \cdot P_p(c) + P_k(k_2) \cdot P_p(a) = \frac{1}{12} + \frac{1}{4} + \frac{1}{12} = \frac{5}{12}$$

$$P(3) = P_k(k_2) \cdot P_p(b) + P_k(k_3) \cdot P_p(a) = \frac{1}{24} + \frac{1}{12} = \frac{3}{12}$$

$$PP(4) = P_k(k_3) \cdot P_p(b) + P_k(k_3) \cdot P_p(c) = \frac{1}{24} + \frac{1}{8} = \frac{1}{6}$$

Now that we have the probability distribution for $P(C)$, we can compute $H(C)$:

$$H(C) = -\sum_{i=1}^{4} p(C = c_1) log_2(p(C = c_1))$$

$$H(C) = -\left[\frac{7}{24} log_2\left(\frac{7}{24}\right) + \frac{5}{12} log_2\left(\frac{5}{12}\right) + \frac{3}{12} log_2\left(\frac{3}{12}\right) + \frac{1}{6} log_2\left(\frac{1}{6}\right)\right]$$

$$H(C) = -[-0.5185 - 0.5263 - 0.5 - 0.4308]$$

$$H(C) \approx 1.9765$$

Now that all the values have been computed, We can compute:

$$H(K|C) = H(K) + H(P) - H(C)$$

$$H(K|C) \approx 1.5 + 1.4591 - 1.9765$$

$$H(K|C) \approx 0.9826$$