Solution Sheet #11

Advanced Cryptography 2021

Solution 1 Decorrelation

Recall that the a-norm of the matrix M is defined as

$$\max_{x_1}\left\{\sum_{y_1}\max_{x_2}\left\{\sum_{y_2}|M_{(x_1,x_2),(y_1,y_2)}|\right\}\right\}.$$

We consider first the most internal maximum.

- When $x_1 = 0$ and $y_1 = 0$, we have $\max\{5+4, 6+8\} = 14$.
- When $x_1 = 0$ and $y_1 = 1$, we have $\max\{1+2, 0+4\} = 4$.
- When $x_1 = 1$ and $y_1 = 0$, we have $\max\{2+4, 10+0\} = 10$.
- When $x_1 = 1$ and $y_1 = 1$, we have $\max\{4+5, 0+1\} = 9$.

Hence, $\|M\|_a = \max\{14+4, 10+9\} = 19$.

Solution 2 Decorrelation and Differential Cryptanalysis

Let $a \neq 0$ and b be such that

$$\mathrm{EDP}^C_{\max} = E(\mathrm{DP}^C(a,b)).$$

As $\mathrm{DP}^C(a,b) \geq 0$ and as $\mathrm{DP}^C(a,0) = 0$, we can assume that $b \neq 0$. We consider the distinguisher described in Algorithm 1. This distinguisher is limited to two queries. Its advantage must thus be less than $\mathrm{BestAdv}_{\mathrm{Cl}^2_a}(C, C^*)$.

We now look for an expression of this advantage. When the oracle implements C, the probability that the distinguisher outputs 1 is $\mathrm{E}(\mathrm{DP}^C(a,b))$. When it implements C^*, the probability that it outputs 1 is $1/(2^m-1)$ since y_1 and y_2 are different random elements and $b \neq 0$. Therefore, the advantage of the distinguisher is equal to

$$\mathrm{E}(\mathrm{DP}^C(a,b)) - \frac{1}{2^m-1}.$$

This leads to the inequality.

Hence, we deduce that studying the order two decorrelation of C is a good way to find an upper-bound on the best differential property and, thus, can be used to prove the resistance of a cipher against differential attacks.

Algorithm 1 A differential distinguisher between C and C^*

Input: an oracle $\mathcal{O}$ implementing either C or C^*, two masks a and b such that $a \neq 0$ and $b \neq 0$
Output: 0 (if the guess is that $\mathcal{O}$ implements C^*) or 1 (if the guess is that $\mathcal{O}$ implements C)
Processing:

pick x uniformly at random
submit x and $x \oplus a$ to $\mathcal{O}$ and get y_1 and y_2
if $y_1 = y_2 \oplus b$ **then**
  output 1
else
  output 0
end if

Solution 3 Decorrelation (2)

1. By definition,
$$|||[C]^d - [C^*]^d|||_\infty = 2 \cdot \mathrm{Adv}_{\mathsf{Cl}^d_{\mathsf{na}}}.$$
As this “measure” represents the advantage of the best non-adaptive distinguisher using d queries, it is rather clear that
$$|||[C]^{d-1} - [C^*]^{d-1}|||_\infty \leq |||[C]^d - [C^*]^d|||_\infty$$
since the best non-adaptative distinguisher using $d-1$ queries can be considered as a non-adaptative distinguisher using d queries, including one which is not taken into account.

2. By definition, an advantage is given by
$$|\Pr[\mathcal{A}^C \to 1] - \Pr[\mathcal{A}^{C^*} \to 1]|.$$
As a probability measure returns always a result in the interval $[0, 1]$, we have
$$|\Pr[\mathcal{A}^C \to 1] - \Pr[\mathcal{A}^{C^*} \to 1]| \leq 1$$
which implies that
$$|||[C]^d - [C^*]^d|||_\infty \leq 2.$$
Furthermore, as $|||.|||_\infty$ is a norm, we have
$$|||[C]^d - [C^*]^d|||_\infty \geq 0.$$

3. The property $\mathrm{Dec}^d(C) = 0$ means that the distance between $[C]^d$ and $[C^*]^d$ is zero. By definition of a distance, this happens if and only if $[C]^d = [C^*]^d$. Obviously this does not depend on the choice of the distance.

4. The above property with $d = 1$ means that $[C]^1 = [C^*]^1$. The coefficient of these matrices are the probabilities $\Pr[C(x) = y]$. Therefore, this property means that for any x and y, we have
$$\Pr[C(x) = y] = \Pr[C^*(x) = y].$$
Since $\Pr[C^*(x) = y] = 2^{-m}$, the property means that for any x and y we have $\Pr[C(x) = y] = 2^{-m}$. In this case we can prove that we have perfect secrecy.

For any x and y, we have

$$\Pr[X = x|C(X) = y] = \frac{\Pr[X = x]}{\Pr[C(X) = y]} \Pr[C(x) = y].$$

The probability $\Pr[C(X) = y]$ can be computed as follows

$$\Pr[C(X) = y] = \sum_{x'} \Pr[C(x') = y|X = x'] \Pr[X = x'].$$

Since C and X are independent, we have

$$\Pr[C(x') = y|X = x'] = \Pr[C(x') = y] = 2^{-m}.$$

Thus $\Pr[C(X) = y] = 2^{-m}$. Therefore we obtain that

$$\Pr[X = x|C(X) = y] = \Pr[X = x]$$

for any distribution of X.

5. $\mathrm{Dec}^d(f_K) = 0$ means that for any pairwise different $x_1, \ldots, x_d$ and any $y_1, \ldots, y_d$, we have $\Pr[f_K(x_i) = y_i \text{ for } i = 1, \ldots, d] = 2^{-md}$.

   Let us pick random pairwise different $x_1, \ldots, x_d$. We obtain that for any $y_1, \ldots, y_d$, the above probability is non-zero. This implies that there exists at least one key k such that $f_k(x_i) = y_i$ for all $i = 1, \ldots, d$. Therefore we must have at least 2^{md} keys, i.e., K must at least have a bit length of md. The purpose of the exercise is to show how to achieve this minimal key size.

6. For any $x, y \in \{0,1\}^m$ we have

   $$[f_K]^1_{x,y} = \Pr[f_K(x) = y] = \Pr[K = x \oplus y] = 2^{-m} = [F^*]^1_{x,y}.$$

   Therefore $[f_K]^1 = [F^*]^1$ which clearly implies that f_K is at distance 0 from F^*, i.e.,

   $$\mathrm{Dec}^1(f_K) = 0.$$

   We notice that we achieve the minimal length for the key here.

7. We take $K = (K_1, \ldots, K_d) \in (\mathrm{GF}(2^m))^d$ (which achieves the minimal length). We define $f_K(x) = K_1 + K_2 x + K_3 x^2 + \ldots + K_d x^{d-1}$ in the sense of $\mathrm{GF}(2^m)$ operations. For pairwise different $x_1, \ldots, x_d$ and any $y_1, \ldots, y_d$, we can find a unique polynomial P such that $P(x_i) = y_i$ by interpolation. The coefficients of this polynomial define a unique key K such that the polynomial is actually f_K. This proves that $\Pr[f_K(x_i) = y_i \text{ for } i = 1, \ldots, d] = 2^{-md}$.