TD 4 – Hash functions

Exercise 1. *Compression functions*

Let $E : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ a block cipher with key size and block size n. We define two compression functions $f_1, f_2 : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ by $f_1(h,m) = E_h(m) \oplus h$ and $f_2(h,m) = E_m(h) \oplus h$ (that is, f_2 is obtained using the Davies-Meyer construction).

1. Describe and analyze the complexity of a first preimage attack against f_1: given t *and* h, it computes m such that $f_1(h,m) = t$.
2. Explain why the previous attack does not apply to f_2. *Which supposedly hard problem on the block cipher does it require to solve?*

Exercise 2. *Rabin's hash function*

Let $E : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ be a block cipher where the block size is the same as the key size. Define the compression function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ by $f(h,m) = E_m(h)$, that is, m is used as key and h as block in the block cipher.

1. Prove that given h^* and m, one can compute h such that $f(h,m) = h^*$.

Given a message $m = m_1 \| \cdots \| m_k \in \{0,1\}^{kn}$, let h_0 be some initialization vector, and $h_i = f(h_{i-1}, m_i)$ for all $i > 0$. Let $H(m) = h_k$. (Note that the value of $H(m)$ depends on h_0.)

2. Prove that given h^*, an adversary can compute an initialization vector h_0 and a k-block message m such that $H(m) = h^*$.

We now assume that h_0 is known but fixed, not chosen by the adversary. We describe an attack that still allows an adversary to find an m such that $H(m) = h^*$. The adversary samples $2k$ blocks $m_1, \ldots, m_k, m'_1, \ldots, m'_k$. For $i = 1$ to k, it computes $h_i = f(h_{i-1}, m_i)$ on the one hand, and h'_i such that $f(h'_i, m'_i) = h'_{i-1}$ where $h'_0 = h^*$ on the other hand.

3. Assume that there exists i and j such that $h_i = h'_j$. Prove that $m = m_1 \| \cdots \| m_i \| m'_j \| \cdots \| m'_1$ satisfies $H(m) = h^*$.

To fully specify the attack and estimate its complexity, we need to find appropriate values for k. The goal is that the probability that there exists i and j such that $h_i = h'_j$ is large enough. **In the rest of the exercise, we assume that f behaves as a random function.**

4. For $i, j \in \{1, \ldots, k\}$, let $X_{ij} = 1$ if $h_i = h'_j$ and 0 otherwise.

- **i.** Compute $\mathbb{E}[X_{ij}]$.
- **ii.** Compute the expectation of the number of pairs (i, j) such that $h_i = h'_j$. *Hint. Use the linearity of expectation.*
- **iii.** Deduce a value of k for which the attack is probably successful.

5.

- **i.** What is the complexity of the generic first preimage attack for hash functions? Is it much larger than the complexity of the attack presented here?
- **ii.** Suppose that we want a collision-resistant hash function. Has the previous attack consequences on the choice of the security parameter n?