Cryptology complementary
Exercises#2

2020-03

Exercise 1: Hash functions *(CS Exam '19)*

In the following questions, $\mathcal{H} : \mathcal{I} \to \{0,1\}^n$ is a cryptographic hash function, where $\mathcal{I} = \bigcup_{\ell=0}^{2^N} \{0,1\}^\ell$. We recall the two following definitions:

— A *second preimage attack* on $\mathcal{H}$ is an algorithm that on input $m \in \mathcal{I}$ returns $m' \neq m \in \mathcal{I}$ s.t. $\mathcal{H}(m') = \mathcal{H}(m)$.

— A *collision attack* on $\mathcal{H}$ is an algorithm that returns m, $m' \neq m \in \mathcal{I}$ s.t. $\mathcal{H}(m) = \mathcal{H}(m')$.

Q. 1:

1. Give an algorithm for a second preimage attack. What is its expected running time (in function of n) for a perfectly random function $\mathcal{H}$ (no justification is necessary)?
2. What is the average complexity of a collision attack for a perfectly random function $\mathcal{H}$?
3. Give the specifications of a hash function $\mathcal{H}' : \mathcal{I} \to \{0,1\}^n$ for which every pair of distinct messages forms a collision. Is it possible to efficiently find second preimages for this function?

We informally call a hash function $\mathcal{H}$ *preimage-resistant* (resp. *collision-resistant*) if there is no "efficient" (first or second) preimage attack (resp. collision attack) on $\mathcal{H}$.

Q. 2:

1. Show that an adversary having a black box access to an efficient second preimage attack can perform a "similarly efficient" collision attack[1]. Is the converse true?
2. Is it possible for a hash function to be collision-resistant but not preimage-resistant?
3. Let $\mathcal{H}$ be such that the best collision attack on it is a generic attack. What can you say about the complexity of preimage attacks on $\mathcal{H}$?

Exercise 2: Coupon collector's problem *(a.k.a.: "gotta catch em' all")*

Let $\mathcal{H} : \{0,1\}^* \to \{0,1\}^n$ be a random oracle.

[1] If this statement were expressed formally, what we want would be a reduction whose time complexity is polynomial in the inputs.

Q. 1: How many calls to $\mathcal{H}$ are expected to be necessary to "collect" all the 2^n possible outputs (i.e. so that one has found a preimage for all $x \in \{0,1\}^n$)?

Hint 1: Try first to express the probability that no preimage was found for a fixed (arbitrary) image, and extend this to the entire co-domain.

Hint 2: We give the following approximation: $\lim_{x\to\infty}(1-\frac{1}{x})^x = e^{-1}$.