Merkle-Damgård Transformation

Merkle-Damgård Transformation: Let $h : \{0,1\}^{n+t} \to \{0,1\}^n$ be a fixed input length compression function. Recall that using the Merkle-Damgård transformation we can construct a hash function $H : \{0,1\}^* \to \{0,1\}^n$ from h as follows:

- Let x be the input.
- Let y_0 be an $n-$bit IV.
- Let $x_{k+1} = L$, where $L = |x|$ written as a t bit binary string.
- Split x into pieces $x_1, x_2, \ldots, x_k$, where each x_i is t bits. The last piece x_k should be padded with zeroes if necessary.
- For $i = 1$ to $k+1$, set $y_i = h(y_{i-1}||x_i)$.
- Output y_{k+1}.

Claim 1. *H is collision resistant, if h is collision resistant.*

Proof. Let us assume for the sake of contradiction that H is not collision resistant. Then there exists a PPT adversary $\mathcal{A}$ who can find a pair x, x', where $x \neq x'$ such that $H(x) = H(x')$ with a non-negligible probability. We will now construct another PPT adversary $\mathcal{B}$ who can break the collision resistance of h. This adversary $\mathcal{B}$ internally runs $\mathcal{A}$ as follows:

- Let x, x' be a collision returned by $\mathcal{A}$ in H.
- $\mathcal{B}$ defines $x_1, \ldots, x_{k+1}$, $y_0, \ldots, y_{k+1}$ and $x'_1, \ldots, x'_{k'+1}$, $y'_0, \ldots, y'_{k'+1}$ as in the Merkle Damgård transformation (here k may or may not be equal to k').

$$\begin{aligned}(H(x) = H(x')) &\Rightarrow (y_{k+1} = y'_{k'+1})\\ &\Rightarrow h(y_k||x_{k+1}) = h(y'_{k'}||x'_{k'+1})\end{aligned}$$

  - If $|x| \neq |x'|$:

$$\begin{aligned}&x_{k+1} \neq x'_{k'+1}\\ \Rightarrow &y_k||x_{k+1} \neq y'_{k'}||x'_{k'+1}\end{aligned}$$

    $\mathcal{B}$ outputs $y_k||x_{k+1}$ and $y'_{k'}||x'_{k'+1}$ as a collision in h.

– If $|x| = |x'|$: For $i = k+1$ to 1, $\mathcal{B}$ checks if $y_{i-1}||x_i$ and $y'_{i-1}||x'_i$ is a collision in h. Since $x \neq x'$, $\mathcal{B}$ is guaranteed to find such an i. It outputs $y_{i-1}||x_i$ and $y'_{i-1}||x'_i$ as a collision in h.

Since $\mathcal{A}$ finds a valid collision x, x' with non-negligible probability, $\mathcal{B}$ can also find a collision in h with non-negligible probability. But since h is collision resistant, such an adversary cannot exist. Hence our assumption is wrong and H is collision resistant. $\square$