6.857 Recitation 3: Merkle-Damgard

TA: Leo de Castro

Friday February 22, 2019

Today

- More Merkle
  - Reversible length extension
  - Merkle-Damgard hash function construction
- Security of ML Review

1 Reversible Length Extension

We're going to start with a very useful trick to pad a message to desired length. It is often the case that our block ciphers require messages to have a length that is a multiple of the block size. We want to pad the length of our message a way such that another person who sees the padded message can know exactly what is the pad and what is the message.

We do this by taking the message m and appending a 1 to the end. We then add zeros until the message is the desired length. Our padded message becomes $m \,||\, 10^*$. Anyone who sees this message can know that the last 1 and all subsequent zeros are part of the pad. This is a nice trick that will be continually useful to us.

2 Merkle-Damgard Hash Function Construction

Very good notes on this construction can be found in last year's (2018) lecture 6 notes on pages 6-7: `http://courses.csail.mit.edu/6.857/2018/files/L06-hash-functions-II.pdf`.

I've included these notes below.

Hash function construction ("Merkle-Damgard" style)

- Choose output size d (e.g. $d = 256$ bits)
- Choose "chaining variable" size c (e.g. $c = 512$ bits)

  [Must have $c \geq d$; better if $c \geq 2 \cdot d$ ...]
- Choose "message block size" b (e.g. $b = 512$ bits)
- Design "compression function" f

  $$f : \{0,1\}^c \times \{0,1\}^b \rightarrow \{0,1\}^c$$

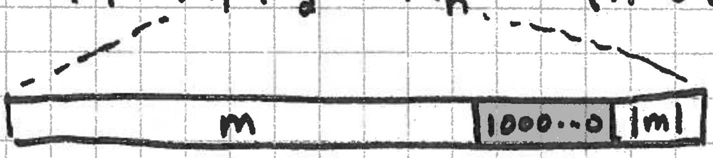
  [f should be OW, CR, PR, NM, TCR, ...]
- Merkle-Damgard is essentially a "mode of operation" allowing for variable-length inputs:

  * Choose a c-bit initialization vector IV, c_0

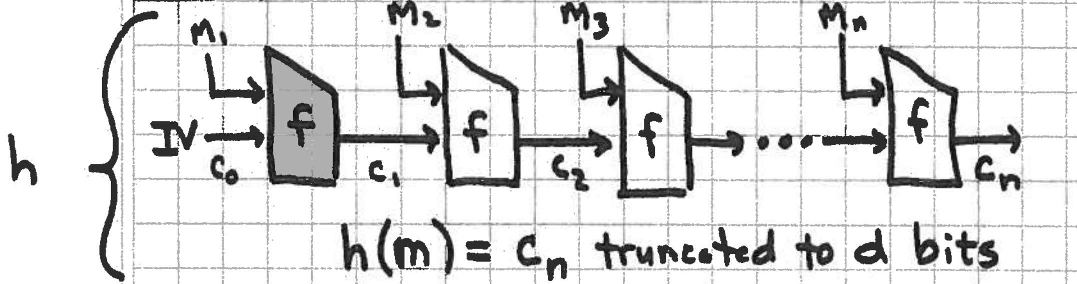
    [Note that c_0 is fixed & public.]
  * [Padding] Given message, append
    - 10^* bits
    - fixed-length representation of length of input

  so result is a multiple of b bits in length:

  $$M = M_1 \, M_2 \cdots M_n \qquad \text{(n b-bit blocks)}$$



Then:



$h(m) = c_n$ truncated to d bits

<u>Theorem:</u> If f is CR, then so is h.

<u>Proof:</u> Given collision for h, can find one for f by working backwards through chain. ▨

<u>Thm:</u> Similarly for OW.

<u>Common design pattern for f:</u>

$$f(c_{i-1}, M_i) = c_{i-1} \oplus E(M_i, c_{i-1})$$

where $E(K,M)$ is an encryption function (block cipher) with b-bit key and c-bit input/output blocks.

(Davies-Meyer construction)