Faster multicollisions

Jean-Philippe Aumasson

University of Applied Sciences Northwestern Switzerland
School of Engineering

Agenda

problem addressed in this talk:

computing multicollisions of iterated hash functions

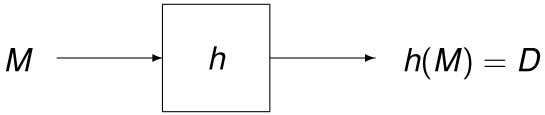
we'll start with definitions:

hash function, multicollision, iterated hash, fixed-point

then we'll describe multicollision attacks:

Joux's, fixed-point based, Kelsey/Schneier's, binary

Hash function

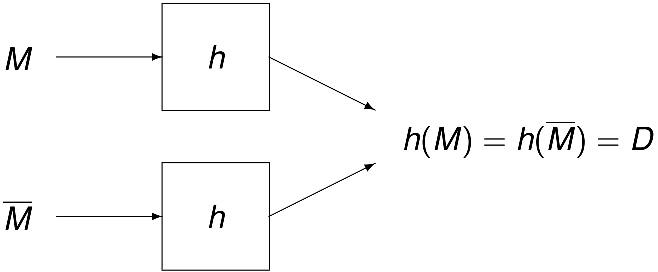


mapping $h : \{0, 1\}^{\star} \mapsto \{0, 1\}^{n}$

message $M \in \{0, 1\}^{\star}$ (arbitrary-length string)

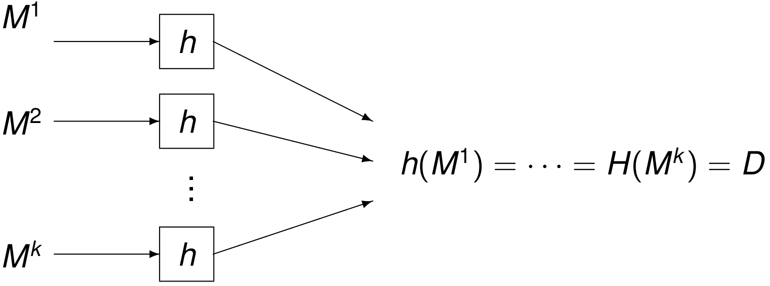
digest $D \in \{0, 1\}^{n}$ (fixed-length string)

Collision



pair of distinct messages $(M, \overline{M})$ such that $h(M) = h(\overline{M})$

Multicollision



$M^1, \ldots, M^k$ distinct

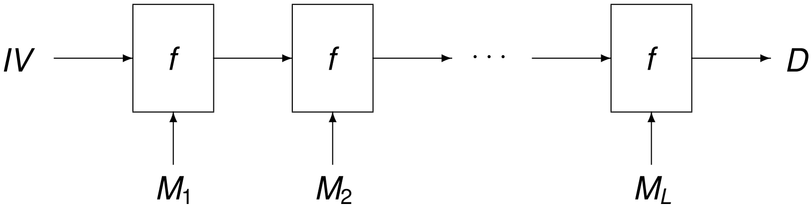
call $(M^1, \ldots, M^k)$ a <u>*k*-collision</u>

Iterated hash

parse $M = M_1 \| M_2 \| \ldots \| M_L$ into m-bit blocks

use a compression function $f : \{0,1\}^n \times \{0,1\}^m \mapsto \{0,1\}^n$

compute $h_{IV}(M) = D$ as



input of the intitial value (IV)

Collision oracle $\mathcal{O}_{\text{col}}$

IV

random $(M, \overline{M})$

such that $f(IV, M) = f(IV, \overline{M})$

given an IV, returns a random pair of colliding messages

Collision oracle $\mathcal{O}_{\text{col}}$

in practice, $\mathcal{O}_{\text{col}}$ models collision attacks

cost of a "query to $\mathcal{O}_{\text{col}}$":

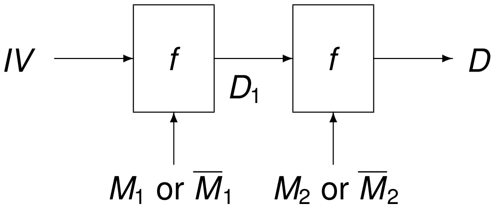
- $\approx 2^{n/2}$ if f is "ideal"
- 2^{63} for SHA-1
- 2^{128} for SHA-256
- 2^{256} for SHA-512
- a minute for MD5
- etc.

Joux's attack (example)

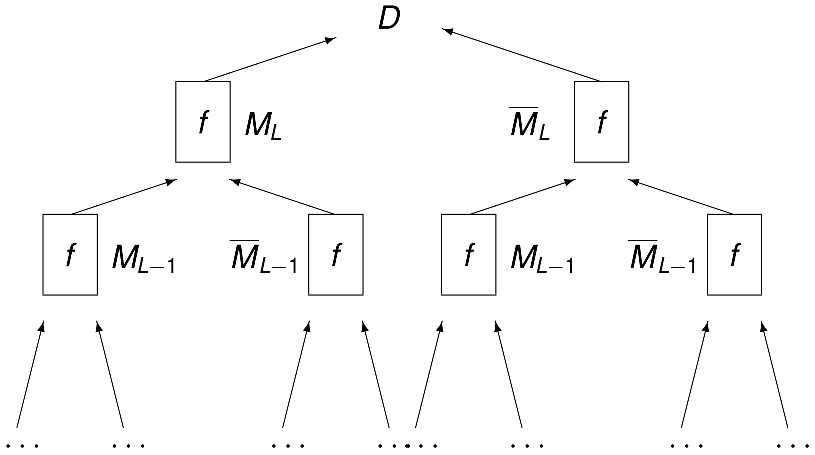
query $\mathcal{O}_{\text{col}}$ with IV, receive $f(IV, M_1) = f(IV, \overline{M}_1) = D_1$

query $\mathcal{O}_{\text{col}}$ with D_1, receive $f(D_1, M_2) = f(D_1, \overline{M}_2) = D$

4-collision $(M_1 \| M_2,\ M_1 \| \overline{M}_2,\ \overline{M}_1 \| M_2,\ \overline{M}_1 \| \overline{M}_2)$

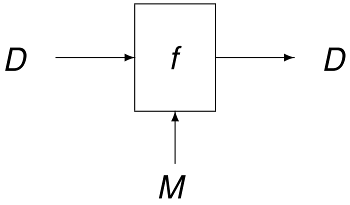


Joux's attack



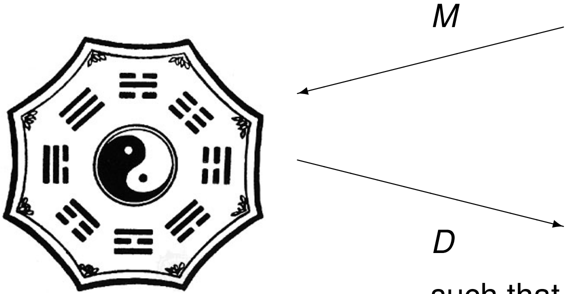
L queries to $\mathcal{O}_{\text{col}} \Rightarrow 2^L$-collision

Fixed point



D fixed point for M iff $f(D, M) = D$

Fixed point oracle $\mathcal{O}_{\text{fp}}$



such that $f(D, M) = D$

given M, returns a fixed point (deterministic!)

such oracle "exists" for MD5, SHA-1, SHA-256

Multicollision attack

query $\mathcal{O}_{\mathrm{fp}}$ to get a fixed-point $f(D, M) = D$

multicollision for h_D:

M

$M \| M$

$M \| M \| M$

$M \| M \| M \| M$

$M \| M \| \ldots \| M \| M$

intermediate values and digests all equal D

Strengthening the iterated hash

(a.k.a. MD-strengthening)

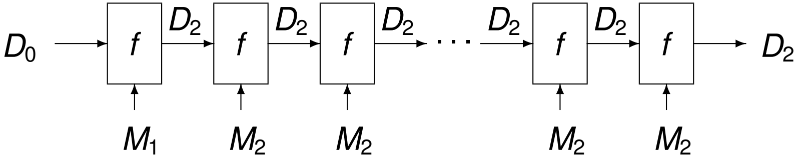
given message M apply iterated hash to

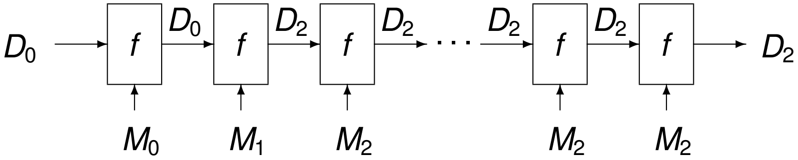
$$M \| \langle \text{bitlength of } M \rangle$$

- defeats the previous attack
- Joux's attack still possible
- collision resistance preservation:

  f collision resistant $\Rightarrow$ h collision resistant

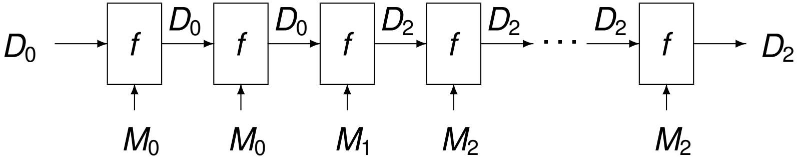
Kelsey/Schneier attack

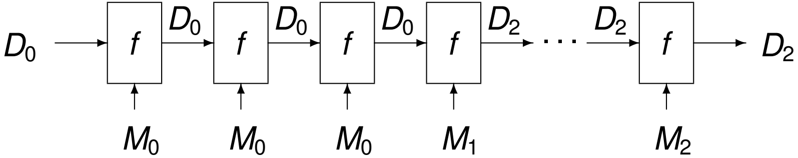




2 fixed points $f(D_0, M_0) = D_0$ and $f(D_2, M_2) = D_2$

Kelsey/Schneier attack





2 fixed points $f(D_0, M_0) = D_0$ and $f(D_2, M_2) = D_2$

Kelsey/Schneier attack

details:

- ask $\mathcal{O}_{\text{fp}}$ for a fixed point $f(D_0, M_0) = D_0$
- meet-in-the-middle to find M_1, M_2, D_2 such that

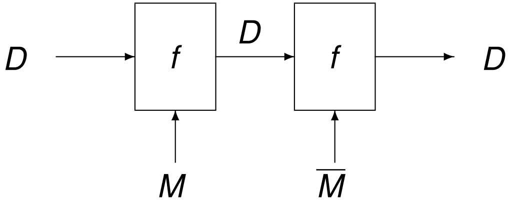
$$f(D_0, M_2) = D_2$$

- return messages of the form

$$M_0 \| M_0 \ldots \| M_0 \| M_1 \| M_2 \| M_2 \| \ldots \| M_j$$

with $\#blocks - 1 = (\#M_0 + \#M_2)$ constant

Fixed-point collision



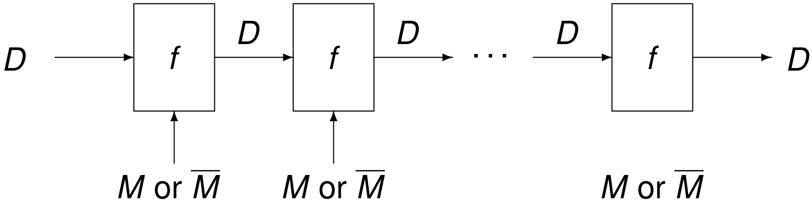
triplet $(D, M, \overline{M})$ such that

$$f(D, M) = f(D, \overline{M}) = D$$

by birthday paradox, can be found with $2^{n/2}$ queries to $\mathcal{O}_{\text{fp}}$

Binary multicollisions

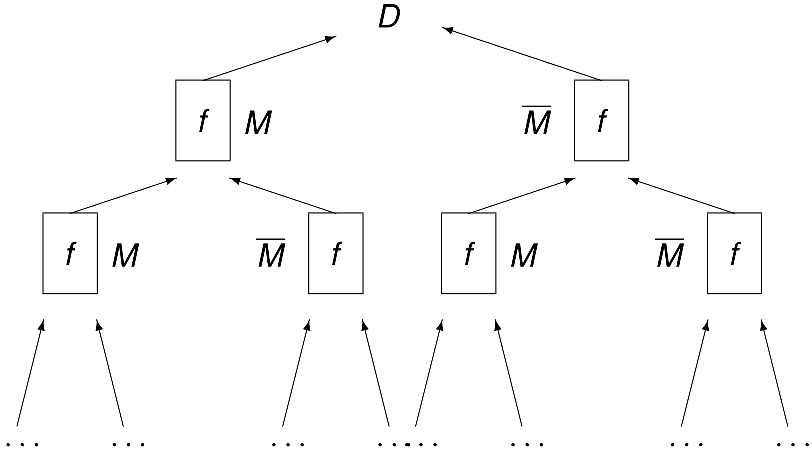
find a fixed-point collision $f(D, M) = f(D, \overline{M}) = D$



L blocks $\Rightarrow 2^L$-collision

costs $2^{n/2}$ queries to $\mathcal{O}_{\text{fp}}$ (independent of k!)

Binary multicollisions



for SHA-256, $2^{2^{10000}}$-collisions cost 2^{128} compressions

(2^{10128} with Joux's)

Cost of a k-collision

Joux

- $\log k$ queries to $\mathcal{O}_{\text{col}}$
  $\equiv \log k \cdot 2^{n/2}$ queries to f ideally

Kelsey/Schneier

- $2^{n/2}$ queries to f + $2^{n/2}$ queries to $\mathcal{O}_{\text{fp}}$
- memory $2^{n/2}$

Binary

- $2^{n/2}$ queries to $\mathcal{O}_{\text{fp}}$
  $\equiv 2^{n/2}$ queries to f for Davies-Meyer schemes
  $\Rightarrow$ optimal

Messages length

Joux: $\log k$

Kelsey/Schneier: $\approx k$

Binary: $\log k$

Summary

Binary multicollisions have

- optimal cost (same as a single collision)
- short messages

but...

- need easily found fixed points (as in SHA-256)
- only work for a chosen *IV*

More in the paper

variants on Kelsey/Schneier and “convergence” to Joux’s

application to concatenated hash functions

distinct-length multicollisions

countermeasures

More in the paper

variants on Kelsey/Schneier and “convergence” to Joux’s

application to concatenated hash functions

distinct-length multicollisions

countermeasures

Dhanyabad / Thank you !