Lecture 19: Preprocessing Attacks

MIT - 6.893
Fall 2020
Henry Corrigan-Gibbs

Plan

- Preprocessing generally
- Hellman/Rainbow tables
- What's next

Logistics

* Winter reading online
* AMA today 3pm-4:30pm
* Course evals
* Anonymous feedback form

Preprocessing Attacks

As far as I know, Study initiated by Hellman (1980) in regard to DES (56-bit key)

Key Recovery on DES block cipher: 56-bit key.

Problem: Given encryption of 2 known plaintexts, find corresponding ciphertext.

Given: $ct_0 = E(k, \text{"0000\cdots0"})$

$ct_1 = E(k, \text{"0000\cdots1"})$

for $k \xleftarrow{R} \{0,1\}^{56}$

Find: $K \in \{0,1\}^{56}$

Brute-force search of the keyspace takes $\approx 2^{56}$ time.

↳ In 1980, this was quite expensive

(now it's cheaper.. see crack.sh)

HOWEVER, attacker wants to decrypt many DES ctexts. It's a standard cryptosystem.

Idea:

- Attacker precomputes a data structure (takes $\approx 2^{56}$ time)
- Using this precomputed data structure, attacker can mount key-recovery in time $\ll 2^{56}$

→ If you want to decrypt many cts, this dramatically lowers cost (See: online cracking tools)

Why I love preprocessing attacks

(again inspired by Spielman's talk)

1. Beautiful theory
2. Work in practice
3. Solve a problem that people care about.

→ What more can you ask for? ←

Problem Statement

Given Function: $f : [N] \to [N]$

$f_{DES}(k) := \underbrace{E(k, \text{"0000}\cdots\text{0"})}_{\text{truncated to 56 bits ... ignoring some details...}}$ $\quad N = 2^{56}$.

Parameters: Space S, Time T

Preprocessing Phase

* Look at f
* Spend as much computation as you like
* Output an S-bit data structure.

Online Phase

* Given (a) your s-bit data structure
  (b) a value $y \in [N]$
* Make $\leq T$ queries to f.
* Output x s.t. $f(x) = y$.

How much space do you need?

Attack	S	T	(Ignoring $\log N$ factors!)
Brute force search in online phase	O	N	
Store all answers in a big lookup table ahead of time	N	O	
Hellman's attack	$N^{2/3}$	$N^{2/3}$	For function chosen at random from $[N] \to [N]$
Possible? No one knows... Very good Q...	$N^{1/2}$?	$N^{1/2}$?	Yao '90 shows that $N^{1/2}$ is best possible in this model.

Concretely, Hellman gives a data structure of size $S = 2^{40}$ that lets you break DES in time $T = 2^{40}$.

↳ Store a hundred GB of data and get a $64,000\times$ speed-up!

↳ Consequences for mass surveillance.

↳ Easy answer: use 256-bit keys, 256-bit block.

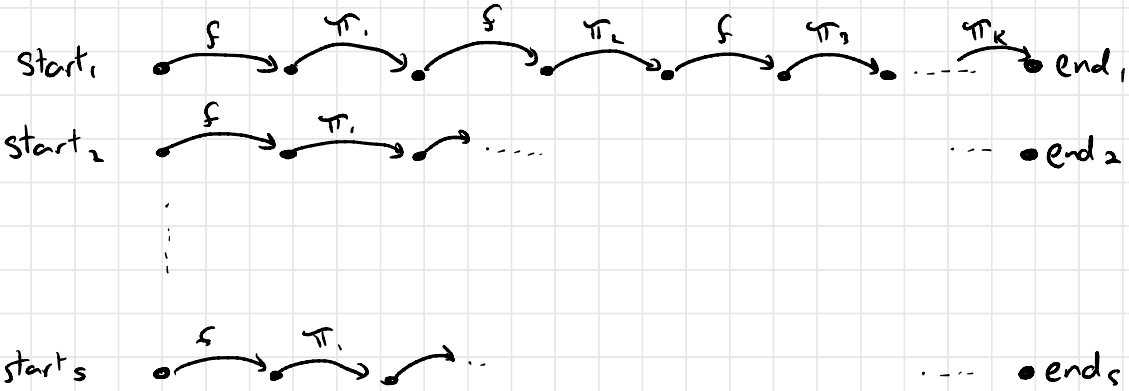
Hellman's Attack

+ practical improvement by Oechslin '03 ← we show.

Function $f : [N] \to [N]$

Random permutations $\pi_1, \ldots, \pi_k : [N] \to [N]$

Preprocessing Phase



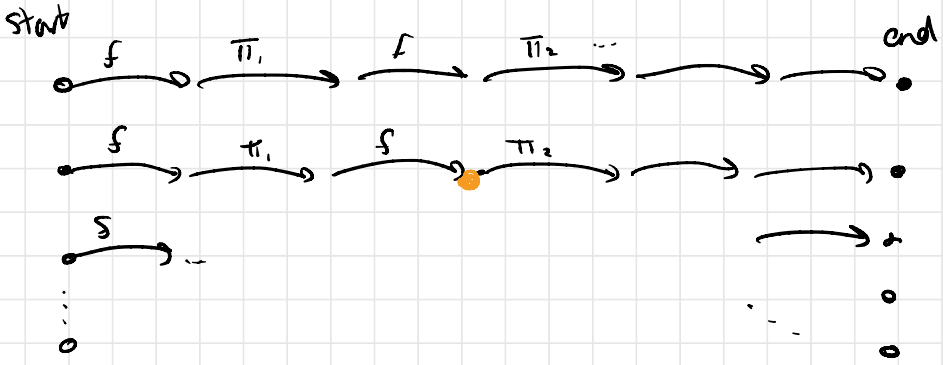
- Store all (start, end) pairs in table in $O(S \log N)$ bits.

- Idea: We take $S = N^{2/3}$ $k = N^{1/3}$ and hope that a good fraction of values in $[N]$ appear in the table.

Online Phase

Given: (start, end) pairs and $y \in [N]$

Goal: Find x s.t. $f(x) = y$.



Strategy

* Guess which column in table y is in ($k \approx N^{1/3}$ possibilities)
* Apply Π and f until reaching endpoint.
  - → If hit endpoint. Success! Can invert by following chain from start.
  - → If don't hit endpoint. Fail. Point y isn't in the table.

Running time: k evals of f $= k^2 = N^{2/3}$ evals of f.
k guesses

Nice property of Rainbow tables: few table lookups (aka, very easy to parallelize w/ k procs)

The only work is to show that the attack succeeds with good probability.

→ Just need to show that table covers a constant fraction of points.

Main idea

As you're building the ith chain, what's the probability that it "collides" with chains $\{1, \ldots, i-1\}$

$$\text{Probability} \approx \left[1 - \left(\frac{S}{N}\right)\right]^k \approx e^{-\frac{Sk}{N}}$$

↳ Taking $Sk = N$ makes this constant.

Great question: Is a better generic preprocessing attack on f inversion possible?

e.g. Data structure $S = 2^{64}$ that breaks AES 128 in time $T = 2^{64}$???

Looking forward

If you're interested in crypto & security...

Classes: 6.875, 6.857 ... advanced topics
6.045, 6.841 (complexity)
6.858 (security)
6.856J (randomized algs)

Harvard too! Law, HKS, ...
"Comparative Digital Privacy"
"Technology, Privacy, and the Trans-National Nature of the Internet"
"Fairness and Privacy" (Dwork)
"Legal Problems in Cybersecurity"

Conferences: Since they're all virtual, they're easy & cheap to access.

- Real-World Crypto (Jan 2021)
- IACR Eurocrypt (May 2021)
- IEEE Security & Privacy (May 2021)
- ⋮

Seminars: Crypto seminar, Security seminar, many more...

Research, summer, PhD, ...: Talk today in OH or email any time.

→ Be in touch! Hearing from my former students always makes my day.