Course Business

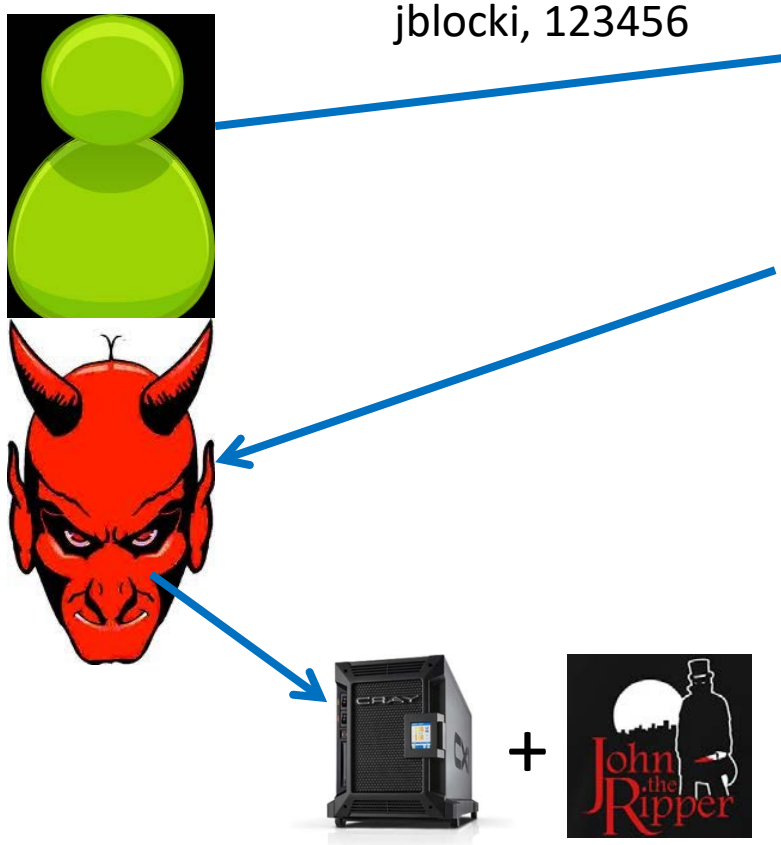
- I am traveling April 25-May 3rd
  - Will still be available by e-mail to answer questions
- Final Exam Review on Monday, April 24th
- Guest Lectures on April 26 and 28 (TBD)

- Final Exam on Monday, May 1st (in this classroom)
  - Adib will proctor
- Practice Final Exam released soon

Cryptography
CS 555

Topic 39: Password Hashing

Password Storage



Username	Salt	Hash
jblocki	89d978034a3f6	85e23cfe0021f584e3db87aa72630a9a2345c062

SHA1(12345689d978034a3f6)=85e23cfe0021f584e3db87aa72630a9a2345c062

Offline Attacks: A Common Problem

- Password breaches at major companies have affected millions of users.

Offline Attacks: A Common Problem

- Password breaches at major companies have affected millions of users.

A Dangerous Problem

Can we increase guessing costs for the attacker?

Attempt 1: Hash Iteration

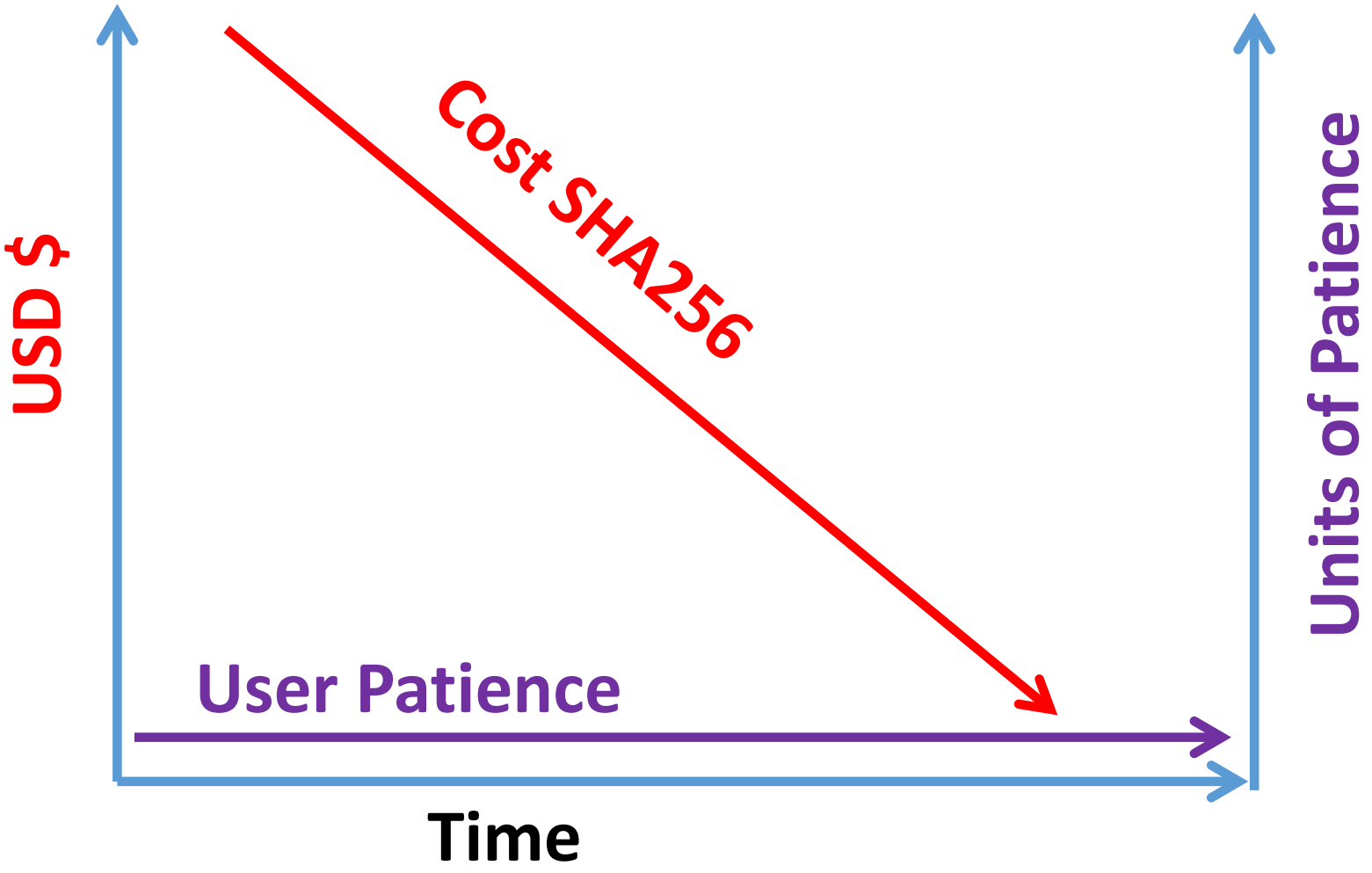
- BCRYPT

- PBKDF2

100,000 SHA256 computations (iterative)

Estimated Cost on ASIC: $1 per billion password guesses [BS14]

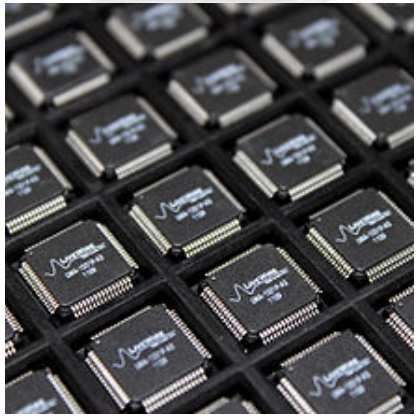
The Challenge



Disclaimer: This slide is entirely for humorous effect. Don’t take it too seriously

Goal: Moderately Expensive Hash Function

Fast on PC and
Expensive on ASIC?



Memory Costs: Equitable Across Architectures

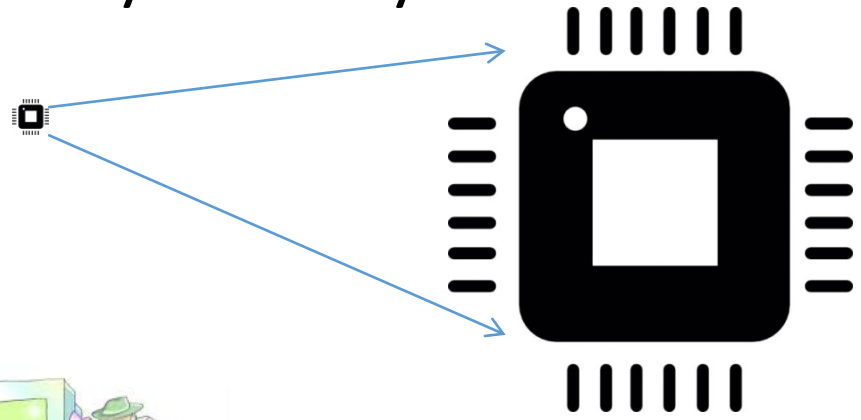
Outline

- Motivation
- **Data Independent Memory Hard Functions (iMHFs)**
  - Graph Pebbling
  - Measuring Pebbling Costs
  - Desiderata
- Attacks on iMHF Constructions
- Constructing iMHFs
- Open Questions

Memory Hard Function (MHF)

- Intuition: computation costs dominated by memory costs

vs.



sCrypt

- Data Independent Memory Hard Function (iMHF)
  - Memory access pattern should not depend on input

iMHF Candidates

- Catena [FLW15]
  - Special Recognition at Password Hashing Competition
  - Two Variants: Dragonfly and Double-Butterfly

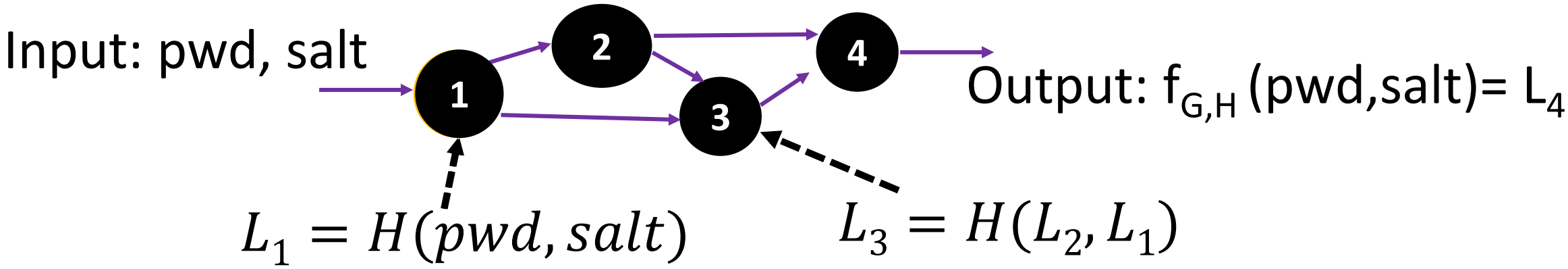
- Argon2 [BDK15]
  - Winner of the Password Hashing Competition
  - Argon2i (data-independent mode) is recommended for Password Hashing

- Balloon Hashing [BCS16]
  - Newer proposal (three variants in original proposal)

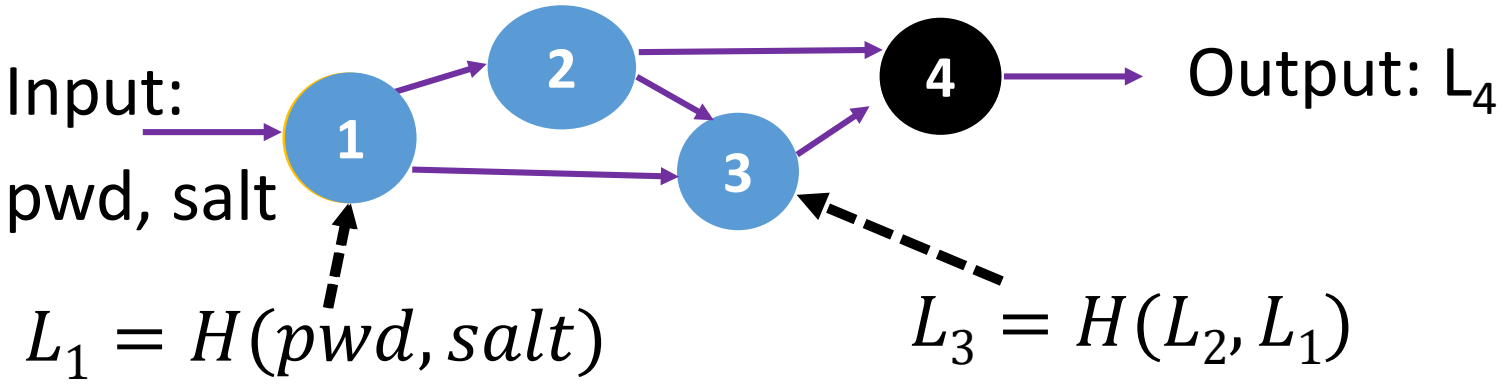
iMHF ($f_{G,H}$)

Defined by

- H:$\{0,1\}^{2k} \rightarrow \{0,1\}^{k}$ (Random Oracle)
- DAG G (encodes data-dependencies)
  - Maximum indegree: $\delta = O(1)$



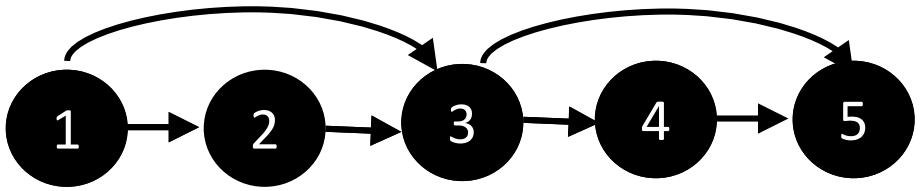
Evaluating an iMHF (pebbling)



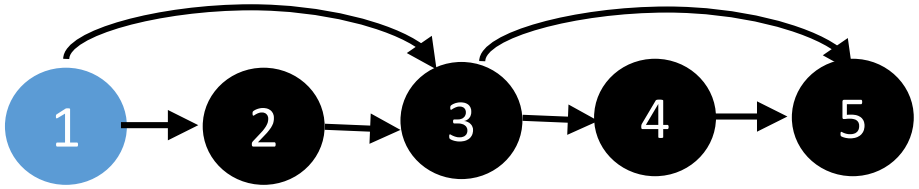
Pebbling Rules : $\vec{P}$=P_1,…,P_t⊂ V s.t.

- $P_{i+1} \subset P_i \cup \{x \in V | \text{parents}(x) \subset P_{i+1}\}$ (need dependent values)
- n∈ P_t (must finish and output L_n)

Pebbling Example

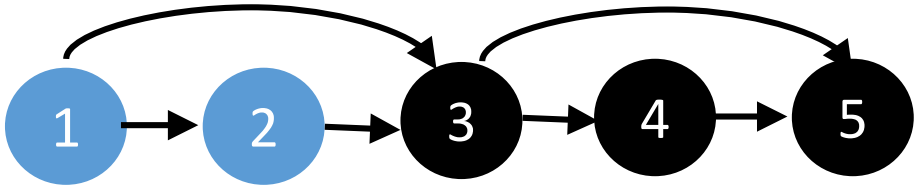


Pebbling Example



$P_1 = \{1\}$

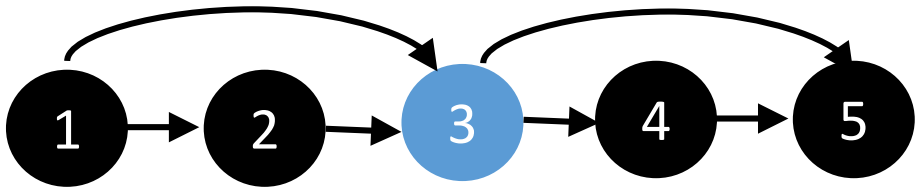
Pebbling Example



$P_1 = \{1\}$

$P_2 = \{1,2\}$

Pebbling Example

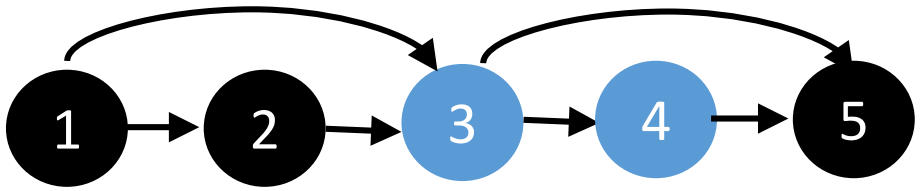


$P_1 = \{1\}$

$P_2 = \{1,2\}$

$P_3 = \{3\}$

Pebbling Example



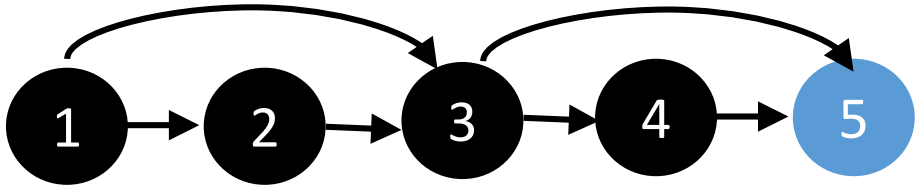
$P_1 = \{1\}$

$P_2 = \{1,2\}$

$P_3 = \{3\}$

$P_4 = \{3,4\}$

Pebbling Example



$P_1 = \{1\}$

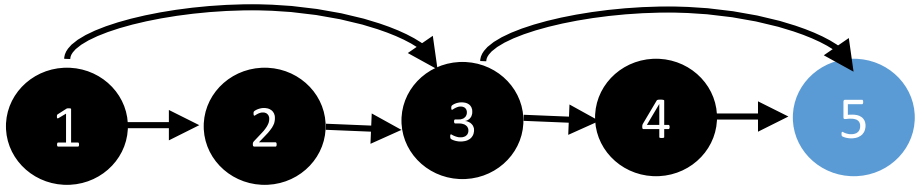
$P_2 = \{1,2\}$

$P_3 = \{3\}$

$P_4 = \{3,4\}$

$P_5 = \{5\}$

Pebbling Example (CC)



$P_1 = \{1\}$

$P_2 = \{1,2\}$

$P_3 = \{3\}$

$P_4 = \{3,4\}$

$P_5 = \{5\}$

$$\mathrm{CC}(G) \leq \sum_{i=1}^{5} |P_i|$$

$$= 1 + 2 + 1 + 2 + 1$$

$$= 7$$

Measuring Cost

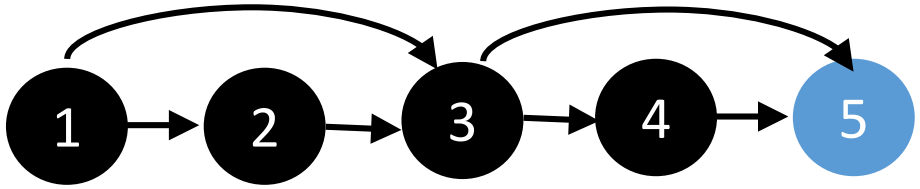
- Cumulative Complexity (CC)

$$\mathrm{CC}(G) = \min_{\vec{P}} \sum_{i=1}^{t_{\vec{P}}} |P_i|$$

- Amortization [AS15]

$$\mathrm{CC}(G, G) = 2 \times \mathrm{CC}(G)$$

Pebbling Example (CC)



$P_1 = \{1\}$

$P_2 = \{1,2\}$

$P_3 = \{3\}$

$P_4 = \{3,4\}$

$P_5 = \{5\}$

$$\mathrm{CC}(G) \leq \sum_{i=1}^{5} |P_i|$$

$$= 1 + 2 + 1 + 2 + 1$$

$$= 7$$

Pebbling Equivalence

Theorem [AS15] (Informal): High pebbling complexity of G implies high amortized memory complexity for the iMHF $f_{G,H}$.

Implication: Structure of the graph G is key to iMHF security

Desiderata

Find a DAG G on n nodes such that

1. Constant Indegree ($\delta = 2$)

2. CC(G) $\geq \frac{n^2}{\tau}$ for some small value τ.

Maximize costs for fixed n (Users are impatient)

Depth-Robustness: The Key Property

Necessary [AB16] and sufficient [ABP16] for secure iMHFs

Naïve Pebbling Algorithms

- Sequential Algorithm (Naïve)
  - Constraint: One new pebble per round
  - Every iMHF is defined via its Naïve algorithm

- Example Naïve (Pebble in Topological Order)
  - Never discard pebbles
  - Time: n
  - Average #pebbles: n/2.
  - $E_R(\text{Naïve}) = \theta(Rn + n^2)$

Amortized Attack Quality

$$\text{Quality}_R(A) = \frac{\text{E}_\text{R}(\text{Naïve})}{\text{E}_\text{R}(A)} \times \#inst(A)$$

Example: Algorithm A evaluates 5 iMHF instances with total cost $\text{E}_\text{R}(A) = 100$ and $\text{E}_\text{R}(\text{Naïve}) = 40$

$$\text{Quality}_R(A) = \frac{40}{100} \times 5 = 2$$

Desiderata

Find a DAG G and a sequential pebbling algorithm N with

1. Constant Indegree ($\delta = 2$)

2. $\text{Quality}_R(A) \leq c$ for every adversary A (c small).

3. $\text{E}_R(\text{Naive}) \geq \frac{n^2}{\tau} + \text{R}n$ for some small value τ.

Memory costs should dominate

Desiderata

Find a DAG G and a sequential pebbling algorithm N with

1. Constant Indegree ($\delta = 2$)

2. $\text{Quality}_R(A) \leq c$ for every adversary A (c small).

3. $E_R(\text{Naive}) \geq \frac{n^2}{\tau} + Rn$ for some small value τ.

Maximize costs for fixed n (Users are impatient)

c-Ideal iMHF

Find a DAG G and a sequential pebbling algorithm N with

1. Constant Indegree ($\delta = 2$)

2. $\text{Quality}_R(A) \leq c$ for every adversary A (c small).

3. $\text{E}_R(\text{Naive}) \geq \frac{n^2}{\tau} + \text{R}n$ for $\tau = O(1)$.

Outline

- Motivation
- Data Independent Memory Hard Functions (iMHFs)
- **Our Attacks**
  - General Attack on Non Depth Robust DAGs
  - Existing iMHFs are not Depth Robust
  - Ideal iMHFs don’t exist
- Subsequent Results (Depth-Robustness is Sufficient)
- Open Questions

Depth-Robustness: The Key Property

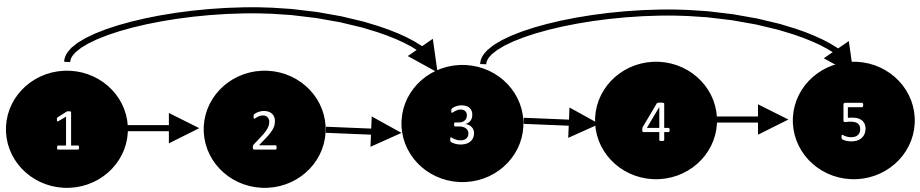
Necessary [AB16] and sufficient [ABP16] for secure iMHFs

Depth Robustness

Definition: A DAG G=(V,E) is (e,d)-reducible if there exists $S \subseteq V$ s.t. $|S| \leq e$ and depth(G-S) $\leq$ d.

Otherwise, we say that G is (e,d)-depth robust.

Example: (1,2)-reducible

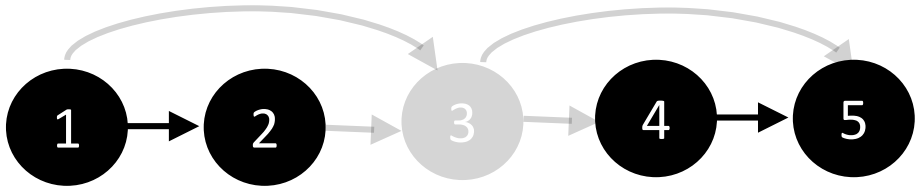


Depth Robustness

Definition: A DAG G=(V,E) is (e,d)-reducible if there exists $S \subseteq V$ s.t. $|S| \leq e$ and depth(G-S) $\leq$ d.

Otherwise, we say that G is (e,d)-depth robust.

Example: (1,2)-reducible



Attacking (e,d)-reducible DAGs

- Input: $|S| \leq e$ such that depth(G-S) = d, g > d

- **Light Phase (g rounds)**: Discard most pebbles!
  - **Goal:** Pebble the next g nodes in g (sequential) steps
  - Low Memory (only keep pebbles on S and on parents of new nodes)
  - Lasts a ``long" time

- **Balloon Phase (d rounds):** Greedily Recover Missing Pebbles
  - **Goal:** Recover needed pebbles for upcoming light phase
  - Expensive, but quick (at most d steps in parallel).

Depth Robustness is Necessary

Theorem (Depth-Robustness is a necessary condition): If G is not (e,d)-node robust then $\mathrm{CC}(G) = \mathrm{O}\left(en + \sqrt{n^3 d}\right)$.

In particular, $\mathrm{CC}(G) = \mathrm{o}(n^2)$ for e,d=o(n).

Are existing iMHF candidates based on depth-robust DAGs?

Answer: No

- Catena [FLW15] is $\left(e, \tilde{O}\left(\frac{n}{e}\right)\right)$-reducible

$$CC = O(n^{1.62})$$

- Balloon Hashing and Argon2i (old version) are $\left(e, \tilde{O}\left(\frac{n^2}{e^2}\right)\right)$-reducible

$$CC = O(n^{1.71})$$

- Argon2i (latest version) is $\left(e, \tilde{O}\left(\frac{n^3}{e^3}\right)\right)$-reducible

$$CC = O(n^{1.77})$$

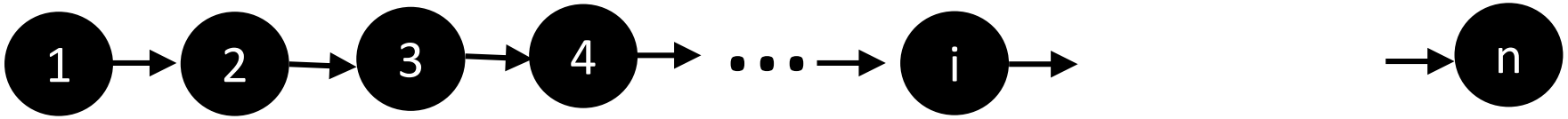
- Similar picture for most other iMHF candidates [AGKKOPRRR16]

Argon2i [BDK]

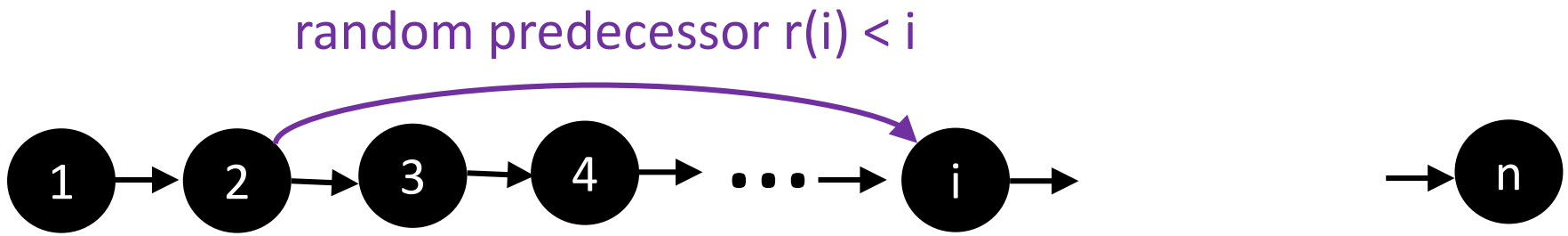
- Argon2: Winner of the password hashing competition[2015]

- Authors recommend Argon2i variant (data-independent) for password hashing.

Argon2i

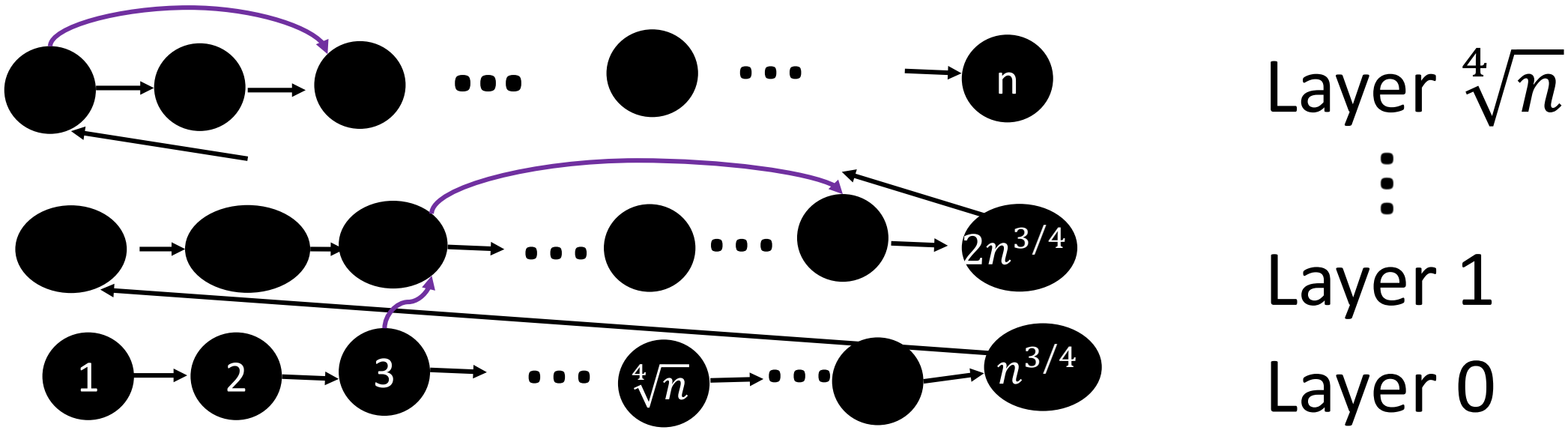


Argon2i



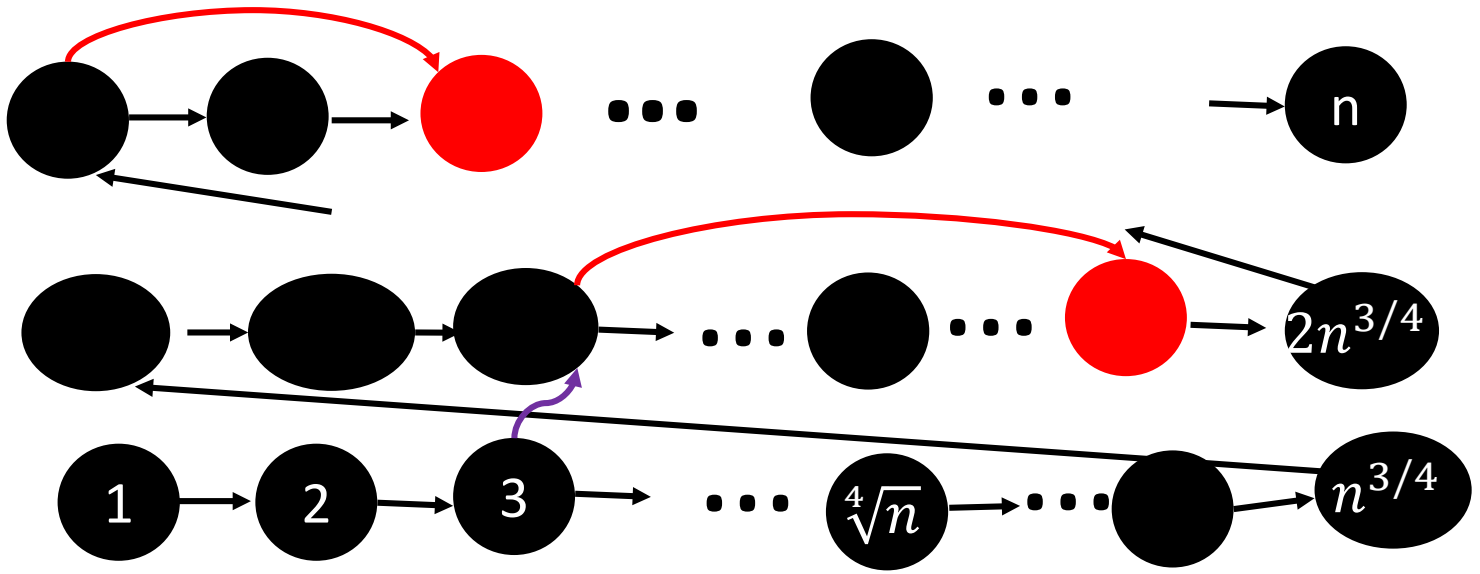
Indegree: $\delta = 2$

Argon2i: Reducing depth to $\sqrt{n}$



Argon2i: Reducing depth to $\sqrt{n}$

Definition: $S_2 = \{v_i | \; v_{r(i)} \text{and } v_i \text{ in same layer}\}$



Layer $\sqrt[4]{n}$

⋮

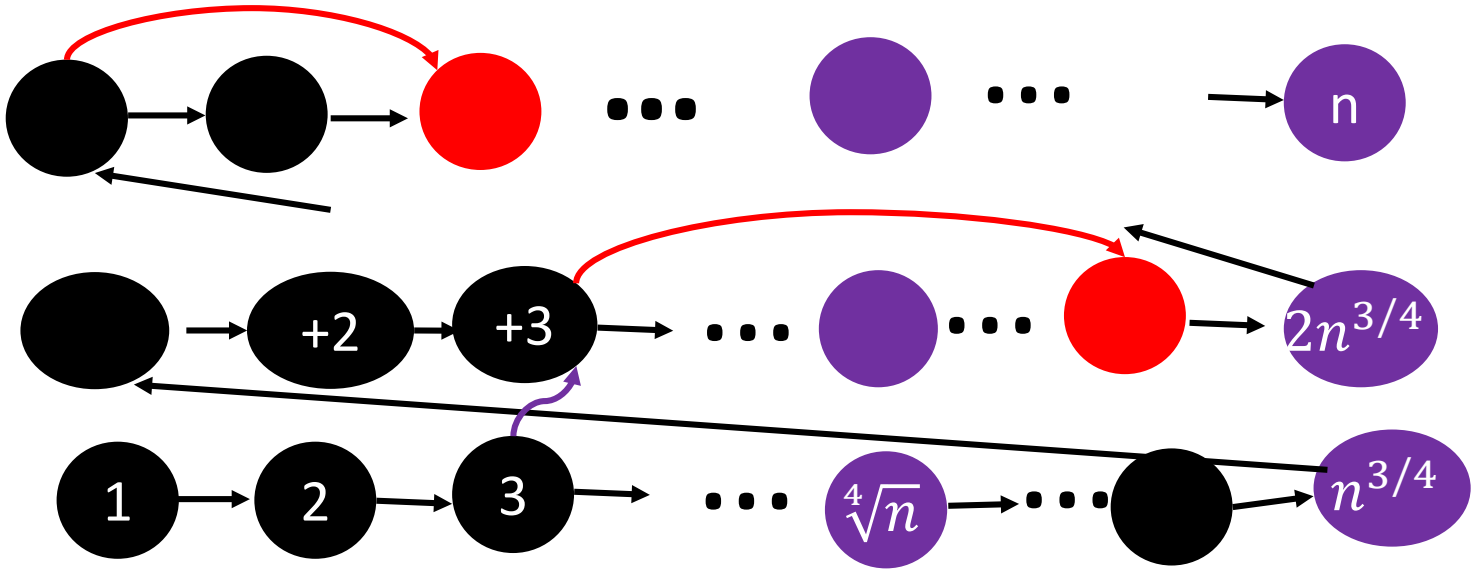
Layer 1

Layer 0

Claim: S_2 is small

Argon2i is a layered DAG (almost)

Let S = S_1+S_2



Layer $\sqrt[4]{n}$

⋮

Layer 1

Layer 0

Fact: Easy to reduce the depth of a path

Attack Simulation ⋮ [AB16b]

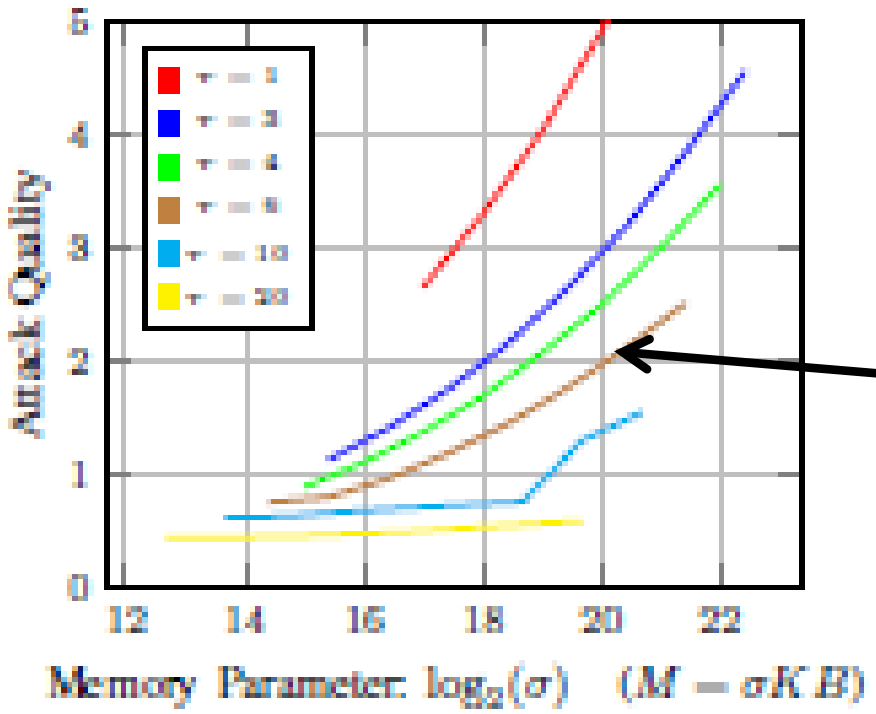


Figure 1: Argon2i-B Attack Quality

Parameter setting could easily be chosen when following Argon2i-B guidelines

Pessimistic Argon 2i-B parameter

Attack on Argon 2i-B is practical even for pessimistic parameter ranges (brown line).

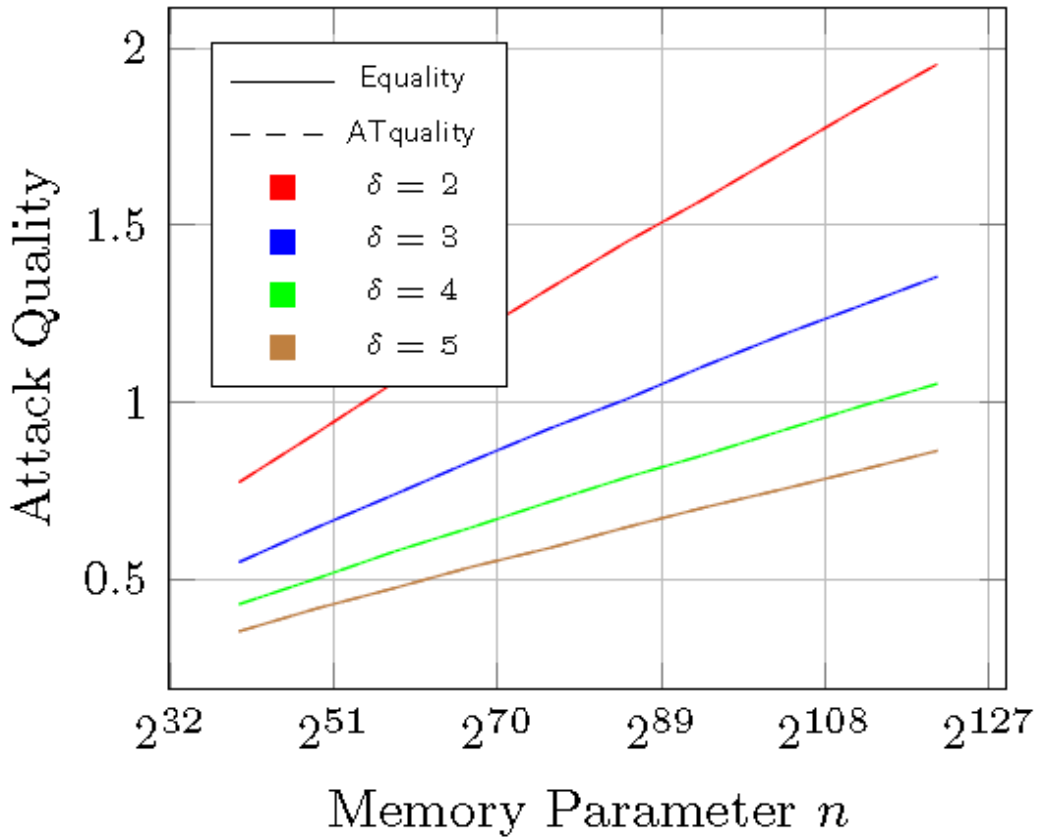
Ideal iMHFs Don't Exist

Thm[AB16]: Any graph G (with constant in-degree) is at least somewhat depth-reducible.

Implication: If CC(G)= $\Omega(n^2)$ there is an attack A with high quality:

$$\text{Quality}_{\text{R}}(A) = \Omega\left(\frac{\log(n)}{\log\log(n)}\right)$$

But, we cannot rule them out in practice



(b) Ideal iMHF

Outline

- Motivation
- Data Independent Memory Hard Functions (iMHFs)
- Attacks
- **Constructing iMHFs (New!)**
  - **Depth-Robustness is *sufficient***
- Conclusions and Open Questions

Depth-Robustness is Sufficient! [ABP16]

Key Theorem: Let G=(V,E) be (e,d)-depth robust then CC(G)$\geq ed$.

Proof: Let $P_1,\ldots P_t$ denote an (optimal) pebbling of G. For 0< i < d define

$$S_i = P_i \cup P_{d+i} \cup P_{2d+i} \cup \cdots$$

one of the sets S_i has size at most CC(G)/d. Now we claim that

$$d \geq \text{depth}(G\text{-}S_i)$$

because any path in G-S_i must have been completely pebbled at some point. Thus, it must have been pebbled entirely during some interval of length d. Thus, G (CC(G)/d,d)-reducible. It follows that CC(G)$\geq ed$.

Depth-Robustness is Sufficient! [ABP16]

Key Theorem: Let G=(V,E) be (e,d)-depth robust then CC(G)$\geq ed$.

Implications: There exists a constant indegree graph G with

$$\mathrm{CC(G)} \geq \Omega\left(\frac{n^2}{\log n}\right).$$

Previous Best [AS15]: $\Omega\left(\frac{n^2}{\log^{10} n}\right)$

[AB16]: We cannot do better (in an asymptotic sense).

Summary

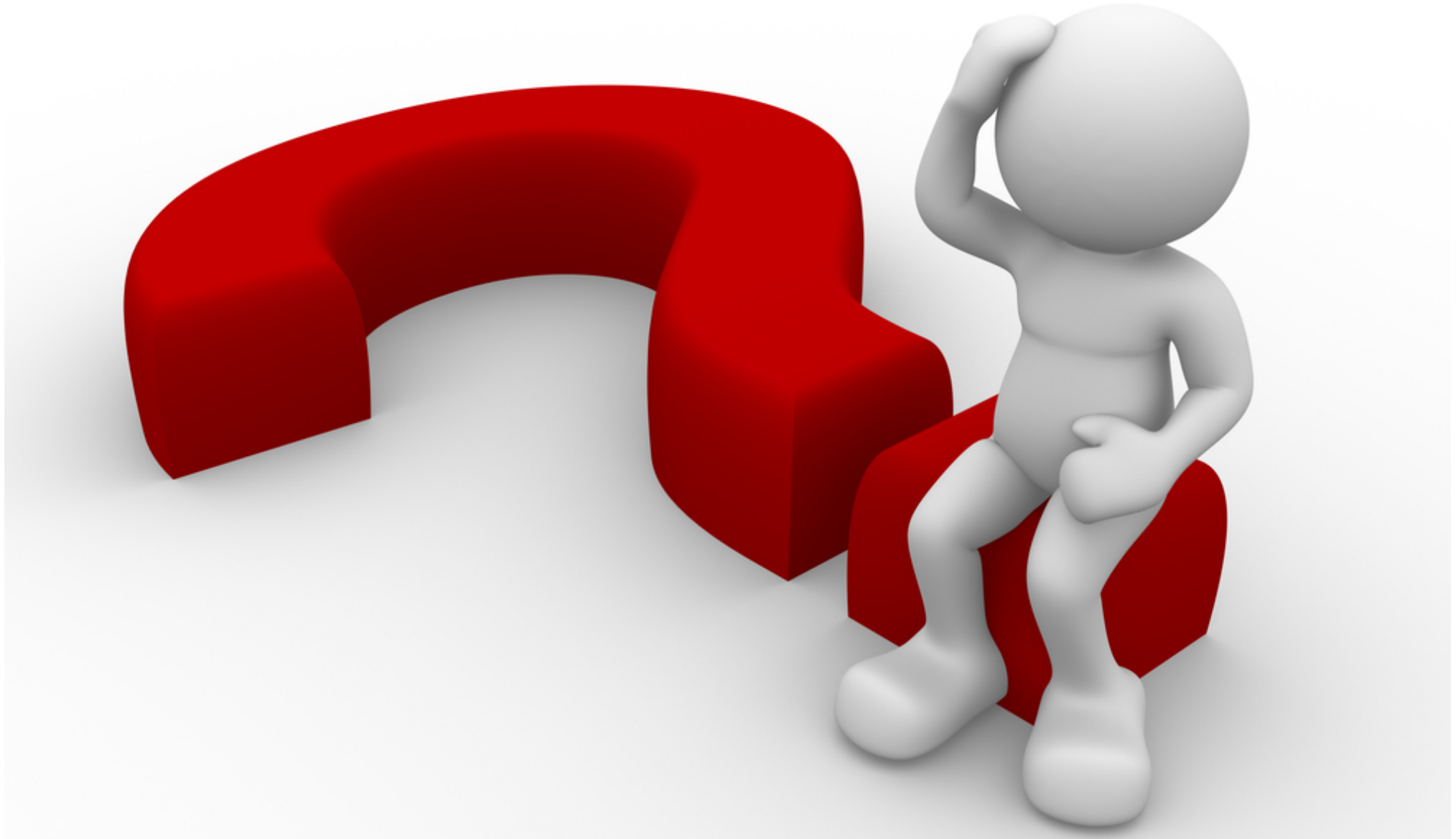
- BCRYPT and PBKDF2 are no longer sufficient for password hashing
- Argon2i is an improvement over BCRYPT and PBKDF2
  - But still has its flaws [AB16,AB17]
- Current Recommendation: Argon2id
  - No side channel attacks? Resists known attacks
  - Side channel attacks reduce security to Argon2i
- Look for improvements in the near future using depth-robust graphs [ABP17]

Conclusions

- Depth-robustness is a necessary and sufficient for secure iMHFs
  - [AB16] [ABP16]

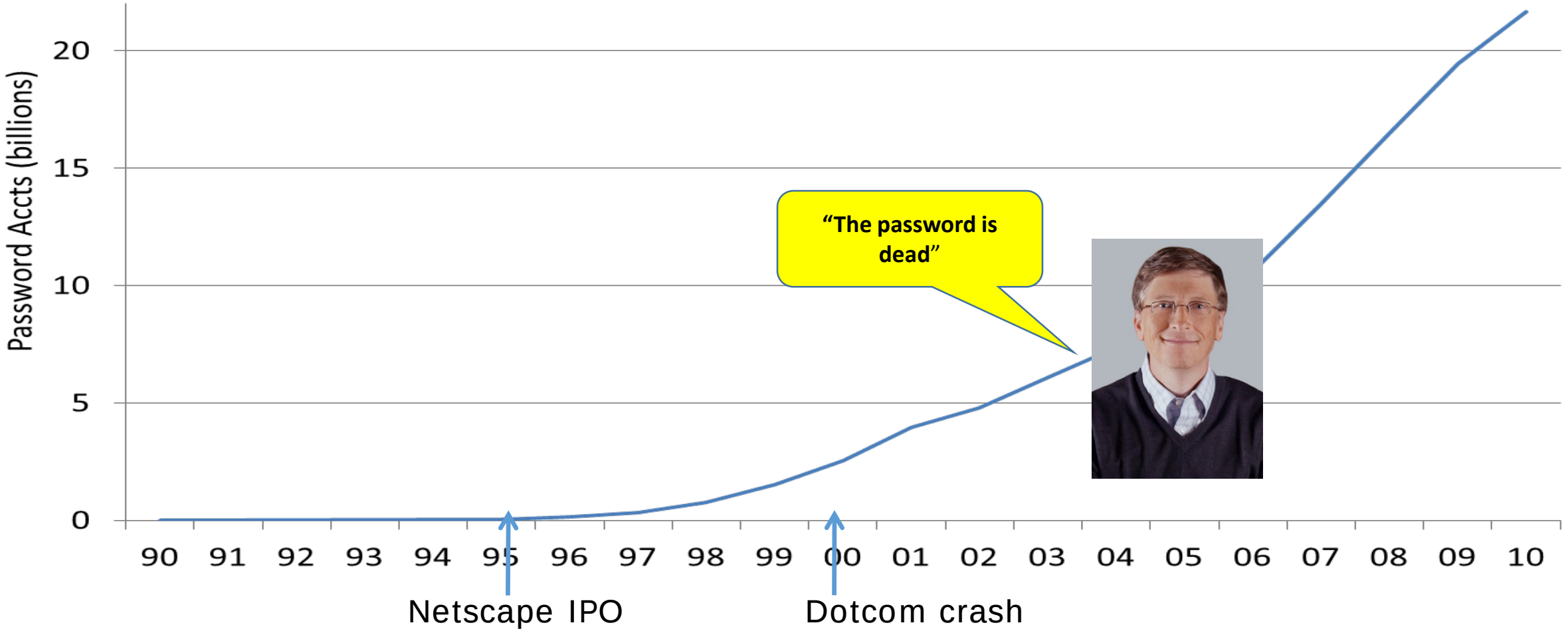
- Big Challenge: Improved Constructions of Depth-Robust Graphs
  - We already have constructions in theory [EGS77, PR80, …]
  - But constants matter!

Thanks for Listening

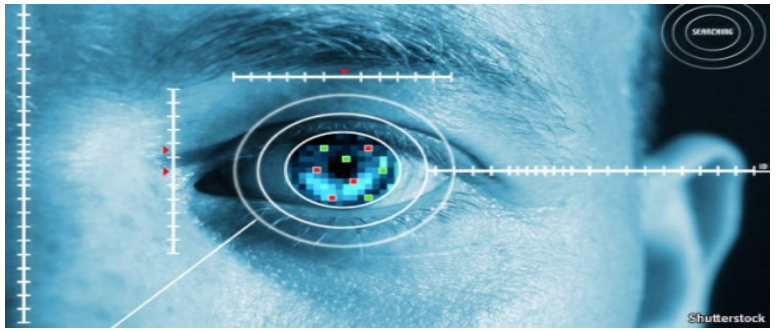


Passwords vs time: Look how far we've come

Source: Cormac's estimate

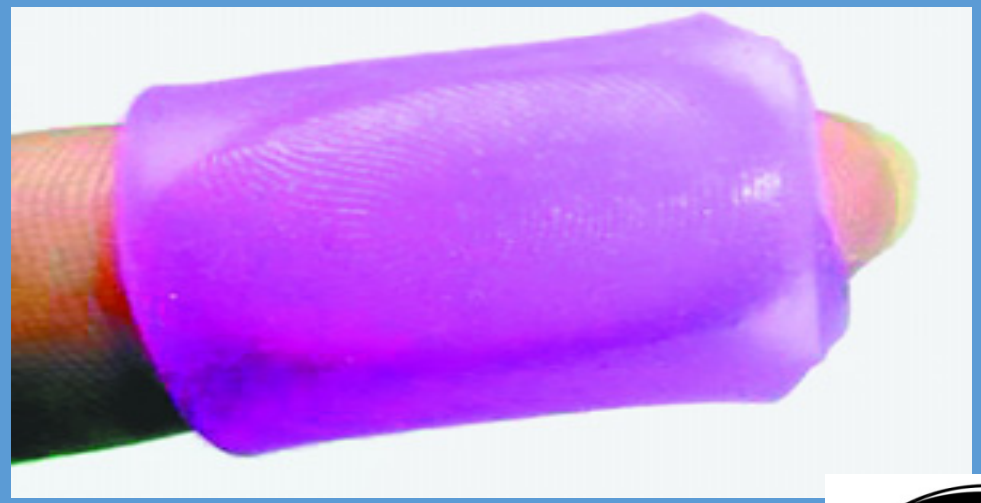


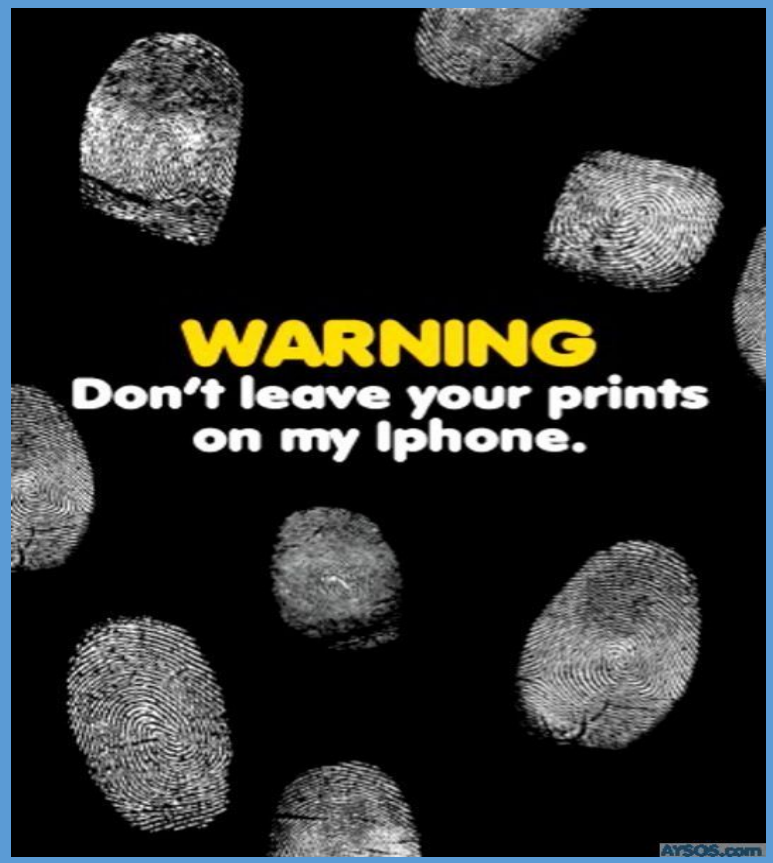
Biometrics



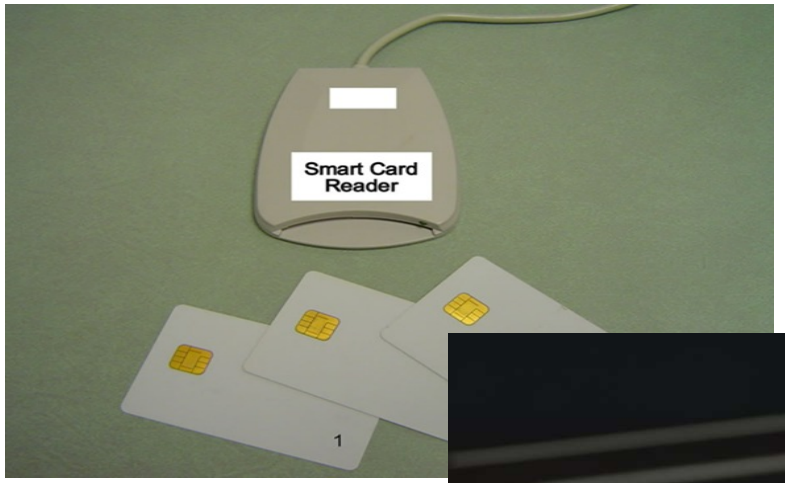
Biometrics

Challenges: Revoke? Secrecy?





Hardware Tokens

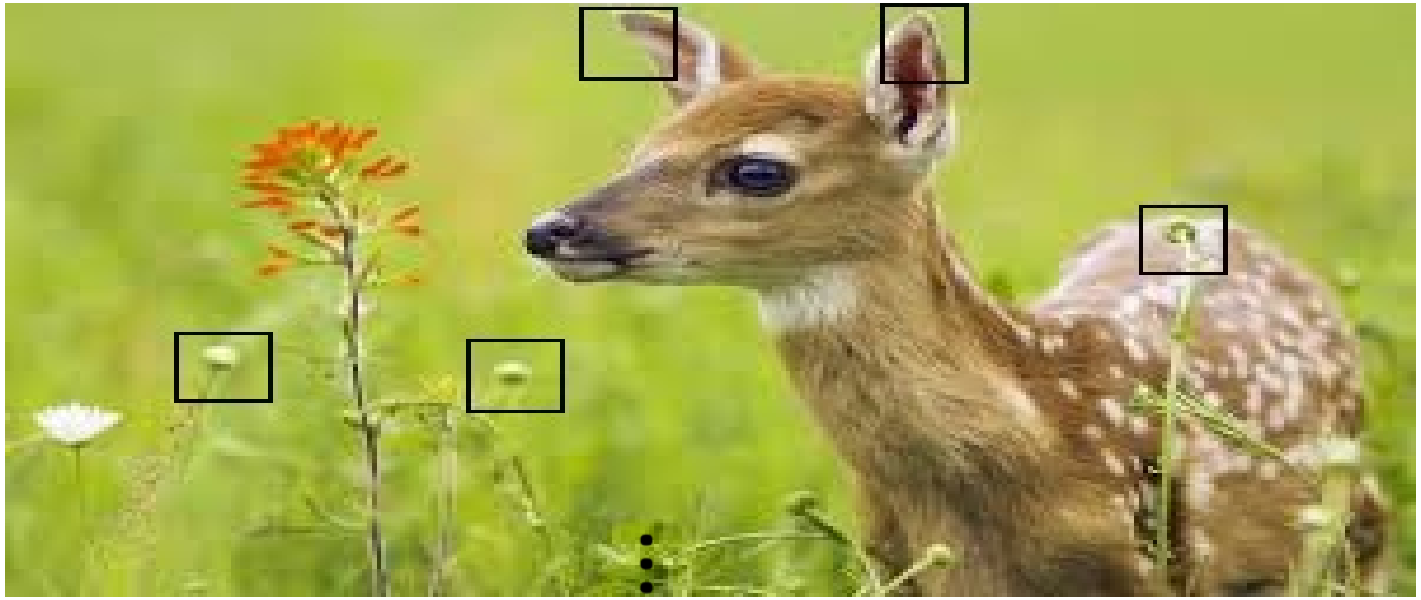


Hardware Tokens

Challenge: $$$ + more stuff to carry around

Graphical Passwords

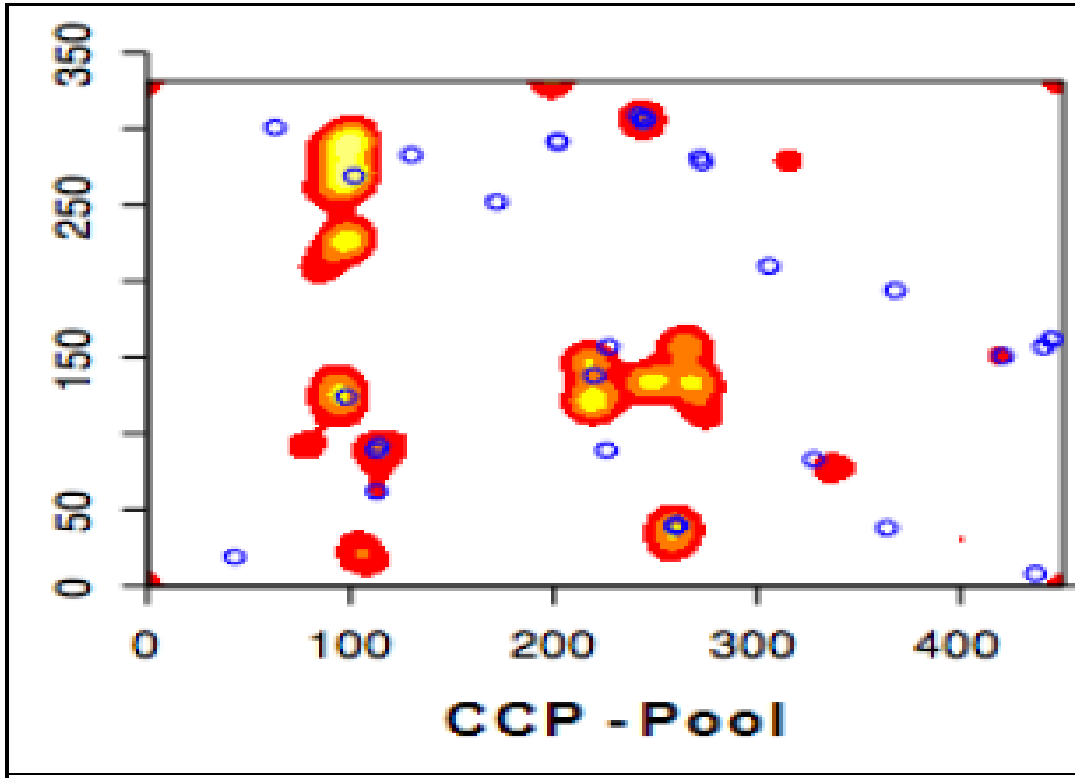
- Examples:
  - Passfaces, Cued Click Points, Windows 8



Graphical Passwords

Challenge: Multiple Passwords

Graphical Passwords: Hotspots



Graphical Passwords: Hotspots

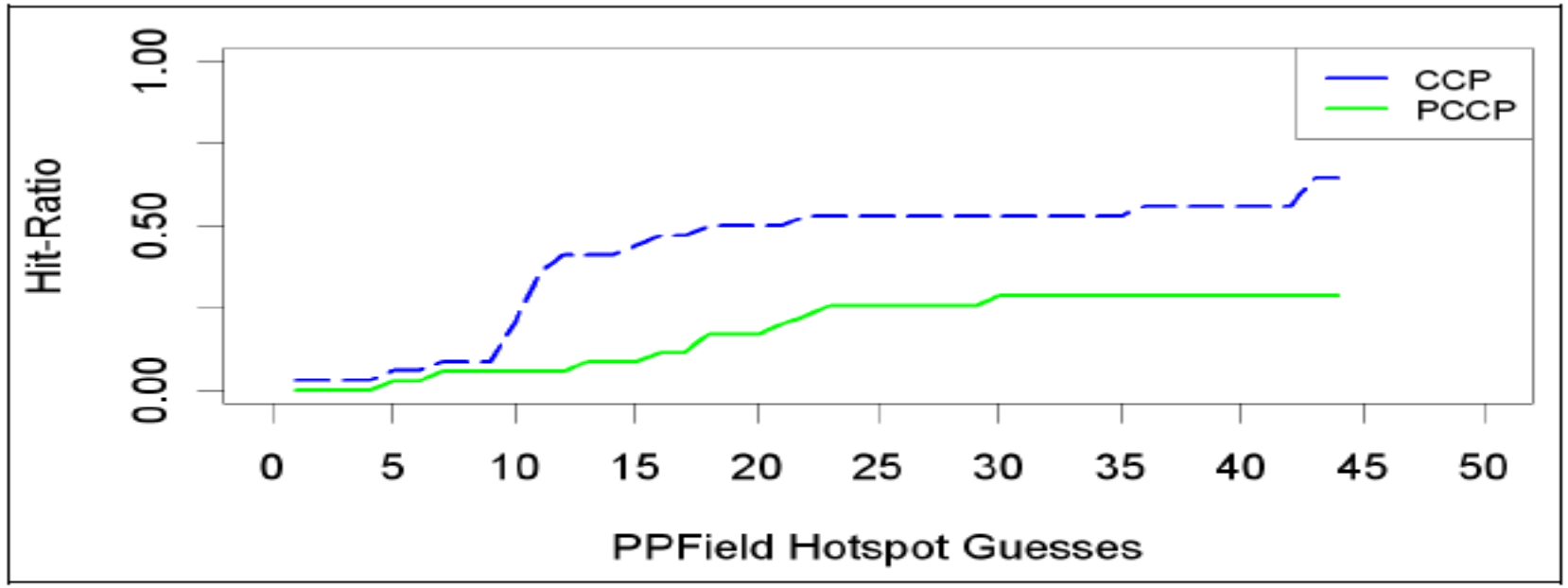


Figure 7: Individual click-points "guessable" using hotspots from the PassPoints-field study on the Pool image

Password Managers

- Password Management Software

LastPass ****

The Last Password You'll Ever Need.

Related Work

Challenge: Single point of failure

References

- **Depth-Robust Graphs and Their Cumulative Memory Complexity**. with Joel Alwen and Krzysztof Pietrzak. EUROCRYPT 2017 (to appear). [ePrint]
- **On the Computational Complexity of Minimal Cumulative Cost Graph Pebbling**. with Samson Zhou. (Working Paper). [arXiv]
- **Towards Practical Attacks on Argon2i and Balloon Hashing**. with Joel Alwen. EuroS&P 2017 (to appear). [ePrint]
- **Efficiently Computing Data Independent Memory Hard Functions**. with Joel Alwen. CRYPTO 2016. [Full Version]