Passwords: a Guide to the Ruins and Lessons for Improvement

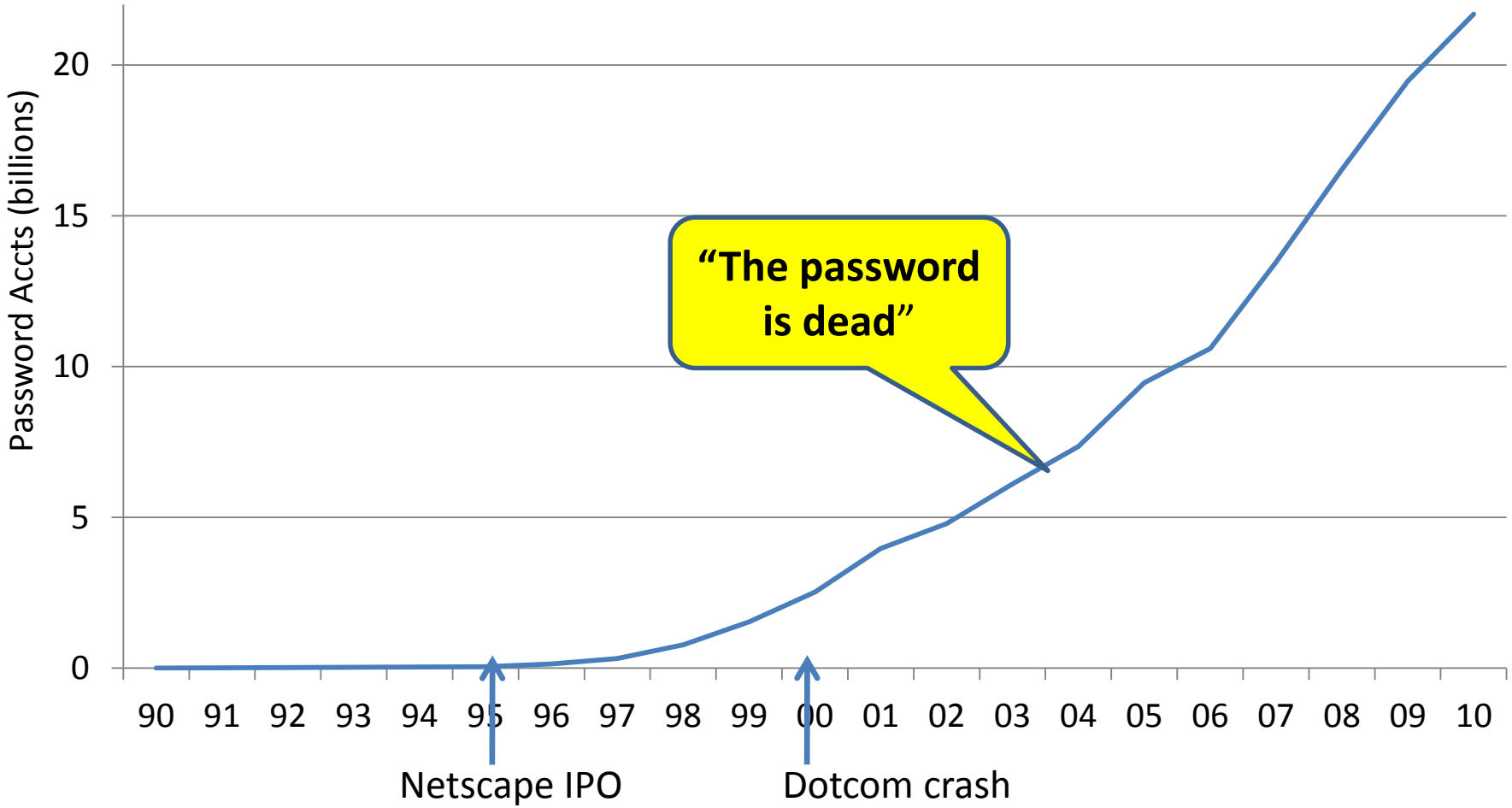
Cormac Herley
Microsoft Research, Redmond

Joint work with Dinei Florêncio and Paul C. van Oorschot


Why are we
here?

Password-protected accts vs time

Source: my estimate



Passwords in 2014

- About 6 million new password-protected accts/day
- 28% compounded growth
  - When do we hit “peak passwords” ?
- Assume 2bln users type one password/day
  - $2x10^9x5/(365x24x60x60)$= 1388 person-years spent typing ***per day.***

Two recent papers:

1. Managing a ***portfolio*** of passwords
   - DF, CH, PvO: Usenix Security, 2014

2. Administering password-protected site
   - DF, CH, PvO: Usenix LISA, 2014.

Ch1: Password Portfolios:
Sustainably Managing Large Numbers of Accounts

Choosing a password

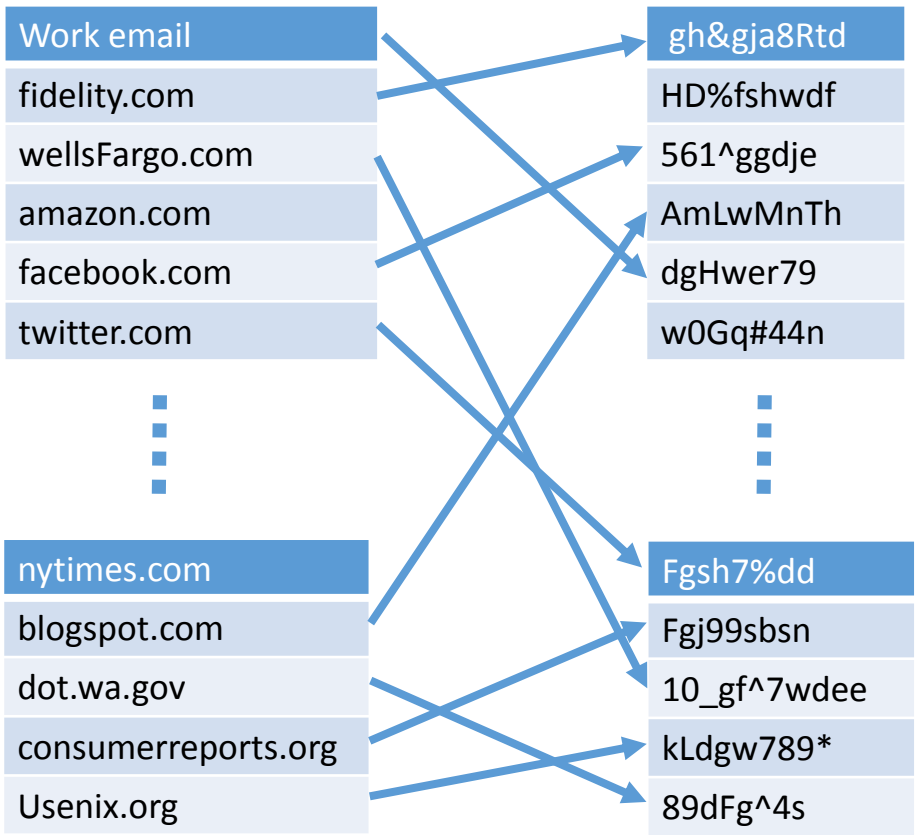
Everyone knows

A1: Passwords should be random and strong

A2: Passwords should not be re-used across accounts

But no-one does.

Portfolio of N random, unique passwords lg(S) each



lg(N!) + N·lg(S)

Must remember:

- N passwords = N·lg(S)
- NxN pwd-to-acct assignment = lg(N!)

$$E(N) = N \cdot \lg(S) + \lg(N!)$$

N=100 *random* passwords of lg(S) bits

$$E(N) = N \cdot \lg(S) + \lg(N!)$$

$$= 4000 + 524 = 4524 \text{ bits}$$

Depends how remember passwords

Random bits

$E(N) = 100 \cdot \lg(S) + 524$

0000111010101000100010010111010000001010101001000100
0101101101111111000010101011010001011011001001110 11
1000110001110100111001001010010010000010000100111011
1110111000001101000001100100001110110000100111011000
1111110011011010100011111000011010010001001100010110
1001000101100101010101010110100110111010100000100110
1000111011101101001111110101100011011110111110011001
1111011001111110011001100010101011100110011110101101 0
0010000001111111100011100000000000111011011100001100
1000111111101011100011011100001101111101001101011101
1111

Claim: memorization task is impossible

N accounts in G groups

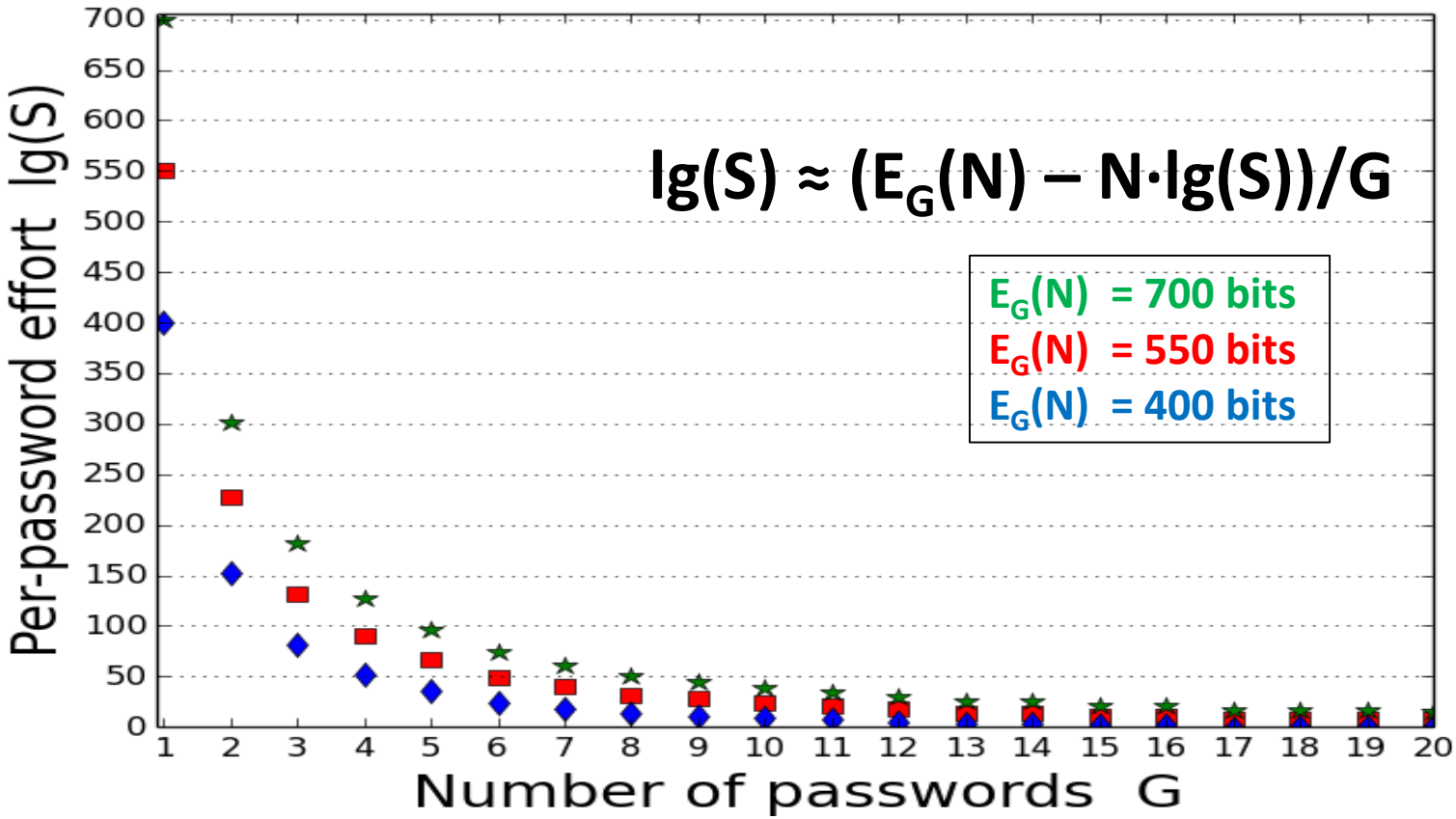
$$E_G(N) \approx G \cdot \lg(S) + N \cdot \lg(G)$$

$$\Rightarrow \quad \lg(S) \approx \frac{(EG(N) - N \cdot \lg(G))}{G}$$

Tradeoff between strength, lg(S), and re-use, G

N = #accts
G = #unique pwds
lg(S) = pwd strength

Tradeoff between strength/re-use:



Fixed effort:

- $\lg(S) \alpha 1/G$
- Stronger pwd => more re-use

Optimize the right thing: Loss or (Loss + Effort)?

$P_i(E)$ = Pr. Compromise
L_i = i-th acct. value

To Minimize:

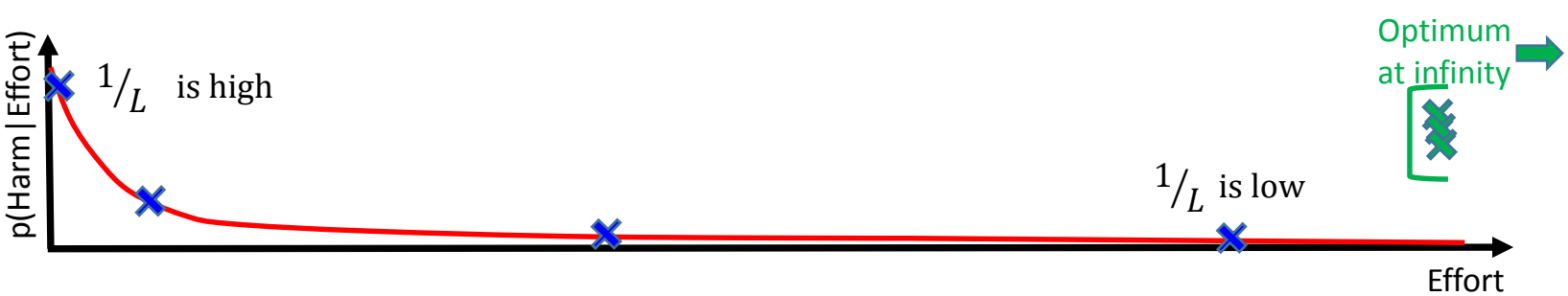
$$L = \sum_{i=1}^{N} P_i\,(E_i)\,L_i$$

$$L + E = \sum_{i=1}^{N} [P_i\,(E_i)\,L_i + E_i]$$

Optimality when:

$$\frac{dP_i\,(E_i)}{dE_i} = 0$$

$$\frac{dP_i\,(E_i)}{dE_i} = {}^{-1}/_{L_i}$$



Re-use Complicates things

Risk is not:

- Independent across accounts
- Dependent only on strength

- Risk to i-th acct also depends on
  - Effort for other accts that share the password
  - Effort to protect from keyloggers, malware
- Soln: divide into system/group/acct attacks

Without this simplification: set of N non-linear eqns

Take-aways on Chap.I

- One password/account impossible as N grows.
- A strategy that rules out weak passwords is sub-optimal
- User practice closer to optimal than advice

Ch2: Administering a password-protected site

Four Basic Questions:

1. **When is strength a factor?**
2. **How do we measure strength?**
3. **How much strength do we need?**
4. **How achieve that needed amount?**

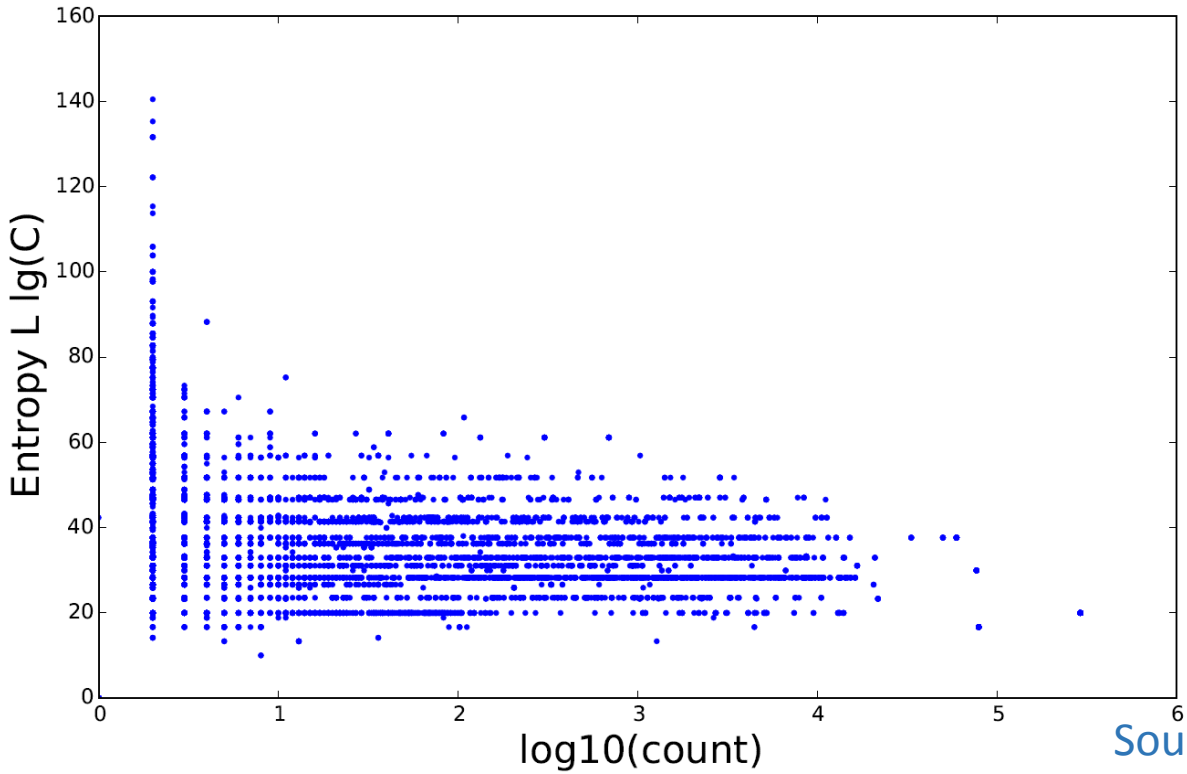
1. When is strength a factor?

“The success of database breaches, client-side malware, phishing and network-sniffing are entirely unaffected by password choice.”

E.g. Rockyou database breach: password choice had no effect on the outcome

2. How do we measure strength?

Don't use "entropy" = Len x lg(Size[CharSet])



Source: Rockyou leak

- L·lg(C) not even approximately monotonic in frequency
- Partial Guess numbers: #guesses to get fraction α of accts (e.g. JohnTheRipper, Bonneau measure)

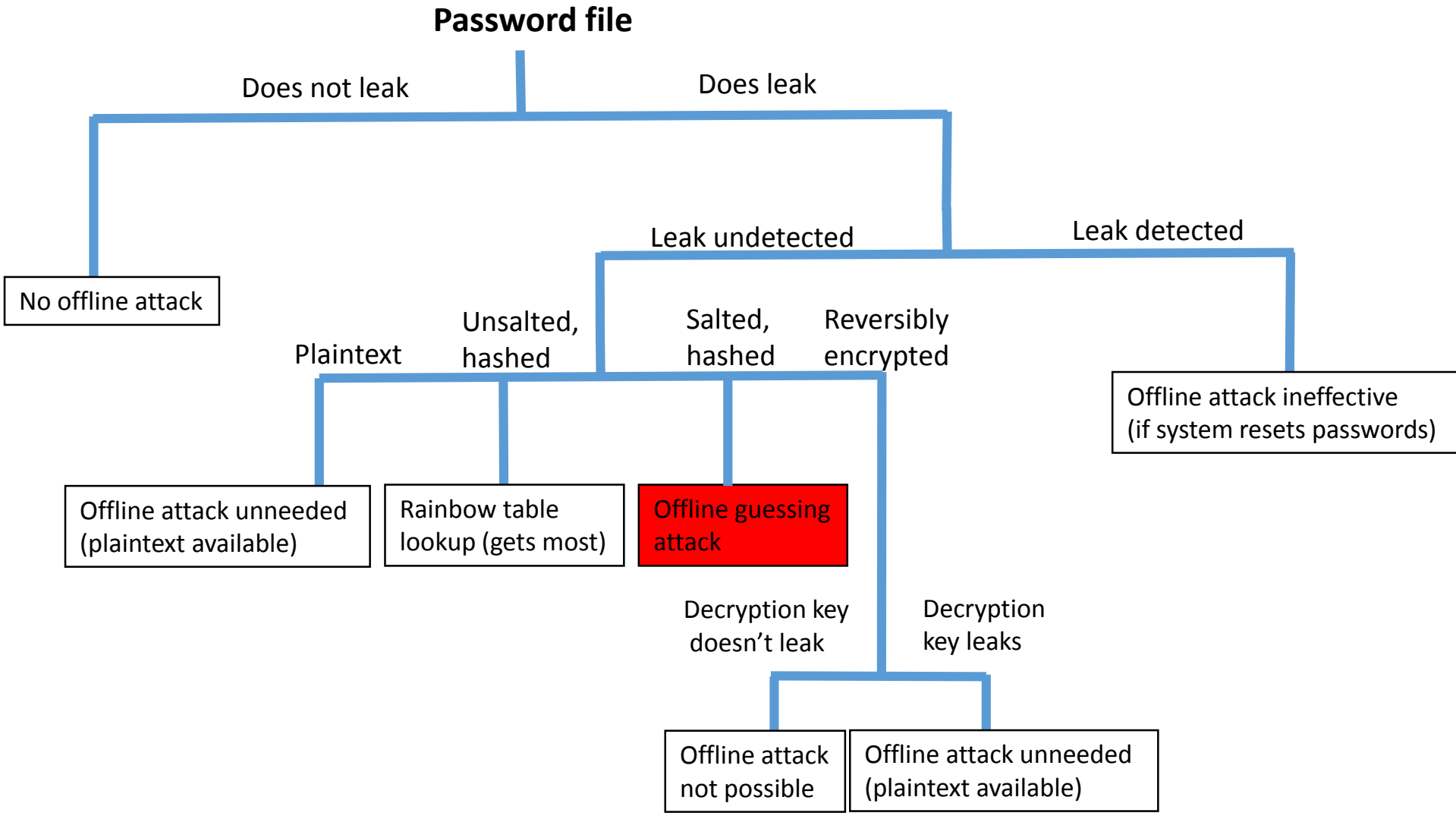
3. How much strength do we need?

Two very different guessing attacks

- **Online:** computed on defender's HW
  - Lockout, rate-limiting, forensics, …..

- **Offline:** computed on attacker's HW
  - Limited only by hardware
  - *Needs to steal the file and ……*

When is *offline* guessing a factor?



Recent breaches

Site	Year	# Accounts	Hashed	Salted	Reversibly Encrypted	Offline guessing attack beyond rainbow tables needed and possible
Rockyou [64]	2009	32m				N
Gawker	2010	1.3m	✓	✓		Y
Tianya	2011	35m				N
eHarmony	2012	1.5m	✓			N
LinkedIn	2012	6.5m	✓			N
Evernote	2013	50m	✓	✓		Y
Adobe	2013	150m			✓	N
Cupid Media	2013	42m				N

- August 2014: 1.2 billion CyberVor set: plaintext
- **In only 2 leaks (Evernote, Gawker) and 51.3mln ex 1.5bln passwords was offline a threat (if breach not discovered).**

Plaintext & Rev. Encryption more common than you think (especially at universities)

- RADIUS dial-up protocol: CHAP required plaintext password
  - Supported until recently.

Carnegie Mellon University

"The university was using a legacy credential management system (since abandoned), which, to meet certain functional requirements, reversibly encrypted user passwords, rather than using salted, hashed records." Mazurek et al, 2013 [CMU]

Leibniz Universität Hannover

"The IDM system stored up to five unique passwords per user using asymmetric cryptography, so it would be possible to decrypt the passwords to do a security analysis." Fahl et al., 2013 [Leibniz University]

Plaintext or reversibly encrypted: composition policies unjustifiable—no offline guessing attack.

Carnegie Mellon University

"Passwords had to be greater than length 8 and include lower, upper, special characters and digits." Mazurek etal, 2013 [CMU]

Large-scale waste of effort based on confused thinking.

How many guesses?

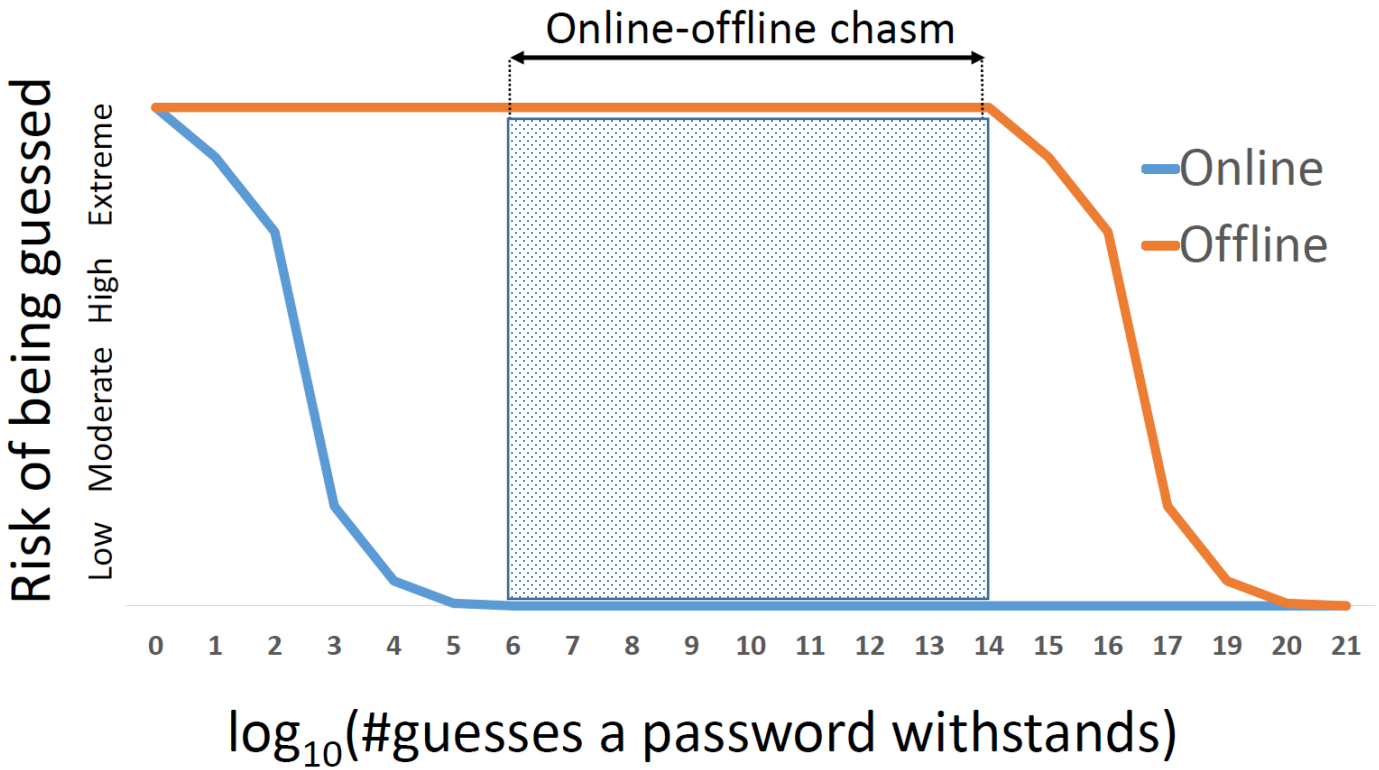
Attack	Type	Guesses	Example
Online	Breadth-first	10^4	“6387”
Online	Depth-first	10^6	“tincan24”
Offline	Breadth-first	10^{14}	“7Qr&2Mu”
Offline	Depth-first	10^{20}	“eTh^D#aW3a8”

Note the enormous difference needed to withstand online/offline

Reasoning:

- **Online Breadth-first: 10^4**
  - Over 4 mos. 17300x more fail events than legit pop. (assuming 1 legit login/user/day)
- **Online Depth-first: 10^6**
  - Lockout or Rate-limit requests, IP blocking
- **Offline Breadth-first: 10^{14}**
  - 1000 GPUs @ 10^{10} guess/sec against 10^6 accts for 4 mos
- **Offline Depth-first: 10^{21}**
  - 1000 GPUs @ 10^{10} guess/sec against 10 accts for 4 mos

No gain in exceeding online threshold while falling short of offline one.



Chasm is 8 orders of magnitude wide!!

4. How achieve needed amount of strength?

Composition Policies: very poor RoI

- Unjustifiable when no offline risk
- Inadequate protection even against online!!!!
- Many LUDS(8) passwords in top 10^4 Rockyou

Blacklisting:

- Block the most common choices, e.g. 10^4
- Good protection for online
- Inconvenience only those who need it.

Case against consuming user effort to defend against offline

- We don’t know how to do it
  - Composition policies, advice and meters are failures
- It’s generally unnecessary
- Also
  - Entire waste if plaintext or reversibly encrypted
  - Exceeding online threshold, but short of offline is waste
  - Task gets harder each year
  - Zero-user burden solutions exist

- Online: defensible goal, currently poorly defended
- Offline: hopelessly remote goal

Ch3: Lessons for the future?

Denying the obvious

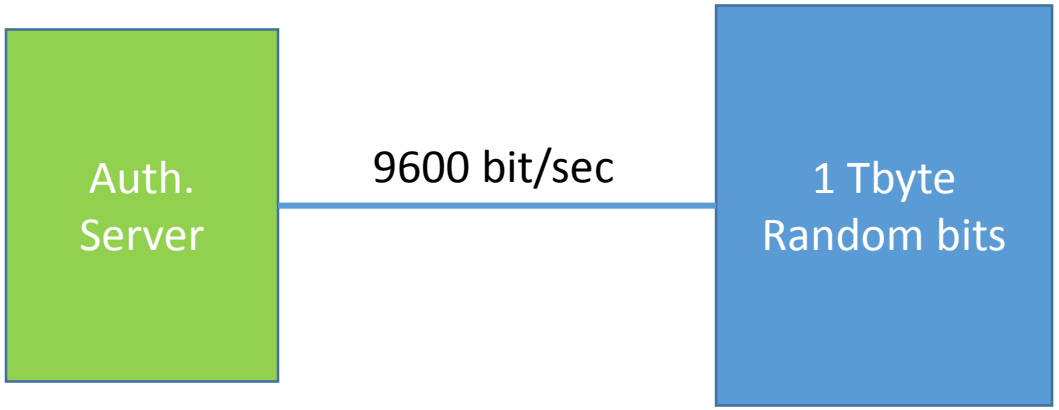
- Plaintext breaches evidence of:
  - Lazy user password habits?
  - Excess user effort?
- Breaches are evidence of:
  - “Passwords are dead”?
  - Important of Revocation/reset?

- Important
  - Revocation/reset is important
  - Protecting databases

Addressing attacks that scale with defenses that don't

- User effort to address attacks is O(N)
  - Do const-time solns exist?
  - HSMs?

Crescenzo et al Hashing Scheme



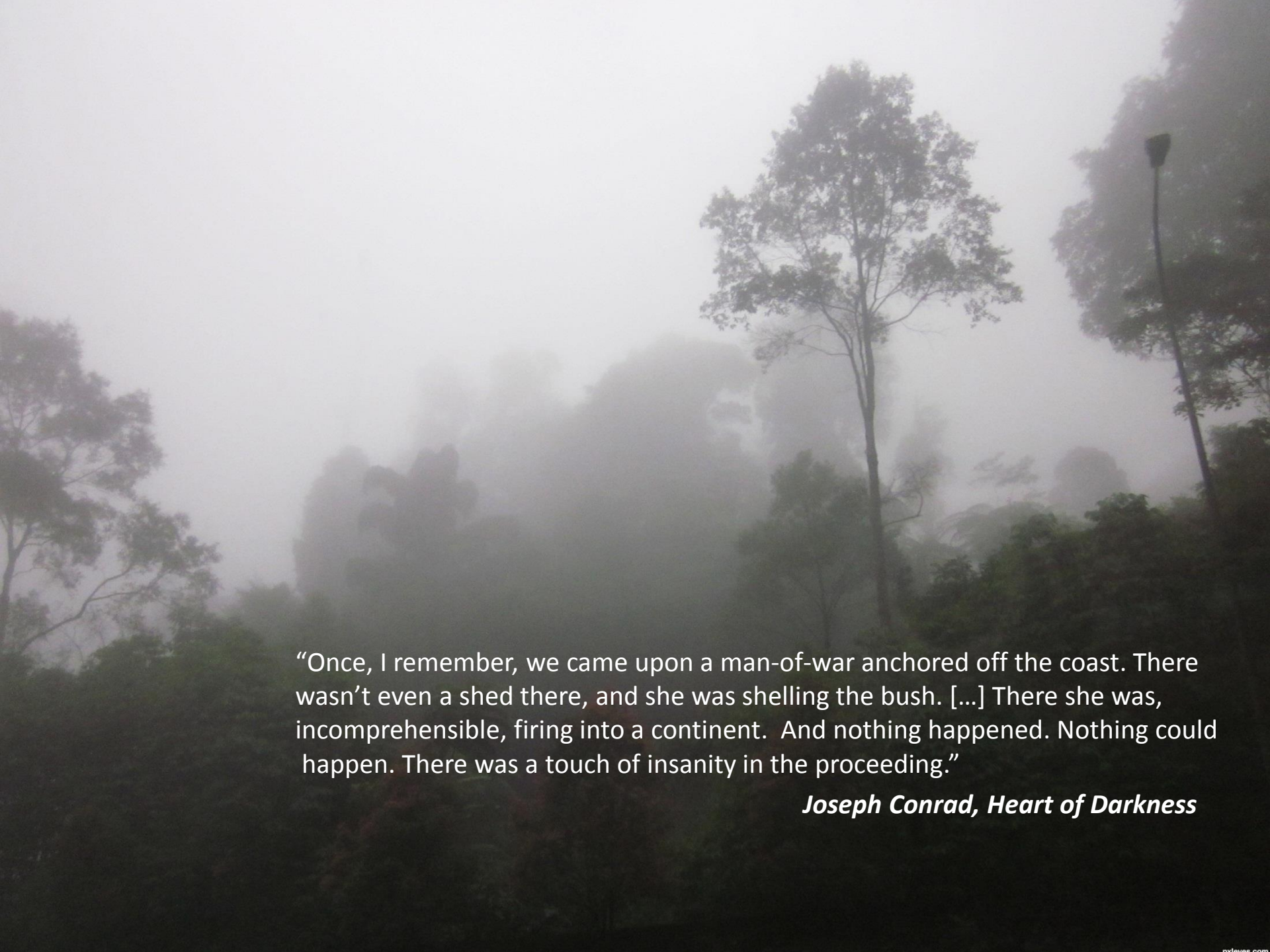
Over-constrained Problems

- Quest to replace passwords
  - Intersection of requirements appears empty
- Password Portfolios
  - Insisting on the necessity of impossible things

- How end up over-constrained?

A	Is re-use a real threat vector?	Y
B	Do bad things happen because of re-use?	Y
C	Can we eliminate that risk by avoiding re-use?	Y
D	Does it follow that you should not re-use?	N

$\mathbf{X} \Rightarrow \mathbf{Y}$ **does not mean** $\overline{\mathbf{X}} \Rightarrow \overline{\mathbf{Y}}$



“Once, I remember, we came upon a man-of-war anchored off the coast. There wasn’t even a shed there, and she was shelling the bush. […] There she was, incomprehensible, firing into a continent. And nothing happened. Nothing could happen. There was a touch of insanity in the proceeding.”

Joseph Conrad, Heart of Darkness

Conclusions

"There are no unclaimed pools of user effort to be had"

- Reclaim the waste

Something can't be both necessary and impossible

- Distinct password per account is not possible

Addressing attacks that scale with defenses that don't does more harm than good.

- User effort against offline attacks is a lost cause

References:

- D. Florencio, C. Herley and P.C. van Oorschot, "An Administrator's Guide to Internet Password Research", Proc. Usenix LISA, 2014
- D. Florencio, C. Herley and P.C. van Oorschot, "Password Portfolios and the Finite-Effort User: Sustainably Managing Large Numbers of Accounts", Proc. Usenix Security, 2014
- J. Bonneau, C. Herley, P.C. van Oorschot and F. Stajano, "The quest to replace passwords: A framework for comparative evaluation of web authentication schemes", IEEE Symp. Security & Privacy 2012.
- C. Herley and P.C. van Oorschot, "A Research Agenda Acknowledging the Persistence of Passwords," IEEE Security and Privacy magazine, Jan. 2012.
- D. Florencio and C. Herley, "Is Everything We Know About Password Stealing Wrong?" IEEE Security and Privacy magazine, Dec 2012.
- S. Schechter, C. Herley and M. Mitzenmacher, "Popularity is Everything: a new approach to protecting passwords from statistical-guessing attacks," Proc. HotSEC 2010
- D. Florencio and C. Herley, "Where Do Security Policies Come From?", SOUPS 2010 [Best paper award at SOUPS]
- C. Herley, P.C. van Oorschot and A.S. Patrick, "Passwords: If We're So Smart Why Are We Still Using Them?" Financial Crypto 2009
- D. Florencio and C. Herley, "A Large Scale Study of Web Password Habits," WWW 2007, Banff.
- D. Florencio, C. Herley and B. Coskun,"Do Strong Web Passwords Accomplish Anything?," Usenix HotSEC '07, Boston.