Software and Security Engineering

Lecture 3: Attitudes to Risk

Richard Mortier

rmm1002@cam.ac.uk

With many thanks to Ross Anderson and Alastair Beresford

Warm-up: Write down your own top three pieces of password advice

- Talk to your neighbour
- What password advice would you give and why?

John Podesta email compromise by Fancy Bear (allegedly Russia)

- White House chief-of-staff; chair of Hillary Clinton's 2016 US Presidential Campaign
- Gmail account was compromised
- 20,000 emails subsequently published by WikiLeaks
- Authenticity of some emails questioned

3

"In March 2016, the personal Gmail account of John Podesta, a former White House chief of staff and chair of Hillary Clinton's 2016 U.S. presidential campaign, was compromised in a data breach, and some of his emails, many of which were work-related, were stolen. Cybersecurity researchers as well as the United States government attributed responsibility for the breach, which was accomplished via a spear-phishing attack, to the hacking group Fancy Bear, allegedly affiliated with Russian intelligence services.

"Some or all of the Podesta emails were subsequently obtained by WikiLeaks, which published over 20,000 pages of emails, allegedly from Podesta, in October and November 2016. Podesta and the Clinton campaign have declined to authenticate the emails. Cybersecurity experts interviewed by PolitiFact believe the majority of emails are probably unaltered, while stating it is possible that the hackers inserted at least some doctored or fabricated emails. The article then attests that the Clinton campaign, however, has yet to produce any evidence that any specific emails in the latest leak were fraudulent. A subsequent investigation by U.S. intelligence agencies also reported that the files obtained by WikiLeaks during the U.S. election contained no "evident forgeries"."

https://en.wikipedia.org/wiki/Podesta_emails

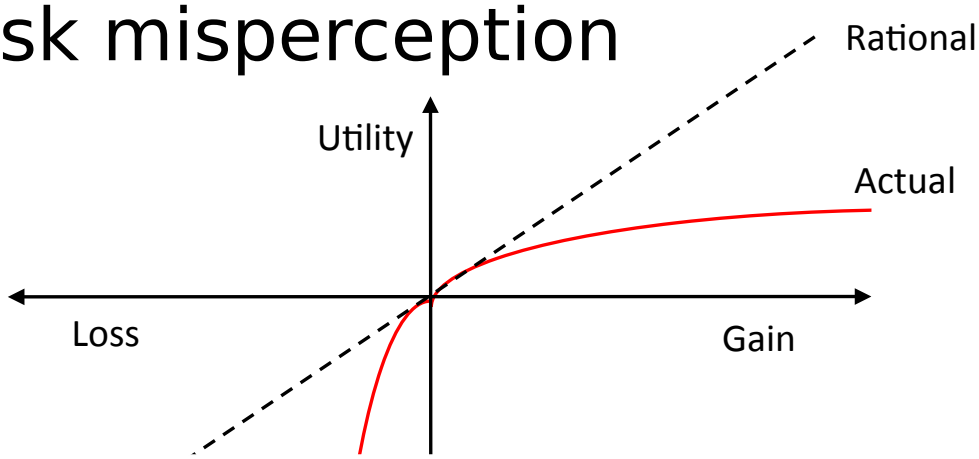
Cognitive factors

- Many errors arise from our highly adaptive mental processes
  - We deal with novel problems in a conscious way
  - Frequently encountered problems are dealt with using rules we evolve, and are partly automatic
  - Over time, the rules give way to skill
- Our ability to automate routine actions leads to absent-minded slips, or following the wrong rule
- There are also systematic limits to rationality in problem solving – so called *heuristics* and *biases*

4

It turns out the psychology is really important in this space. Once you become skilled at something, such as playing the Piano, you start to do things automatically (e.g. play the D major scale). The ability to automate familiar actions can be used against us. For example, if you're the Chief Clark and you get a request to pay an invoice from the CEO, and it has similar phrasing and so as the last N emails, then you arrange for payment of the invoice as requested; you don't stop and consciously consider whether this is in fact part of a spear-phishing attack.

Risk misperception



People offered £10 or a 50% chance of £20 usually prefer the former; if offered a loss of £10 or a 50% chance of a loss of £20 they tend to prefer the latter!

5

"Asymmetry between gains and losses: People are risk averse with respect to gains, preferring a sure thing over a gamble with a higher expected utility but which presents the possibility of getting nothing. On the other hand, people will be risk-seeking about losses, preferring to hope for the chance of losing nothing rather than taking a sure, but smaller, loss (e.g. insurance).

"Threshold effects: People prefer to move from uncertainty to certainty over making a similar gain in certainty that does not lead to full certainty. For example, most people would choose a vaccine that reduces the incidence of disease A from 10% to 0% over one that reduces the incidence of disease B from 20% to 10%."

https://en.wikipedia.org/wiki/Risk_perception

Framing decisions about risk – *Asian disease problem*

Scenario A, choose between:

a) "200 lives will be saved"

b) "with p=1/3, 600 saved; with p=2/3, none saved"

Here 72% choose (a) over (b).

Scenario B, choose between:

1) "400 will die"

2) "with p =1/3, no-one will die, p=2/3, 600 will die"

Here 78% prefer (2) over (1)

6

Risk misperception: empirical studies have shown that "a bird in the hand is worth two in the bush". Modern prospect theory explains the irrationalities that humans have when it comes to risk. This can be used to manipulate. Decisions are heavily influenced by framing. The Asian disease problem is one of the most famous. Here 600 people are infected with a deadly, fictional disease. The numbers and percentages come from Tversky and Kahneman (1981) with a summary here: https://en.wikipedia.org/wiki/Framing_(social_sciences)

[Present the details on the slide]

The ability to framing decisions to change perception is why marketeers talk about a 'discount' or 'saving' while fraudsters exploit the fact that people facing losses take more risks. There is more on this in the Economics, Law and Ethics course next year.

Social psychology

- Authority matters: Milgram showed over 60% of all subjects would torture a 'student'
- The herd matters: Asch showed most people could deny obvious facts to please others
- Reciprocation is built-in: give a gift, to increase your chance of receiving one

7

Milgram showed that the lab setting with a white coat gave authority and many participants would do as directed, even when they ask participants to electrocute students who get answers wrong. The student was actually an actor; no electricity involved. https://en.wikipedia.org/wiki/Milgram_experiment

Most people will follow the herd: here seven actors and one subject look at two lines, A obviously longer than line B. Yet when the seven actors say that line B is longer, the subject will follow them and confirm B is longer. https://en.wikipedia.org/wiki/Asch_conformity_experiments

Reciprocation: even monkeys do tit-for-tat. Further information: https://en.wikipedia.org/wiki/Reciprocity_(social_psychology)

For further information on these and other areas, see Robert B. Cialdini, Influence: Science and Practice (ISBN 0-321-18895-0). https://en.wikipedia.org/wiki/Influence:_Science_and_Practice

Fraud psychology

All the above plus:

- Appeal to the mark's kindness
- Appeal to the mark's dishonesty
- Distract them so they act automatically
- Arouse them so they act viscerally

Note: the *mark* is the person being defrauded

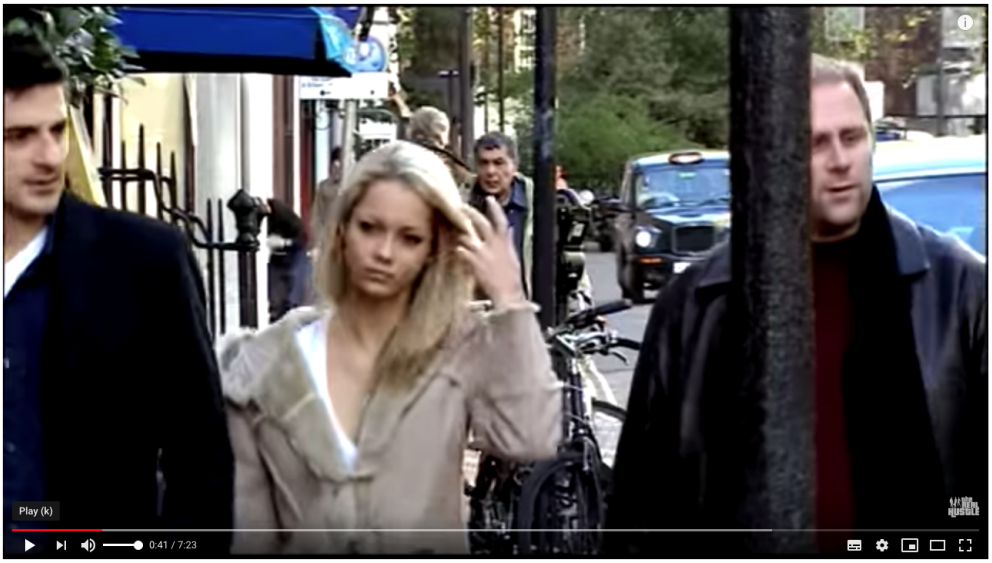
8

Note: the mark is someone who is destined to be defrauded.

- Down a pub: "I'm a bit short of cash, so I wonder whether you could do me a favour and buy this a TV for £40?" (Clearly TVs cost more than this.)
- Via email: "I need help safeguarding $400 million from …"
- Sales training school: "if you need someone to sign on the line for something, you put the pen on the clipboard, push the clipboard towards the person who is doubtful and "accidentally" drop the pen off the clipboard towards the mark, who then catches it. Now the mark has the pen in their hand and they are more likely to sign.

For further reading see: Stajano and Wilson, Understanding scam victims: seven principles for systems security, University of Cambridge Computer Laboratory Technical Report 754. https://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-754.pdf
Also, search for "The Real Hustle" videos on YouTube.

The Lottery Scam



https://www.youtube.com/watch?v=oI2gBLn6CU8

9

[Start playing at 40 seconds in. Run for around two minutes, then explain the remainder.]

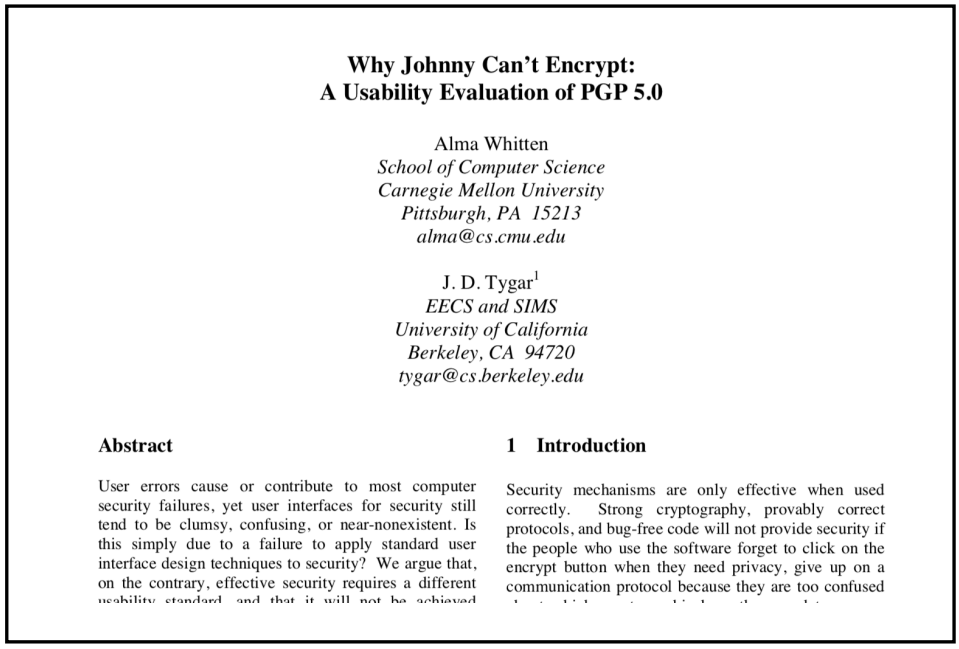
People only follow advice which confirms their own world view

- Users have different mental models. Explore how your users see the problem – the ‘folk beliefs’
- Given a model of their world view, target approach to appeal to it.

10

Therefore you often need to offer one than more piece of advice in order to find the one which fits their own view.

Affordances: Johnny Can't Encrypt



Why Johnny Can't Encrypt:
A Usability Evaluation of PGP 5.0

Alma Whitten
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213
alma@cs.cmu.edu

J. D. Tygar[1]
EECS and SIMS
University of California
Berkeley, CA 94720
tygar@cs.berkeley.edu

Abstract

User errors cause or contribute to most computer security failures, yet user interfaces for security still tend to be clumsy, confusing, or near-nonexistent. Is this simply due to a failure to apply standard user interface design techniques to security? We argue that, on the contrary, effective security requires a different

1 Introduction

Security mechanisms are only effective when used correctly. Strong cryptography, provably correct protocols, and bug-free code will not provide security if the people who use the software forget to click on the encrypt button when they need privacy, give up on a communication protocol because they are too confused

11

Alma Whitten asked 12 subjects who had no previous experience of public-key cryptography to use PGP to conduct a simple (encrypted) email exchange. Results were poor: participants in the experiment could not get things right. They sent the private key to the corresponding person by mistake. They forgot to encrypt and sign. And so on. The essence of the problem is that PGP was simply not useable by the non-expert (and likely some experts too).

See: A Whitten, JD Tygar. Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0. USENIX Security Symposium, 1999. https://www.usenix.org/legacy/events/sec99/full_papers/whitten/whitten.ps

The power of default

Most people don’t opt in or out; they go with default

Can exploit this for good (or evil):

- Pensions
- Privacy settings in an online service
- Use of crypto
- …

Therefore defaults may be contentious

12

In the past, many people didn't enrol in a pension scheme because they never got around to going to payroll to set it up or sign up with a third-party provider. Pensions are now offered by default, but you can opt out, which sets a safer default for everyone (less poverty in old age which requires support from the state).

There are also conflicts of interest: engineering defaults might suggest one approach (no crypto means less CPU load) and security requirements might suggest another (crypto protects passwords as we shall see later). Similarly, advertising performance, and therefore revenue, might suggest no HTTPS so adverts work better. There are tensions and these are hard to resolve.

Economics versus psychology

Most people don't worry enough about computer security, and worry too much about terrorism

How could we fix this, and why is it not likely to be?

13

Two approaches to terrorists: 1) play it up as much as possible (e.g. George Bush Jr) and say "woe is us, this is terrible, we must invade these countries, ..." Or 2) this is terrible but we will get them in the end (e.g. George Bush Sr).

If we want to reduce the effect of terrorism, then you need to make it less salient: remove the guns and visible security in airports, etc. Replace guns with pastel sofas, and so on.

The compliance budget

- ‘Blame and train’ as an approach is suboptimal
- It’s often rational to ignore warnings
- People will spend only so much time obeying rules, so choose the rules that matter
- Violations of rules also matter: they’re often an easier way of working, and sometimes necessary
- The ‘right’ way of working should be easiest: look where people walk, and lay the path there

14

Further reading: Beautement, Adam, M. Angela Sasse, and Mike Wonham. "The compliance budget: managing security behaviour in organisations." Proceedings of the 2008 New Security Paradigms Workshop. ACM, 2009. https://dl.acm.org/citation.cfm?id=1595684

Where should the path be?

15

Make the easiest path also the one which is safe and secure. Otherwise, people will do this...

Differences between people

- Ability to perform certain tasks varies widely across subgroups of the population, including by
  - Age
  - Gender
  - Education
  - …
- Yet all customers receive complex password rules and anti-phishing advice

16

When you work at a new tech start-up, it's very easy to assume that everyone is 20, has 20-20 vision and has a degree in computer science. This leads to the situation where you say "use a randomly generated password on each website; don't write them down". However most of the population will struggle with this guidance. Indeed performance at tasks varies significantly across the population. Sometimes there is correlation with age (e.g. due to physical mobility or vision requirements) or gender (in societies with gendered interest in IT).

More accidents with Volvos?

Volvo ÖV 4, April 1927

17

Volvos have a reputation for safety. So, why are there more accidents involving more Volvo drivers? Two possible explanations: (1) Bad drivers buy Volvos; (2) Volvo drivers drive faster because they think think that they are protected and safe in a Volvo. It's really hard to tell. This is called risk compensation.

Other examples: “It has been observed that motorists drove faster when wearing seatbelts and closer to the vehicle in front when the vehicles were fitted with anti-lock brakes. By contrast, shared space is a highway design method which consciously aims to increase the level of perceived risk and uncertainty, thereby slowing traffic and reducing the number of and seriousness of injuries.” https://en.wikipedia.org/wiki/Risk_compensation

Understanding error helps us build better systems

- Significant psychology research into errors
- Slips and lapses
  - Forgetting plans, intentions (strong habit intrusion)
  - Misidentifying objects, signals
  - Retrieval failures (“its on the tip of my tongue”)
  - Premature exits from action sequences (using the ATM)
- Rule-based mistakes; applying the wrong procedure
- Knowledge-based mistakes; heuristics and biases

18

Human brains exhibit a number of different errors. We need to understand these if we are to build robust, human-centred systems.

Strong habit intrusion: When I cycle to the train station from the Computer Lab, I often find myself turning into Queens’ College gates on the way there. The reason for this is because I frequently cycle from the CL to Queens’ so I do this by default. I’m ”on autopilot”.

When you go to the cash machine should you give the customer the cash then their card (US); or card then cash (UK). Cash second is best: that's why you went to the machine in the first place, therefore you will leave once you have the cash (and leave the card behind). Men are more likely to be goal focused then women; which means that men are more likely to leave their card in the ATM in the US than women.

Training and practice reduce errors

Inexplicable errors, stress free, right cues	10
Regularly performed simple tasks, low stress	10
Complex tasks, little time, some cues needed	10
Unfamiliar task dependent on situation, memory	10
Highly complex task, much stress	10
Creative thinking, unfamiliar complex operations, time short & stress high	~1

19

The automotive industry carried out an analysis into whether training and practice reduce errors. While training does help, there are inherent limits on the ability to reduce the probability of an error. Somewhat predictably, the hardest tasks to perform without error are those which involve creative thinking and unfamiliar operations when time is short.

Passwords are cheap, but...

- Will users enter passwords correctly?
- Will they remember them?
- Will they choose a strong password?
- Will the write them down?
- Will the password be different in each context?
- Can the user be tricked into revealing passwords?

20

Passwords are great for companies and implementers. They are cheap and users (think) they know how to use them. Important for innovation, since you can grow the user base of an online platform with tiny marginal cost and without the requirement to provide an additional hardware.

It's not helpful to (effectively) say "Choose something you can't remember, and don't write it down". The alternative, which is now considered best practice, is to use a password manager integrated into the web browser and smartphone and able to generate strong random passwords; it has its own drawbacks however – a significant issue is backup. Two factor makes things significantly stronger, however it is less usable. If password manager is not a usable solution, separating accounts into (the few) high-value ones (e.g. bank, email) and the (many) low-value ones and ensuring each class of accounts has a separate password is better than using the same password everywhere.

For further information read: Joseph Bonneau, Cormac Herley, Paul C. van Oorschot and Frank Stajano "The Quest to Replace Passwords: A Framework for Comparative Evaluation of Web Authentication Schemes" In Proc. IEEE Symposium on Security and Privacy 2012. Extended version available as University of Cambridge Computer Laboratory Technical Report UCAM-CL-TR-817. See: https://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-817.html

User studies are important

Experiment to see if first-year NatScis could be trained to use passwords effectively. Three groups:

- Control group of 100 (+100 more observed)
- Green group: use a memorable phrase
- Yellow group: choose 8 chars at random

Expected strength: Y > G > C; got Y = G > C

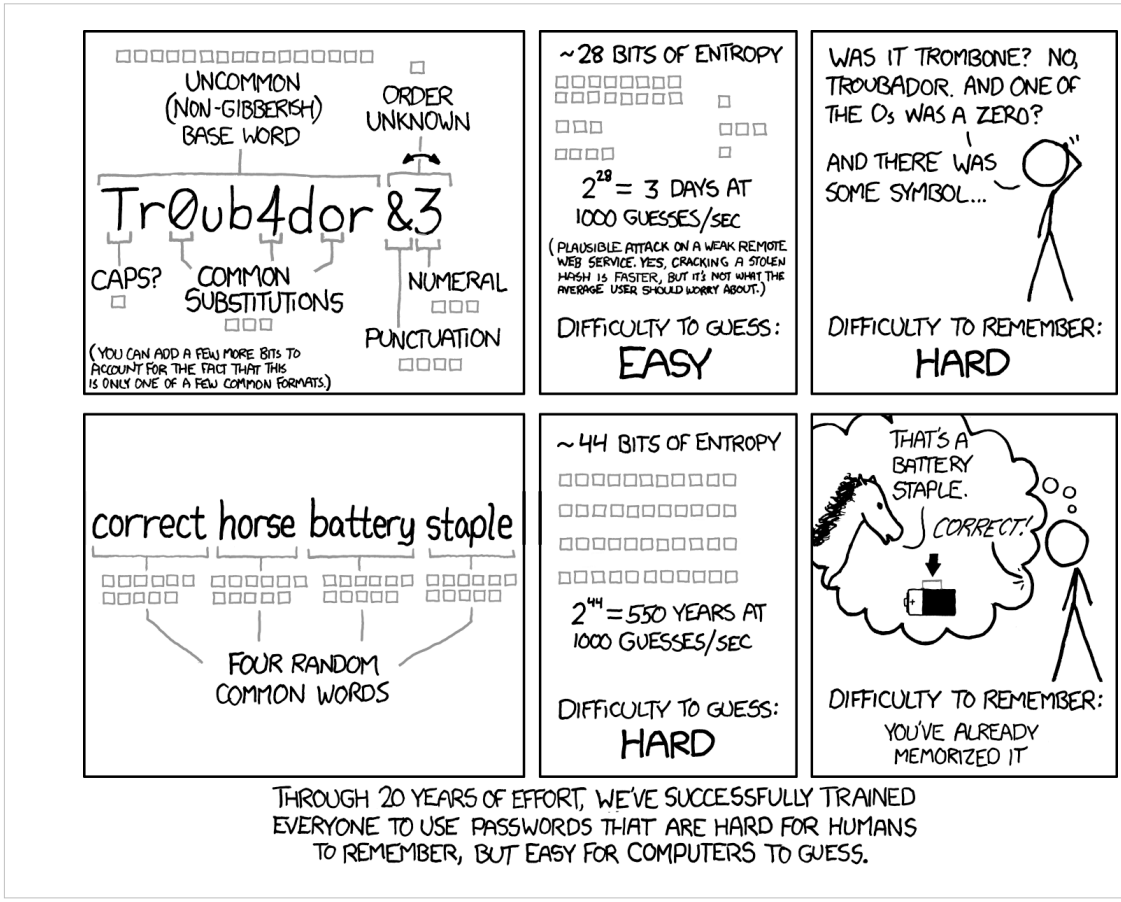
Expected resets: Y > G > C; got Y = G = C

We had 10% non-compliance

21

Twenty years ago Ross Anderson and Alan Blackwell ran a simple experiment. Split NatSci students into three groups: a control group, a group told to use a memorable phrase, and a group told to choose 8 characters at random. 10% non-compliance is amazing: these students volunteered to take part in an experiment, they are scientists, keen, and yet they didn't do as instructed. Take-home message: if you want to find things out you need to do a proper randomised control trial with real people. We would never have guessed that 10% would be in non-compliance.

Further reading: Jianxin Yan, Alan Blackwell, Ross Anderson and Alasdair Grant. The memorability and security of passwords – some empirical results. University of Cambridge Computer Laboratory Technical Report 500, September 2000. https://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-500.pdf



NB: NIST only recently rescinded the advice to change passwords regularly.

Hardware and online support to limit brute force is challenging

- Online services and tamper-proof hardware can be used to limit brute-force guessing, such as
  - Bank card PIN (3 guesses on card; 3 online)
  - iPhone PIN (timeouts)
  - Login attempts to webservices (timeouts; care required)
  - …

If the typical person has five cards with the same PIN, how many wallets do you need to find before you get lucky?

23

If passwords are easy to brute-force by repeated guessing, can we limit the number of possible attempts? Note that this is simply note possible in some settings (e.g. encrypted data stored on a stolen harddisk). Success or otherwise depends on the distribution of PINs and passwords (likely distinctly not random) and the number of accounts, cards or devices you have at your disposal.

For example, if you find a wallet in the street with five cards in it, assuming all PINs are the same, what is the chance that you can guess the PINs before cards start getting blocked (e.g. inside 25 attempts). People tend to pick PINs which are easy to remember! Could be 1234, could also be year of birth of children (e.g. 2002). Therefore you need many fewer wallets in reality than random guessing would suggest; 11-18 wallets turns out to be enough.

For more information see: Bonneau, Joseph, Sören Preibusch, and Ross Anderson. "A birthday present every eleven wallets? The security of customer-chosen banking PINs." International Conference on Financial Cryptography and Data Security. Springer, 2012. https://link.springer.com/chapter/10.1007/978-3-642-32946-3_3

Mitigate worst effects of a stolen password file

- Use key stretching techniques such as PDBKF2:

```
public PBEKeySpec(
        char[] password, byte[] salt,
        int iterCount, int keyLength
)
```

- Establish breach reporting laws
- Externalise the problem with Oauth
- Use other factors to determine whether login legit

Defence in depth is important (see Swiss Cheese Model earlier). What can we do to limit harm if the password file is stolen? This is important since users often reuse passwords across websites and apps, and email addresses are typically used as the username and therefore are also likely to be the same across multiple sites. Example: use javax.crypto.spec.PBEKeySpec

Rather than storing any passwords in plaintext, use a cryptographically secure one-way hash function. This means that, given a hash of the password, you cannot determine the plaintext. To check whether a password is valid, simply hash the user-supplied password and compare with the hash version previously stored. Given that there are a small number of potential passwords that many people use, a hash function alone is not very secure – an attacker could pre-compute the hashes of many common passwords to allow easy inversion at scale. Therefore, store a per-user cryptographic salt (random number) along with the hashed value. This means any inversion table needs to be built per-user, which does not offer any performance benefit.

A breach reporting laws says that the breach must be reported to the individuals who have been compromised. Users can then take action. This also means that other companies can find out about it when the breach is large (they are individuals too). Therefore these third-party companies can take appropriate action as required.

Oauth offers a potential solution since you no longer have to store passwords. Sounds great in principle, but it then means the website's operation is reliant on a third party. No third-party, no access to any accounts. It also means that if the Oauth vendor is hacked, your site is compromised. A related example: banks rely on SMS as second factor, so go to phone company, pretend to be the customer, and get a new SIM issued. Privacy is a problem with Oauth: Facebook knows how many customers you have if you use Oauth for authentication; bad for users and also bad for you when you try and sell the company to Facebook -- they know how often customers log in and how long there on the site (e.g. with "like" buttons).

Authentication is no longer a binary yes/no, but good systems use lots of side-information (e.g. location of login, speed of typing, etc) . Authentication systems get benefits of scale, thus encouraging use of centralisation (e.g. with Oauth) since smaller sites simply don't have the expertise, data and dedicated security team.

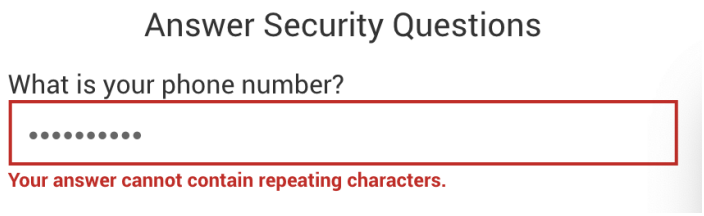
Password recovery is a weak point

- Password recovery often involves basic info which doesn't change:
  - What was the name of your first school?
  - What was the name of your first pet?
  - …
- Little ability to change this information
- Accounts for public figures are especially vulnerable

25

Famously Sarah Palin's AOL account got hacked because password recovery was poor: the answers to her recovery questions were in the public domain, so access to her email was obtained through public data.

A poor implementation of password recovery…



"I did it. I found the all-time dumbest security question answer requirement. Good job @fedex."
Luke Millar (@ltm on Twitter), 28th April 2019

26

Source: https://twitter.com/ltm/status/1122290624940560385

Externalities need consideration

- One firm's action has side-effects for others
- Password sharing a conspicuous example; we have to enter credentials everywhere
- Everyone wants recovery questions too
- Many firms train customers in unsafe behaviour from clicking on external links or redirecting the browser to third-party domains for payment
- Much 'training' amounts to victim blaming

27

It's not enough to look at things in isolation. For example, people are on Facebook because their friends are on Facebook – the so-called network effect. Similarly, compromise of one website results in a compromise of another website because the passwords for many users are the same.

Iterative guessing of card details with botnet on websites works

- Of Alexa top 500 websites, 26 use Primary Account Number (PAN) and expiry date
- 37 use PAN + postcode (numeric digits only for some, add door number for others)
- 291 ask for PAN, expiry date and CVV2

There is enough variation in requirements across websites that you can iteratively generate valid credentials

28

This is an externalities issue because you can first go to sites which require an account number and expiry date and use these to find valid combinations of these by brute force. Some sites require some of the postcode, so you can then guess this by using several such sites, and so on.

“We came to the important observation that the difference in various websites' security solutions introduces a practically exploitable vulnerability in the overall payment system. An attacker can exploit these differences to build a distributed guessing attack that generates usable card payment details (card number, expiry date, card verification value [CVV2], and postal address) one field at a time. Each generated field can be used in succession to generate the next field by using a different merchant's website.” For further information, see: Ali, Mohammed Aamir, Budi Arief, Martin Emms, and Aad van Moorsel. "Does the online card payment landscape unwittingly facilitate fraud?." IEEE Security & Privacy 15, no. 2 (2017): 78-86. https://ieeexplore.ieee.org/abstract/document/7891527

MAT HONAN GEAR 08.06.12 08:01 PM

HOW APPLE AND AMAZON SECURITY FLAWS LED TO MY EPIC HACKING

29

Mat Honan. “In the space of one hour, my entire digital life was destroyed. First my Google account was taken over, then deleted. Next my Twitter account was compromised, and used as a platform to broadcast racist and homophobic messages. And worst of all, my AppleID account was broken into, and my hackers used it to remotely erase all of the data on my iPhone, iPad, and MacBook. In many ways, this was all my fault. My accounts were daisy-chained together. Getting into Amazon let my hackers get into my Apple ID account, which helped them get into Gmail, which gave them access to Twitter. Had I used two-factor authentication for my Google account, it’s possible that none of this would have happened, because their ultimate goal was always to take over my Twitter account and wreak havoc.”

https://www.wired.com/2012/08/apple-amazon-mat-honan-hacking/

He lost all of the data stored on his laptop, including all the photos of his one-year old daughter.

Amazon ⇢ Apple ID ⇢ Gmail ⇢ Twitter

(And all they wanted was his three letter Twitter handle!)

- Twitter: find personal website, then Gmail, home address
- Gmail: account recovery gave “m••••n@me.com”
- Amazon: call with name, address, email to associate a new credit card number (fake) to the account
- Amazon: call (again) with name, address, credit card number and associate new email address with the account
- Amazon: Use web password reset to new email address; get last four digits of all credit cards in the account
- Apple: Call with billing address and last four digits credit card to get temp password for “m••••n@me.com”
- Gmail: reset password sent to “m••••n@me.com”
- Twitter: reset password sent to Gmail

30

Attack worked back from Twitter. Twitter profile listed his personal website, which listed his Gmail address. Whois record against the website provided home address, which is also the billing address for his credit card. The attacker used account recovery on Gmail to reveal backup email address as “m••••n@me.com” which is then guessable given his name.

Then the attacker called Amazon as Hanon and requested to add a credit card to his account. For this the attacker needed the name of the account holder, home address and email address. The attacker had all this. The attacker then provided the credit card information (a fake one will do) which is added to the account; hang up. Attacker called Amazon back and said he’d lost control of Hanon’s account, provided the account holder name, billing address and credit card number (the one just added). Amazon then associated a new email address with the account. The attacker then went to the Amazon website, used the reset password link and got the reset password sent to the new email address. The attacker then viewed the last four digits of all credit cards associated with the account.

The attacker then called AppleCare and gave them name, address and last four digits of credit card. This then allowed the attacker to gain access to Matt’s Apple ID and control of the iCloud account. From here the attacker reset the Gmail password to the backup email address (“m••••n@me.com”) which the attacker then controlled. From there reset the Twitter account via the Gmail account.

Image source: https://edition.cnn.com/2019/04/21/us/iowa-social-media-influencer-domain-name-trnd/index.html

You can avoid the difficulties of a technical attack by simply using force. While this doesn't scale well, it might be an effective means of carrying out a targeted attack. Thankfully it's harder than it might at first appear. The same approach has been used for stealing Cryptocurrency: "Robbers order gunpoint Bitcoin transfer after Moulsford break-in. Four robbers broke into a house and demanded at gunpoint the occupants transfer Bitcoins into another account. Thames Valley Police said the aggravated burglary happened in Reading Road, Moulsford, Oxfordshire, at about 09:40 GMT on 22 January. Four men broke in and threatened the two men and a woman inside with what appeared to be a firearm. One was told to transfer an amount of the digital currency but the transaction failed, police said." https://www.bbc.co.uk/news/uk-england-oxfordshire-42864053