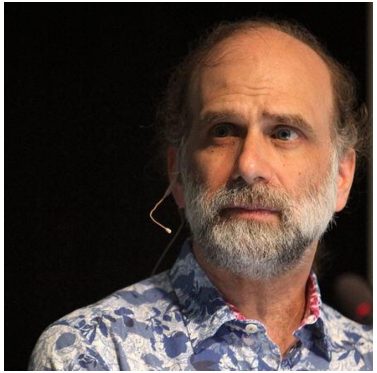
# CS915/435 Advanced Computer Security

# - Users and Security

# Quote of the day

“Only amateurs attack machines; professionals target people.”

- Bruce Schneier

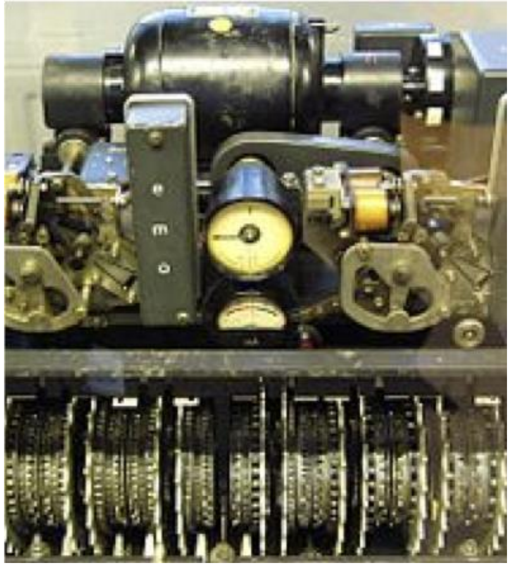


# Outline

- Introduction
    - Usability and psychology
- Example of attacking human weakness: phishing
    - What is phishing?
    - How does it work?
    - Why it works?
    - Defence
- Example of leveraging human strength: CAPTCHA
    - What is CAPTCHA?
    - How Does it work?
    - What could go wrong
    - CAPTCHA design

# Introduction

- Why first study the human?
  - Human is often the weakest link
- Lesson from the history
  - Lorenz cipher was used in WWII for top secret communication by Nazi German
  - In 1941, a German operator used the same key to send (almost) the same message twice: a forbidden practice
  - One human error led to the total break of Lorenz cipher



Lorenz cipher in WWII (also known as "Tunny" by British intelligence)

# Two-time pad attack

- Tunny code works like a one-time pad.

$M1 \oplus K = C1$

$M2 \oplus K = C2$

$0 \oplus 0 = 0$

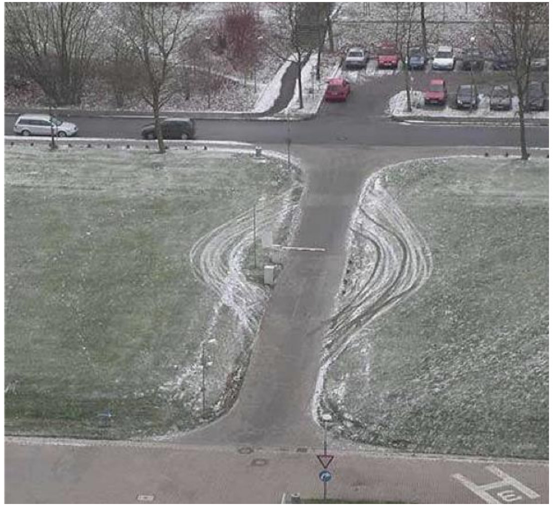
$0 \oplus 1 = 1$

$1 \oplus 0 = 1$

$1 \oplus 1 = 0$

Hence, $C1 \oplus C2 = M1 \oplus M2$, which can break into M1 and M2 based on language redundancies

# Importance of understanding human



“We assume users are always careless, usually incompetent and sometimes dishonest.”

-- Ross Anderson (Security Engineering)

# Usability

- How to create asymmetry in usability: good use is easy while bad use is difficult.
- For example: potato peeler

# Psychology

- Many real attacks exploit psychology as much as technology, e.g., phishing, pretexting
- Insights from psychology research
  - Capture error: a kind of error where a more frequent and more practiced behaviour takes place when a similar but less familiar action was intended. E.g., User Account Control (e.g., in Windows Vista)
  - Post completion error: once people have accomplished the immediate goal, they are easily distracted from tidying-up actions.
  - E.g., should an ATM machine be designed to give money first or card first?

# Road map

- Introduction
  - Usability and psychology
- **Example of attacking human weakness: phishing**
  - What is phishing?
  - How does it work?
  - Why it works?
  - Defence
- Example of leveraging human strength: CAPTCHA
  - What is CAPTCHA?
  - How Does it work?
  - What could go wrong
  - CAPTCHA design

# What is phishing?

- It’s not a typo!
- “Phishing” originated from “fishing”, and “phisher” from “fisher”
- But where is the sea?
    - Largely the community of billions of Internet users
- What is the bait?
    - Often emails (you win a lottery! Problems with your account, password etc.)
    - Instant message, SMS, etc.
- What is the “phish”?
    - Personal information, password, credit card number, bank account no, etc.

# A little history (I)

- Hackers usually replace “F” with “Ph”
  - “Ph” originated from an early form of hacking, known as “phreaking”, which started in the early 1970s when one of the earlier hackers, John Draper, created his blue box and used it to hack telephone systems
- Phishing was coined in 1996 by hackers who used to steal passwords from unsuspecting America On-Line (AOL) users
  - At the time, hacked accounts were called “phish”

# A little history (II)

- In 1997, phish (or hacked accounts) became a form of currency between hackers
  - E.g. a hacker would trade 10 working America Online (AOL) accounts for hacking software they needed
- Since then, phishing attacks developed from simply stealing AOL accounts to targeting online banking, payment services such as PayPal, and online e-commerce sites
  - By August 2003, most major banks have been targeted by phishing attacks

# Phishing is a business

- Billions of dollars are being made by criminals
- Gangs of phishers all over the world, primarily in Eastern Europe, Asia, Africa and the Middle East.
- Phishing also used extensively by organized crime groups
  - An Al-Qaeda group in Spain used stolen credit cards to set up their crimes and make purchases. They also used stolen calling cards for communication.

# Road map

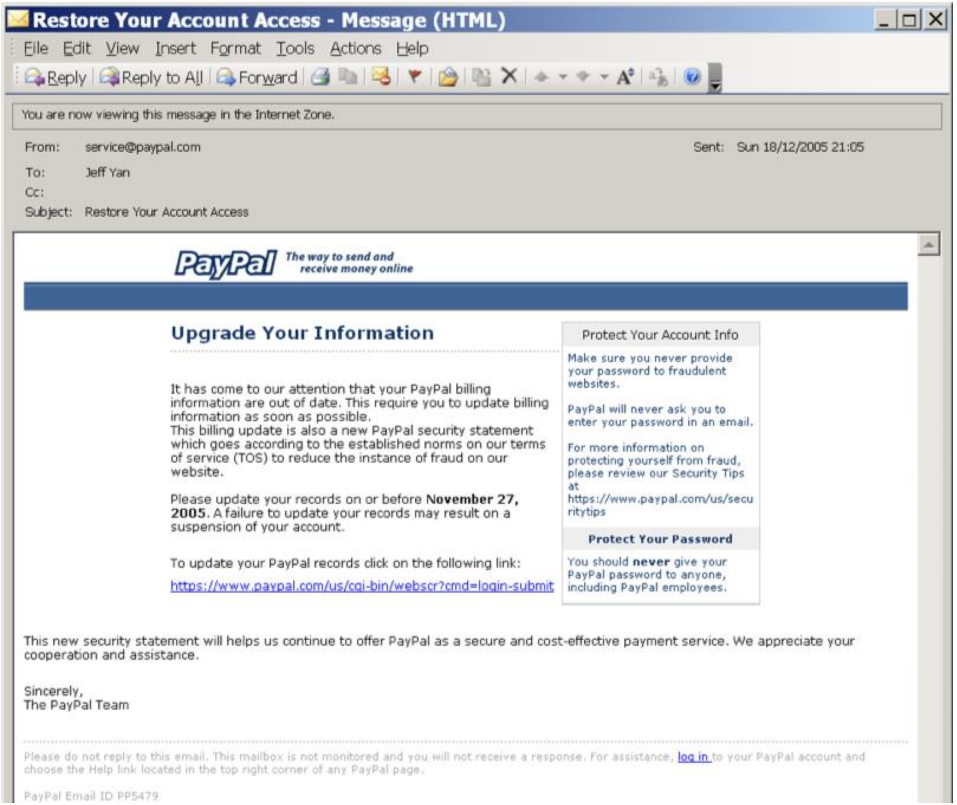
- Introduction
  - Usability and psychology
- Example of attacking human weakness: phishing
  - What is phishing?
  - **How does it work?**
  - Why it works?
  - Defence
- Example of leveraging human strength: CAPTCHA
  - What is CAPTCHA?
  - How Does it work?
  - What could go wrong
  - CAPTCHA design

# So, how is a typical attack?

- Phishing emails are sent to thousands of users, hoping to reach the clients of an online company, like paypal, online banks etc.
- The email tries to convince the user that something is wrong with their account and asks the user to confirm his data (password, credit card number, mother’s maiden name)
- There is a link in the email that points to a spoof site that resembles the legitimate one
- When the user clicks on the link, he is directed to the spoof site. The fraudsters try to fool the user by creating an illusion of security.

# A real example: phishing email

- Looks more than being real!
- What is the trick?
- Clicking the link, you will be directed to a spoof site



Restore Your Account Access - Message (HTML)
File Edit View Insert Format Tools Actions Help
Reply Reply to All Forward

You are now viewing this message in the Internet Zone.

From: service@paypal.com

Sent: Sun 18/12/2005 21:05

To: Jeff Yan

Cc:

Subject: Restore Your Account Access

PayPal The way to send and receive money online

**Upgrade Your Information**

It has come to our attention that your PayPal billing information are out of date. This require you to update billing information as soon as possible.
This billing update is also a new PayPal security statement which goes according to the established norms on our terms of service (TOS) to reduce the instance of fraud on our website.

Please update your records on or before **November 27, 2005**. A failure to update your records may result on a suspension of your account.

To update your PayPal records click on the following link:
https://www.paypal.com/us/cgi-bin/webscr?cmd=login-submit

Protect Your Account Info

Make sure you never provide your password to fraudulent websites.

PayPal will never ask you to enter your password in an email.

For more information on protecting yourself from fraud, please review our Security Tips at https://www.paypal.com/us/securitytips

**Protect Your Password**

You should **never** give your PayPal password to anyone, including PayPal employees.

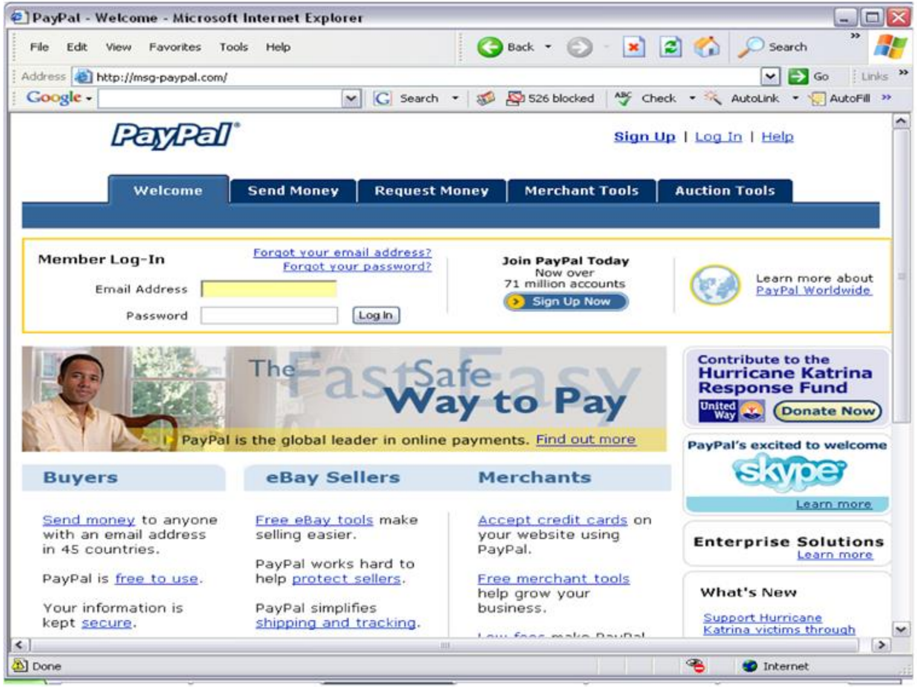
This new security statement will helps us continue to offer PayPal as a secure and cost-effective payment service. We appreciate your cooperation and assistance.

Sincerely,
The PayPal Team

Please do not reply to this email. This mailbox is not monitored and you will not receive a response. For assistance, log in to your PayPal account and choose the Help link located in the top right corner of any PayPal page.

PayPal Email ID PP5479

# Real vs phishing pages



http://msg-paypal.com/

Member Log-In

Forgot your email address?
Forgot your password?

Email Address

Password

Join PayPal Today
Now over
71 million accounts

Learn more about PayPal Worldwide

The Fast Safe Easy Way to Pay

PayPal is the global leader in online payments. Find out more

Contribute to the Hurricane Katrina Response Fund

United Way

PayPal's excited to welcome skype Learn more

Enterprise Solutions Learn more

What's New

Support Hurricane Katrina victims through

Buyers

Send money to anyone with an email address in 45 countries.

PayPal is free to use.

Your information is kept secure.

eBay Sellers

Free eBay tools make selling easier.

PayPal works hard to help protect sellers.

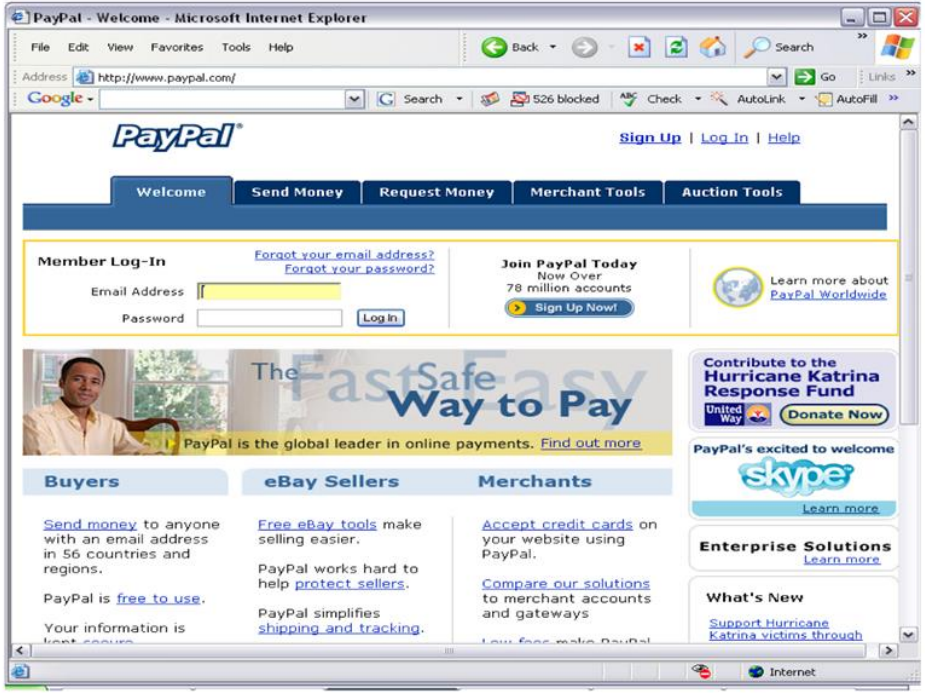
PayPal simplifies shipping and tracking.

Merchants

Accept credit cards on your website using PayPal.

Free merchant tools help grow your business.

Low fees make PayPal



http://www.paypal.com/

Member Log-In

Forgot your email address?
Forgot your password?

Email Address

Password

Join PayPal Today
Now Over
78 million accounts

Learn more about PayPal Worldwide

The Fast Safe Easy Way to Pay

PayPal is the global leader in online payments. Find out more

Contribute to the Hurricane Katrina Response Fund

United Way

PayPal's excited to welcome skype Learn more

Enterprise Solutions Learn more

What's New

Support Hurricane Katrina victims through

Buyers

Send money to anyone with an email address in 56 countries and regions.

PayPal is free to use.

Your information is kept secure.

eBay Sellers

Free eBay tools make selling easier.

PayPal works hard to help protect sellers.

PayPal simplifies shipping and tracking.

Merchants

Accept credit cards on your website using PayPal.

Compare our solutions to merchant accounts and gateways

Low fees make PayPal

# Properties of spoof sites

- Uses legitimate logos from the spoofed site to convince the user that it is real
- Disguise the URL by
  - Typejacking (using a hostname similar to the real hostname): **paypal** vs **paypaI**
  - Replacing the URL with an IP address
  - Containing the honest site’s URL as a substring: www.ebayfoo.com

# Properties of spoof sites

- Spoof sites are usually online for a few hours or days
- Fraudsters copy html code from the victim site, so spoof sites contain links to images on the honest site
- Spoof sites rarely use SSL. They try to give the impression they are using encryption by using various tricks. But with Let’s Encrypt, this is changing.
- Some spoof sites have inconsistencies, grammatical errors etc.

# Where did the money go?

- Once fraudsters get hold of your account
- They quickly transfer money to a series of local/overseas accounts
- “Mules” are employed as intermediate accounts
- Recovering the transfer across borders is rather difficult

# Road map

- Introduction
  - Usability and psychology
- Example of attacking human weakness: phishing
  - What is phishing?
  - How does it work?
  - **Why it works?**
  - Defence
- Example of leveraging human strength: CAPTCHA
  - What is CAPTCHA?
  - How Does it work?
  - What could go wrong
  - CAPTCHA design

# What was exploited by phishers? (I)

- Human tendency to trust brands, logos and trust indicators
- Users can not reliably and correctly determine the sender identity in emails
- Cannot reliably distinguish legitimate email and website content from illegitimate content that has the same “look and feel”
    - Users cannot reliably parse domain names Hsbc.mybank.com

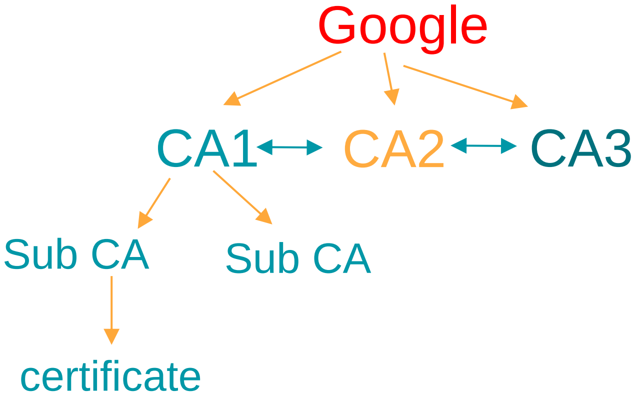
# What was exploited by phishers? (II)

- Cannot reliably distinguish actual hyperlinks from images of hyperlinks
- Cannot reliably distinguish an image of a window in the content of a webpage from an actual browser window
- Do not notice the absence of a security indicator
- Do not understand the meaning of the SSL lock icon
- Do not understand SSL certificates

# Security failures

- “Social engineering” attacks
  - Deception
  - Everything predictable can be spoofed
- Authentication failure
  - Cannot reliably authenticate a bank site
  - PKI is theoretically perfect, however …

# PKI failure



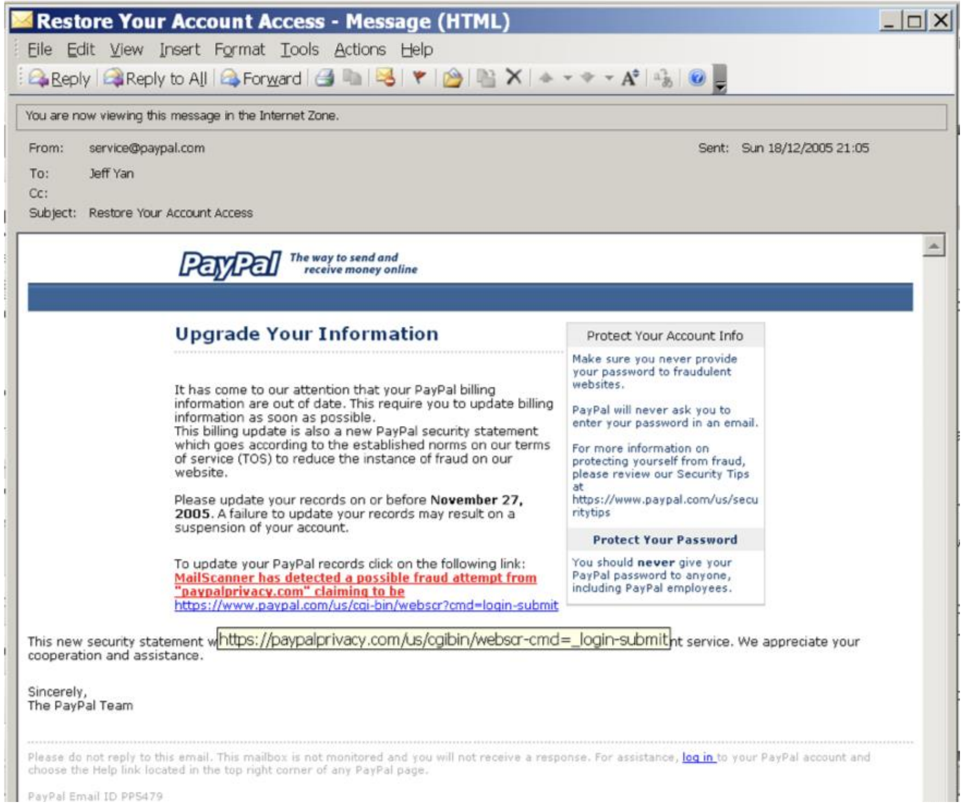
- About 400 CAs pre-installed in IE
- Breaking one CA means breaking entire PKI
- In 2011, DigiNotar was breached and fraudulent certificates were issued for Google, Yahoo! etc.

# Road map

- Introduction
  - Usability and psychology
- Example of attacking human weakness: phishing
  - What is phishing?
  - How does it work?
  - Why it works?
  - **Defence**
- Example of leveraging human strength: CAPTCHA
  - What is CAPTCHA?
  - How Does it work?
  - What could go wrong
  - CAPTCHA design

# Phishing defence 1

- Anti-spam solutions
  - Preventing phishing email from reaching the end user



Restore Your Account Access - Message (HTML)
File Edit View Insert Format Tools Actions Help
Reply Reply to All Forward

You are now viewing this message in the Internet Zone.

From: service@paypal.com
Sent: Sun 18/12/2005 21:05
To: Jeff Yan
Cc:
Subject: Restore Your Account Access

PayPal The way to send and receive money online

**Upgrade Your Information**

It has come to our attention that your PayPal billing information are out of date. This require you to update billing information as soon as possible.
This billing update is also a new PayPal security statement which goes according to the established norms on our terms of service (TOS) to reduce the instance of fraud on our website.

Please update your records on or before **November 27, 2005**. A failure to update your records may result on a suspension of your account.

To update your PayPal records click on the following link:
MailScanner has detected a possible fraud attempt from "paypalprivacy.com" claiming to be
https://www.paypal.com/us/cgi-bin/webscr?cmd=login-submit

Protect Your Account Info

Make sure you never provide your password to fraudulent websites.

PayPal will never ask you to enter your password in an email.

For more information on protecting yourself from fraud, please review our Security Tips at https://www.paypal.com/us/securitytips

**Protect Your Password**

You should **never** give your PayPal password to anyone, including PayPal employees.

This new security statement w https://paypalprivacy.com/us/cgibin/webscr-cmd=_login-submit nt service. We appreciate your cooperation and assistance.

Sincerely,
The PayPal Team

Please do not reply to this email. This mailbox is not monitored and you will not receive a response. For assistance, log in to your PayPal account and choose the Help link located in the top right corner of any PayPal page.

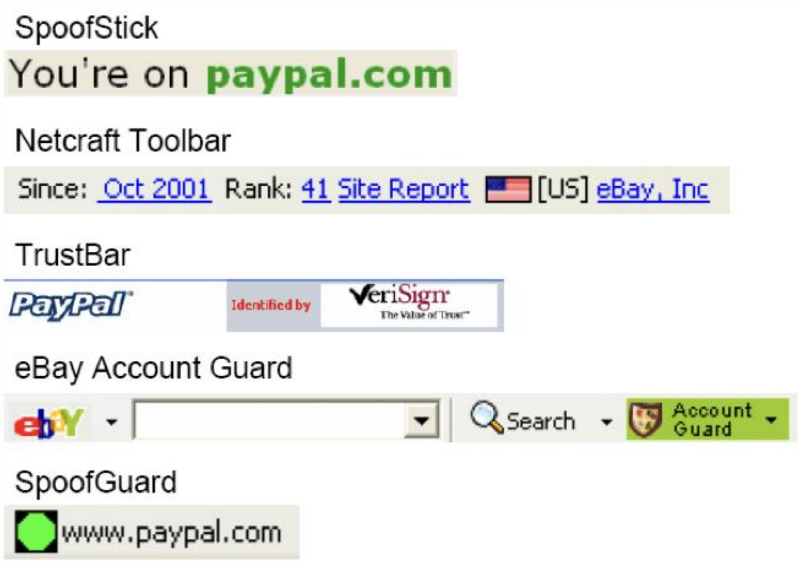
PayPal Email ID PP5479

# Phishing defence 1

- Difficulties
  - Anti-spam: no perfect solution (yet)
  - **False positive** could lead the users to miss important messages and they might disable anti-spam software completely
  - **False negative** annoy users
  - Which is more important for a user/business: false positive or false negative?

# Phishing defence 2

- Browser side solution: security toolbar
  - Implemented as browser plug-ins
  - Examples:

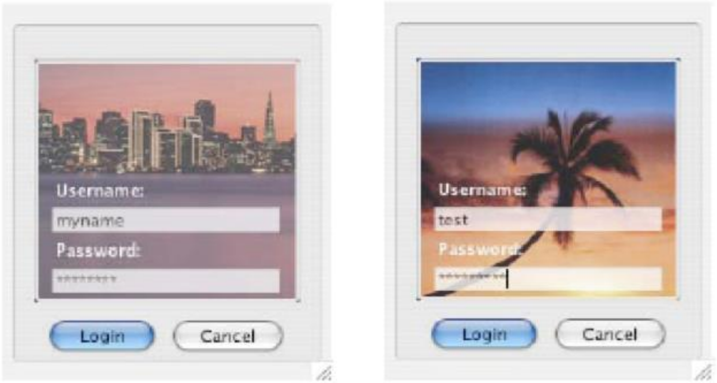


# Phishing defence 2

- Difficulties
  - How many plug-ins should the users install?
  - Plug-ins have to updated to stay one step ahead

# Phishing defence 3

- Trusted path (on Windows)
    - By pressing Ctrl + Alt + Del, an authentic Windows logon prompt appears
    - This enables secure sign in on Windows
- Use a background image that is easily recognisable for the user, but hard to predict for phishers.



Username: myname

Password: ********

Username: test

Password: *********

**Figure 1: The trusted password window uses a background image to prevent spoofing of the window and textboxes.**

# Phishing defence 4

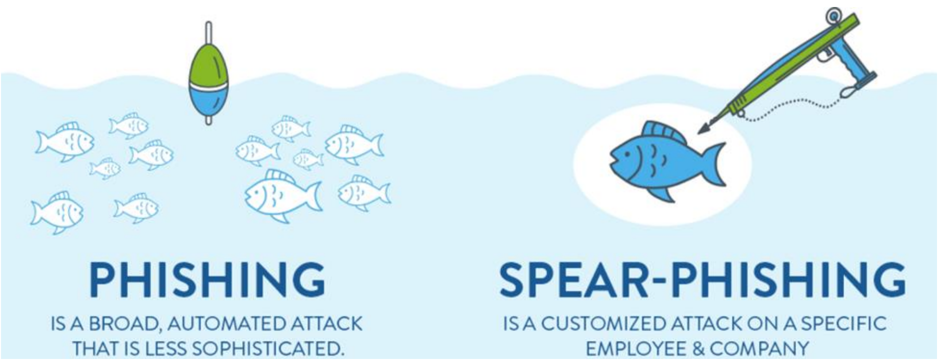
- Two factor authentication
- But doesn’t solve the problem.
- Only push phishing attacks to real-time

# Phishing defence 5

- Probably the best defence for the bank is not to be a soft touch.
    - Pursue fraudsters viciously and relentlessly (e.g., following the phishing links and go after the phishers)
    - Tighten internal control on money transfer
    - That won’t solve the phishing problem per se
    - But will push the phishing towards other banks

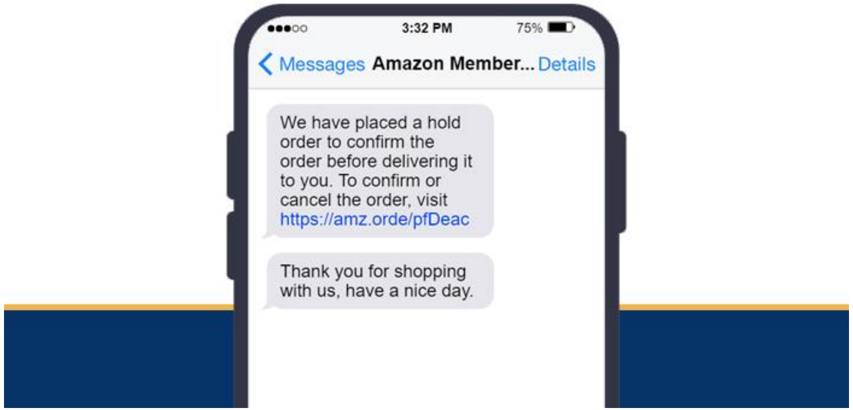
# Current trends

- Emails now dominated by few providers (Google, MS, AOL, Apple, Yahoo)
- This allows the better use of machine learning to detect spamming emails
- Hence, attackers are changing tactics
  - Random to targeted attacks (Spear Phishing)
  - Move to a different platform: mobile phones (smishing)


PHISHING
IS A BROAD, AUTOMATED ATTACK THAT IS LESS SOPHISTICATED.
SPEAR-PHISHING
IS A CUSTOMIZED ATTACK ON A SPECIFIC EMPLOYEE & COMPANY

# Smishing = SMS + phishing

Sample Smishing Attack



We have placed a hold order to confirm the order before delivering it to you. To confirm or cancel the order, visit https://amz.orde/pfDeac

Thank you for shopping with us, have a nice day.

- Attacks send phishing links via SMS instead of emails
- Why it works so effectively?
- People tend to act on links quicker on phones: aged 18 to 24 send 67 messages per day on average — and receive 1,831
- An attacker can spoof an arbitrary string as the SMS sender (e.g., police)
- Companies are moving away from SMS and using authenticators instead

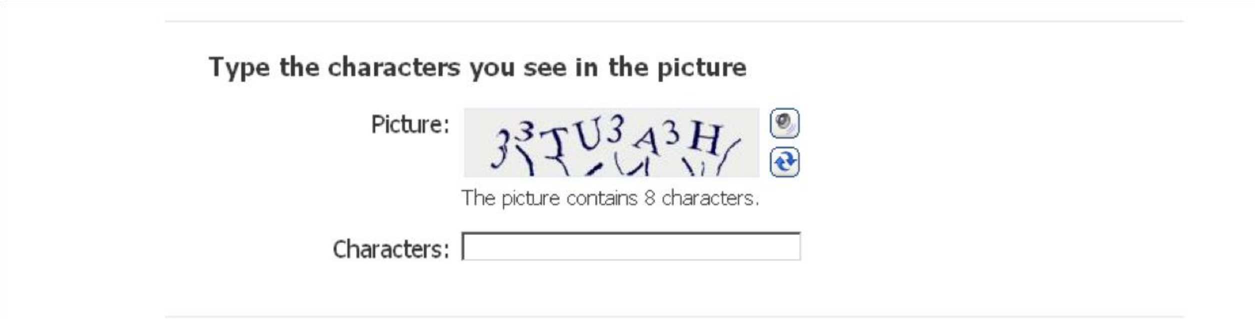
# Road map

- Introduction
  - Usability and psychology
- Example of attacking human weakness: phishing
  - What is phishing?
  - How does it work?
  - Why it works?
  - Defence
- **Example of leveraging human strength: CAPTCHA**
  - What is CAPTCHA?
  - How Does it work?
  - What could go wrong
  - CAPTCHA design

# What the Human Does Better Than the Computer?

- Human is not always doom and gloom in the security analysis.
- There are tasks that the human brain performs much better than a computer.
- For example, we are extremely good at recognizing other people’s faces and their voices – a hard problem for computer.
- We could make use of this asymmetry to improve security – for example, CAPTCHA

# What is CAPTCHA?



**Type the characters you see in the picture**

Picture:

The picture contains 8 characters.

Characters:

- **C**ompletely **A**utomated **P**ublic **T**uring Test to Tell **C**omputers and **H**umans **A**part
  - The term coined by Luis von Ahn, Manuel Blum, Nicholas J. Hopper, and John Langford (all from Carnegie Mellon University) in 2000

# Why was it invented?

"On the Internet, nobody knows you're a dog."

# Is a real human behind a computer? – a real story

- Slashdot.com 1999 online poll: “What is the best graduate school in CS?”
- Turned out to be a contest between bots written at Carnegie Mellon University (CMU) and MIT: 21,032 vs 21,156; every other school: < 1,000 votes
  - Voting bots vs human voters

# Bots: more examples

- Email account registration bots: could sign up for thousands of accounts every minute with free email service providers
- Weblog bots: could post numerous comments in weblogs pointing both readers and search engines to irrelevant sites

# CAPTCHA

- CAPTCHA was invented as a solution to tell humans and machines apart
- A CAPTCHA test is also called a CAPTCHA challenge
- CAPTCHA often makes use of a hard, open problem in AI
    - Optical character recognition
    - Speech recognition

# Turing test vs CAPTCHA (I)

- Famous questions by Alan Turing in 1950
    - Can a machine think?
    - If a computer could think, how could we tell?
- Turing test: if the responses from the computer were indistinguishable from that of a human, the computer could be said to be thinking (intelligent).

# Turing test vs CAPTCHA (II)

- Both aim to distinguish human from computers
- Main difference
  - Turing test: a human judge
  - CAPTCHA: a computer judge
- CAPTCHA is also called automated Turing Test.

# Road map

- Introduction
  - Usability and psychology
- Example of attacking human weakness: phishing
  - What is phishing?
  - How does it work?
  - Why it works?
  - Defence
- Example of leveraging human strength: CAPTCHA
  - What is CAPTCHA?
  - **How Does it work?**
  - What could go wrong
  - CAPTCHA design

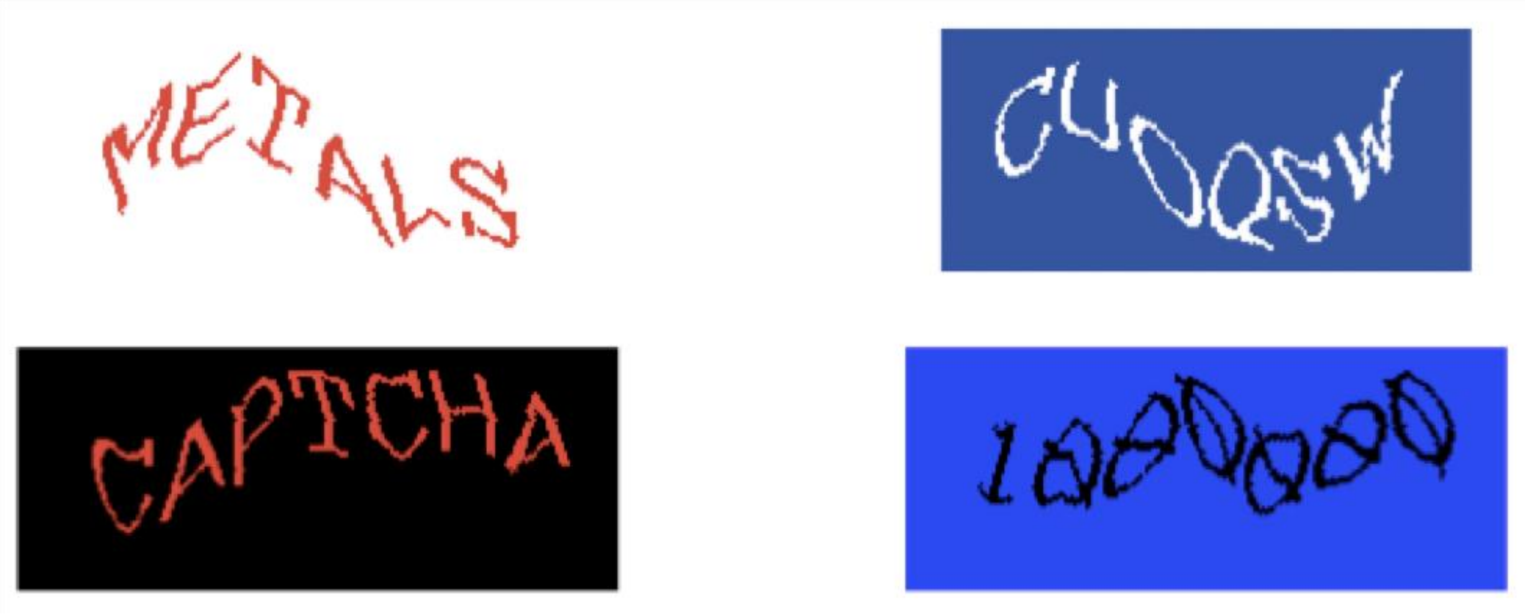
# As an authentication means

- Verifies that there is a human behind a computer
  - S -> C: a CAPTCHA challenge
  - C -> S: response
- Possible CAPTCHA challenges
  - Sound based: difficult to recognize human speech with a noisy background
  - Text based: difficult to read distorted text
  - Image based: difficult to recognize visual pattern
  - Any other hard problems could be used for CAPTCHA?

# Road map

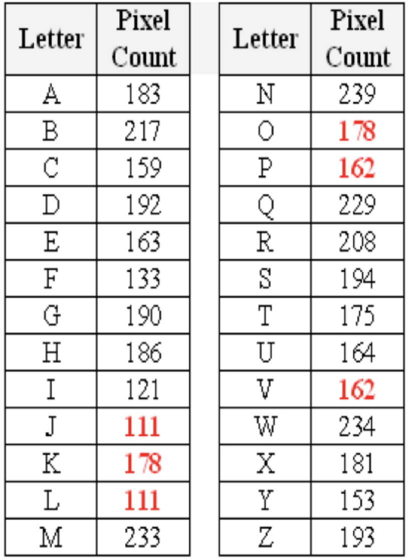
- Introduction
  - Usability and psychology
- Example of attacking human weakness: phishing
  - What is phishing?
  - How does it work?
  - Why it works?
  - Defence
- Example of leveraging human strength: CAPTCHA
  - What is CAPTCHA?
  - How Does it work?
  - **What could go wrong**
  - CAPTCHA design

# CAPTCHAs from Captchaservice.org

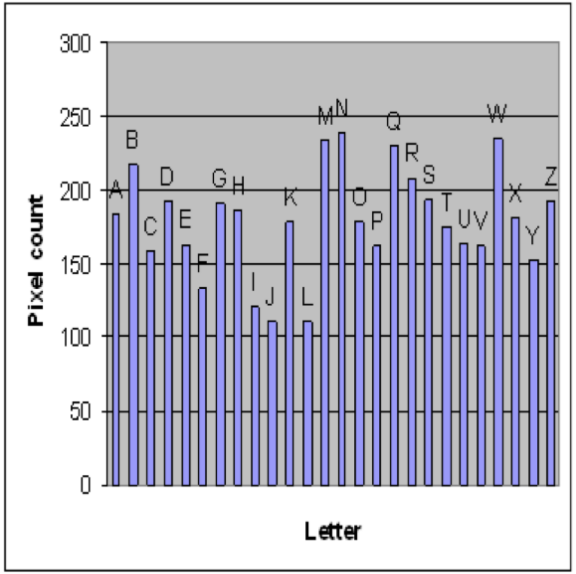


# Finding the weakest link …

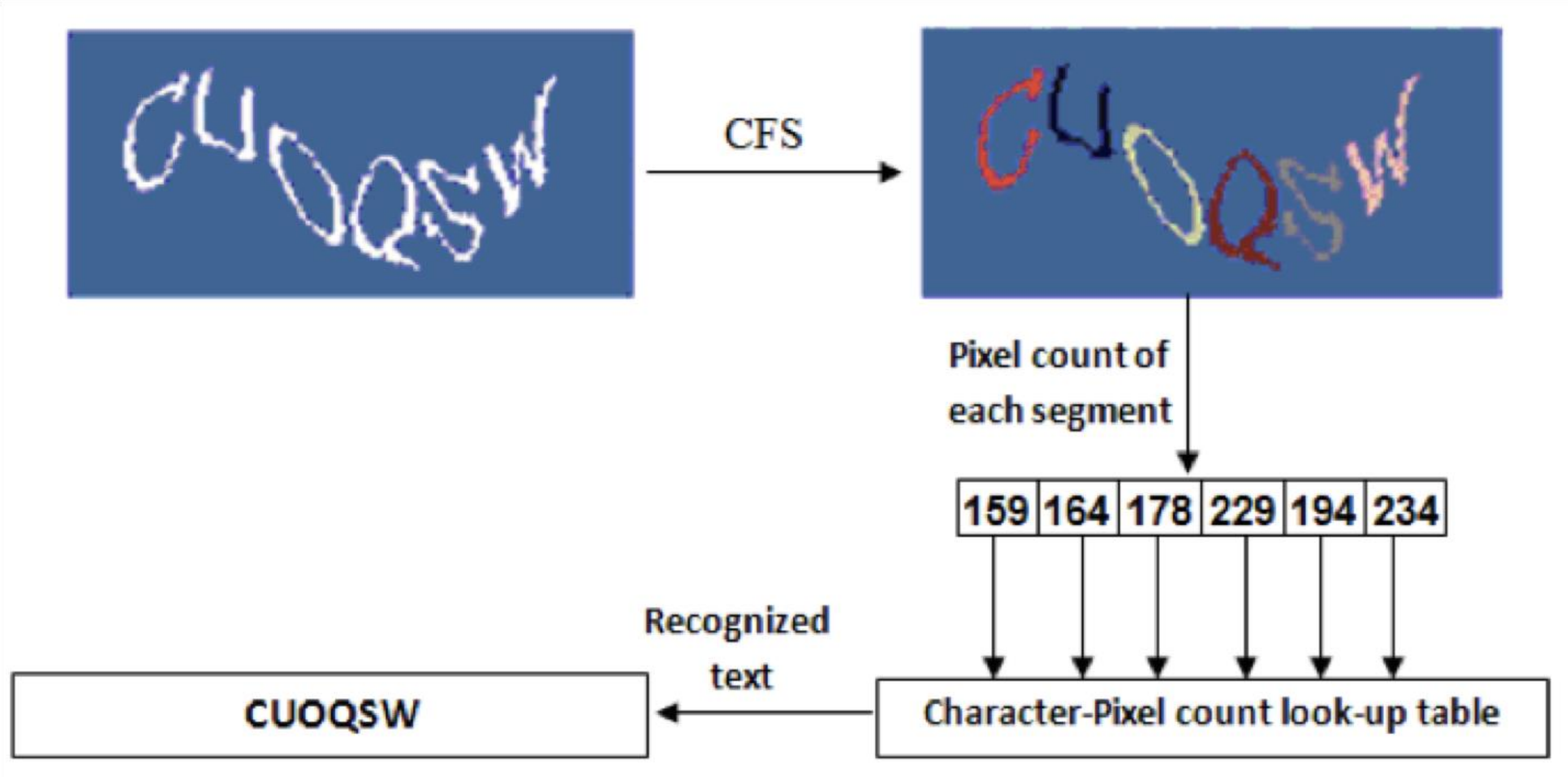
- Critical flaw: no matter how the characters are distorted, the pixel counts remain the same!



| Letter | Pixel Count |
|---|---|
| A | 183 |
| B | 217 |
| C | 159 |
| D | 192 |
| E | 163 |
| F | 133 |
| G | 190 |
| H | 186 |
| I | 121 |
| J | 111 |
| K | 178 |
| L | 111 |
| M | 233 |
| N | 239 |
| O | 178 |
| P | 162 |
| Q | 229 |
| R | 208 |
| S | 194 |
| T | 175 |
| U | 164 |
| V | 162 |
| W | 234 |
| X | 181 |
| Y | 153 |
| Z | 193 |



# Example of attack



CFS

Pixel count of each segment

| | | | | | |
|---|---|---|---|---|---|
| 159 | 164 | 178 | 229 | 194 | 234 |

Character-Pixel count look-up table

Recognized text

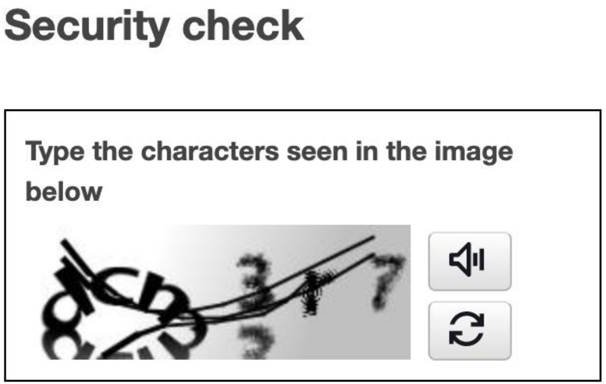
CUOQSW

# Attack results

- Almost 100% successful at breaking each CAPTCHA with little processing time.
- By design, the security of a CAPTCHA scheme should be based on the AI problem
- But breaking this scheme doesn’t involve solving any AI problem.
- Lesson from the attack: predictable patterns greatly weaken the security!

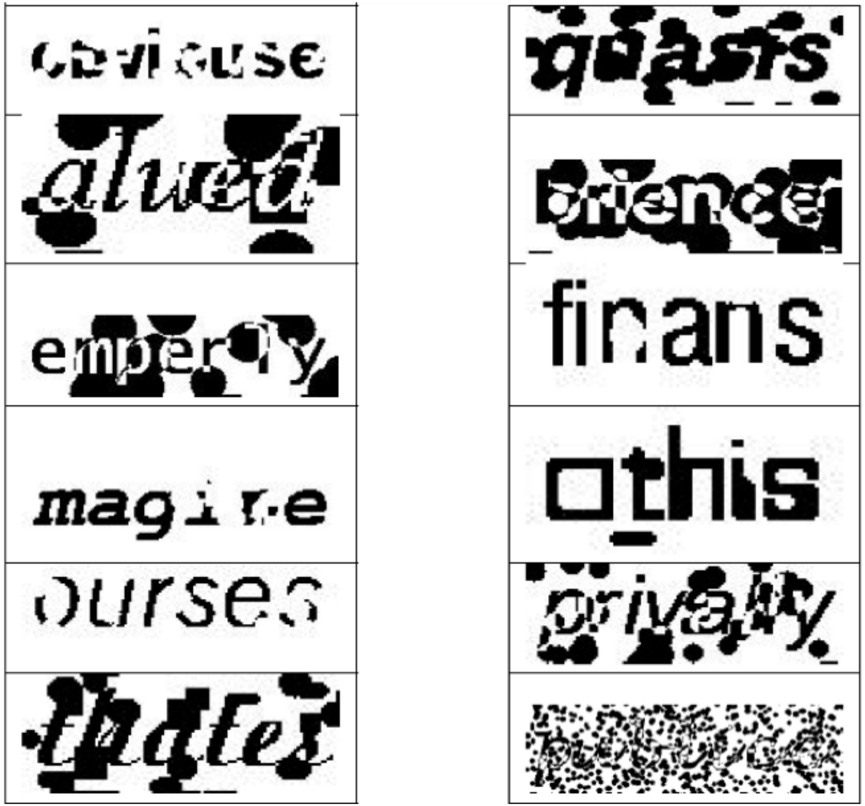
# More examples of CAPTCHA

- More creative ways to distort the text to make it difficult for the machine
- But, the also make it very difficult for the human.



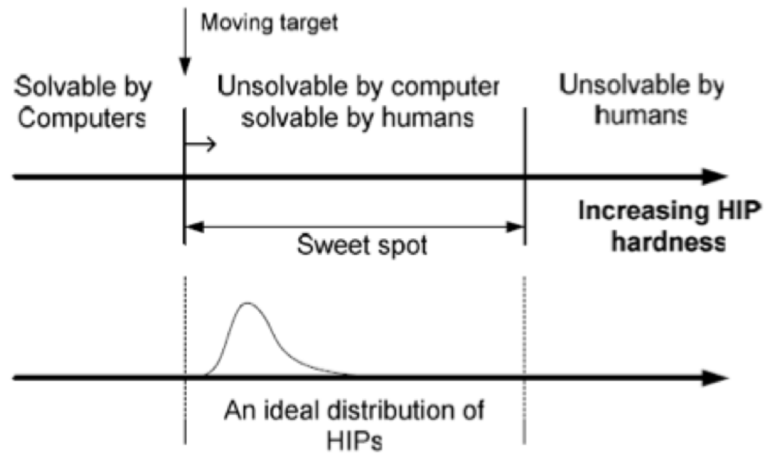
**Security check**

Type the characters seen in the image below



# CAPTCHA: sweet-spot

- The spot where CAPTCHA balances between strong security and good usability.
- Finding the sweet-spot is still an open problem.



# Protocol attack

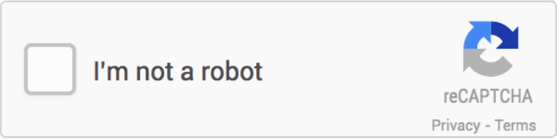
- Even if we have a perfect CAPTCHA scheme in the sweet-spot, there is still a fundamental problem in using CAPTCHA for authentication
- A generic attack for all CAPTCHA schemes
  - **Man in the middle attack**: adversary could shift the load of solving CAPTCHA challenges to porn site visitors
  - **Outsourcing attack**: bot users could outsource the task of solving CAPTCHA challenges to people in low-paying countries

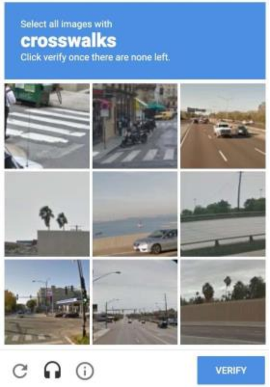
# Current trends

- The latest CAPTCHA technologies
    - **reCAPTCHA** v2
    - **reCAPTCHA** v3
- Regarded as state-of-the-art
    - Google is behind the reCAPTCHA
- Widely adopted by millions of websites

# reCAPTCHA v2

- Launched in 2014
- Using “risk analysis” that is heavily based on Google cookies
- Many users just need to tick a box (e.g., those using Chrome or logged into Google account)
- But others need to solve an image/audit recognition problem (e.g., Firefox user who disabled third-party cookies)
- Attacking reCAPTCHA v2:
  - Use AI technology to perform image recognition
  - Redirect challenges to human ($1-3 to solve 1000 reCAPTCHA v2)




Select all images with
**crosswalks**
Click verify once there are none left.

# reCAPTCHA v3

- reCAPTCHA v3 developed to improve user experience over v2
- Unlike v2, v3 is invisible for website visitors
- However, the burden is now shifted to be on the website administrators
- V3 returns a score between 0 and 1
- The website admin need to decide the threshold and what to do
  - Grant access
  - Ask the user a challenge (downgrading to v2)
  - Block access
- Difficult decisions even for expert webmasters

# Limitations of reCAPTCHA v2 and v3

- Solution strongly tied to Google (best user experience on Chrome, and less satisfactory experience on other browsers)
- Free only up to one million API calls a month - reCAPTCH Enterprise typically charges $1 for every thousand calls

# Research problems

- Systematic analysis of phishing websites
- Protocol attack against CAPTCHA - is a solution ever possible?
- New ways to build CAPTCHA for *mobile devices*
- Can we have better solutions than CAPTCHA v2 and v3? (not tied to any specific company and can be freely available)