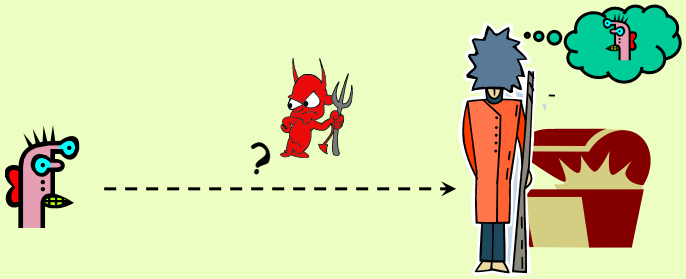
CS 203 / NetSys 240

User Authentication

0

Basic Problem



How do you prove to someone that you are who you claim to be?

Any modern system (distributed or not) with access control must solve this problem

1

Many Ways to Prove Who You Are

- What you know
  - Passwords, answers to "personal" questions
  - Secret keys
- Where you are
  - IP address, home phone, cell phone
  - What about MAC addresses?
- What you are
  - Biometrics, e.g., face or fingerprint
- What you have
  - Secure tokens or smartphone (e.g., DUO)

2

Password-Based Authentication

- User has a secret password
  System checks it to authenticate the user
- How is is the password initialized (set)?
- How is the password communicated at login time?
  - Eavesdropping risk
- How is the password stored on the "other side"?
  - In the clear? Encrypted? Hashed?
  - In sw? In hw? Tamper-resistant?
- How does the system check the password?
- How easy is it to guess the password?
  - Easy-to-remember passwords are easy-to-guess
  - Password file is difficult to keep secret

3

Other Aspects

- Usability
  - Hard-to-remember passwords?
  - Carry a physical object (with passwords)
  - Password hints?
  - Password vault
- Denial of Service (DoS)
  - Stolen wallet, destroyed wallet
  - Attacker tries to authenticate as you, account locked after three failures
  - "Suspicious" credit card usage
- Social engineering (works often, exploits gullibility)
  - e.g., attacker who knows your name, SSN, DoB, etc. calls your bank to re-set account password
  - Or, attacker calls your employer's IT dept., pretends to be VP of janitorial affairs, asks for VPN access as an emergency

4

Passwords Breaches in the Real World

[PasswordResearch.com]

- From high school pranks…
  - Student in Texas changes school attendance records
  - Students in California change grades
    - Different authentication for network login and grade system, but teachers were using the same password (very common)
- …to serious cash
  - British accountant uses co-workers' password to steal $17 million for gambling
- …to identity theft
  - Helpdesk employee uses passwords of a credit card database to sell credit reports to Nigerian scammers

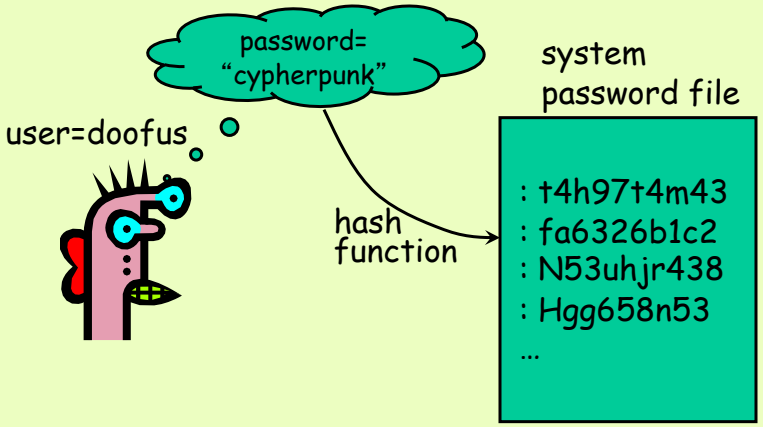
6

First passwords and then what?

- First step after any successful intrusion: install sniffer or keylogger to steal more passwords
- Second step: run cracking tools on **password files**
  - Usually on other hijacked computers
- In Kevin Mitnick's "Art of Intrusion", 8 out of 9 exploits involve password stealing and/or cracking
  - Excite@Home: usernames and passwords stored in the clear in troubleshooting tickets
    - Note: anyone ever bring a MAC in for repair? I did…
  - "Dixie bank" hack: use default router password to change firewall rules to enable incoming connections

7

UNIX-Style Passwords



8

Password Hashing

- Instead of user password, store H(password)
- When user enters password, compute hash and compare with entry in password file
  - System does not store actual passwords!
  - Difficult to go from hash to password!
    - Do you see why hashing is better than encryption here?
- Hash function H must have some properties:
  - One-way: given H(password), hard to find password
    - No practical algorithm better than simple trial and error
  - Is collision resistance needed?
    - Weak and/or strong?

9

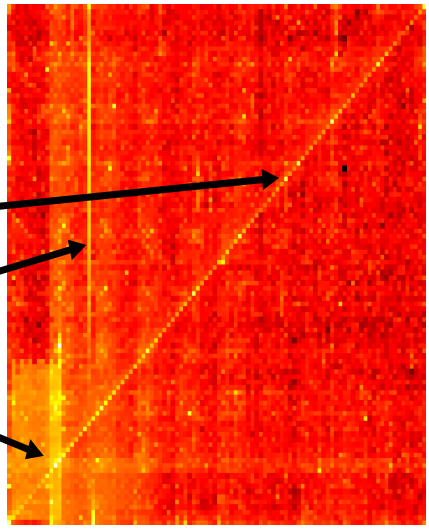
UNIX Password System

- Uses DES encryption as a hash function
  - Encrypts NULL string (repeatedly) using password as the key
    - Truncates passwords to 8 characters!
  - Artificial slowdown: runs DES 25 times
- Problem: passwords are not truly random
  - With 52 upper- and lower-case letters, 10 digits and 32 punctuation symbols, there are $94^8 = 2^{52} \approx$ 6 quadrillion possible 8-character passwords
  - Humans like to use dictionary words, human and pet names, which brings it down to only ≈ 1 million = 2^{20} common passwords
  - Are PINs better? 4-8 decimal digits = max. 2^{27}

10

A PIN heatmap

Other interesting insights into how people think about PINs…



Source: N. Berry. Datagenetics. 3 Sept. 2012.

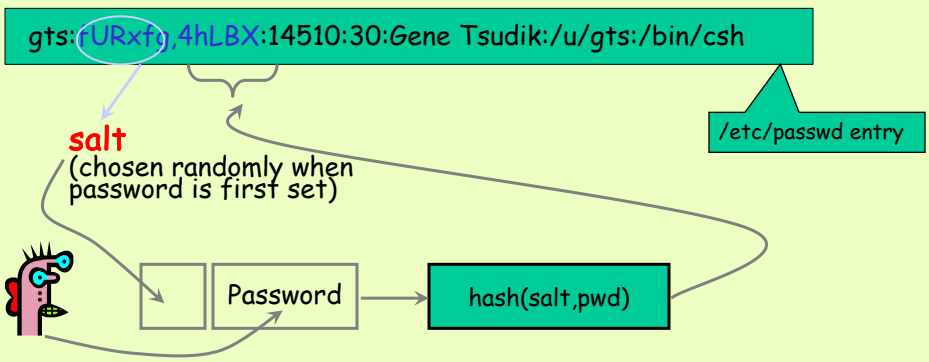
11

Dictionary Attack

- Unix password file /etc/passwd is world-readable
  - Contains user IDs and group IDs which are used by **many** system programs
- Dictionary attack is possible because most passwords come from a small dictionary
  - Attacker can pre-compute **H(word)** for every word in the dictionary – this only needs to be done once!!
    - This is an example of an offline (passive) attack
    - Once password file is obtained, cracking is instantaneous
  - With 1,000,000-word dictionary and assuming 10 (interactive!) guesses per second, brute-force online attack would take 50,000 seconds (14 hours) on average

12

Salt is good for you…



- Users with the same password have <u>different</u> entries in the password file
- **Offline** dictionary attack becomes much harder

13

Advantages of Salting

- Without salt, attacker can pre-compute hashes of all dictionary words once for <u>all</u> password entries
  - Same hash function on all UNIX machines; identical passwords hash to identical values
  - One table of hash values works for all password files
- With salt, attacker must compute hashes of all dictionary words once for <u>each</u> combination of salt value and password
  - With 12-bit random salt, same password can hash to 4096 different hash values

14

Shadow Passwords

gts:x:14510:30:Gene Tsudik:/u/gts:/bin/csh

/etc/passwd entry

Indicates that hashed password is not stored in a world-readable file

- Store hashed passwords in /etc/shadow file which is only readable by system administrator (root)
- Add expiration dates for passwords

15

Password Security Risks

- Keystroke loggers
  - Hardware
    - KeyGhost, KeyShark, others
  - Software (spyware)
- Shoulder surfing
- Video cameras
- Thermal cameras (post factum)
- Keyboard acoustic emanations (even via VoIP)
- Good vibrations? ☺

- Same password at multiple sites
- Broken implementations
- Social engineering

16

How People Use Passwords

- Write them down
- Use same password at multiple sites
  - Do you use the same password for Amazon, X, Facebook, Instagram, and bank account? UCI net access? Do you remember them all?
- Make them easy to remember
  - "password1!", "ZotZot123", "Antmuncher"
- Some services use "secret questions" to reset passwords, e.g.:
  - What is your favorite pet's name?
    - Early 2000-s Paris Hilton's T-Mobile cellphone hack
  - When did you graduate from high school?
  - What is your mother's maiden name?
  - What was the make of your first car?

17

Hotmail Passwords (2009)

- 10,000 Hotmail passwords posted to Pastebin
  - Obtained from a phishing attack
    - All account names start with "A" or "B"
  - Two more lists found later (Gmail, AOL, Yahoo)
- Most common password: "123456"
- 20% are only 6 characters long
- 42% use only lower-case letters
- Only 6% mix in numeric and other characters

18

Also in 2009: 32,000,000 Leaked Passwords from RockYou Database:

- "Social gaming" company database with 32 million user passwords from partner social networks
- Passwords stored in the clear
- December 2009: entire database hacked using a SQL injection attack and posted on the Internet

http://www.theregister.co.uk/2010/01/21/lame_passwords_exposed_by_rockyou_hack/
http://www.pcmag.com/article2/0,2817,2358273,00.asp

19

2009: 320,000,000 Leaked Passwords from RockYou Database:

Password Popularity – Top 20 [Imperva]

Rank	Password	Number of Users with Password (absolute)
1	123456	290731
2	12345	79078
3	123456789	76790
4	Password	61958
5	iloveyou	51622
6	princess	35231
7	rockyou	22588
8	1234567	21726
9	12345678	20553
10	abc123	17542

Rank	Password	Number of Users with Password (absolute)
11	Nicole	17168
12	Daniel	16409
13	babygirl	16094
14	monkey	15294
15	Jessica	15162
16	Lovely	14950
17	michael	14898
18	Ashley	14329
19	654321	13984
20	Qwerty	13856

http://www.theregister.co.uk/2010/01/21/lame_passwords_exposed_by_rockyou_hack/
http://www.pcmag.com/article2/0,2817,2358273,00.asp

20

Memorability vs. Security

- Beware of "clever" tricks…
- One bank's idea for making PINs "memorable"
  - If PIN is 2256, write your favorite word in the grid

1	2	3	4	5	6	7	8	9	0
	b								
	l								
				u					
					e				

Normally 9,999 choices for PIN – hard to guess

Now only a few dozen possible English words – easy to guess!

  - Fill the rest with random letters

[Ross Anderson, "Why Cryptosystems Fail?"]

21

Heuristics for Guessing Attacks

- Dictionary with words spelled backwards
- First and last names, streets, cities
- Same with upper-case initials
- All valid license plate numbers in your state
- Room numbers, telephone numbers, etc.
- Foreign languages?
- Letter substitutions and other tricks
  - If you can think of it, attacker will, too
- Enforcing selection of upper-/lower-case letters, numbers, special characters. How effective is this?

22

What about Password Vaults or Password Managers?

The Good:

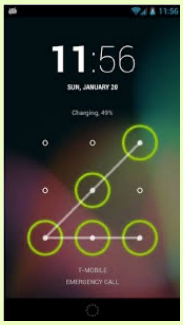
- Really strong random passwords, no re-use!
- One place to go for all of them
- Only one good/strong password to remember

The Bad:

- Super-duper attractive target for attackers
  - Also, keep in mind insider attacks!
- Single point of failure: outages, DoS attacks
- Internet (dis)connectivity

24

Screen Unlock Patterns



The Good:

- Easy to remember
- Easy to use/enter
- More intuitive than passwords?

The Bad:

- Specific to phone/screen (i.e., not universal)
- Screen smudges (leak info)
- Acoustic Leakage (like from keyboards+passwords)
- More susceptible to shoulder-surfing. Why?
- People tend to pick easy/predictable patterns

25

Strengthening Passwords

- Add biometrics
  - e.g., fingerprints, iris scans, keystroke dynamics or voiceprints
  - **Revocation** is a problem
- Graphical passwords
  - Goal: increase size of memorable password space
- Example: image recognition passwords
  - Rely on difficulty of computer vision
  - Image recognition is easier for humans, harder for machines
  - Present user with a sequence of images, ask to pick the right one(s) several times in a row, in order to log in

26

Graphical Passwords

- Images are easy for humans to remember
  - Especially if you invent a memorable story to go along with the images
- Dictionary attacks on graphical passwords are believed to be difficult
  - Images are "random" (is this true?)
- Still not a perfect solution
  - Need infrastructure for displaying and storing images (PINs/passwords don't need graphics)
  - Shoulder surfing

27

An example:
"Passfaces"
Meets the Challenge

"For every problem, there is a solution that is simple, elegant, and wrong."
(H. L. Mencken)

28

The Brain Deals with Faces Differently than Any Other Image

Face recognition is a dedicated process which is different from general object recognition.

Source: Face Recognition: A Literature Survey. US National Institute of Standards and Technology (NIST)

29

Recall vs. Recognition

You must **RECALL** a password

You simply **RECOGNIZE** a face

Remember High School …. ***What kind of test did your prefer?***

Fill in the Blank

1 2 3 g f w y _____

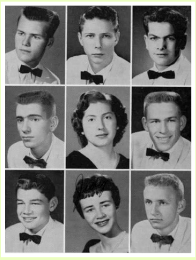
Multiple Choice

30

We Never Forget a Face

Think about how many people you already recognize.

Why wouldn't you remember your Passfaces?



- *"Haven't used Passfaces in 6 months. I decided to take another look at it and, amazingly, I logged right in!"*
- *In one major government installation, there have been no forgotten Passfaces in over three years. The more it's used, the easier it gets.*

31

The approach

Familiarize the user with a randomly-selected set of faces and check if they can recognize them when they see them again

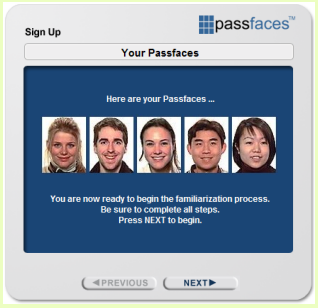
It's as easy as recognizing an old friend

32

How Passfaces Works

Library of Faces

User Interface



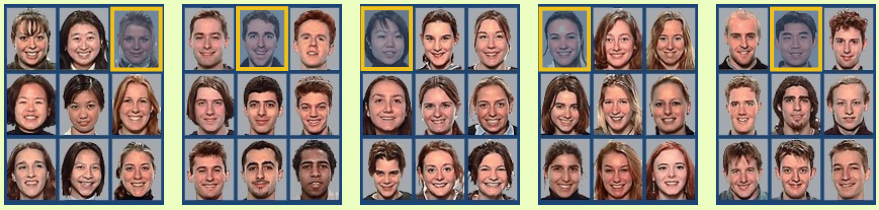
Users Are Assigned a Set of 5* Passfaces

* Typical implementation – 3 to 7 possible as standard

33

How Passfaces Works

- 5 Passfaces are Associated with 40 associated decoys
- Passfaces are presented in five 3 by 3 matrices each having 1 Passface and 8 decoys



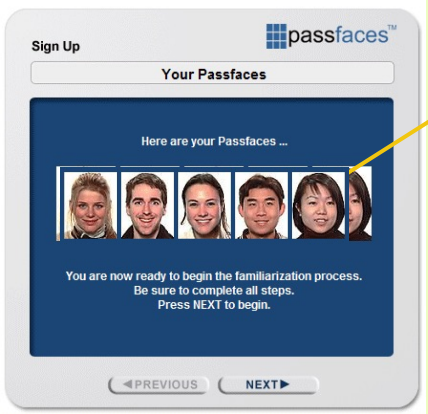
34

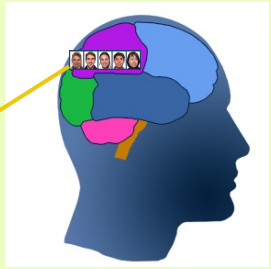
New Users are Familiarized with their Passfaces

- Users familiarize via 2-4 minute familiarization process
- Using instant feedback, encouragement, and simple dialogs, users are *trained* until they can easily recognize their Passfaces
- The process is optimized and presented like an easy game

35

Familiarization Puts Cookies in the Brain





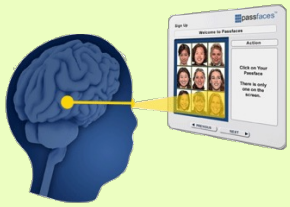
Like a *mindprint* or *brain cookie*

But, unlike fingerprints,
Passfaces require no special hardware

And, unlike browser cookies,
Passfaces authenticate the actual user

36

Cognometrics: A New Class of Authentication



- Passfaces represents a new, 4th class of authentication:

Cognometrics

Recognition-Based Authentication

37

Two Variants:

- System picks passfaces for user
  - Much harder to remember
  - Requires more/longer training
- User picks own passfaces from library
  - Easier
  - Less training (some needed anyhow)

What if user uploads his own passface choices?

38

Empirical Results

- Experimental study of 154 *computer science* students at Johns Hopkins and Carnegie Mellon
- User-selected Passfaces variant (where users pick)
- Conclusions:
  - "… faces chosen by users are highly affected by the race of the user… the gender and attractiveness of the faces bias password choice… In the case of male users, we found this bias so severe that we do not believe it possible to make this scheme secure against an online attack…"
- 2 guesses enough for 10% of male users
- 8 guesses enough for 25% of male users

See: http://awildduck.com/?p=2268

39

So…

- Should users choose their own faces?
- Or should system do it for them?
- Some people are dysfunctional in terms of face recognition
- Need good-enough/large-enough display
  - Unlike PINs/PWs
  - Not good for smartphones
- How cumbersome is changing one's Passfaces set?

42

Shoulder Surfing

- What is it?
- Graphical password schemes are perceived to be more vulnerable to "shoulder surfing" and video recording
- Experimental study with graduate students at UMBC
  - 4 types of passwords: Passfaces with mouse, Passfaces with keyboard, dictionary text password, non-dictionary text password (random words and numbers)
- Result: non-dictionary text password most vulnerable to shoulder surfing
  - Why do you think this is the case?

45

Why?

The 25 Most Popular Passwords of 2015: We're All Such Idiots

http://gizmodo.com/the-25-most-popular-passwords-of-2015-were-all-such-id-1753591514

46

Biometric Authentication

- Nothing to remember
- Nothing to compute
- Low-burden: no devices to carry around
- Can't share biometrics (usually)
- Can be fairly unique
  - … if measurements are sufficiently accurate

47

Problems with Biometrics

- Identification vs. authentication
  - Identification = associating an identity with an event or a piece of data
    - Example: fingerprint at a crime scene
  - Authentication = verifying a claimed identity
    - Example: fingerprint scanner of a live finger to enter a building
- How hard is it to forge biometric readings?
  - Difficulty of forgery is routinely overestimated
  - Analysis often doesn't take into account the possibility of computer-generated forgery
- Revocation is difficult or impossible
- Potentially expensive, trusted infrastructure

48

Desired Properties

- Universality: Usable by (almost) everyone who needs to it, i.e., most people.
- Uniqueness: Sufficient degree of distinction within target population.
- Permanence: Consistent over time, i.e., stable.
- Unobtrusiveness: Does not disrupt normal work-flow / behavior, i.e., minimal burden.
- Circumvention Difficulty: Hard mimic/modify or impersonate others.
- Equipment Cost: low-cost added features for end-user platform
- Bootstrapping: Easy/short enrollment phase

49

Biometric Error Rates

- "Fraud rate" = false positive vs. "insult rate" = false negative
  - Fraud = system accepts a forgery
  - Insult = system rejects valid user
- Increasing acceptance threshold increases fraud rate, decreases insult rate
  - Pick a threshold so that fraud rate = insult rate
- For example, banks often set target fraud rate of 1%, insult rate of 0.01%
  - Common hand-written signature recognition systems achieve both error rates of around 1% - not good enough!

50

Biometrics (1)

- Face recognition (by an algorithm)
  - Error rates about 1% for best algorithms on powerful hw (down from 5% 10 years ago) given reasonable variations in lighting, viewpoint and expression
  - Rates go sharply up for weak hw, often up to 40%
  - Injuries, hair, medicines/drugs, plastic surgery, acne, etc.
- Fingerprints
  - Traditional method for identification
  - 1911: first US conviction on fingerprint evidence
  - UK traditionally requires 16-point match
    - Probability of false match is 1 in 10 billion
    - No successful challenges until 2000
  - Fingerprint damage impairs recognition
    - Eczema, scars, missing fingers?

51

Biometrics (2)

- Iris scans
  - Irises are very random and unique, but stable throughout one's life
    - Different for two eyes of the same individual
  - 256-byte iris code based on concentric rings between the pupil and the outside of the iris
  - Error rate lower than 1 in a million
  - Best biometric mechanism currently known
  - Expensive… Not well-liked
  - Can be done on smartphone but accuracy goes down
- Hand geometry
  - Used in nuclear premises entry control, INSPASS (discontinued in 2002)
- Others: voice, ear shape, wrist vein pattern, voice, DNA, keystroke dynamics

- Best I've ever seen: IBM on-line challenge-based handwriting recognition system.
  http://www.computer.org/csdl/trans/tp/1990/08/i0787-abs.html

52

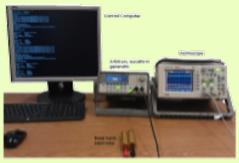
Pulse Response Biometric (NDSS'14)

home-grown ☺
https://dl.acm.org/doi/abs/10.1145/3023359

- Pulse signal applied to the palm of user's one hand.
- Biometric is captured by measuring response in the user's other hand.

53

User Safety



Voltage (V)	1	1.5
Max Current (mA)	0.1	500+
Exposure	100ns	~500ms

Also: touch-lamps and touch-switches
NOTE: IRB authorizations obtained at UCI and Oxford.

54

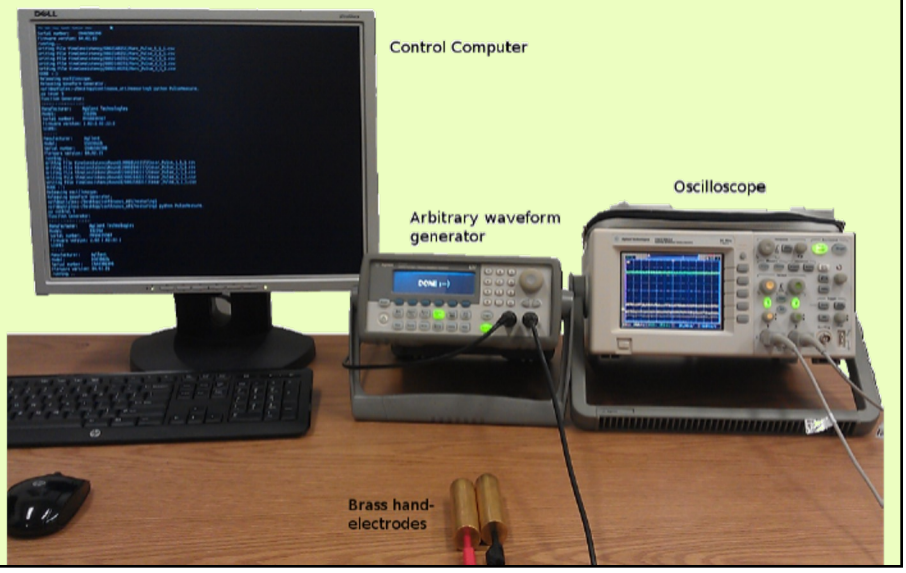
Scenario 1: PIN Entry

55

Scenario 2: Continuous Authentication

56

Experimental Setup @ UCI



58

Subjects

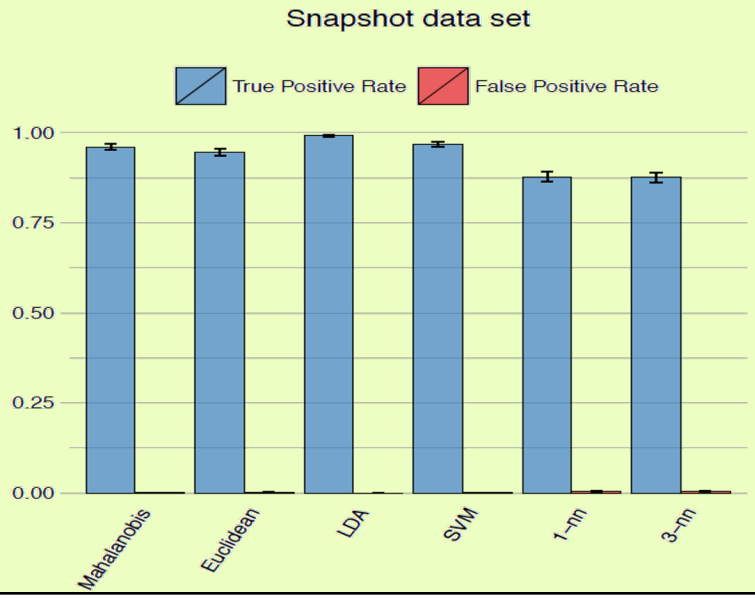
Table I. Test subject population and sample size. The age band of the subject population ranges from 24 to 38.

Data set	Test subjects	Females	Males	Samples per subject
Snapshot	30	9	21	20
Over-time†	16	2	14	25

†Test subjects were measured in five different sessions over time.

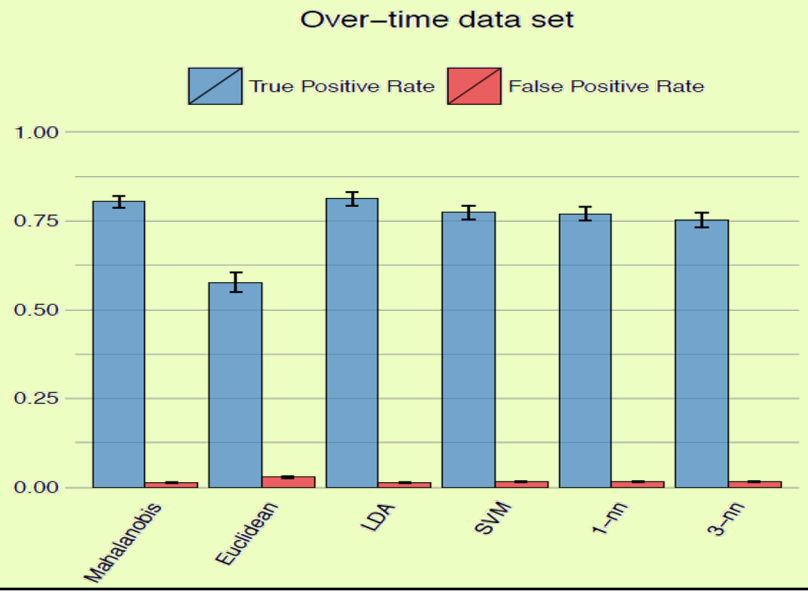
59

Identification (snapshot)



60

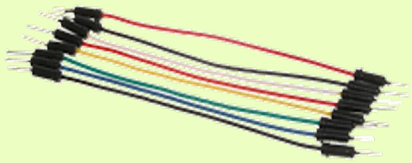
Identification (over time)



61

How to subvert this biometric?

- Measure victim's pulse response
- Build a contraption mimicking it exactly



62

Risks of Biometrics

- DoS: Criminal gives an inexperienced policeman fingerprints in the wrong order
  - Record not found; gets off as a first-time offender
- Voice-Prints: Can be attacked using recordings
  - Trivial to subvert voice-recognition systems
- In countries where fingerprints are used to pay pensions, there are persistent tales of "Granny's finger in the pickle jar" being the most valuable property she bequeathed to her family
- Birthday paradox
  - With false accept rate of 1 in a million, probability of false match is above 50% with only 1609 samples

64

Play-Doh Fingers

[Schuckers]

- Alternative to gelatin
- Play-Doh fingers fool 90% of fingerprint scanners
  - Clarkson University study
- Suggested perspiration measurement to test "liveness" of the finger

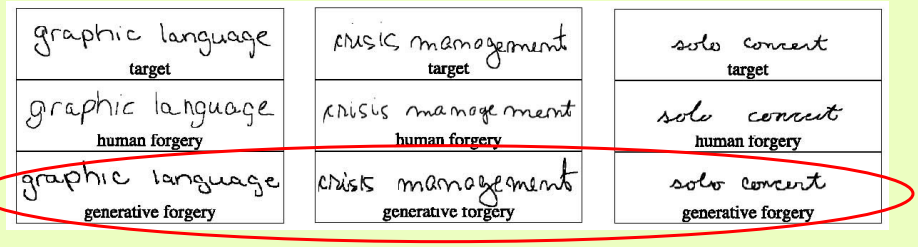
65

Bypassing Biometrics

66

Forging Handwriting

[Ballard, Monrose, Lopresti]



Generated by computer algorithm trained on handwriting samples

67

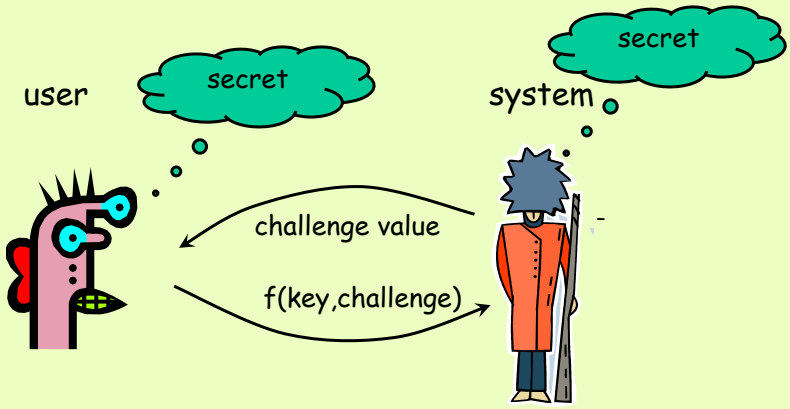
Biometrics: summary

- Only partially effective
- Tricky to use on a large scale
- Require in-person enrollment
- Hard to revoke
- Require pervasive infrastructure
  - E.g., iris scanners at each point-of-access?

- Biometrics are about "what you ARE"
- What about "what you have"?

68

Challenge-Response



Why is this better than a password over a network?
Can the user compute f()??

69

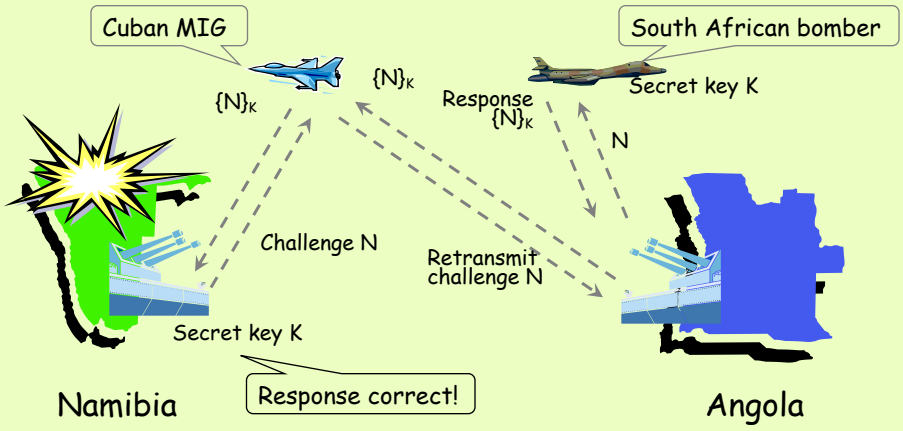
Challenge-Response Authentication

- User and system share a secret key
- Challenge: system presents user with a random number
- Response: user computes response based on secret key and challenge
  - Secrecy: difficult to recover key from response
    - One-way hashing or symmetric encryption work well
  - Freshness: if challenge is long-enough (>=160 bits), fresh, and unpredictable, attacker cannot replay an old response
- Good for systems with pre-installed secret keys
  - Car keys; military friend-or-foe identification;
  - What about humans? Badges? Dongles? Fobs? Smartphones?

70

MIG-in-the-Middle Attack

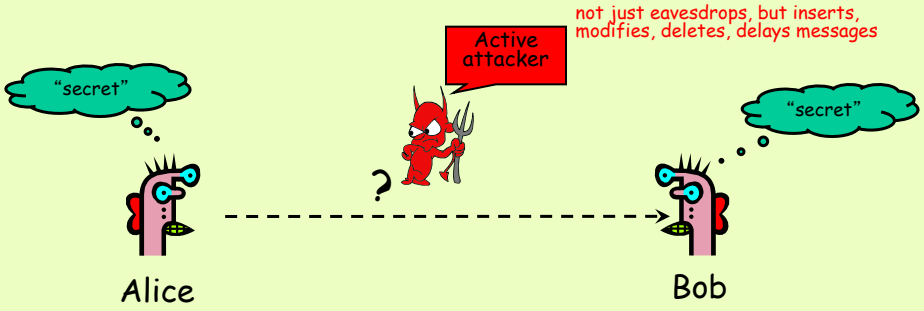
[Ross Anderson]



Any ideas on how to fix this?

71

Authentication with Shared Secret

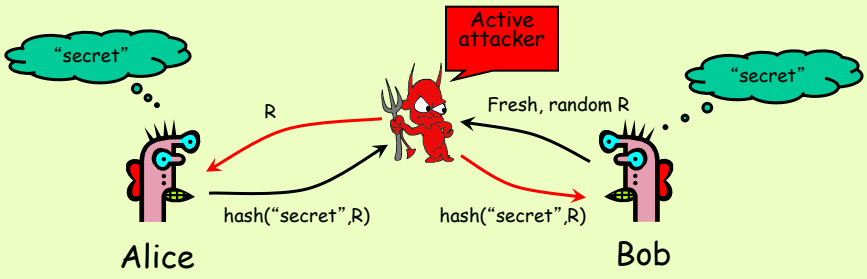


- Alice and Bob share a secret.
- How can they identify each other over a network?

What have we learned from the systems we've seen?

72

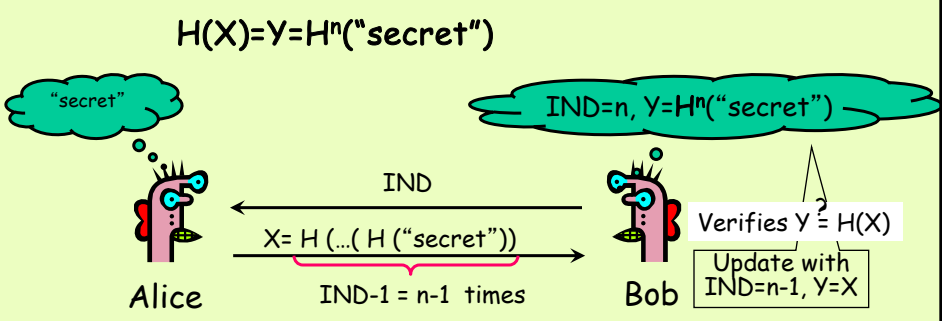
Challenge-Response: MiTM



- Man-in-the-middle (MiTM) attack on challenge-response
  - Attacker successfully authenticates as Alice by simple replay
- This is an attack on authentication, not secrecy
  - Attacker does not learn the shared secret
  - However, response opens the door to a dictionary attack

73

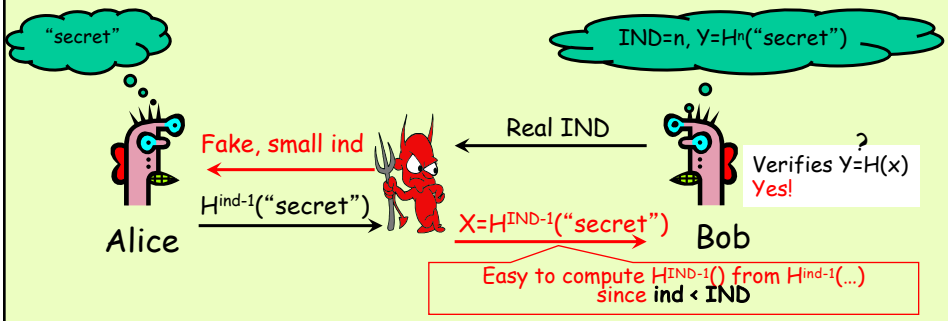
Lamport's Hash (used in S/Key OTP system)



- Main idea: "hash chain"
  - Moving up the chain (computing next hash) is easy, moving down the chain (inverting the hash) is hard
  - n should be large (can only use it for n authentications)
- For verification, only need the "root" Y (top link) of chain

74

"Small ind" Attack

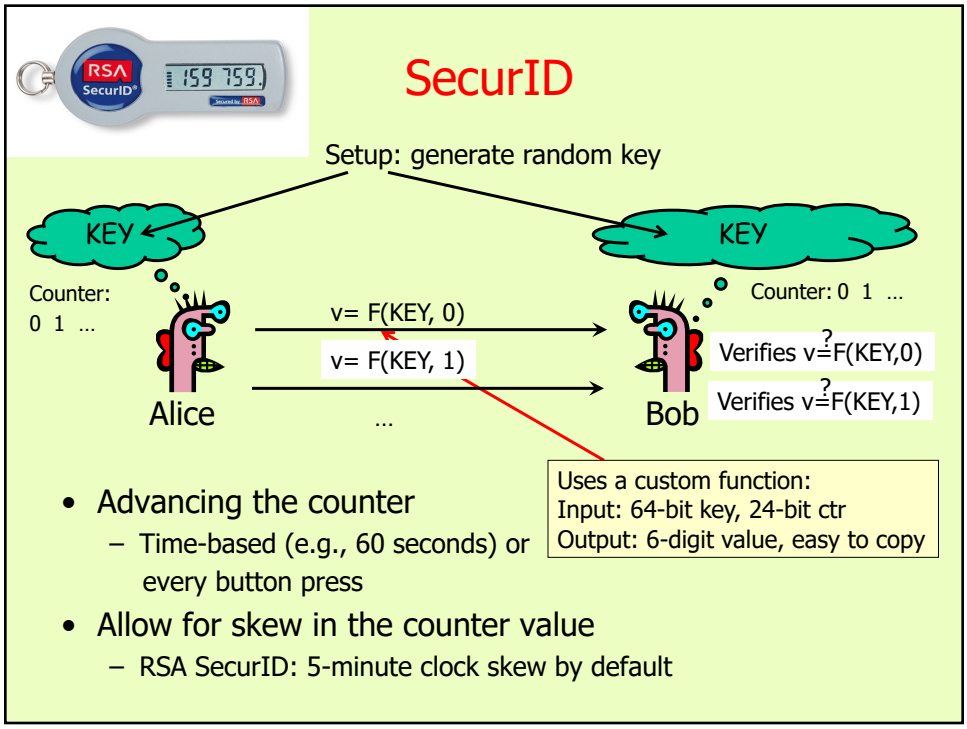


$H^{(IND-1)-(ind-1)}\left(H^{ind-1}(\text{"secret"})\right) = H^{IND-1}(\text{"secret"})$

Problems:

- First message from Bob is not authenticated!
- Alice should remember/store current value of IND

75


SecurID
Setup: generate random key
KEY
KEY
Counter:
0 1 …
Counter: 0 1 …
v= F(KEY, 0)
v= F(KEY, 1)
…
Alice
Bob
Verifies v=?F(KEY,0)
Verifies v=?F(KEY,1)
Uses a custom function:
Input: 64-bit key, 24-bit ctr
Output: 6-digit value, easy to copy
• Advancing the counter
– Time-based (e.g., 60 seconds) or every button press
• Allow for skew in the counter value
– RSA SecurID: 5-minute clock skew by default

76