CSE 484 (Winter 2008)

User Authentication

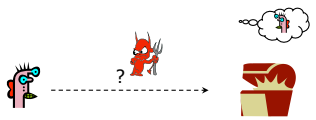
Tadayoshi Kohno

Thanks to Dan Boneh, Dieter Gollmann, John Manferdelli, John Mitchell, Vitaly Shmatikov, Bennet Yee, and many others for sample slides and materials ...

Goals for Today

- User Authentication
  - Conventional Passwords
  - Graphical Passwords
  - Biometrics
  - More

Basic Problem



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

Any system with access control must solve this problem

Many Ways to Prove Who You Are

- What you know
  - Passwords
  - Secret key
- Where you are
  - IP address
  - Physical location
- What you are
  - Biometrics
- What you have
  - Secure tokens
- All have advantages and disadvantages

Why Authenticate?

- To prevent an attacker from breaking into our account
  - Co-worker, family member, ...
- To prevent an attacker from breaking into any account on our system
  - Unix system
    - Break into single account, then exploit local vulnerability or mount a "stepping stones" attack
  - Calling cards
  - Building
- To prevent an attacker from breaking into any account on any system

Also Need

- Usability!
  - Remember password?
  - Have to bring physical object with us all the time?
- Denial of service
  - Stolen wallet
  - Try to authenticate as you until your account becomes locked
  - What about a military or other mission critical scenario
    - Lock all accounts - system unusable

Password-Based Authentication

- User has a secret password.
  System checks it to authenticate the user.
  - May be vulnerable to eavesdropping when password is communicated from user to system
- How is the password stored?
- How does the system check the password?
- How easy is it to remember the password?
- How easy is it to guess the password?
  - Easy-to-remember passwords tend to be easy to guess
  - Password file is difficult to keep secret

Common usage modes

Amazon = t0p53cr37

UWNetID = f0084r#1

Bank = a2z@m0$;

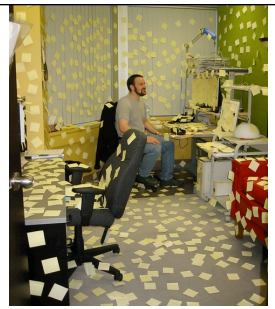


Image from http://www.interactivetools.com/staff/dave/damons_office/

Common usage modes

- Write down passwords
- Share passwords with others
- Use a single password across multiple sites
  - Amazon.com and Bank of America?
  - UW CSE machines and MySpace?
- Use easy to remember passwords
  - Favorite <something>?
  - Name + <number>?
- Other "authentication" questions
  - Mother's maiden name?

Some anecdotes [Dhamija and Perrig]

- Users taught how to make secure passwords, but chose not to do so

- Reasons:
  - Awkward or difficult
  - No accountability
  - Did not feel that it was important

Social Engineering

- "Hi, I'm the CEO's assistant. I need you to reset his password right away. He's stuck in an airport and can't log in! He lost the paper that he wrote the password on.
- "What do you mean you can't do it!? Do you really want me to tell him that you're preventing him from closing this major deal?
- "Great! That's really helpful. You have no idea how important this is. Please set the password to ABCDEFG. He'll reset it again himself right away.
- "Thanks!"

University of Sydney Study [Greening '96]

- 336 CS students emailed message asking them to supply their password
  - Pretext: in order to "validate" the password database after a suspected break-in
- 138 students returned their password
- 30 returned invalid password
- 200 changed their password
- (Not disjoint)

- Still, 138 is a lot!

Awkward

- How many times do you have to enter your password before it actually works?
  - Sometimes quite a few for me! (Unless I type extra slowly.)
- Interrupts normal activity
  - Do you lock your computer when you leave for 5 minutes?
  - Do you have to enter a password when your computer first boots? (Sometimes it's an option.)
- And memorability is an issue!

Memorability [Anderson]

- Hard to remember many PINs and passwords
- One bank had this idea
  - If pin is 2256, write your favorite 4-letter word in this grid
  - Then put random letters everywhere else

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

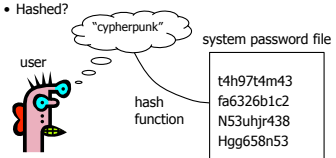
Memorability [Anderson]

- Problem!
- Normally 10000 choices for the PIN --- hard to guess on the first try
- Now, only a few dozen possible English words --- easy to guess on first try!

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

UNIX-Style Passwords

- How should we store passwords on a server?
  - In cleartext?
  - Encrypted?
  - Hashed?



hash function

system password file

t4h97t4m43
fa6326b1c2
N53uhjr438
Hgg658n53
…

Password Hashing

- Instead of user password, store H(password)
- When user enters password, compute its hash and compare with entry in password file
  - System does not store actual passwords!
  - System itself can't easily go from hash to password
    - Which would be possible if the passwords were encrypted
- Hash function H must have some properties
  - One-way: given H(password), hard to find password
    - No known algorithm better than trial and error
    - It should even be hard to find any pair p1,p2 s.t. H(p1)=H(p2)

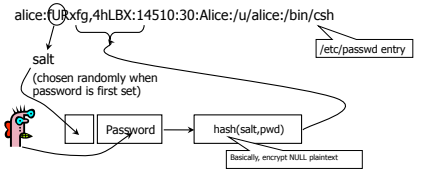
UNIX Password System

- Uses DES encryption as if it were a hash function
  - Encrypt NULL string using password as the key
    - Truncates passwords to 8 characters!
  - Artificial slowdown: run DES 25 times
    - Why 25 times? Slowdowns like these are important in practice!
  - ("Don't use DES like this at home.")
  - Can instruct modern UNIXes to use MD5 hash function
- Problem: passwords are not truly random
  - With 52 upper- and lower-case letters, 10 digits and 32 punctuation symbols, there are $94^8 \approx 6$ quadrillion possible 8-character passwords (around 2^{52})
  - Humans like to use dictionary words, human and pet names ≈ 1 million common passwords

Dictionary Attack

- 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 many passwords come from a small dictionary
  - Attacker can compute H(word) for every word in the dictionary and see if the result is in the password file
  - With 1,000,000-word dictionary and assuming 10 guesses per second, brute-force online attack takes 50,000 seconds (14 hours) on average
    - This is very conservative. Offline attack is much faster!
    - As described, could just create dictionary of word-->H(word) once!!

Salt



- Users with the same password have different entries in the password file
- Dictionary attack is still possible!

Advantages of Salting

- Without salt, attacker can pre-compute hashes of all dictionary words once for all password entries
  - Same hash function on all UNIX machines
  - Identical passwords hash to identical values; one table of hash values can be used for all password files
- With salt, attacker must compute hashes of all dictionary words once for each password entry
  - With 12-bit random salt, same password can hash to 2^{12} different hash values
  - Attacker must try all dictionary words for each salt value in the password file
- Pepper: Secret salt (not stored in password file)

Other Password Issues

- Keystroke loggers
  - Hardware
  - Software / Spyware
- Shoulder surfing
  - It's happened to me!
- Online vs offline attacks
  - Online: slower, easier to respond
- Multi-site authentication
  - Share passwords?