Lecture 8: Security

Forms:

- parts of websites that collect users' personal information
- used extensively in eCommerce websites
- some info you might be nervous about passing to a website:
  - Credit card number
  - Social Security number
  - Other personal information (phone number, address, etc)
- there is a "secure connection" icon in most browsers (e.g. a yellow locked padlock icon on the right side of the status bar in IE)
- you can't be entirely trustworthy of the icon as guaranteeing security
  - e.g. it matters *which server you connect to*, as indicated by the URL
  - URLs (as given by links in phishing emails) can be deceiving
    - e.g. www.bankofthevvest.com instead of www.bankofthewest.com
- Internet banking websites are the most dangerous type of target since, if accessed with a stolen password, they allow immediate transfer of funds (as opposed to, say, credit cards where you are often not liable for charges placed by someone else in case of theft)

Cookies:

- text files installed on your computer by some websites you visit
- a website can put anything in your cookie
  - usually the timestamp indicating when you visited
  - usually a unique number establishing your identity
- by design, your browser sends the cookie back to the website the next time you visit it
  - by looking at the unique number in the cookie, the website you're visiting "remembers" your identity, comparing it perhaps to some data about your preferences in its own database
- cookies are a useful thing:
  - allow websites to remember your preferences
    - e.g. remember your user name (but not password) for Internet banking sites
    - remember your username *and* password (most webmail)
    - your shopping preferences (e.g. Amazon.com)
    - your viewing preferences (e.g. Weather.com)
- but cookies can be dangerous:
  - many websites share cookies with third-party advertising websites, which allows them to get at least *some* information about you (e.g. your buying preferences)
- privacy solutions:
  - in Internet Explorer: Tools → Internet Options → Privacy tab → "Advanced" button: select the checkbox "Override automatic cookie handling" and choose desired settings

- deleting cookies through a browser (e.g. in IE, Tools: Tools → Internet Options → General tab, click on "delete cookies" and/or "clear history")
- eliminating cookies manually: e.g., in Windows, all files in C:\Documents and Settings*username*\Cookies

Logs:

- most web servers keep logs of usage:
  - who visited (by IP address or host name)
  - when
  - what browser was used
- example: logs for E-1
  - analysis of daily traffic in the month of October shows a dramatic spike of page visits just before Exam 1 (14,000 page requests)
  - a "request" here is defined as *any* piece of a webpage, such as an image (most pages consist of several such pieces)
  - Domain report: a pie chart with percentages of TLDs
    - ~50% from .net, 20% from .com, and half a percent each from .uk (England), .de (Germany), .jp (Japan), etc.
  - Search Query report:
    - "roman rubinstein" is the top search query with 26 requests last month
    - "ppt hard disk concepts" with 24
    - David's phone number (searching from Problem Set 1)
    - Etc

Data Recovery:

- we already mentioned it in our discussion of hard drives: it is possible to extract even seemingly "erased" data from an old hard drive
- data is typically left even after formatting a hard drive
- formatting a hard drive takes a lot of time b/c formatting utilities try to verify the hard drive, *not* because they erase all information
- properly "sanitized" hard drive: all data is changed to 0s or 1s
- there are also secure-erase programs that wipe specific files (but many of them, especially the free ones, are buggy and do not do the job adequately)
- *Handout*: an MIT research study that looked at ~300 used hard drives
- Demonstration: from David's work for the DA's office:
  - "write-blocker": a piece of hardware for data recovery companies and forensic analysis institutions – connects to a hard drive (through an IDE port) and allows to read data *from* the hard drive, but does not allow to write *to* the hard drive.
    - Plugs into a USB port on the other end
  - "Encase": a program that allows you to browse a user's hard drive while also performing different searches on it
    - the *de-facto* standard in computer forensic analysis today

        - shows the preserved directory structure (this one's from a Windows machine)
        - "Gallery" mode: shows all images in icon format
        - may be scripted to collect all URLs visited
        - also shows files that were deleted – even if they're not there, there's evidence that they were there
        - may un-erase (if partly) many files

Passwords:

- difficult to remember many secure passwords
    - "notepad" technique: writing down your passwords (defeats the purpose)
    - new ways of authentication are appearing, e.g. biometric authentication, but they are slow to emerge and adopt
- some malicious programs, called *keyloggers*, if installed on your computer as part of a malicious software kit, may record everything you type, including passwords

Packet Sniffing:

- remember that hubs connect computers in such a way that any computer on the network may "listen" to traffic from other machines
    - This is called "promiscuous network mode"
    - Switches alleviate some danger, but if a message goes out to the Internet, it is bound to pass through various networks
- only an *end-to-end* connection, usually achieved by establishing an encrypted channel over the Internet, is secure
- TCP and IP packets conform to specific strict formats and are easy to decode anywhere along the path of the packets.

Hacking:

- *hacker* vs *cracker*: historically, "hackers" were the "good guys" and "crackers" were the "bad guys"
- these days, "hacker" is often used to denote a malicious user
- *"trashing"* or *"dumpster diving"* (physically looking for information in trash) is the most mundane way to look for sensitive info
- *phishing* (discussed in an earlier lecture)

Viruses and Worms:

- range widely in the amount of damage they cause
- the goal of a *worm* is to propagate itself automatically, entering computers through vulnerabilities in the host operating system
- a *virus*, in contrast, requires human interaction to be unleashed (by opening, e.g., an email attachment)

- pop-ups may bring you to websites that invite you to download "anti-virus" software that itself is a virus
    - Alt+F4 is a way to close a window in IE (and most other Windows programs)
- *hoaxes*: chain letters warning of viruses that do not actually exist
    - often ask you to delete a system file
- avoiding viruses and worms:
    - "smart computing": not opening suspicious attachments, even if they are from your friends and family (may not protect you from worms)
    - running anti-virus software
    - spyware removal tools
- authors of biggest viruses are often caught
- *zero-day attack*: e.g. Microsoft announces a fixpack, but not all users and IT professionals update their systems immediately, allowing hackers to expose the vulnerability
- *firewall*: a device or program that filters network traffic (incoming and outgoing)
    - Protects against network-based attacks (usually worms)
- Demonstration: SpyBot and HijackThis
    - HijackThis allows you to scan your Windows Registry for programs that load at start up, and eliminate entries of malicious programs (spyware)
    - SpyBot searches the hard drive for spyware
    - "hijacking" Internet connection: redirecting your browser queries to a different website than what you asked for