# Lecture Notes: Internet Security and Privacy

These notes accompany this week's Internet Security and Privacy Techniques handout.

## Why is Security Needed?

The Internet was originally designed for sharing information between collaborating scientists. And at the time (early 70s) personal computers didn't yet exist. Nor, for that matter, were there widespread data network lines. So, security didn't much enter into the thinking of the Internet's founders; they were far more concerned with robust message delivery in the face of an anticipated World War III.

In fact, the very open nature of the Internet protocols likely contributed to its enormous success through the 1990s and today.

But when so many people can connect so easily, there are bound to be a few bad apples trying to ruin the party. In earlier days, a tremendous amount of technical know-how was required to crack into computers through the network. Gradually the software used by these `l33t hax0rs` ("elite hackers" in the lingo) made it onto the Internet and could be wielded by any old pimply-faced high-school punk, the aptly-named "script kiddies". (The self-titled MafiaBoy who cracked CNN, eBay, Yahoo!, and others in 2000, was a 15-year old Montreal brat who fit nicely into this category.)

These software tools are now run automatically by programs that search the Internet for weak spots. It is said that an unprotected computer will be targeted within 15 minutes of its initial network connection. Some security companies set up such computers (called "honeypots") from time to time to attract just this sort of malicious attention in order to study the means and methods of the attackers.

Writing viruses and breaking into computer systems was once an ego-driven activity. Now it's a business. Taking over personal computers (called zombies) can be profitable. Control of these zombies is traded and sold in blocks of thousands. The victimized computers are then used to send out spam e-mail messages, or attack a certain computer in coordination with one another. The latest nefarious idea is to break into a computer, encrypt all of its files, and then demand a ransom payment from the computer owner: no money, no files.

Even with all of this technical acumen so widely available, the number one way to gain illegal entry to a compute system is through "social engineering". That is, calling up the target organization and weaseling your way into their confidence, perhaps enough to get a naive desk jockey to cough up a password. Walking purposefully through a building with a clipboard is an awfully good way to get into restricted areas.

Disgruntled employees are also a great source of trouble to security-conscious organizations. But these unhappy folks are unlikely to cause the Internet population as a whole much harm.

## Computer Viruses

Computer viruses existed long before the popularization of the Internet, but they can now proliferate far more rapidly in the connected world.

There are actually three types of infectious programs: viruses, worms, and trojan horses.

A computer virus is a remarkable analogy of its biological counterpart. Like a bio-virus, computer viruses are not "alive", in the sense that a computer virus must first infect a normal program before its code can be run. A bio-virus must infect a cell before it can hijack the DNA replication in the cell to create copies of itself. Once a computer virus has infected a program, the next time the program (possibly even part of the operating system) runs, the virus's payload will be activated. Typically a virus will try to copy itself to other programs, or across the network, but most also contain some form of harmful instructions: delete files, format drives, send personal files across the Internet, and so forth.

A computer worm, on the other hand, is usually a fully-functional program that depends exclusively on a network for its reproduction. The MS Blaster worm that attacked, and continues to attack, Microsoft Windows installations is an example of such a worm. Once a computer is infected, the worm will create copies of itself with instructions for each to seek out another nearby Windows computer and infect it also. In this way, the MS Blaster worm (and variants) circled the globe in less than 24 hours. (The "Slammer" worm is estimated to have reached every Internet-connected computer in the world within 10 minutes.)

Finally, a trojan horse is a program that looks normal (even inviting, like the original Trojan Horse), but that carries a terrible payload. Some trojan horses may masquerade as cute little dancing frogs on the screen, or whatever. Others may set up a screen that looks exactly like the normal Windows login screen and as soon as you type in your user name and password, the trojan horse sends that information off to someone on the Internet, and then pops up the regular login screen (just as if you'd accidentally typed an incorrect password).

## Firewalls

A "network firewall" is like a filter for Internet messages: it lets some pass through, but blocks others.

A firewall selectively allows certain messages to pass through based on its type (e-mail, web, file sharing). All other messages are stopped at the gates. In this manner, you can protect yourself to a certain degree by simply not allowing Internet traffic to contaminate your machine unless it arrives through a recognized port (e.g. the one used for the web, or for e-mail, or for file transfers).

Obviously, a firewall can't protect you from malicious messages that arrive on ports that you need to remain open, so a firewall is never a complete security solution.

However, firewalls are pretty good at fending off network worms.

Internet security companies, journalists, and even Microsoft, are now saying that every computer should have not only an anti-virus scanner, but an installed and maintained firewall as well. Some popular firewall products are listed on this week's handout.

## Identity Theft

Protecting your online privacy is the best step you can take in preventing the new but popular crime of Identity Theft: stealing someone's "identity" in order to fraudulently carry out monetary transactions in the name of the stolen identity.

The more pieces of information such a thief can piece together, the easier it is for him or her to request credit cards, a SIN number, driver's licenses, and so forth, all in your name. The thief then usually max-es out all of your lines of credit and skedaddles, leaving you responsible for the charges, and facing a very difficult process to re-establish your credit rating.

Of the 10,000 or so arrests made by the US Secret Service in 2002 (the agency responsible for such financial crimes), 94% of them were related to identity theft.

If you suspect you've been the target of identity theft, follow the list of instructions found on the Privacy Commissioner of Canada's website [www.privcom.gc.ca/fs-fi/02_05_d_10_e.asp].

## Spam

Spam is the e-mail equivalent of the junk mail (pamphlets, menus, shopping flyers) that fill your non-Internet mailbox every day. Spam is simply any piece of unsolicited e-mail: it may be sent to just you, or it may be sent to a million others--either way it's spam if you didn't ask for it.

The term "spam" originated, as do many geeky concepts, from a Monty Python skit [www.detritus.org/spam/skit.html].

Lists of 100,000s or even millions of e-mail addresses are available for purchase which can then be used to send out that many spam messages in a few minutes. Even if only a minute fraction of people respond, spam is still the most cost-efficient marketing method in existence.

## Authenticating Sites and Using Encryption

There are two main problems with the out-of-the-box Internet: identifying people or computers at the other end of the network, and guarding against eavesdroppers. Authentication solves the first problem, and encryption solves the second.

Authentication allows us to identify people or organizations, and their computers. Normally, when meeting people face-to-face, we can identify them by their appearance, voice, or some form of recognized document, such as a passport or driver's licence. Similarly, we identify organizations by address, or perhaps business licence.

Across the Internet, authentication is much more difficult, and so mechanisms such as user name and password were invented to allow computer systems to recognize (authenticate) valid users. Passwords have a number of weaknesses, so the latest authentication rage is biometric authentication: fingerprint scans, retinal scans, face recognition.

All such authentication mechanisms have a common weakness: before any passwords or biometric information is traded, how can you know if the person you are talking to *really is* the person you think you're talking to?

Web shoppers have another authentication problem: how can you tell for certain that a company's e-commerce website truly is run by that company, and not some con artist who set up a fraudulent site to appear exactly like the legitimate company's website?

The solution to both these problems is something called a "digital certificate". A certificate is issued by a recognized Trust Authority (Verisign being the most well-known TA) that promises to only issue such certificates once the recipient--individual or organization--has credibly proved their identity to the authority. The actual process for applying for these certificates requires faxing all sort of incorporation documents (if one wants a corporate certificate), the names of the organization's president and directors, and a final voice confirmation from the trust authority to the president of the organization to ensure everything is on the level. The application process for individuals is a little less rigorous.

Once these certificates have been granted, the person or organization possessing the certificate can be reliably authenticated (presuming one trusts the Trust Authority). The technical makeup of these certificates (text files filled with a long string of mumbo-jumbo characters--really just a very, very long number) more or less guarantees that they can't be forged, at least not without an overwhelming investment in computer processing power and time, or a revolution in mathematical number theory.

Even once the two parties in an Internet conversation have been authenticated, and the operations that they may perform have been authorized, the messages that are sent back and forth are still readable by anyone able to listen in (at any of the points along the route between the two parties, or even on the local network of one of the parties if a simple broadcast protocol like Ethernet is used). These messages might contain private information such as: passwords, credit card numbers, or sensitive records about a person (e.g. legal, tax, medical).

The Internet was originally designed to allow the free exchange of information. No one at the time realized the potential of the Internet, and its consequent commercialization. Most of the protocols in use hearken back to those early days, and so do nothing to hide the contents of their messages.

In response to this problem, a series of encryption technologies was developed to scramble these messages, and then unscramble them once they reached their proper destination. The most common of these technologies is called SSL (Secure Sockets Layer, although the technology as a whole is now known as Transport Layer Security--TLS). SSL is usually added to an existing protocol to better secure its messages: HTTPS encrypts HTTP web traffic, SSH encrypts both telnet and FTP traffic, and PGP (Pretty Good Privacy) is one way to encrypt e-mail messages.

SSL is a certificate-based system, meaning that using SSL implies both authentication and encryption.