Chapter 10: Security

Cryptography and Network Security

Computer networks can provide great benefits, quickly sending information almost instantaneously across continents and around the world. However, in order for them to reach their full potential, we need to ensure that information sent on the network is secure. For bank transactions, electronic-commerce, and even some types of e-mail we must be certain that the information sent remains confidential. In fact, there are several important issues which must be addressed when sending or receiving information over a computer network. These include:

Confidentiality—We want to ensure that the information we are sending is secure and confidential; that it cannot be read by anyone else on the network.

Authentication—We need to verify that the party we are talking to is indeed who they say they are. If I'm communicating with a website that claims it's my bank, maintaining secure communications with the website does me no good, if the website is, in fact, not actually my bank's but a scam operation.

Integrity—We need to know that the message we are receiving has not been tampered with and that we have received the entire original message.

Non-Repudiation—In some cases, if someone sends us a message, we need to be able to prove that they did indeed send us the message. We do not want them to be able repudiate the message—in other words, we don't want them to be able to deny having sent the message.

In this section we study how computers handle issues of network security and authentication. Cryptography is central to providing secure communications, so we begin our discussion with a brief overview of how cryptography works. We then discuss how secure transactions work on the World-Wide Web. We conclude our discussion with some suggestions that you can take to ensure your own security when working on the Internet.

Cryptography

Cryptography—the science of encrypting and decrypting messages—is the basis for secure network communication.

Basic Cryptography

Let's take a look at a basic cryptography example. Suppose Alice wants to send a message to Bob, but she wants to make sure if it is intercepted by Mallory, that Mallory won't be able to read the message. Alice and Bob need to choose a *cipher*. A cipher is a method for transforming the original unencrypted message or *plaintext* into *ciphertext*, which only Alice and Bob can read. Using the cipher, Alice *encrypts* the message. When Bob receives the ciphertext, he uses the cipher to *decrypt* the message, converting it back into plaintext.

To make the example a bit more concrete, let's try actually encrypting a message using a simple cipher. The cipher we will use is called the *Caesar Cipher* or Caesar Shift. It was purportedly used by Julius Caesar to send secure messages to his generals. The Caesar Cipher takes the original message and replaces each letter in the message with the letter which comes three places later in the alphabet—for example, the letter 'A' will be replaced with the letter 'D', the letter 'B' will be replaced by the letter 'E'. Letters at the end of the alphabet wrap back to the beginning, so the letter 'X' becomes 'A', 'Y' becomes ‘B’, and ‘Z’ becomes ‘C’. We can summarize the relationship between the letters into a simple table like this:

Original Letter	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Encrypted Letter	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

If Alice wants to send the plaintext message:

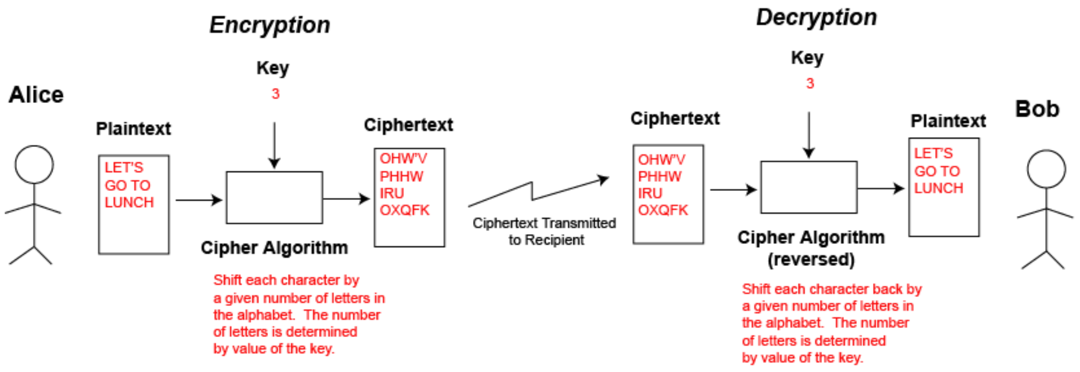
LET’S MEET FOR LUNCH

using the Caesar Cipher this becomes:

OHW’V PHHW IRU OXQFK

When Bob receives the ciphertext, he reverses the Caesar Cipher—shifting back each of the letters the places in the alphabet—revealing the original message.

Our Caesar Cipher example illustrates some important points about cryptography. Our encryption and decryption process actually combines an *algorithm* (that is a step-by-step set of procedures) and a *key*. In this case, the algorithm says that we should shift all the letters in the message a given amount. The key is the number of spaces we need to shift the letters. In our example we are using a key of 3, as we are shifting all the original letters down three positions in the alphabet. We can use the same algorithm with a different key to produce different ciphertext—for example we might choose to shift all the letters by five positions, instead of three.[1] Taking both the cipher algorithm and key into account, we can draw a diagram showing communication between Alice and Bob as shown in figure XXX.


Encryption
Key
3
Alice
Plaintext
LET'S
GO TO
LUNCH
Ciphertext
OHW'V
PHHW
IRU
OXQFK
Cipher Algorithm
Shift each character by a given number of letters in the alphabet. The number of letters is determined by value of the key.
Ciphertext Transmitted to Recipient
Decryption
Key
3
Ciphertext
OHW'V
PHHW
IRU
OXQFK
Plaintext
LET'S
GO TO
LUNCH
Bob
Cipher Algorithm (reversed)
Shift each character back by a given number of letters in the alphabet. The number of letters is determined by value of the key.

A Cipher Algorithm takes Plaintext and a Key and tranforms it to Ciphertext. When the Ciphertext is received the Cipher Algorithm is reversed--the Ciphertext and Key are combined to retrieve the original Plaintext. Note that for this to work both the sender (Alice) and the recipient (Bob) must know the value of the key.

Now suppose Mallory intercepts Alice’s encrypted message he may perform *cryptanalysis* in an attempt to break Alice’s code and determine the original contents of Alice’s message. He has a variety of methods at his disposal. The Caesar shift, for example, is highly susceptible to frequency analysis, in which the frequency of letters in the ciphertext is compared to the frequency of letter usage in the English language—Mallory knows that the letter ‘E’ is the most common letter used in English and the letter ‘H’

[1] Historically the Caesar shift combined both the alphabetic shift cipher algorithm with a specific set key value of 3. Modern cryptologists separate the shifting algorithm from the key value and consider them as separate entities. Thus we could combine the alphabetic shifting algorithm with any of 25 different values.

shows up repeatedly in Alice's ciphertext, therefore 'H' is likely to represent the plaintext character 'E'. As an alternative, if Mallory knows the algorithm used to encrypt the message, but doesn't know the key, he can try a brute force attack. In this case he simply applies the algorithm on all the possible keys—first trying shift by one, then two, then three, until finally he gets lucky and decrypts the message. Since there are only 25 possible keys for our shifting algorithm, if Mallory knows the algorithm used to encrypt the message, using a brute force attack he will be able to decrypt the message.

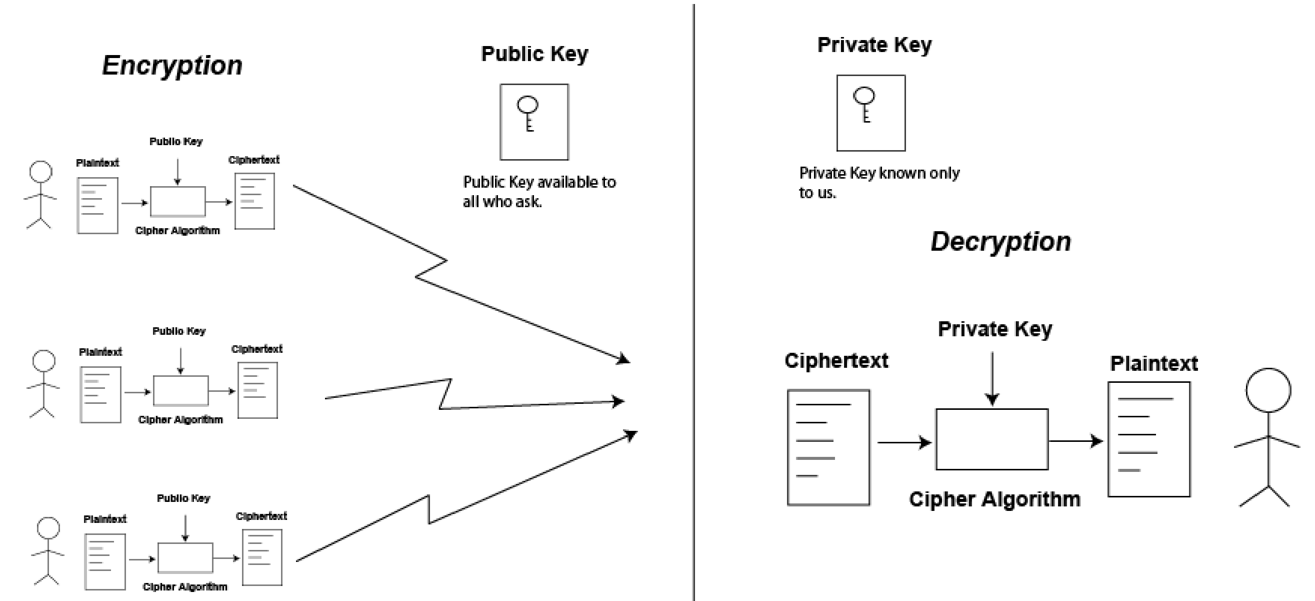
At first glance, the security of Alice's message appears to depend on both the secrecy of the cipher algorithm used and on the key itself. In fact, for such a simple cipher, the message does indeed depend on the security of the cipher, as our shift cipher can easily be broken by simply trying all the keys. However, modern ciphers do not depend on the secrecy of the cipher used. In fact, the algorithm for encrypting secure messages on the World Wide Web is widely published. Modern ciphers depend on the security of the keys alone. The idea behind publishing the cipher algorithm is that we want as many people as possible to study the cipher and determine whether or not there are any weaknesses in the cipher which a cryptanalysis can exploit. If no weaknesses are found, then the only approach for breaking an encrypted message is to try the brute force method, testing all possible keys. Because modern cryptography depends on key security, we can rate how secure an application is by asking how large the key is. This determines how many different key combinations are possible and how many key combinations an attacker would have to try in order to decrypt the message. Two levels of security are currently used on the web. An older version uses 40-bit keys. This provides for 2^{40} or 1,099,511,627,776 different key combinations. While that seems like a fair number of keys, in fact modern computers can now break messages encrypted with 40-bit keys, so 40-bit key use is no longer recommended. Instead the newer standard provides for 128-bit keys. This gives us 2^{128} or 3.4×10^{38}—that's a 3.4 with 38 zeroes after it before the decimal point, which is a very large number. 128-bit keys should remain secure for quite some time.

Symmetric and Asymmetric (Public-Key) Encryption

The encryption method we've described so far is called *symmetric encryption*. In symmetric encryption, both the encrypter and decrypter of our secret message use the same key. While symmetric encryption works for many purposes, it does require that somehow the sender and receiver of our message both have access to the same key. In some cases this can be impractical. For example, suppose before you made a secure purchase at `amazon.com` you and Amazon had to agree on a 128-bit number to use as a key for secure communications. How would you communicate to agree on the key? One party can't simply select a key and then send it to the other party via e-mail because standard e-mail is not secure. The key could be sent via surface mail (also know as *snail mail*), however, that would mean that a purchase could only be made after waiting for a physical mail letter to arrive.

An alternative to symmetric encryption was first discovered in the 1970s. This alternative is called *asymmetric encryption*. In asymmetric encryption, the person encrypting the message and the person decrypting the message use different keys. The keys are mathematically related so that the encrypter can use one key to encrypt the message, and the person decrypting the message can use the matching key to decrypt the message. Although the two keys are related, the decrypting key cannot be easily derived from the encrypting key. Therefore if a criminal or other malefactor was to get a hold of the key used to encrypt the message, he still would not be able to decrypt the message.

Asymmetric encryption is sometimes referred to *public-key encryption*. Using asymmetric encryption, if I want to allow others to send me secure messages, I can give out copies of the key used for encryption. Because this key can only be used to encrypt messages, not decrypt them, I can make it publicly available for anyone to see. They can then take my publicly available key and encrypt messages and send them to me. The decryption key is my *private key*. I am the only one who knows it. I take messages encrypted with my public key and I decrypt them with my private key.



Outsiders use Public Key to encrypt messages.

Although everyone has a copy of the Public Key, encrypted messages are safe because the Private Key is needed to decrypt Ciphertext.

As an alternative, we can make the decryption key public and the encryption key private. This allows us to send messages which must be from us, because we are the only ones who could have encrypted them. However, because the decryption key is public, everyone can read the messages. As we'll see this is used for authentication. It can also be used for non-repudiation.

Web Security

Sending secure information on the web involves both the use of symmetric and asymmetric cryptography. Let's take a look at how web security works.

Suppose you want to purchase something from an online merchant. Before making a purchase, we need to ensure that (1) we really are communicating with the merchant's website, not some hacker's website which is masquerading as the merchant, and (2) that any financial information we send to the merchant is secure.

In order to provide authentication, the web assumes the existence of one or more trusted *certification authorities*. These authorities guarantee the identities of merchants and other organizations or individuals on the web. Each certification authority has a public-private key pair which can be used for asymmetric encryption. Our web browser knows about these certification authorities in advance and knows their public keys (or can obtain them in a secure manner).

When a merchant decides to build an e-commerce website, she generates a public-private key pair. She then contacts a certification authority and verifies her identify to the satisfaction of the authority. The authority takes her information and her public key and generates a certificate combining this information. The certification authority then uses its private key and combines it with the information in the certificate to come up with a unique number. This number is used to sign the certificate. When someone receives the signed certificate they can use the certification authority's public key on the signature. If the information found in the certificate matches the information we find by decrypting the signature then we know the certificate is valid—because only the certification authority has access to their private key, only they would have been able to encrypt the information.

Now, when we go to a merchant's website and try to complete a purchase, they send our web browser a copy of their certificate. Our web browser checks the certificate to determine the certification authority and then checks to verify that the certificate was properly signed by the certification authority. If it is properly signed we have verified that we are at the merchant's legitimate website.

Before sending our credit card information, we need to ensure that our communications are secure. Communication with the merchant's website could continue on using asymmetric public key encryption. However, asymmetric encryption is much slower than symmetric encryption. Because of this, once we've authenticated the identity of the webserver we are visiting, we want to switch to symmetric encryption. However, we need a secret key known by both my personal computer and the merchant's web server. In order to generate a secret key, my web browser chooses a random number which will be used to generate our secret key. I can't send this number directly through the Internet, because it might be intercepted. However, I can encrypt the number using the merchant's public key. The merchant's computer receives the encrypted number and decrypts the number using the merchant's private key. We now have a shared secret key and communication proceeds using symmetric encryption based on our shared key.

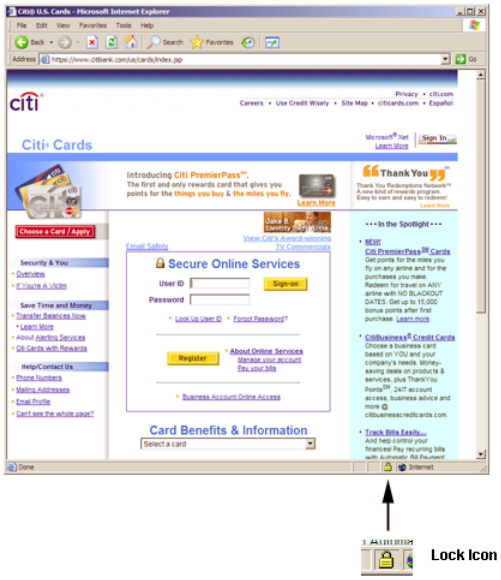
Personal Security

Web security is largely transparent to the end-user. Exchange of certificates and encryption and decryption of messages takes place without any intervention from the user. However, there are some things that you as a user can do to stay safe when working on the Internet. In this section we discuss some of the most important.[2]

Web Browser and E-Mail Issues

Websites and e-mail are now used as tools by con artists. Here are some issues you should think about.

- Pay attention to your web browser. The status bar will display a lock icon as shown below if the website is using encryption.



The first and only rewards card that gives you points for the things you buy & the miles you fly.

Thank You Redemptions Network℠ A new kind of rewards program. Easy to earn and easy to redeem!

Get points for the miles you fly on any airline and for the purchases you make. Redeem for travel on ANY airline with NO BLACKOUT DATES. Get up to 15,000 bonus points after first purchase. Learn more.

Choose a business card based on YOU and your company's needs. Money-saving deals on products & services, plus ThankYou Points℠, 24/7 account access, business advice and more @ citibusinesscreditcards.com.

And help control your finances! Pay recurring bills with Automatic Bill Payment

Also check the URL listed at the top of your web browser, particularly if following a link from a questionable amateur website. Just because a website looks exactly like the amazon.com

[2] The textbook contains a number of other sections on security—see also the section on SPAM, Adware, and Spyware in Chapter 8, the section on Online Fraud in Chapter 12, and the discussion on Worms, Viruses, and Trojan Horses in Chapter 15

website doesn't mean that it is. If the URL listed in the address bar says `http://www.amazon.com/` then you should be okay (however, see footnote below).[3] If the URL says something different or worse has just an IP number listed, then don't enter in your account information—for example, the URL shown below is very likely a fraud website:

```
http://128.12.34.165/exec/obidos/subst/home/home.html
```

Also check and make sure that the URL does not include an @ sign in it. A link to

```
http://www.amazon.com@scammer.com/enterinfo.html
```

is actually a link to scammer.com not to amazon.com.

On a related note, be particularly careful following links on e-mail. There is no authentication on the source of an e-mail message. In fact a criminal can very easily fake the source of an e-mail message.[4] Therefore, just because an e-mail message appears to come from your bank or your credit card company, that doesn't mean it actually does. Even if the source of the e-mail message looks exactly correct, the message may be fabricated. Criminals can easily create an e-mail messages that looks official and which contains hypertext links to a fake website which looks exactly the same as the official website. The e-mail message will often look official right down to the logo and graphics.

To avoid this problem, I would recommend that, rather than following a link on an e-mail message, you either call the company via telephone or enter in the company's standard URL manually into the web browser. For example, if I were to unexpectedly receive an e-mail message purportedly from Bank of America containing a hypertext link to a webpage asking for account information or passwords, I would either (1) handle the transaction over the phone by calling Bank of America using a number from the back of my credit card, found from a the telephone book, or from their official website (but definitely not from the e-mail message), or (2) I would type in the standard `http://www.bankofamerica.com/` URL into my web browser manually, and then login to the website following standard procedures. What I would specifically not do is to click on any link within the e-mail message. The chances of receiving a fake e-mail message with a link which takes me to a scam website is simply too high to ignore.

- As you're probably aware, e-mail is also subject to standard scams not directly related to technology. If you receive an e-mail message telling you about an amazing get rich quick scheme, you can bet it's a scam.

Computer Security

While our discussion has focused on web security, you should also pay attention to the security of your own personal computer. Hackers attack personal computers on the Internet for a variety of reasons. Some hackers break into personal computer and install programs on them. The programs can turn the computer into a "zombie" which can be ordered to attack commercial webservers or other target computers. By attacking a commercial webserver using zombie computers, the hacker multiplies his power (because he can use many zombie computers at once) and obscures his tracks. Programs installed by hackers are also sometimes used to send SPAM (unsolicited advertisement e-mail). Hackers may also break into personal computers to gain access to personal or financial information.

[3] As of this writing, Internet Explorer contains a security hole which allows a website to create a frameless window over the address location listing the URL. This frameless window would then be used to falsely list a URL. This means that checking the URL is not a 100% foolproof method of determining your location. Hopefully this security hole will be closed by the time your read this.

[4] As we've seen we have the technical know how needed to authenticate messages on the computer. In theory we could overhaul the e-mail system used throughout the Internet and require all e-mail messages to authenticate their sender. To date, there simply hasn't been sufficient impetus to force this to occur.

You can take a number of steps to reduce the chances of a hacker breaking into your computer.

- Password protect your computer and use a strong password. Passwords should not be based on real words. Hackers have access to programs which can perform a "Dictionary Attack". These programs will try all the words in the dictionary as possible passwords for your computer account. Passwords should contain both letters and numbers, and should use both upper- and lower-case letters. Remember, in a brute force attack, the attacker must try all possible key combinations. If, as with many people, you simply use lower-case letters in your password, that limits the number of combinations possible. If you use upper- and lower-case letters and numeric digits, that increases the number of key combinations, making your password harder to break. You may also consider adding punctuation characters into your password, if your computer supports them. Make sure your password is at least eight characters long—the longer your password is, the harder it will be to break.

  You should follow the same password guidelines for websites which you have accounts on. Also be careful using the same password for multiple websites. If you use the same password and username for your financial websites as you do for online newspapers, online bulletin boards, and hobby websites and one of the online bulletin boards is actually setup by a hacker to gain usernames and passwords, your financial accounts might be compromised.

- Consider using an Internet Firewall. Businesses actually use hardware firewalls. A hardware firewall is a special purpose computer which regulates all Internet traffic entering or leaving the company. When communicating with computers outside the company, all computers inside the company must send their messages through the firewall computer. All messages received by the company from outside and all messages sent from inside the company to outside must pass through the firewall computer before going to their final destination. The firewall computer analyses the messages sent to and from the company. If it sees a message that it does not believe is legitimate it blocks the message. The rules that a firewall computer uses to distinguish between legitimate and illegitimate traffic are configurable and can be set by the company's network administrator. A firewall might be setup to only allow certain kinds of traffic—for example allowing e-mail, but not web traffic. A firewall might also allow communication with specific computers on the Internet, or might deny access to specific computers.

  Relatively inexpensive hardware firewalls can be purchased to guard your computer or home network. In addition, personal firewall software programs can be installed on a computer providing some of the functionality of a hardware firewall. These programs can block access to your computer from outside intruders and in some cases can prevent programs on your computer from sending information without your approval. The latest versions of both Microsoft Windows and Apple's MacOS include personal firewalls. However, these security features may not initially be turned on when you purchase a computer. A firewall of some kind (either a hardware or software firewall) is particularly important if your computer is always connected to the Internet (as opposed to if you have to explicitly turn on a modem and call to connect to a service provider every time you use the Internet). Computers which are always connected are much more likely to be attacked by a hacker.

- Finally, make sure you keep your operating system, web browser, and e-mail program updated with the latest security patches available. Unfortunately many programs contain security breaches which may be exploited by hackers to attack your computer. When a security breach in a program is discovered a patch for the program is written. If you don't keep your computer programs updated with the latest patches your computer will be particularly vulnerable for attack.

Spam, Adware, and Spyware

Much of the power of the Internet comes from its global reach. Unfortunately while allowing our computer to contact anyone in the country or in the world provides benefits, it also carries some costs. Not only can we contact anyone anywhere on the Internet, they can also contact us. This opens our computer to a variety of potential problems. These problems include direct threats such as hacker attacks and computer viruses. They can also include marketing techniques gone amuck as well as outright scams. We'll explore each of these topics in turn in different sections of this textbook. In this section we take a look at some questionable approaches to marketing performed on the Internet—Spam, Adware, and Spyware.

Spam

Spam is the name used for unsolicited bulk e-mail.[5] This includes unsolicited advertisements (e.g., cheap Viagra, amazing mortgage rates) as well as outright scams (e.g., stock tips designed to pump up stocks). In some sense, Spam can be thought of as simply the equivalent to junk mail received from the US Postal Service. However, the nature of Internet e-mail is such that SPAM creates much more of a problem than traditional junk mail from the post office.

There is almost no cost to sending out bulk e-mail. In contrast, sending a shopping flyer via the US Postal Service may not cost much, but it costs enough to restrict sending the flyer to those who may actually be interested in it. Spam senders (or Spammers) make no attempt to target their e-mail messages, because there is no economic reason to do so. Mortgage refinance advertisements are just as likely to be received by those without a mortgage as those with one. Viagra e-mails are sent to young and old, male and female alike. Pornography spam is sent to children as well as adults. English speakers may receive spam from foreign countries written in languages they can't even read. Because the cost per e-mail is virtually zero, spam senders will try to send to as many e-mail addresses as they can, regardless of whether the message is appropriate to the audience or even if the audience can read the message at all.

Costs of Spam

While spam has very few costs to the sender, it has some very real costs to everyone else on the Internet. Spam costs us all in time and effort as we sort through our e-mail messages. While various spam solutions are available, each has potential costs such as filtering out legitimate e-mail along with spam. Spam reduces the overall effectiveness of Internet e-mail.

Moreover, spam takes up an increasingly large amount of e-mail traffic. Some estimates show spam at over 50% of all e-mail traffic with the percentage showing every sign of continuing to rise. Thus spam forces us to have more e-mail servers than we ought to and it increases the amount of Internet traffic, reducing overall network efficiency.

Getting Spam

So why do you get spam? How does the spammer get your e-mail address? Spammers harvest e-mails from a variety of sources. Some spammers use automatic search programs to search the web for webpages with e-mail addresses on them. When they find an e-mail address on a webpage they add it to their list. Spammers may target Internet news groups or bulletin boards, looking for e-mail addresses. Some unscrupulous websites may ask you to create an account and then sell your e-mail address to spammers.

Spam Solutions

A variety of methods have been suggested to solve the spam problem. These include legal approaches, technical solutions, and economic methods.

[5] Spam is purportedly named after a Monty Python skit which involved the Hormel SPAM meat product.

Legal Approaches

One step in reducing the amount of e-mail is to write legislation in an attempt to eliminate it. While this approach may have merit, it also has some major limitations. Spam legislation is only effective if the actual sender of the spam is known. Because of the way the current Internet e-mail system is currently run, spammers can easily obscure the origins of their e-mail message. While spam legislation may reduce spam advertisements, it is unlikely to eliminate spam e-mail from con artists—the contents of con artist e-mail is in fact already illegal yet such e-mail is sent nevertheless.

Some spam experts believe that some legislation labeled as “anti-spam” may actually increase the amount of spam received. Anti-spam legislation varies on whether spam is legal unless a user specifically requests to be taken off an e-mail list (opt-out) or whether no spam may be sent unless the user specifically requests it (opt-in). Opt-out legislation may legitimize the spam business, by providing clear legal backing for sending spam to users who have not explicitly opted out. This may cause direct marketing groups, currently not involved in spam, to enter the spam business.

Technical Solutions

A variety of technical means can be used to reduce spam. A filter can be added to a mail server or to the e-mail program itself. This filter can scan e-mail before it is presented to the user. If the e-mail contains specific words, say “Viagra” or “Mortgage” the program assumes that it is spam and deletes it. This approach has several shortcomings. First, as with many spam solutions it may filter out legitimate e-mail. A pharmacist running a spam filter which deleted all e-mails with the word “Viagra” in it, might have a problem. Similarly, while we might not want to read yet another e-mail promising incredibly low mortgage rates, we might want to see an e-mail from our real estate agent discussing the status of the mortgage on our new house.

Spam filters can get quite sophisticated. Some rate e-mail messages on the probability that they are spam. Messages with specific keywords are marked as likely to be spam. Messages in foreign languages are likely to be spam. Messages from e-mail servers in foreign countries might be assumed as likely spam. Spam filters also evaluate the message headers in an attempt to determine if the sender listed matches the route the mail has taken. The spam filter then totals up all the factors pointing to the e-mail message being legitimate and compares them to any factors which suggest the e-mail message is spam. If it finds sufficient cumulative evidence suggesting the message is spam it deletes it.

Some spam filters allow us to change ratings. For example, if I’m from France, I should be able to inform the spam filter that it should continue to assume messages from Nigeria are spam, but messages from France aren’t. I also need to tell it that messages written in French should not be considered spam.

A more drastic approach is to only accept e-mail from known senders. Using this approach if we expected an e-mail from someone, we would need to manually add them to our list of verified senders. For those not on the list, we can either simply filter out and automatically delete their messages or we can force the sender to do something which takes time and effort. For example, we might require the sender to go to a webpage and ask them to carry out some manual task—showing them a picture of an animal, for example, and requiring them to type in the name of the animal into a form on the webpage. The idea behind this approach is that a spam sender won’t bother going to the webpage and carrying out the task listed, so they won’t be able to send us e-mail. A legitimate sender who really wants us to get their e-mail message will go to the test webpage enter the information found, and thus their e-mail will get through. This approach can reduce the amount of spam received, but it also greatly increases the amount of hassle involved with sending e-mail.

Economic Approaches

As we previously mentioned, one reason for the prevalence of spam is that e-mail costs virtually nothing to send. If we change the economics of e-mail we can reduce or eliminate spam. Suppose, for example, we charged everyone $0.001 per e-mail message sent. For most of us, this would cost at most a few dollars a month. However, to a spammer sending millions and millions of messages per day, costs would add up quickly.

Recommendations for Dealing with Spam

While there is no way to completely eliminate spam, you can take some steps to reduce the amount of spam received and reduce the amount of time spam takes.

- Having several e-mail addresses, each for a specific purpose can help manage spam. Keep one e-mail address exclusively for work. Be very careful on who you give this e-mail address to. Use a second e-mail address for casual use—for example when you want to register to use a free website, or when you're participating in online discussions or bulletin boards. This technique will limit the amount of spam you receive at your work e-mail address.

- Be careful about posting an actual e-mail address on the web. Remember spammers sometimes cull webpages looking for e-mail addresses. If I list my e-mail address on my webpage as

  psy at statecollege.edu

  rather than

  psy@statecollege.edu

  humans reading the webpage will be able to determine my e-mail address, but automated programs trying to find e-mail addresses for spammers likely won't be able to tell that is is an e-mail address and thus won't add my address to the spam list.

- Use a spam filter on your e-mail. This will reduce the amount of spam you see (but keep in mind it may also occasionally delete a legitimate e-mail message).

Spyware and Adware

While spam brings advertising to your mailbox, spyware and adware take a more direct and even more intrusive approach. Spyware and adware are programs which are surreptitiously installed on your computer. They work behind the scenes to gather marketing information on your interests; they display advertisements; and in some cases they may be used to steal information such as your credit card numbers.

A spyware program is a program which is installed on your computer—typically without your full knowledge—which spies on your activities. A spyware program can study which websites you visit. It can track your keystrokes as you type in both your web browser and in other programs. Spyware may gather information on your habits and send it to a marketing company. It may also be used by criminals to gather financial and other personal information on you.

Adware is a program which is typically installed without your full knowledge which presents advertisements to you when you use the computer. These advertisements may be displayed randomly, often appearing as popup advertisements. Some adware overrides the standard behavior of your web browser, resetting the web browser's startup page and search pages to use those of the adware company.

In addition to their direct effect of gathering information on your actions and presenting advertisements, spyware and adware may adversely affect the performance of your computer. Spyware and adware programs run in the background simultaneously with your other programs. If a computer has a large number of spyware and adware programs installed, they may slow the computer down to a crawl, presenting you from getting your real work done.

Alternate Definitions and Related Issues

Before we get any further, I should mention that there are no "legal" definitions for spyware and adware. While the definition of spyware I have given is largely universal, the term adware is sometimes used to include any program which presents advertisements. This includes both programs running in the background as I have described, but might also include programs which you knowingly install and run which include advertisements as part of their interface. For example an instant message system or a music player might set aside part of its application window to display an advertisement. I don't include these in my definition of adware as in these cases we can clearly see where the advertisement is coming from, and we can make a clear decision to either run the program including advertising or to remove the program with advertising. Using my definition, a program is only adware if the program which is creating the ads is not clearly visible and if the user has not explicitly decided to run the program creating the ads. Both my definition and the alternate definition are in widespread use.

We should also distinguish between popup ads generated by websites we are visiting vs. popup ads generated by adware. While many users find popup ads annoying, if we visit a website, particularly if we are getting to read content on the website for free, that website has the right to display a popup ad. Advertising is one method websites use to gain revenue. This allows them to stay in business and to continue to provide us with free websurfing. Of course, if we find a website's popups sufficiently annoying we can stop patronizing it. We'll explore website revenue models in a later chapter in this book. The point for this discussion is that not all popup advertising is coming from adware. Adware popups are distinguished by coming from the adware program running on your computer. They have nothing to do with the websites you visit and do not help those websites generate revenue or stay in business.[6]

Here are some of the characteristics used by computer experts in determining whether or not a program is a legitimate program or if it is spyware or adware:

- Legitimate programs place an icon on the desktop or (on Microsoft Windows) in the start menu. They include both an icon to run the program as well as an icon which can be used to run an uninstall program. If they display advertisements, they make it very clear that the advertisements are appearing as a result of running the program.

- In contrast, spyware and adware programs typically try to remain as invisible as possible. They run automatically as soon as your computer starts up. They don't include any indication that they are running because they don't want you to know that they are on your computer. Advertisements from adware may appear mysteriously with no indication as to what is causing the ad to appear. Spyware and adware programs do not include an uninstall icon. In fact they generally try to make their removal as difficult as possible.

How Spyware and Adware Get on Your Computer

Clearly no one actually wants spyware and adware, so how does it get on our computers? There are a variety of methods that spyware and adware purveyors use to install their products.

- One of the most common methods used is to piggyback on top of another product. While the user thinks he is installing an improved clock, a weather tracking program, a screensaver, or an animated "helper" in fact both the program the user thinks he's installing and the spyware program are both getting installed at the same time.

  While not every piece of free software found on the Internet includes adware or spyware, many of them do. Be particularly careful of software found on peer-to-peer file sharing networks. Free "fun" software advertised on webpage banner or popup advertisements is another potential source of adware or spyware.

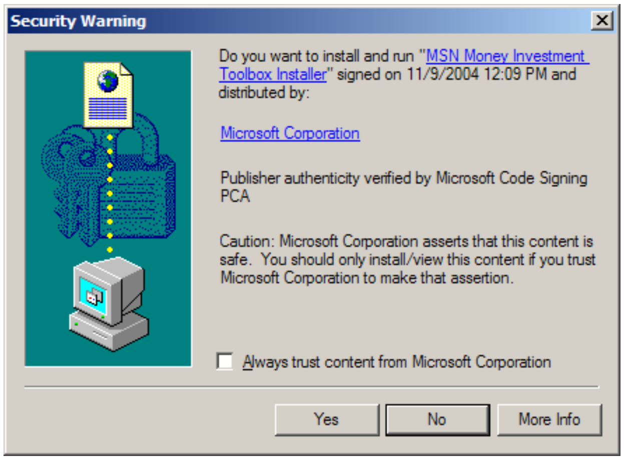
- Web browsers generally limit what a webpage can do on your computer. For example, a webpage can't read your hard disk or track your websurfing habits. For security reasons, most webpage programs are run inside a virtual *sandbox*. They can do whatever they want within

[6] Some adware programs have gone so far as to override advertisements provided by the websites you patronize and replace them with advertisements from the adware company. For example, I might visit an Internet newspaper and instead of seeing the newspaper's banner advertisements, the adware program would secretly replace them with its own ads. This causes a loss of marketing revenue to the newspaper website and instead provides marketing revenue to the adware company.

the sandbox, but they can't do anything outside of the sandbox. The sandbox includes the ability to change what's shown in the web browser. But the files on your computer and the ability to install programs are outside the sandbox. So a program associated with the webpage couldn't normally effect your disk drive or install a program.

Microsoft Internet Explorer supports a feature called ActiveX which uses a different approach to security. ActiveX controls do not work within a sandbox. They potentially have unlimited access to your computer and can do virtually anything they want. To prevent malicious ActiveX controls, Internet Explorer notifies the user before installing an ActiveX control. It also typically requires that an ActiveX control be certified and it authenticates the producer of the ActiveX control.[7] However, nothing prevents an adware company from certifying that they have indeed created the ActiveX control. A user presented with an ActiveX installation message may simply notice that the control is certified and may not look further overlooking the fact that the certificate shows that the control was created by the "Fly-by-Night Direct Marketing Company".

This is an example of an ActiveX installation dialog box.



Do you want to install and run "MSN Money Investment Toolbox Installer" signed on 11/9/2004 12:09 PM and distributed by:

Microsoft Corporation

Publisher authenticity verified by Microsoft Code Signing PCA

Caution: Microsoft Corporation asserts that this content is safe. You should only install/view this content if you trust Microsoft Corporation to make that assertion.

□ Always trust content from Microsoft Corporation

In this case it shows that we are about to install the "MSN Money Investment Toolbox Installer" ActiveX control. This ActiveX control has been verified as coming from the Microsoft Corporation. If we trust Microsoft Corporation enough to give it unlimited access to our computer, we can go ahead and click on "Yes". If we have doubts about the corporation and don't trust them we should click on "No". Again if you have any doubts about the company listed or have never heard of the company listed, you should click on "No".

While ActiveX controls are used by legitimate businesses to enhance their webpage, ActiveX controls can be themselves spyware, or they may install spyware programs. Once a user has agreed to the installation of an ActiveX control, there is virtually no limit to what that control can do.

- While ActiveX may be considered a feature and is part of the Internet Explorer design, other security holes exist in IE and other web browser. While these holes are fixed as soon as they are found, some spyware and adware companies install their programs by taking advantage of these security holes before they are patched.

[7] While IE is typically configured to require authentication of ActiveX controls, this setting may be changed. Some IE installations may be set to allow completely unrestricted installation of ActiveX controls.

Legality of Spyware and Adware

Some spyware is clearly illegal. Tracking a user as she types a credit card number and then sending it surreptitiously to another computer, for example, is clearly illegal. However, much of spyware and adware is in a legal gray zone, and in some cases it may be perfectly legal. The programs which spyware and adware piggyback on top of may include a very brief (and obscure) reference to the installation of the spyware within their End User License Agreement (EULA). Since most users don't read the EULA, they don't realize that they're agreeing to allow collection of marketing data or presentation of ads at the same time that they're installing the program the adware is piggybacked on top of.

Recommendations for Dealing with Spyware and Adware

- Regularly run a spyware/adware removal program. Two of the most popular removal programs are Ad-Aware and Spybot – Search & Destroy. These programs will search your disk drive for spyware and adware and will remove any programs found. You should be aware that anti-virus software often does not search for spyware. You'll want to run both anti-virus and anti-spyware software to keep your computer healthy.

- Don't indiscriminately install programs. Random programs found on the Internet are highly likely to contain adware or spyware. This is particularly true of programs found on peer-to-peer file sharing networks (e.g., Kazaa). If you see a program you really think might be useful, do some research and find out if the program includes adware or spyware.

- Read through the End User License Agreement (EULA) for any software package you install. Pay careful attention to the privacy policy listed in the EULA.

- If you use Internet Explorer, be particularly careful when you see IE's ActiveX installation dialog. You may want to consider turning off ActiveX all together or restricting its use beyond the default settings. Be aware, however, that doing so may prevent some legitimate websites from working properly.