Internet and WWW

p.192

The internet is a system of **inter**connected **net**works that can be accessed globally based on transmission (TCP) and internet (IP) protocols. Users are able to exchange information in a number of different formats. Within the internet, many resources and services are built on top of it such as WWW, electronic mail, file sharing, audio calls etc.

It relies on a physical infrastructure of routers and networks to allow individual devices to connect to other networks and devices.

WWW
The world wide web is a service of the internet based on HTTP (hypertext transfer protocol). It is a collection of media and information in the form of web pages where users can search, retrieve, and share data on the internet. This data can be located via hyperlinks (URL).

The internet is used to send and receive data from web servers. Web browsers are used to access web media.

URL
A uniform resource locator (URL) is a text-based address to access a web page.

Architecture of a URL:

1) Protocol
2) Domain host
3) Domain name
4) Domain type
5) Country code top-level domain

https://**www**.**wisdom**.**wis**.**edu**.**hk**/

Link with paths:

1) Path
2) File name

https://wisdom.wis.edu.hk/**login**/**index.php**

Hypertext Transfer Protocol (HTTP)

HTTP is a set of rules that must be followed when transferring data across the internet.

It is on the application layer, meaning the layer of human-computer interaction. ‘Applications can access network services here’.

To use HTTP, the computers (request, response) communicating must both use HTTP protocol.

HTTP is connectionless:

1. Computer makes a request for a web page
2. HTTP connection is made to the server
3. Computer and server disconnect
4. When the response is ready
5. Computer and server reconnect
6. Data is delivered

HTTP is also stateless as the computer and server only know about each other while they are connected. Previous data is not recorded. When they reconnect, their information must be sent again.

HTTPS
If SSL or TLS is used the protocol changes to HTTPS in the address bar. It means that a method of security is used.

SSL
SSL or secure socket layer is a security protocol that encrypts data transmitted between a user and a web server. SSL uses digital certificates to verify the authenticity of a website. This digital certificate contains the website's public key. When the user wants to communicate with a web server, a digital certificate is requested and then validated by the user's browser. Once authenticated, data transmission occurs, and SSL encrypts the data sent between them.

TLS
TLS is a successor to SSL. It has two layers: Handshake layer, where a secure connection is established between two endpoints. The record layer is where data is securely transmitted using cryptography.

Web browsers

Web browsers are software that is used to display web pages on screens by interpreting the html from the web servers.

Website - structured group of interrelated web pages

Features of a web browser

- Providing navigation tools - forward and backward buttons (must give example)
- Cookie management
    - Stores persistent cookies on hard drive, deletes on expiry date
- Providing an address bar
- Storing bookmarks and favourite websites
- Multiple tabs can be open at the same time
- Recording user history
    - Allows users to access previously viewed websites
- Interprets HTML and converts into viewable text

- Sends URLs to DNS
- Manages protocols - requests digital certificates from web servers

HTML
HTML (hypertext markup language) is a scripting language which is used to build and create the structure and presentation for web pages. It is used to display content visually and can format media such as images, text, videos, vectors, etc.

Domain Name Server (DNS)

Domain Name Server (DNS)
A domain name system negates the need for users to memorise IP addresses. Instead, the user types in the web page's URL into the address bar which is then converted into the corresponding IP address which the computer can understand.

Many domain name servers are located around the world. They each store a database of domain names (URLs) and their corresponding IP addresses.

1) The user types the URL into the address bar
2) The web browser will send the URL to the nearest domain name server (DNS)
3) The DNS will then find the corresponding IP address from its database
4) If the DNS does not have the IP address, it will request it from the next nearest DNS until it is found. It will then be sent back to the DNS which will send the IP address to the user's computer
5) The web browser makes a http request to the web server using its IP address
6) The web server will then re-establish the http connection and transmit the html files from the website to the computer
7) The web browser then interprets the HTML files and displays the information on the user's computer

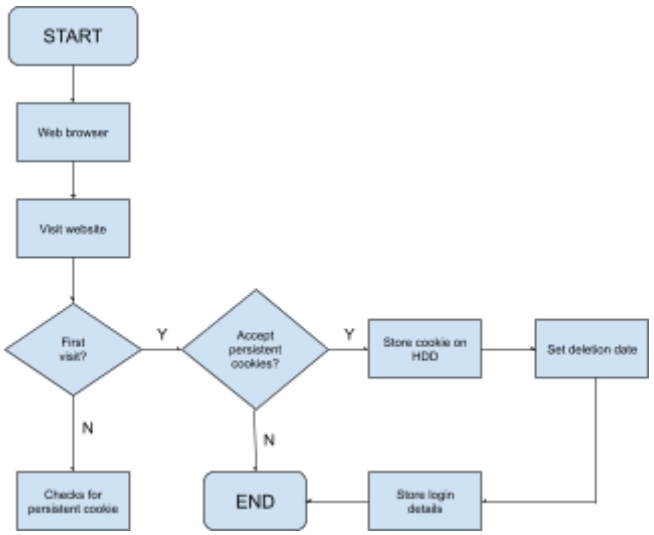
Cookies

Small files stored on a computer's memory sent by a web server. They are in the format of (key, value). Websites use cookies for user tracking and preferences. This allows websites to personalise for the user, such as language or showing relevant recommendations. Additionally, information can be stored temporarily during a session for the website to use.

When a user logs into a website, the website checks if the user has any cookies set on their browser. If so, the browser reads the information in the cookie. If not, the web server will send a cookie to the user.

These can be either session or permanent cookies.

Persistent	Session
Stored on hard drive until an expiry date is reached or the user deletes it	Stored until session ends (website is closed), no expiry date
Legitimate websites encrypt data stored in a persistent cookie	Stored in temporary storage (RAM)
Can store user login details, preferences long-term	Can store virtual shopping cart items
Can collect information from users computer	Cannot personally identify users or collect information from computer
Efficient method of carrying data from one website session to another, reducing the need for web servers to store massive amounts of data with their own resources	



Digital currencies

Fiat currencies are the more common, conventional currencies that may exist in physical forms such as the US dollar. Digital currencies **only exist digitally (in a digital format)**.

Digital currencies allow for online transactions such as PayPal or credit cards. Users can buy goods and services online using a digital currency. This is much more efficient than using physical cash, especially when paying a large sum of money. However it may be converted back to physical currency if necessary.

However, digital currencies use centralisation. This raises issues with confidentiality as the money does not go directly from buyer to seller, but through a central bank. The central banks and governments also control exchange rates.

Cryptocurrency removes the need of a central bank by allowing currencies to be transferred directly. These transactions are all available to the public, and the money can be kept track of and verified easily, secured by cryptography. The rules are also defined by the community itself.

Blockchain is used to secure transactions. The system is held within a blockchain network.

Blockchain

Blockchain is a database in which every computer or network connected to the blockchain has access to the information of every transaction. It is a decentralised system; the networks are not connected to a central server. Once a transaction has been made, every network receives a copy of the transaction, which prevents the data being changed unless every member on the network agrees to it.

Blocks are the basic components of a block chain and are created when a transaction is made. They contain the data about the transaction, its own hash code, and the previous block's hash code.

The chain begins with a genesis block, containing the data, its hash code, and 0000 as its previous hash. Whenever a transaction is made, a new block is created.

If any of the blocks have been altered, the hash code is changed, and the next block's previous hash is invalid, causing the rest of the chain to be invalid.

As a security precaution, to add a new block a miner must complete proof-of-work,an algorithm which validates that the block contains the most recent and accurate information, giving it the correct hash. This algorithm takes a long time (several minutes), meaning that cyber criminals cannot immediately add blocks undetected.

Miners are special network members who receive a commision for every new block created, hence 'miners' as they 'mine' crypto.

Cyber security

Threats

- Brute force attacks
    - The hacker essentially tries a number of combinations to ‘guess’ a password
    - They can either go through a list of common, simple passwords such as 11111 or 12345
    - Or they can go through a ‘word list’ of probable words, which is more efficient than raw trial and error. It is a program that creates a list of many many words that the hacker will try.
    - A strong and regularly changing password can prevent this
- Data interception
    - The cyber criminal intercepts a communication channel and reads the data being transmitted to steal information
    - **Wardriving** is a method in which the hacker would set up a device near the victim which would intercept Wi-Fi signals and read data
    - A **packet sniffer** is a program or hardware device that allows cyber criminals to read private data from packets travelling across networks
    - It examines packets and sends the data back to the hacker - however, if the data is encrypted, the hacker would not be able to read the information.
    - Wired equivalent privacy or **WEP** encrypts data which adds another layer of privacy and can reduce the effects of wardriving
    - Using private Wi-Fi connections and using strong passwords on routers lowers the chances of data interception
- Distributed denial of service (DDoS) attacks
    - A cyber criminal uses bots to flood a machine or network resource such as a web server with spam traffic causing it to crash or be unable to service legitimate requests. This prevents users from accessing a part of a network.
    - The cybercriminal can send malware which will turn computers into bots
    - An up-to-date malware tracker may help as well as a firewall or proxy server to filter unwanted spam
- Hacking
    - Is unauthorised, illegal access to a computer system
    - This then allows the hacker to delete, alter, and share private data
    - Encryption although not allowing the hacker to read or understand the data, the hacker can still delete the files
    - Anti-hacking software, intrusion-detection, firewalls and strong passwords can prevent hacking
    - Ethical hacking is where paid hackers inspect a system’s security and test it with hacking attacks.
- Malware
    - Anti-malware software and firewall can prevent malware from attacking a system by removing or filtering out risky files
    - Viruses are programs that can replicate themselves by attaching themselves to programs that a victim may run
        - Viruses rely on an active host to run and do damage
        - They aim to delete or corrupt files, or cause a computer to malfunction

        - An example of a trigger is an infected email attachment, or infected software the user unknowingly downloads, infected websites
    - Worms
        - These are similar to viruses, however do not require a trigger or active host to replicate themselves
        - Because of this, they can infect an entire network once they have access to a system in the network via an application
        - Like viruses, they can create their own files which allows them to replicate and send copies of themself to others in the network
        - They rely on poor network security to spread unnoticed
    - Trojan horse
        - Is malicious software disguised as legitimate software, tricking the user into downloading the malware
        - Once downloaded, the trojan virus has access to the device and can then send personal information to hackers
        - Firewalls or standard security systems cannot prevent this as this relies on manipulation of the user, and users can overrule anti-malware software
    - Spyware
        - Is a type of malware which illegally monitors a user's activity, and can gain unauthorised access to personal information (credit card details, passwords, back account information)
        - This information is then sent to a third party, which is either maliciously used or sold
        - Anti-spyware software
    - Adware
        - Its purpose is to overload an end-user with unwanted advertisements
        - This may come in different forms such as popups or browser redirects
        - It is typically harder to remove or detect as it is difficult to decide whether or not it is harmful
    - Ransomware
        - Is something hackers can do to get money
        - The hacker will install malware on a victim's computer that blocks access to the victim's personal data until a ransom is paid.
        - This can be encrypting the victims data and decrypting it if a ransom is paid.
- Phishing
    - Is when an attacker sends seemingly legitimate emails to a victim
    - They may appear to be from legitimate companies, banks, or close friends/family
    - Tricks users into **voluntarily** sending private information, or contain fake links to fake websites/attachments the victim may click on
    - Only effective when initiated
    - Spear phishing is targeted at an individual, can be done to conduct corporate espionage, etc.
    - Anti-phishing software can detect phishing emails, authenticating the legitimacy of the sender will prevent phishing, firewalls can also prevent phishing

    - Users should also be wary of phishing scams, look out for https, check for typos in the address bar, do not interact with popups, and do not open suspicious emails
- Pharming
    - Is malicious code unknowingly installed on a victim's computer, possibly from clicking a link
    - It redirects victims to a fake website which resembles a legitimate website
    - This can trick victims into giving personal information; credit card information, login details, etc.
    - This can also be done through DNS cache poisoning which alters the legitimate website's IP address from the DNS and redirects the user to a fake website
    - SSL (https), and special software on browsers or anti-malware software can detect pharming and phishing, checking address bar for changes

The difference between phishing and pharming is that phishing requires action from the victim where pharming unknowingly redirects users to fake websites.

- Social engineering
    - Cyber criminals manipulate the victims human emotions to drop their guard
    - Gain access to private information or lead users to download malicious software
    - Bypasses standard methods of malware prevention

1) A company has offices in four different countries. Communication and data sharing between the offices is done via computers connecting over the internet.

a) Describe three data security issues the company might encounter during their day-to-day communications and data sharing

Data interception (wardriving)
Ransomware
Phishing

b) For each issue described, explain why it could be a threat to the security of the company

Sensitive data about the company, or financial information may be stolen leading to a loss in profits
Ransomware will hinder the company's productivity, as well as cause it to lose money to the hacker
Phishing may gain access to sensitive information about the company

c) For each issue described, describe a way to mitigate the threat that has been posed.

Wired equivalent privacy or WEP can protect the data sent through wireless communication through encryption
Hacking can be prevented through the use of strong passwords and security software such as anti-hacking software or anti-malware.
Phishing can be avoided by validating the sender, having anti-phishing software installed, and by using a firewall.

Worm: type of malware that intends to corrupt computers. It does not require a trigger, and can replicate itself to spread across all devices in a network.

Reducing cyber threats

Access levels
This can protect users' private information in databases. Many systems have a hierarchy of access levels to control who can read, write, amend, or delete data so that data is not misused. This is usually done through a username and password system to verify a person's access level.

One group of people may only be able to read that information, and another group may be able to alter or rewrite the information.

Anti-malware software

- Anti-virus
    - Utility software (system software)
    - Constanting running in the background
    - Checks files before they are run or loaded into a computer
    - Heuristic checking - compares software's behaviour to the behaviour of a virus, which is stored in a database of known viruses.
    - Quarantines files by locking it and preventing it from affecting the computer
    - Completes full-system scans to find dormant viruses
    - Needs to be kept up to date as new viruses are constantly found
    - Deletes files containing viruses
- Anti-spyware
    - Detects and removes spyware installed on the computer
    - Prevents users from downloading viruses
    - Encrypts users files
    - Encrypts keyboard strokes
    - Blocks access to webcams and mics
    - Scans for signs that the users data has been taken
    - Spyware is detected by looking for typical features of spyware or certain file structures that spyware is associated with. If spyware is identified it is deleted or blocked.

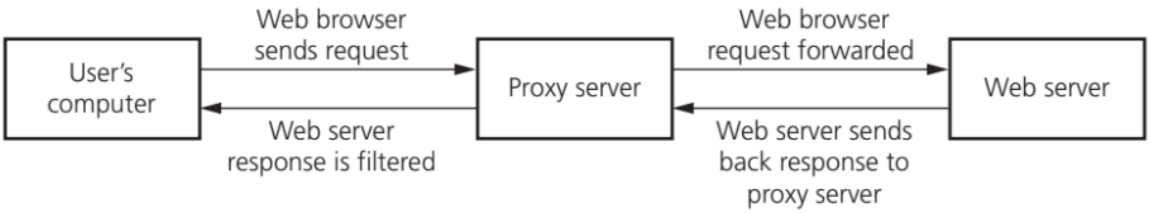
Firewall
A firewall is a network security device either software or hardware based that filters incoming and outgoing network traffic based on a set of rules. It sits between a user's computer and an external network. For example, the West Island firewall prevents students from accessing gaming websites, pornography, and gambling.

Network traffic is examined and compared to a given set of criteria, or a list of blocked IP addresses. Data can be blocked from entering the network or a security warning can be sent to the user. The user can also be warned if software within their computer or internal network is attempting to access an external data source.

The firewall may also log the details of incoming and outgoing traffic, recording the user who made it and who the receiver is. This allows later examination by the user or network manager.

However the firewall may be bypassed by the user, leaving the device vulnerable.

Proxy server
Proxy servers sit in between the user's computer and the web server they are trying to access. Its purpose is to filter possible threats or unwanted content by preventing direct access from the web server to the user's computer. It filters incoming traffic from the web server and can even block access to a web server if necessary.



It also keeps the user's IP address anonymous so that they cannot be tracked.

On top of this, it uses a cache to store information from a website so that when the user accesses it again, it is much faster.

Proxy servers can also act as firewalls (preventing files from an external network that meet a certain criteria from accessing an internal network).

Authentication
Used to verify a user so that their data or systems cannot be altered or read without the user's permission.

- Passwords
    - Strong passwords are long and have variation
    - They are changed regularly
- Two-factor authentication
    - Using two methods to identify a user
    - E.g password and PIN sent to mobile device or user's email

Biometrics

- Fingerprints
    - They also cannot be lost or stolen, as they are physically attached to someone, unlike a keycard or key
    - Fingerprints are identified by the patterns of ridges or valleys
    - Expensive to set up
    - Fast and easy to use
    - Not intrusive
    - Very difficult to replicate
    - Cannot sign in as someone else as the pattern only aligns with one user

- Retina scans
    - Uses infrared light to scan the unique pattern of blood vessels in the retina
    - very intrusive processes
    - Can be slow to verify (around 10-15 seconds)
    - Very accurate
    - Very expensive
    - Impossible to replicate
- Voice and facial recognition
    - Non-intrusive
    - Relatively inexpensive
    - Not very accurate
    - Face recognition can be affected by lighting, age, glasses or masks
    - Voice recognition can be affected by illness, age

Automatic software updates
Software updates are available relatively often, they are known as patches. They fix bugs, improve performance, and also improve anti-malware software. However, sometimes patches can cause issues and disrupt the functioning of the software, and the user must wait for the next patch for the issue to be resolved.

- Device must be plugged in and charged
- Usually occurs overnight while the user is logged off
- May require a system reboot after installation, also automatic

Checking emails and websites
To prevent phishing and pharming, spelling errors in emails, email addresses, and URLs can help the user detect a fake website or fake email address. The tone of the email should also be considered.

Typo-squatting is where websites have names very close to legitimate websites to try to trick the user into giving personal information.