Cambridge

OL- IGCSE

Computer science

CODE: (0478)

Chapter 05

The internet and its uses

5.1 The internet and the World Wide Web (WWW)

5.1.1 The differences between the internet and the World Wide Web (WWW)

The word **internet** comes from INTERconnected NETwork, since it is basically a worldwide collection of interconnected networks.

In contrast, the **World Wide Web (WWW)** is only a part of the internet that users can access using web browser software.

Table 5.1 Summary of differences between the internet and the World Wide Web

Internet	World Wide Web (WWW)
• users can send and receive emails	• it is a collection of multimedia web pages and other information on websites
• allows online chatting (via text, audio and video)	• http(s) protocols are written using hypertext mark-up language (HTML)
• makes use of transmission protocols (TCP) and internet protocols (IP)	• uniform resource locators (URLs) are used to specify the location of web pages
• it is a worldwide collection of interconnected networks and devices	• web resources are accessed by web browsers
	• uses the internet to access information from web servers

5.1.2 Uniform resource locators (URLs)

Web browsers are usually just referred to as browsers

Web browsers are software that allow users to access and display web pages on their device screens. Browsers interpret **hypertext mark-up language (HTML)** sent from websites and produce the results on the user's device. **Uniform resource locators (URLs)** are text addresses used to access websites. A URL is typed into a browser address bar using the following format:

protocol://website address/path/file name

The protocol is usually either http or https.

The website address is:

- domain host (www),
- domain name (website name),
- domain type (.com, .org, .net, .gov, for example),
- and sometimes country code (.uk, .de, .cy, for example).

The path is the web page, but is often omitted and it then becomes the root directory of the website (see example below).

The file name is the item on the web page. For example:

https://www.hoddereducation.co.uk/ict

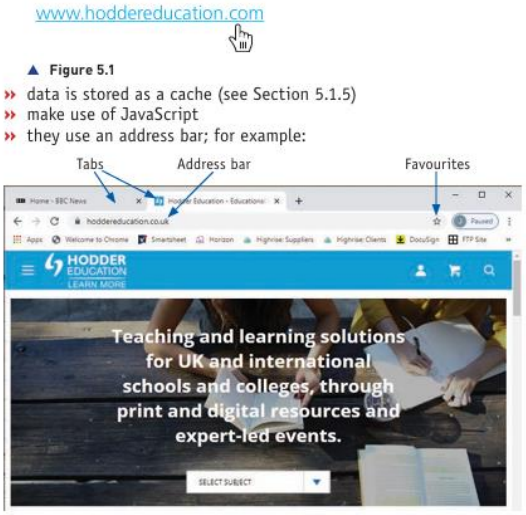
5.1.3 HTTP and HTTPS

Hypertext transfer protocol (http) is a set of rules that must be obeyed when transferring files across the internet. When some form of security (for example, SSL or TLS) is used, then this changes to https (you will often see the green padlock in the status bar as well).

5.1.4 Web browsers

Most browsers have the following features:

» They have a home page
» They can store a user's favourite websites/web pages (referred to as bookmarks)
» They keep a history of websites visited by the user (user history)
» They have the ability to allow the user to navigate forwards and backwards through websites/web pages already opened
» Many web pages can be open at the same time by using multiple tabs » they make use of cookies (see Section 5.1.6)
» They make use of hyperlinks that allow navigation between websites and web pages; links can be opened in one of two ways: either open in a new tab by using + or open in the same tab by simply clicking on the link



▲ **Figure 5.1**

» data is stored as a cache (see Section 5.1.5)
» make use of JavaScript
» they use an address bar; for example:

▲ **Figure 5.2** Browser address bar

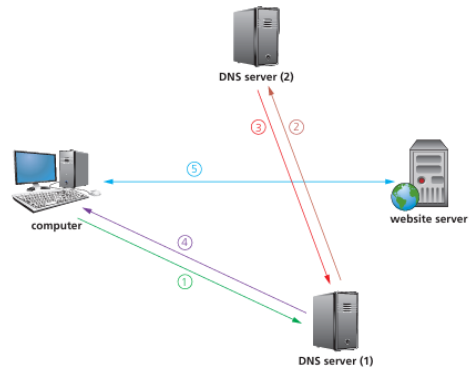
5.1.5 Retrieval and location of web pages

To retrieve pages from a website your browser needs to know this IP address. The **Domain Name Server (DNS)** (also known as domain name system) is a system for finding IP addresses for a domain name given in a URL.

DNS servers contain a database of URLs with the matching IP addresses. Figure 5.3 shows how a web page can be located and then sent back to the user's computer. The DNS plays a vital role in this process:

(1) The user opens their browser and types in the URL (www.hoddereducation.co.uk) and the browser asks the DNS server (1) for the IP address of the website.

(2) In this case, let's assume the DNS server can't find www.hoddereducation.co.uk in its database or its cache, so it sends out a request to a DNS server (2).



▲ **Figure 5.3** How DNS is used to locate and retrieve a web page

(3) The DNS server (2) finds the URL and can map it to 107.162.140.19; this IP address is sent back to the DNS server (1) which now puts this IP address and associated URL into its cache/database.

(4) This IP address is then sent back to the user's computer.

(5) The computer now sets up a communication with the website server and the required pages are downloaded. HTML files are sent from the website server to the computer. The browser interprets the HTML, which is used to structure content, and then displays the information on the user's computer.

5.1.6 Cookies

this tracks data about users, such as IP addresses and browsing activity

Cookies are small files or code stored on a user's computer. They are sent by a web server to a browser on a user's computer. Each cookie is effectively a small look-up table containing pairs of (key, data) values, for example, (surname, Jones) (music, rock). Every time a user visits a website, it checks if it has set cookies on their browser before. If so, the browser reads the cookie which holds key information on the user's preferences such as language, currency and previous browsing activity. Cookies allow user tracking and maintain user preferences. Collected data can also be used to customise the web page for each individual user. For example, if a user buys a book online, the cookies remember the type of book chosen by the user and the web page will then show a message such as "*Customers who bought Hodder IGCSE ICT also bought Hodder IGCSE Computer Science*".

There are two types of cookie:

- session cookie
- persistent (or permanent) cookie.

If a cookie doesn't have an expiry date associated with it, it is always considered to be a session cookie. So what are the basic differences?

There are two types of cookie:
» Session cookie » Persistent (or permanent) cookie

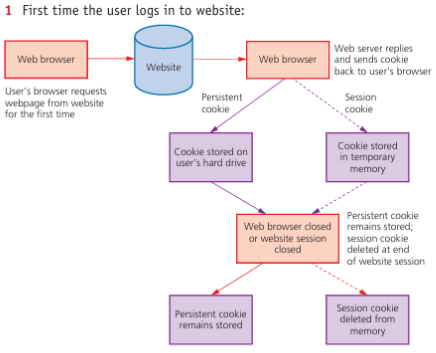
Session cookies

Session cookies are used, for example, when making online purchases. They keep a user's items in a **virtual shopping basket**. This type of cookie is stored in temporary memory on the computer, doesn't actually collect any information from the user's computer and doesn't personally identify a user.

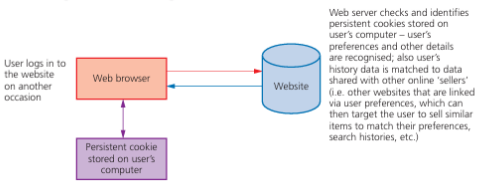
Persistent (permanent) cookies

Persistent cookies remember a user's log in details (so that they can authenticate the user's browser). They are stored on the hard drive of a user's computer until the expiry date is reached or the user deletes it. These cookies remain in operation on the user's computer even after the browser is closed or the website session is terminated.

Figures 5.4 and 5.5 summarise what happens when a website is first visited and then what happens in subsequent visits:



▲ **Figure 5.4** Cookies (first login)



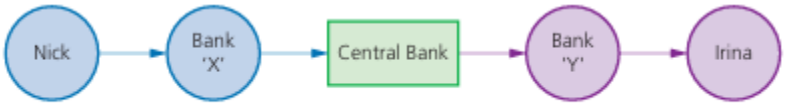
▲ **Figure 5.5** Cookies (subsequent logins)

5.2 Digital currency

5.2.1 What is digital currency?

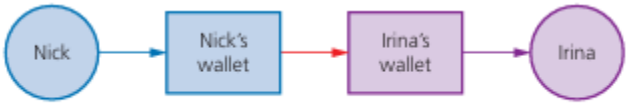
Digital currency exists purely in a digital format. It has no physical form unlike conventional **fiat currency** (for example, $, £, €, and ¥)

Digital currency relies on a **central banking system**. For example, suppose Nick wishes to send Irina some money; Nick uses bank 'X' and Irina uses bank 'Y':



▲ **Figure 5.6** Digital currency

The problem with centralisation is maintaining confidentiality and security; these have always been issues with digital currency systems. However, one example of digital currency, known as cryptocurrency, has essentially overcome these issues by introducing decentralisation:



▲ **Figure 5.7** Cryptocurrency and decentralisation

» Cryptocurrency uses cryptography to track transactions; it was created to address the problems associated with the centralisation of digital currency.

» Traditional digital currencies are regulated by central banks and governments (in much the same way as fiat currencies). This means all transactions and exchange rates are determined by these two bodies. Cryptocurrency has no state control and all the rules are set by the cryptocurrency community itself.

» Unlike existing digital currencies, cryptocurrency transactions are publicly available and therefore all transactions can be tracked and the amount of money in the system is monitored.

» The cryptocurrency system works by being within a blockchain network which means it is much more secure.

5.2.2 Blockchaining

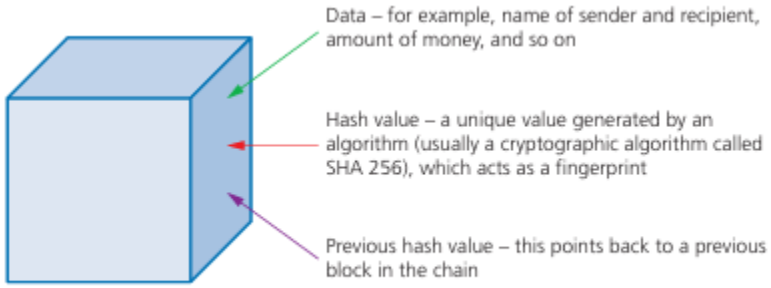
Essentially, the blockchain consists of a number of interconnected computers but they are **not connected** to a central server. **All** transaction data is stored on all computers in the blockchain network.

Whenever a new transaction takes place, all the networked computers get a copy of the transaction; therefore **it cannot be** changed without the consent of **all** the network members.

This effectively removes the risk of security issues such as hacking. Blockchain is used in many areas, such as:

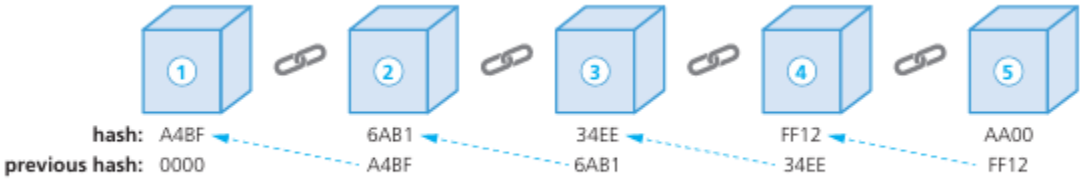
» Cryptocurrency (digital currency) exchanges
» Smart contracts
» Research (particularly within pharmaceutical companies)
» Politics
» Education

How blockchain works Whenever a new transaction takes place, a new **block** is created



▲ **Figure 5.8** Block description

A new hash value is created each time a new block is created. This hash value is unique to each block and includes a **timestamp**, which identifies when an event actually takes place. We will now consider what happens when a chain of blocks is created. Figure 5.9 shows part of a typical blockchain:



▲ **Figure 5.9** Part of blockchain (showing 5 blocks)

It is clear from Figure 5.9 how these blocks are connected. Block '1' is known as the **genesis block** since it doesn't point to any previous block. Now suppose block '2' is changed in some way.

This is prevented by **proof-of-work**, which makes sure it takes ten minutes to determine the necessary proof-of-work for each block **before** it can be added to the chain. This is 'policed' by **miners**, which are special network users that get a commission for each new block created.

5.3 Cyber security

5.3.1 Cyber security threats

Data can be corrupted or deleted either through accidental damage or malicious acts. There are also many ways data can be intercepted leading to cyber security threats. The following list shows the cyber threats which will be considered in this section:

» Brute force attacks
» Data interception
» Distributed denial of service (DDoS) attacks
» Hacking
» Malware (viruses, worms, Trojan horse, spyware, adware and ransomware)
» Phishing
» Pharming
» Social engineering.

Brute force attacks

If a hacker wants to 'crack' your password, they can systematically try all the different combinations of letters, numbers and other symbols until eventually they find your password. This is known as a **brute force attack** and there isn't a lot of sophistication in the technique.

One way to reduce the number of attempts needed to crack a password is to first go through a series of logical steps:

1. Check if the password is one of the most common ones used (the five most common are: 123456, password, qwerty, 111111 and abc123); since these simple passwords are seen so many times it's a good place for the hacker to start.

2. If it isn't in the common password list, the next thing to do is to start with a strong word list (this is a text file containing a collection of words that can be used in a brute force attack); some programs will generate a **word list** containing a million words. Nonetheless this is still a faster way of cracking a password than just total trial and error.

Data interception

Data interception is a form of stealing data by **tapping** into a wired or wireless communication link. The intent is to compromise privacy or to obtain confidential information.

Interception can be carried out using a **packet sniffer**, which examines data packets being sent over a network. The intercepted data is sent back to the hacker. This is a common method when wired networks are used.

Wi-Fi (wireless) data interception can be carried out using wardriving (or sometimes called **Access Point Mapping).**

Therefore, to safeguard against wardriving, the use of a **wired equivalency privacy (WEP**) encryption protocol, together with a firewall, is recommended.

Distributed Denial of Service (DDoS) attacks

A denial of service (DoS) attack is an attempt at preventing users from accessing part of a network, notably an internet server. This is usually temporary but may be a very damaging act or a large breach of security. The attacker may be able to prevent a user from:

» Accessing their emails
» Accessing websites/web pages
» Accessing online services (such as banking). One method of attack is to flood the network with useless spam traffic. How does this cause a problem?

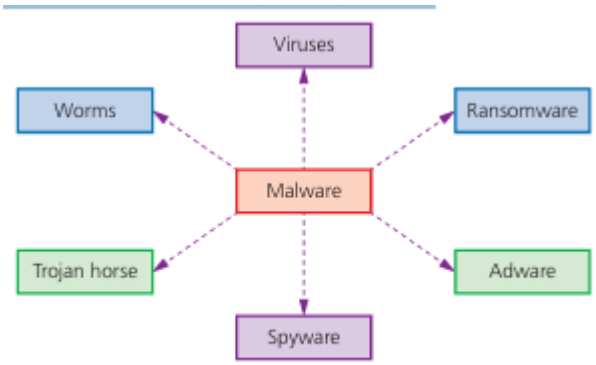
In a **distributed denial of service (DDoS**) the spam traffic originates from many different computers, which makes it hard to block the attack.

Hacking

Hacking is generally the act of gaining illegal access to a computer system without the user's permission. However, universities and companies now run courses in **ethical hacking**.

Malware

Malware is one of the biggest risks to the integrity and security of data on a computer system. There are many forms of malware; this chapter will only consider the following in any detail:



▲ **Figure 5.10** Malware types

Viruses

Viruses are programs or program code that replicate (copies themselves) with the intention of deleting or corrupting files, or causing a computer to malfunction.

Viruses need an **active host** program on the target computer or an operating system that has already been infected, before they can actually run and cause harm (that is, they need to be executed by some trigger before starting to cause any damage).

Worms

Worms are a type of stand-alone malware that can self-replicate. Their intention is to spread to other computers and corrupt whole networks; unlike viruses, they don't need an active host program to be opened in order to do any damage.

Trojan horse

A Trojan horse is a program which is often disguised as legitimate software but with malicious instructions embedded within it. A Trojan horse replaces all or part of the legitimate software with the intent of carrying out some harm to the user's computer system.

Spyware

Spyware is software that gathers information by monitoring a user's activities carried out on their computer. The gathered information is sent back to the cybercriminal who originally sent the spyware.

Adware

Adware is a type of malware. At its least dangerous it will attempt to flood an end-user with unwanted advertising.

Although not necessarily harmful, adware can:
» Highlight weaknesses in a user's security defences
» Be hard to remove – it defeats most anti-malware software since it can be difficult to determine whether or not it is harmful
» Hijack a browser and create its own default search requests

Ransomware

Essentially, ransomware are programs that encrypt data on a user's computer and 'hold the data hostage'. The cybercriminal waits until the ransom money is paid and, sometimes, the decryption key is then sent to the user. It has caused considerable damage to some companies and individuals.

Summary of malware

Table 5.2 summarises the six types of malware described in Section 5.3.1.

Table 5.2 Summary of types of malware

Viruses – programs (or program code) that can replicate/copy themselves with the intention of deleting or corrupting files, or causing the computer to malfunction. They need an active host program on the target computer or an operating system that has already been infected before they can run
Worms – these are types of standalone viruses that can replicate themselves with the intention of spreading to other computers; they often networks to search out computers with weak security that are prone to such attacks
Trojan horses – these are malicious programs often disguised as legitimate software; they replace all or part of the legitimate software with the intent of carrying out some harm to the user's computer system
Spyware – software that gathers information by monitoring, for example, all the activity on a user's computer; the gathered information is then sent back to the person who sent the software (sometimes spyware monitors key presses and is then referred to as key logging software)
Adware – software that floods a user's computer with unwanted advertising; usually in the form of pop-ups but can frequently appear in the browser address window redirecting the browser to a fake website which contains the promotional adverts
Ransomware – programs that encrypt the data on a user's computer; a decryption key is sent back to the user once they pay a sum of money (a ransom); they are often sent via a Trojan horse or by social engineering

Phishing

Phishing occurs when a cybercriminal sends out legitimate-looking emails to users. The emails may contain links or attachments that, when initiated, take the user to a fake website; or they may trick the user into responding with personal data

There are numerous ways to help prevent phishing attacks:

- users need to be aware of new phishing scams; those people in industry or commerce should undergo frequent security awareness training to become aware of how to identify phishing (and pharming) scams
- it is important not to click on any emails links unless totally certain that it is safe to do so; fake emails can often be identified by 'Dear Customer .....' or 'Dear email person@gmail.com .....' and so on
- it is important to run anti-phishing toolbars on browsers (this includes tablets and mobile phones) since these will alert the user to malicious websites contained in an email
- always look out for https or the green padlock symbol in the address bar
- regular checks of online accounts are also advisable as well as maintaining passwords on a regular basis
- ensure an up-to-date browser is running on the computer device (which contains all of the latest security upgrades) and run a good firewall in the background at all times; a combination of a desktop firewall (usually software) and a network firewall (usually hardware) considerably reduces the risk of hacking, pharming and phishing on network computers
- be very wary of pop-ups and use the browser to block them; if pop-ups get through your defences, don't click on 'cancel' since this can ultimately lead to phishing or pharming sites – the best option is to select the small ✖ in the top right-hand corner of the pop-up window which closes it down.

Note: another term connected to phishing is **spear phishing**; this is where the cybercriminal targets **specific** individuals or companies to gain access to sensitive financial information or industrial espionage – regular phishing is not specific regarding who the victims are.

Pharming

Pharming is malicious code installed on a user's computer or on an infected website. The code redirects the user's browser to a fake website **without** the user's knowledge.

Why does pharming pose a threat to data security?

As mentioned above, pharming redirects internet users to a fake or malicious website set up by, for example, a hacker; redirection from a legitimate website to the fake website can be done using **DNS cache poisoning**

Every time a user types in a URL, their browser contacts the DNS server; the IP address of the website will then be sent back to their browser. However, DNS cache poisoning changes the real IP address values to those of the fake website; consequently, the user's computer will connect to the fake website.

It is possible to mitigate against the risk of pharming:
» Use of anti-virus software can detect unauthorised alterations to a website address and warn the user of the potential risks.
» However, if the DNS server itself has been infected (rather than the user's computer) it is much more difficult to mitigate the risk.
» Many modern browsers can alert users to pharming and phishing attacks.
» It is very important to check the spelling of websites to ensure the web address used is correct.
» As with phishing, use of https or the green padlock symbol bar is an additional form of defence. in the address

Social engineering

Social engineering occurs when a cybercriminal creates a social situation that can lead to a potential victim dropping their guard. It involves the manipulation of people into breaking their normal security procedures and not following best practice. There are five types of threat that commonly exist:

Instant messaging	Malicious links are embedded into instant messages; for example, an important software upgrade (relies on user's curiosity)
Scareware (for example, fake anti-virus that looks real)	This is often done using a pop-up message that claims that the user's computer is infected with a virus; the user is told they need to download the fake anti-virus immediately (relies on user's fear)
Emails/phishing scams	The user is tricked by the apparent genuineness of an email and opens a link in the email; this redirects their browser to a fake website (relies on user's trust of well-known companies)
Baiting	The cybercriminal leaves a malware-infected memory stick somewhere where it can be found; the finder picks up the memory stick and plugs it into their computer (just to see who it belongs to) and unwittingly downloads malicious malware (relies on user curiosity)
Phone calls	For example, a so-called IT professional calls the user on their mobile claiming their device has been compromised in some way; the user is advised to download some special software that allows the cybercriminal to take over the user's device giving them access to personal information (relies on fear)

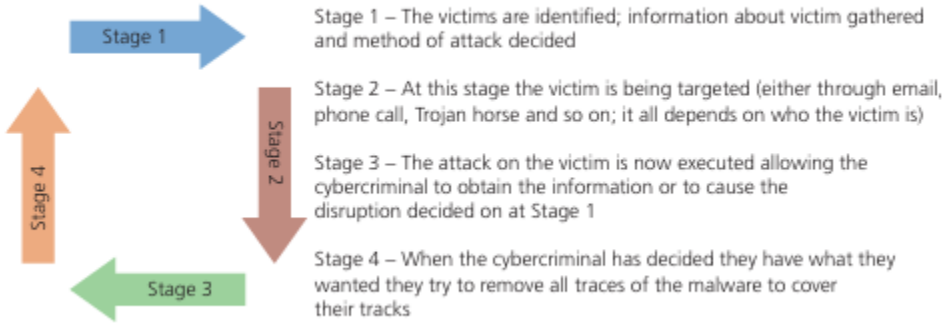
▲ **Figure 5.11** Social engineering

The three most common ones to exploit are:
» Fear – the user is panicked into believing their computer is in immediate danger and isn't given time to logically decide if the danger is genuine or not; fear is a very powerful emotion that can easily be exploited by a cybercriminal

» Curiosity – the user can be tricked into believing they have won a car or they find an infected memory stick lying around; their curiosity gets the better of them and they give their details willingly to win the car

» Empathy and trust – a real belief that all genuine-sounding companies can be trusted, therefore emails or phone calls coming from such companies must be safe; a dangerous assumption that the cybercriminal can exploit fully

Figure 5.12 shows the course of action taken by a cybercriminal in targeting their victim:



▲ **Figure 5.12** Stages in a typical social engineering scam

5.3.2 Keeping data safe from security threats

Access levels

In many computer systems, user accounts control a user's rights. This often involves having different **levels of access** for different people.

In this type of application, users are allowed to use **privacy settings** rather than passwords to decide the level of access (for more on this see later in this section).

▲ **Figure 5.13** Access level log in screen

Anti-malware

The two most common types of anti-malware are anti-virus and anti-spyware.

Anti-virus

Anti-virus has already been described in great detail in Chapter 4.

Anti-spyware

Anti-spyware software detects and removes spyware programs installed illegally on a user's computer system. The software is based on one of the following methods:

» Rules – in this case, the software looks for typical features which are usually associated with spyware thus identifying any potential security issues
» File structures – in this case, there are certain file structures associated with potential spyware which allows them to be identified by the software

The general features of anti-spyware are:

» Detect and remove spyware already installed on a device
» Prevent a user from downloading spyware
» Encrypt files to make the data more secure in case it is 'spied' on
» Encryption of keyboard strokes to help remove the risk posed by the keylogging aspects of some spyware
» Blocks access to a user's webcam and microphone (the software stops the spyware taking over the control of a user's webcam and microphone which can be used to collect information without the user's knowledge)
» Scans for signs that the user's personal information has been stolen and warns the user if this has happened

Authentication

Authentication refers to the ability of a user to prove who they are. There are three common factors used in authentication:
» Something you know (for example, a password or PIN code)
» Something you have (for example, a mobile phone or tablet)
» Something which is unique to you (for example, biometrics).

Passwords and user names

Passwords are used to restrict access to data or systems. They should be hard to crack and changed frequently to retain any real level of security. Passwords can also take the form of biometrics.

Biometrics

Biometrics can be used in much the same way as passwords as a way of identifying a user

▲ **Figure 5.14** Fingerprint scan

Fingerprint scans

Images of fingerprints are compared against previously scanned fingerprint images stored in a database; if they match, then a user has been correctly recognised. The system compares patterns of 'ridges' and 'valleys' that are unique. The accuracy of the scan is about around 1 in 5000. Fingerprint scanning techniques have the following benefits as a form of security:

» Fingerprints are unique, therefore this technique can improve security since it would be difficult to replicate a person's fingerprints

» Other security devices (such as magnetic cards to gain entry to a building) can be lost or even stolen which makes them less effective

» It would be impossible to 'sign in' for somebody else since the fingerprints would match with only one person on the database

» Fingerprints can't be misplaced; a person always has them!

Retina scans

Retina scans use infrared light to scan the unique pattern of blood vessels in the retina (at the back of the eye); it is a rather unpleasant technique requiring a person to sit totally still for 10 to 15 seconds while the scan takes place;

it is very secure since nobody has yet found a way to duplicate the blood vessels patterns. The accuracy is about 1 in 10 million.

▼ **Table 5.3** Comparison of biometric devices

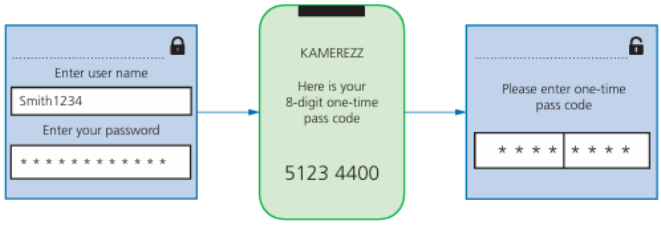
Biometric technique	Benefits	Drawbacks
fingerprint scans	it is one of the most developed biometric techniques very easy to use relatively small storage requirements for the biometric data created	for some people it is very intrusive, since it is still related to criminal identification it can make mistakes if the skin is dirty or damaged (e.g. cuts)
retina scans	very high accuracy there is no known way to replicate a person's retina	it is very intrusive it can be relatively slow to verify retina scan with stored scans very expensive to install and set up
face recognition	non-intrusive method relatively inexpensive technology	it can be affected by changes in lighting, the person's hair, change in age, and if the person is wearing glasses
voice recognition	non-intrusive method verification takes less than 5 seconds relatively inexpensive technology	a person's voice can be easily recorded and used for unauthorised access low accuracy an illness such as a cold can change a person's voice, making absolute identification difficult or impossible

▲ **Figure 5.15** Retina scan

Biometric applications

Two-step verification

Two-step verification requires two methods of authentication to verify who a user is. It is used predominantly when a user makes an online purchase using a credit/debit card as payment method.



▲ **Figure 5.17** Two-step verification using a mobile phone

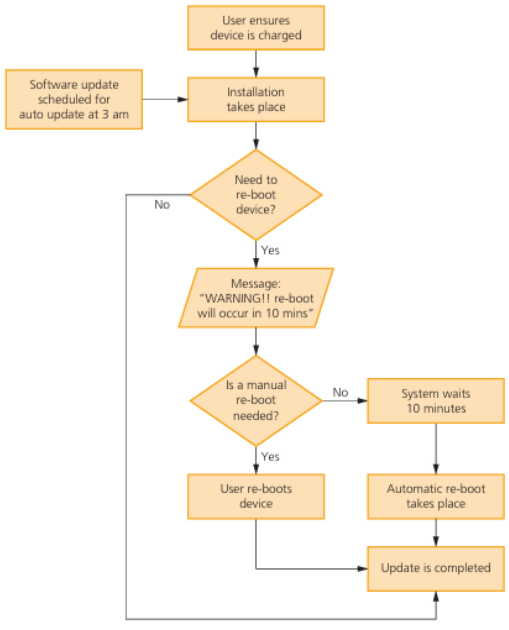
Automatic software updates

These updates are vital since they may contain **patches** that update the software security (to protect against malware) or improve the software performance. The only downside to this is the potential for updates to disrupt your device following installation.

Checking the spelling and tone of communication and URL links When emails are sent to you, there are three actions you always need to take before opening them or activating any links in them.

» Check out the spellings in the email and in the links; professional, genuine organisations will not send out emails which contain spelling or major grammatical errors (for example, Amazzon.com)

» Carefully check the tone used in the email message; if it is rushing you into doing something or if the language used seems inappropriate or incorrect, then it could be a phishing email or worse



▲ **Figure 5.18** Automatic software update flow chart

There are five things to look out for:

1 The email address itself; no legitimate company will use an email address such as: @gmail.com
Carefully check the part of the address after the '@' symbol which should match the company's name; for example:

account-update@amazon.com

2 The tone of the email and bad spelling of words is a clear indication of a potential scam. Look at this message that claimed it came from PayPal. See if you can find the ten errors in the email that should set off alarm bells.

From: PayPal <paypal@customer-notices55.com>
To: PayPal user 551-121-998
Sent: Feb 1st 2021 @ 10:55
Subject: Compremised Account [CaseID Nr: KX-003-551-121-998]

Dear Customer

We need you help to resolve issue with account. We have temporarily stop account due to problem's.
Unusual account activity on PayPal account means action need be taken immediately. If your not sure this was you, an unauthorized user might be trying to access your accounts. Please to log in here to change your password:

LOG IN HERE

▲ **Figure 5.19** Sample scam email

Did you find all the errors? An email like this looks official but there are many clues that it didn't come from a legitimate company; such as, many spelling mistakes, grammatical errors and the domain name in the email address. An email like this should be regarded as phishing; by clicking on the 'LOG IN HERE' box, you will divulge passwords and other key information since you will be sent to a fake 'PayPal' website.

3 Misspelling of domain names in a link are very common errors found in emails sent by scammers and fraudsters. The authors of this book have seen these incorrect spellings:

www.gougle.com
www.amozon.com

This is known as **typo squatting** where names close to the genuine names are used to fool you.

4 Suspicious links; destination addresses should match the rest of the email. Look at this message that claims to be from Netflix:

NETFLIX

Failed subscription renewal notice.

Your bank has rejected our last attempt to collect your monthly subscription. Please click on the link below to review your billing details.Thank you.

CONTINUE>>

When you hover over this link you can see that the destination is a website with the address: http://billing.com/id1234121XA3

there is no mention anywhere of the company in this URL– very high chance it is a scam to try and collect your personal details

▲ **Figure 5.20** Second example of probable scam

5 Other errors to look out for are just plain spelling mistakes. Look at this address from TKMaxx; find the three errors:

http://www.tkmax.co.ie

» since the company involve online payments, it's very likely to use secure links therefore you would expect to see `https`
» the spelling of the company is incorrect
» it is more likely to see `.com` since they are a large company.

Firewalls

A firewall can be either software or hardware. It sits between the user's computer and an external network

▲ **Figure 5.21** Typical firewall set up

The main tasks carried out by a firewall include:

» To examine the 'traffic' between user's computer (or internal network) and a public network (for example, the internet)

» Checks whether incoming or outgoing data meets a given set of criteria » if the data fails the criteria, the firewall will block the 'traffic' and give the user (or network manager) a warning that there may be a security issue

» The firewall can be used to log all incoming and outgoing 'traffic' to allow later interrogation by the user (or network manager)

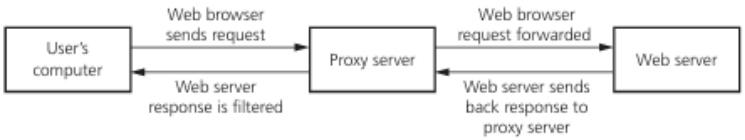
» Criteria can be set so that the firewall prevents access to certain undesirable sites; the firewall can keep a list of all undesirable IP addresses

» It is possible for firewalls to help prevent viruses or hackers entering the user's computer (or internal network)

» the user is warned if some software on their system is trying to access an external data source the user is given the option of allowing it to go ahead or request that such access is denied

Proxy servers

Proxy servers act as an intermediate between the user and a web server:



▲ **Figure 5.22** Proxy server

Features of proxy servers:

» allows internet traffic to be filtered; it is possible to block access to a website if necessary

» keeps users' IP addresses secret which improves security

» if the internet traffic is valid, access to the web server is allowed

» if the internet traffic is invalid, access to the web server is denied

» it is possible to block requests from certain IP addresses

» prevents direct access to a web server by sitting between the user and the web server

» if an attack is launched, it hits the proxy server instead – this helps to prevent hacking, DoS, and so on » used to direct invalid traffic away from web servers which gives additional protection

» by using the feature known as a cache, it is possible to speed up access to information/data from a website; when the website is first visited, the home page is stored on the proxy server; when the user next visits the website, it now comes from the proxy server cache instead, giving much faster access

» proxy servers can also act as firewalls.

Privacy settings

Privacy settings are the controls available on web browsers, social networks and other websites that are designed to limit who can access and see a user's personal profile. They were discussed earlier in the section on access rights. Privacy settings can refer to:

» a 'do not track' setting; the intention here is to stop websites collecting and using browsing data which leads to improved security

» a check to see if payment methods have been saved on websites; this is a useful safety feature which prevents the need to type in payment details again (every time you have type in financial details, there will be a risk of data interception)

» safer browsing; an alert is given when the browser encounters a potentially dangerous website (the undesirable website will be in a 'blacklist' stored on the user's computer)

» web browser privacy options (e.g. storing browsing history, storing cookies)

» website advertising opt-outs; a website may be tracked by any number of third parties who gather information about your browsing behaviour for advertising purposes

» apps; for instance, the sharing of location data in map apps can be switched off.

Secure sockets layer (SSL)

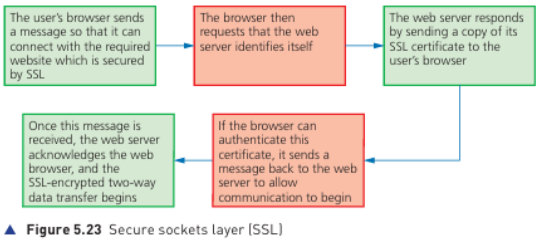
Secure Sockets Layer (SSL) is a type of protocol – a set of rules used by computers to communicate with each other across a network. This allows data to be sent and received securely over the internet.

The term SSL certificate was mentioned in Figure 5.23. An SSL certificate is a form of digital certificate which is used to authenticate a website. This means any communication or data exchange between browser and website is secure provided this certificate can be authenticated.

The address window in the browser when https protocol is being applied, rather than just http protocol, is quite different:

using https:	secure	https://www.xxxx.org/documents
using http:		http://www.yyyy.co.uk/documents

Figure 5.23 shows what happens when a user wants to access a secure website and receive and send data to it:



▲ **Figure 5.23** Secure sockets layer (SSL)

Key terms used throughout this chapter

internet – the world-wide interconnection of networks; the internet makes use of TCP and IP protocols

World Wide Web – a massive collection of web pages and is based on hypertext transfer protocols (http and https)

(web) browser – software that connects to a domain name server (DNS) to locate IP addresses; a browser interprets HTML web pages sent to a user's computer so that the user can read documents and watch multimedia

hypertext mark-up language (HTML) – the language used to design, display and format web pages, and to write http(s) protocols

uniform resource locator (URL) – a text-based address for a web page

hypertext transfer protocol secure (https) – http with extra security (such as SSL) applied

hyperlink – highlighted text or an image that is activated by clicking and links to further text, images, a web page or a website

domain name server (DNS) – a server that looks up domain names for websites (for example, www.hoddereducation.com) in order to find the IP addresses that a computer needs to locate the web servers (for example, 107.162.140.19)

cookie – a text file sent from a website to a user's browser; it is used to remember user preferences each time they visit the website

user preferences – settings or options stored in cookies that can remember customised web pages or indicate browsing history to target adverts

session cookie – a cookie that is stored temporarily on a computer; it is deleted when the browser is closed or the website session ends

persistent cookies – a cookie that is stored on the user's hard drive and only deleted when the expiry date is reached or the cookie is deleted by the user

virtual shopping basket – an area of memory in a website where items a user wishes to purchase are temporarily stored; items remain in the basket until payment is made or the session has ended

digital currency – currency (a system of money) that exists in electronic form only; it has no physical form and is essentially data on a database

cryptocurrency – a form of digital currency that uses a chain of decentralised computers to control and monitor transactions

cryptography – the protection of data/information by use of coding; it usually involves encryption and decryption

blockchain – a decentralised database where all transactions are stored; it consists of a number of interconnected computers but not a central server

timestamp – a digital record of the date and time that a data block is created in blockchain networks

proof-of-work – the algorithm used in blockchain networks to confirm a transaction and to produce new blocks to add to the chain; special users called miners complete and monitor transactions on the network for a reward

brute force attack – a 'trial and error' method used by cybercriminals to crack passwords by finding all possible combinations of letters, numbers and symbols until the password is found

word list – a text file containing a collection of words used in a brute force attack

data interception – an attempt to eavesdrop on a wired or wireless network transmission; cybercriminal often use packet sniffing or access point mapping / wardriving to intercept data

packet sniffing – a method used by a cybercriminal to examine data packets being sent over a network and to find the contents of a data packet, which are sent back to the cybercriminal

wardriving – using a laptop, antenna, GPS device and software to intercept Wi-Fi signals and illegally obtain data; sometimes called Access Point Mapping

wired equivalency privacy (WEP) encryption protocol security – an algorithm for wireless networks to protect them against data interception

denial of service (DoS) attack – a cyberattack in which cybercriminals seek to disrupt the normal operation of a website by flooding it with requests; also used to clog up a user's mailbox by sending out thousands of spam emails

distributed denial of service (DDoS) attack – a denial of service (DoS) attack in which the fake requests come from many different computers, which makes it harder to stop

spam – unsolicited emails sent to a user's mailbox

hacking – the act of gaining illegal access to a computer system without the owner's permission

malware – programs (such as viruses, worms and Trojan horses) installed on a user's computer with the aim of deleting, corrupting or manipulating data illegally

virus – a program or program code that replicates itself with the intention of deleting or corrupting files or by causing the computer system to malfunction

active host – functioning software that a virus can affect by attaching itself to the code or by altering the code to allow the virus to carry out its attack

worm – a stand-alone type of malware that can self-replicate; unlike viruses, worms don't need an active host; they can spread throughout a network without the need for any action by an end-user

Trojan horse – a type of malware that is designed to look like legitimate software but contains malicious code that can cause damage to a computer system

spyware – a type of malware that gathers information by monitoring a user's activities on a computer and sends the gathered information back to the cybercriminal who sent out the spyware

adware – a type of malware that attempts to flood the end-user with unwanted advertising

ransomware – a type of malware that encrypts data on a user's computer and 'holds the data hostage' until a ransom is paid

phishing – sending out legitimate-looking emails designed to trick the recipients into giving their personal details to the sender of the email

spear phishing – similar to phishing but targeting specific people or organisations rather than carrying out a blanket attack

pharming – redirecting a user to a fake website in order to illegally obtain personal data about the user without their knowledge; unlike phishing, pharming is initiated without needing any action by the user

DNS cache poisoning – altering IP addresses on a domain name server (DNS) with the intention of redirecting a user's browser to a fake website; carried out by a pharmer (see pharming) or hacker (see hacking)

social engineering – manipulating people into breaking normal security procedures (such as giving away their password) in order to gain illegal access to computer systems or to place malware on their computer

access levels – different levels of access in a computer system allowing a hierarchy of access levels depending on user's level of security

anti-spyware – software that detects and removes spyware programs installed on a system; the software is based on typical spyware rules or known file structures

authentication – the process of proving a user's identity by using something they know, something they have or something unique to them

biometrics – type of authentication that uses a unique human characteristic, such as fingerprints, voice or retina blood vessel pattern

two-step verification – a type of authentication that requires two methods of verification to prove the identity of a user

patch – an update for software that is developed to improve the software and/or to remove any bugs

typo squatting – the use by cybercriminals of subtle spelling errors in website addresses used to trick users into visiting their fake websites

firewall – software or hardware that sits between a computer and an external network (for example, the internet); the firewall monitors and filters all incoming and outgoing traffic

proxy server – a server that acts as an intermediary server through which internet requests are processed; it often makes use of cache memory to speed up web page access

privacy settings – controls available on social networking and other websites which allow users to limit who can access their profile or what they are allowed to see

secure sockets layer (SSL) – a security protocol used when sending data over a network (such as the internet)

SSL certificate – a form of digital certificate which is used to authenticate a website; providing the SSL certificate can be authenticated, any communication or data exchange between browser and website is secure

Revision questions

1. (a) Four statements about cookies are shown in the table below. Study each statement. Tick (✓) to show whether the statement is true or false.

Statement	True	False
they are a form of spyware		
they are used only in advertising		
they are used to track browser use		
they act in the same way as a virus		

b) Five descriptions and five security issues are shown below. Draw a line to connect each description to the correct security issue.

Description	Security issue
malicious code installed on the hard drive of a user's computer or on the web server; this code will re-direct user to a fake web site without their consent	hacking
software that gathers information by monitoring key presses on a user's computer and relays the information back to the person who sent the software	pharming
program or code that replicates itself and is designed to amend, delete or copy data and files on a user's computer without their consent	phishing
the act of gaining illegal access to a computer system without the owner's consent	spyware
creator of code sends out a legitimate-looking email in the hope of gathering personal and financial data; it requires the recipient to follow a link in the email or open an attachment	virus

2. Six security issues and six descriptions are shown below. Draw a line to link each security issue to its correct description.

Security issue	Description
Pharming	illegal access to a computer system without the owner's consent or knowledge
Phishing	software that gathers information by monitoring key presses on a user's keyboard; the data is sent back to the originator of the software
Viruses	malicious code installed on the hard drive of a user's computer or on a web server; this code will re-direct the user to a fake website without the user's knowledge
Hacking	creator of code sends out a legitimate-looking email in the hope of gathering personal and financial information from the recipient; it requires the user to click on the link in the email or attachment
Spyware	a message given to a web browser by a web server; it is stored in a text file; the message is then sent back to the server each time the browser requests a page from the server
Cookies	program or code that replicates itself; designed to amend, delete or copy data or files on a user's computer; often causes the computer to crash or run slowly

3. (a) Check digits are used to ensure the accuracy of input data. A 7-digit code number has an extra digit on the right, called the check digit

Digit position	1	2	3	4	5	6	7	8
Digit	–	–	–	–	–	–	–	–

The check digit is calculated as follows:

• each digit in the number is multiplied by its digit position

• the seven results are then added together

• this total is divided by 11

• the remainder gives the check digit (if the remainder = 10, the check digit is X) (i) Calculate the check digit for the following code number.

Show all your working.

4 2 4 1 5 0 8 …

(ii) An operator has just keyed in the following code number:
3 2 4 0 0 4 5 X
Has the operator correctly keyed in the code number?

Give a reason for your answer.

b) When data are transmitted from one device to another, a parity check is often carried out on each byte of data. The parity bit is often the leftmost bit in the byte.

(i) If a system uses even parity, give the parity bit for each of the following bytes:

parity bit							
	1	1	0	0	1	1	0

parity bit							
	0	0	0	0	0	0	1

(ii) A parity check can often detect corruption of a byte. Describe a situation in which it cannot detect corruption of a byte.

4. (a) Computer ethics involves a number of different topics.

(i) A student made the following statement on an examination paper: "It allows a user to have the freedom to run, copy, change and adapt the software and then pass it on to a colleague, friend or family member." Identify which computer term the student was describing.

(ii) Explain what is meant by computer ethics

(b) The four statements below refer to firewalls and proxy servers. Study each statement. Tick the appropriate column(s) to indicate whether the statement refers to a firewall and/or a proxy server.

Statement	Firewall	Proxy server
Speeds up access of information from a web server by using a cache		
Filters all Internet traffic coming into and out from a user's computer, intranet or private network		
Helps to prevent malware, including viruses, from entering a user's computer		
Keeps a list of undesirable websites and IP addresses		

(c) Explain three ways of preventing accidental loss or corruption of data.

5. a) One of the key features of von Neumann computer architecture is the use of buses. Three buses and three descriptions are shown below. Draw a line to connect each bus to its correct description.

Bus	Description
address bus	this bus carries signals used to coordinate the computer's activities
control bus	this bi-directional bus is used to exchange data between processor, memory and input/ output devices
data bus	this uni-directional bus carries signals relating to memory addresses between processor and memory

(b) The seven stages in a von Neumann fetch-execute cycle are shown in the table below. Put each stage in the correct sequence by writing the numbers 1 to 7 in the right hand column. The first one has been done for you.

Stage	Sequence number
the instruction is then copied from the memory location contained in the MAR (memory address register) and is placed in the MDR (memory data register)	
the instruction is finally decoded and is then executed	
the PC (program counter) contains the address of the next instruction to be fetched	1
the entire instruction is then copied from the MDR (memory data register) and placed in the CIR (current instruction register)	
the address contained in the PC (program counter) is copied to the MAR (memory address register) via the address bus	
the address part of the instruction, if any, is placed in the MAR (memory address register)	
the value in the PC (program counter) is then incremented so that it points to the next instruction to be fetched	

6. Five computing terms are described below.

Write the name of the term being described.

Software that anyone can download for free from the Internet and then use without having to pay any fees. The usual copyright laws apply and a user license is important.

Software that gives the user the chance to try it out free of charge before actually buying it. The software is subject to the usual copyright laws. As a rule, not all the features found in the full version are available at this stage.

Software where users have freedom to run, copy, change and adapt it. This is an issue of liberty and not of price since the software guarantees freedom and the right to study and modify the software by having access to the actual source code.

Set of principles that regulates the use of computers in everyday life. This covers intellectual property rights, privacy issues and the effects of computers on society in general. The taking of somebody's idea or software and claim that the idea or software code were created by the "taker"

7. Six security issues and six descriptions are shown below. Draw a line to link each security issue to its correct description.

Security issue	Description
Pharming	illegal access to a computer system without the owner's consent or knowledge
Phishing	software that gathers information by monitoring key presses on a user's keyboard; the data is sent back to the originator of the software
Viruses	malicious code installed on the hard drive of a user's computer or on a web server; this code will re-direct the user to a fake website without the user's knowledge
Hacking	creator of code sends out a legitimate-looking email in the hope of gathering personal and financial information from the recipient; it requires the user to click on the link in the email or attachment
Spyware	a message given to a web browser by a web server; it is stored in a text file; the message is then sent back to the server each time the browser requests a page from the server
Cookies	program or code that replicates itself; designed to amend, delete or copy data or files on a user's computer; often causes the computer to crash or run slowly

8. (a) Five statements and three types of software are shown below. Draw lines to connect each statement with the correct type of software

(b) Describe three ethical issues that should be considered when using computers

(c) Security of data is very important. Three security issues are viruses, pharming and spyware. Explain what is meant by each issue

(d) Describe three tasks carried out by a firewall.

9. A company has a number of offices around the world.

(a) Data is transmitted between the offices over the Internet. In order to keep the data safe the company is using Secure Socket Layer (SSL) protocol and a firewall at each office. Explain how SSL protocol and a firewall will keep the company's data safe

(b) A company stores personal details of its customers on a computer system behind a firewall. Explain, with reasons, what else the company should do to keep this data safe.

c) The supermarket uses secondary storage and off-line storage to store data about its stock. Explain what is meant by secondary storage and off-line storage.

10. A business wants to use a biometric security system to control entry to the office. The system will use a biometric device and a microprocessor. Explain how the biometric security system will make use of the biometric device and the microprocessor to control entry to the office.