Cambridge IGCSE

Computer Science

Section 1

Encryption

Unit 2:
Data Transmission

Objectives

- Understand the need for and purpose of encryption when transmitting data
- Understand how data is encrypted using symmetric and asymmetric encryption

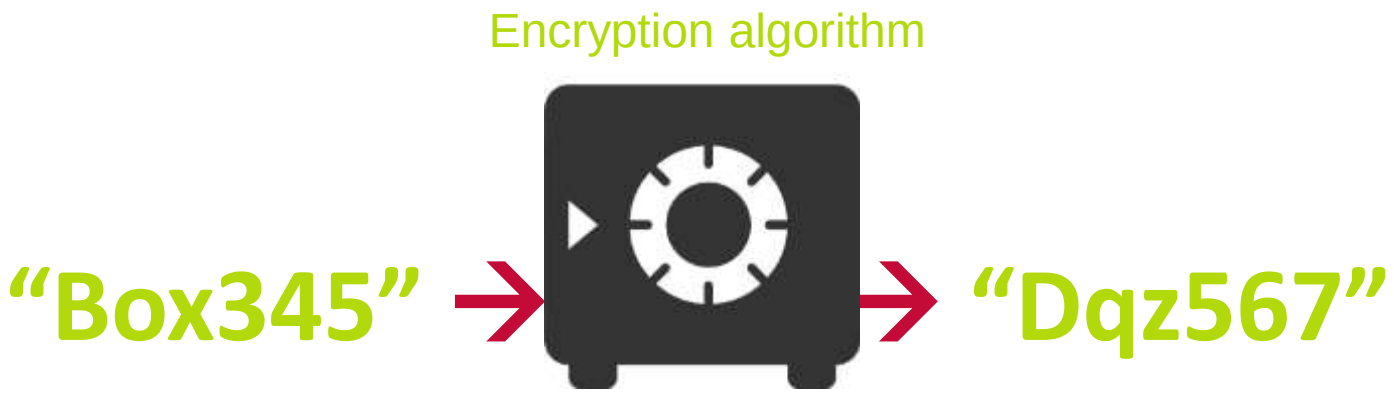
Intercepting data

- Data that is transmitted over a network can be intercepted
  - Any intercepted data can be read and understood unless measures are taken to prevent it from being interpreted
  - These measures are known as **encryption**

Encryption (加密)

- Data that is transmitted over a network can be intercepted - 拦截
- Encryption is the encoding of data so that it cannot be easily understood by someone who intercepts it

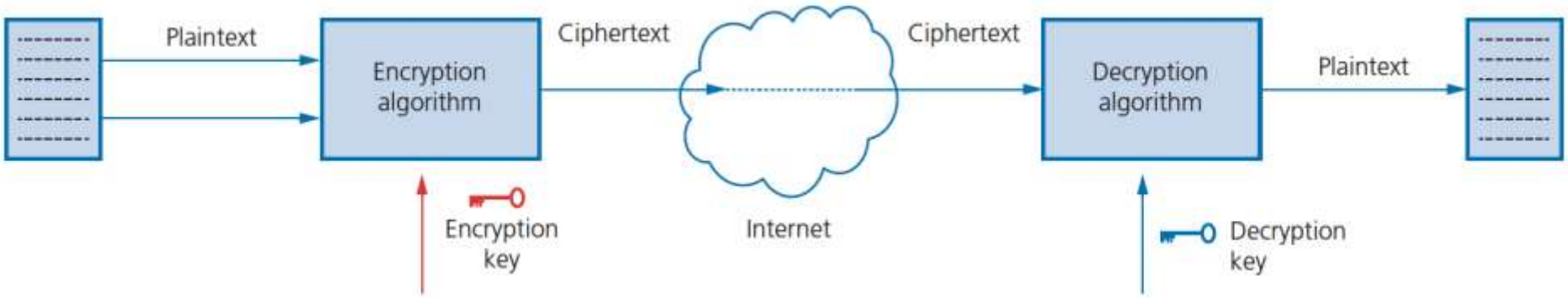
- A simple shift cipher might encode "*Box345*" as follows:



Encryption (加密)

- Encryption is needed in order to **keep secure the data we transmit between computers**
- Encryption does not stop an unauthorised user (hacker) from getting the data, but the encrypted data does **make it difficult for them to understand what it means**
- Encryption works by using a **key** to encrypt (加密) and decrypt (解密) the data

Encryption (加密)



- **Plaintext**: the original message to be encrypted
- **Ciphertext**: the encrypted message
- **Encryption**: the process of converting plaintext into ciphertext
- **Key**: a sequence of numbers used to encrypt or decrypt, often data using a mathematical formula
- **Encryption algorithm**: the formula for encrypting the plaintext - two inputs: **plaintext** and a **secret key**

Encryption terminology

- **Plaintext**: the original message to be encrypted
- **Ciphertext**: the encrypted message
- **Encryption**: the process of converting plaintext into ciphertext
- **Key**: a sequence of numbers used to encrypt or decrypt, often data using a mathematical formula
- **Encryption algorithm**: the formula for encrypting the plaintext - two inputs: **plaintext** and a **secret key**

Encryption techniques

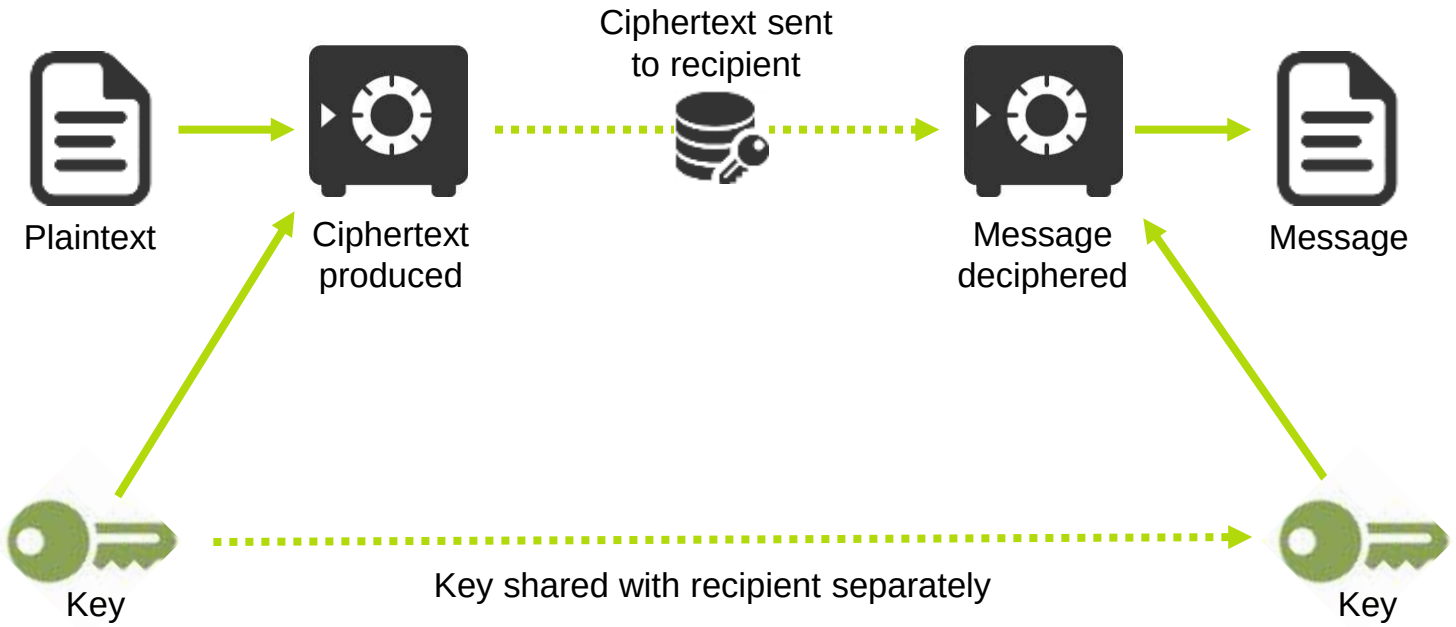
- **Symmetric encryption**
  - A **single key** is used to **encrypt and decrypt** a message and must be given to the recipient of your message to decrypt the data

- **Asymmetric encryption**
  - **Two keys** are used - one to **encrypt** and the other to **decrypt** data
  - This is more secure as it means that you never have to send or reveal your decryption key

Encryption (加密)

Symmetric encryption

- **Same key used to encrypt and decrypt a message**



Symmetric encryption

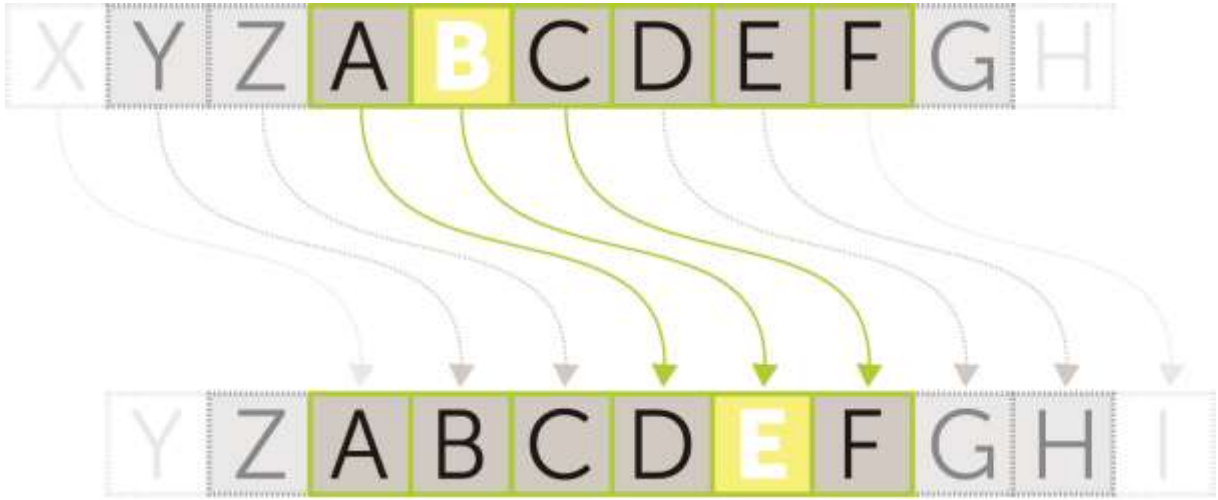
- **Same key used to encrypt and decrypt a message**
- This means if two users both have the key they can share the message.
- The problem is if anyone else gets this key they can also access the message.
- Messages can be intercepted when sent over a network, and if they also have the key it makes the encryption useless.
- As a single encryption key is required for both sender and recipient, **security** is the main problem for symmetric encryption.

The Caesar shift cipher

- The earliest known substitution cipher was invented by Julius Caesar
  - Each letter is replaced by the letter *n* positions further on in the alphabet
  - *n* is the key and is used to encrypt and decrypt the message
  - This is an example of **symmetric encryption**

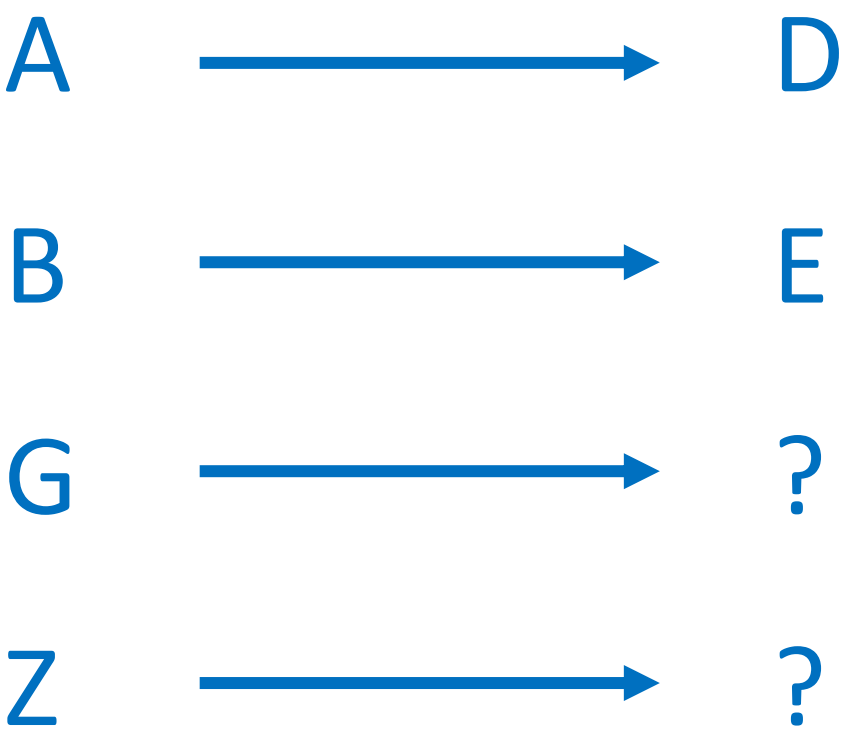
Caesar cipher

- The Caesar cipher is most basic type of encryption and the most insecure
- Letters of the alphabet are shifted by a given number
- **Example**: Each character in a message is replaced by the character 3 places ahead of it in the alphabet



Caesar cipher

- **Example**: Each character in a message is replaced by the character 3 places ahead of it in the alphabet



Deciphering the code

- **Key = Shift → 3**
- Decode "**DWWDFN DW GDZQ**"

- Using a different key, crack the code for "**PCRPCYR**"

Deciphering the code

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

Decode the following message:

FURVV WKH ULYHU WRQLJKW

Deciphering the code

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

It should say:

CROSS THE RIVER TONIGHT

Key strength

Example

Plaintext into ciphertext using 10-digit encryption key.

This time, the encryption key is **4 2 9 1 3 6 2 8 5 6**.

So, the message ‘COMPUTER SCIENCE IS EXCITING’ becomes ...

C	O	M	P	U	T	E	R
4	2	9	1	3	6	2	8
G	Q	V	Q	X	Z	G	Z

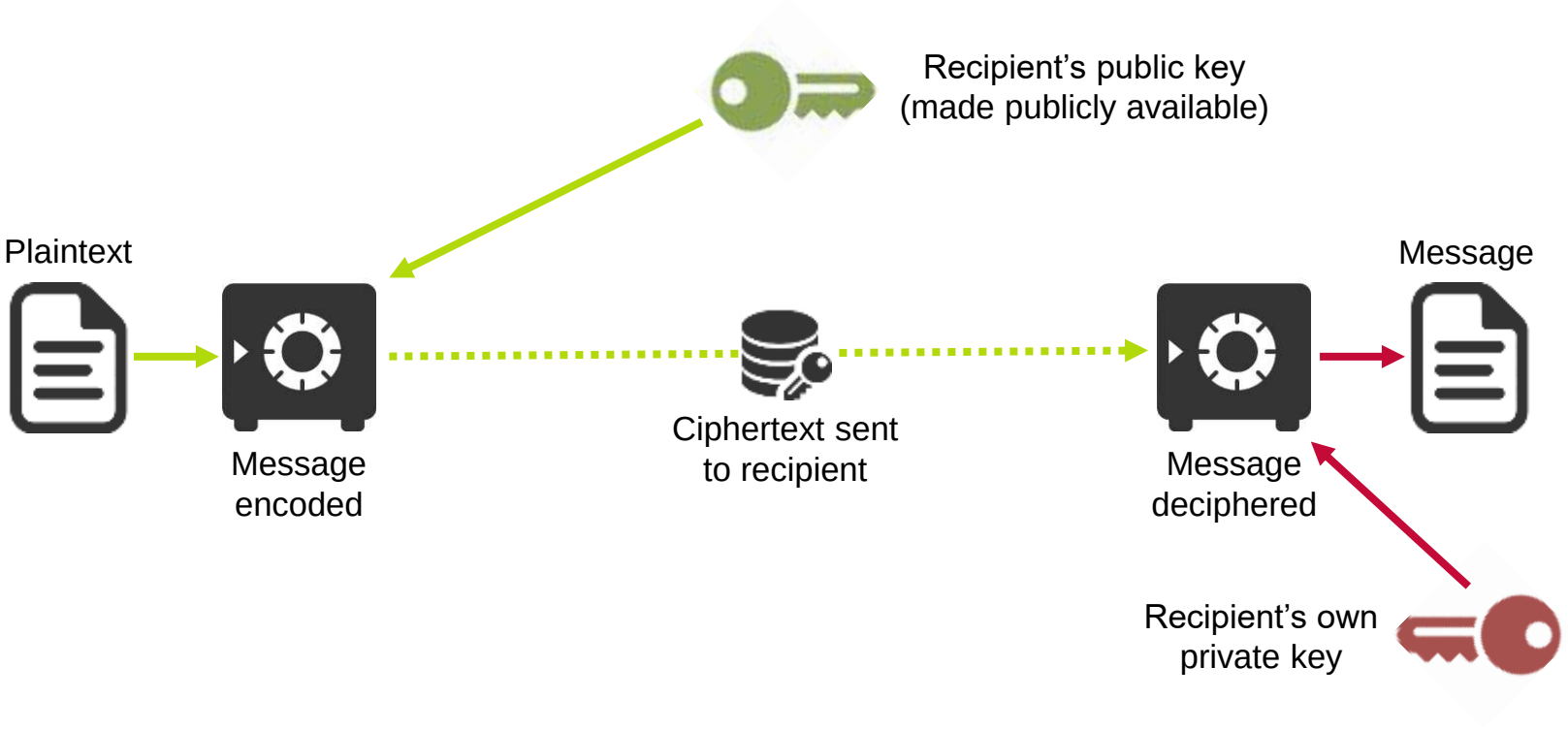
S	C	I	E	N	C	E
5	6	4	2	9	1	3
X	I	M	G	W	D	H

I	S
6	2
O	U

E	X	C	I	T	I	N	G
8	5	6	4	2	9	1	3
M	C	I	M	V	R	O	J

Asymmetric encryption

- **Two keys**! A **public key** known to everyone for **encrypting** and a **private, secret key for decrypting**

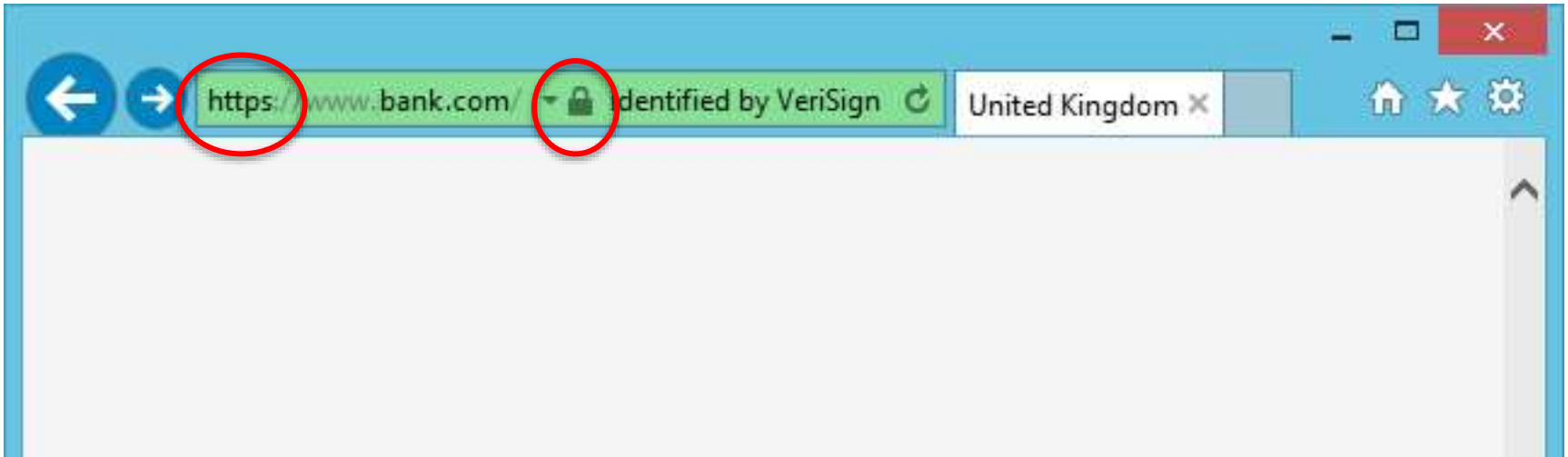


Asymmetric encryption

- **Two keys**! A **public key** known to everyone for **encrypting** and a **private, secret key for decrypting**
- The first key - the **public key** - is used to encrypt the message, it cannot decrypt it. It can be sent to anyone the user wants to receive a message from.
- The second key is kept secret by the user. It is called the **private key**. It is used to decrypt the message.
- The source of a message can be trusted as they need to know the encryption keys to be able to communicate.

Asymmetric encryption

- **Two keys**! A **public key** known to everyone for **encrypting** and a **private, secret key for decrypting**
- **SSL** (Secure Socket Layer) is a protocol for transmitting private documents over the Internet
- It uses asymmetric (public key) encryption to encrypt data before transmission

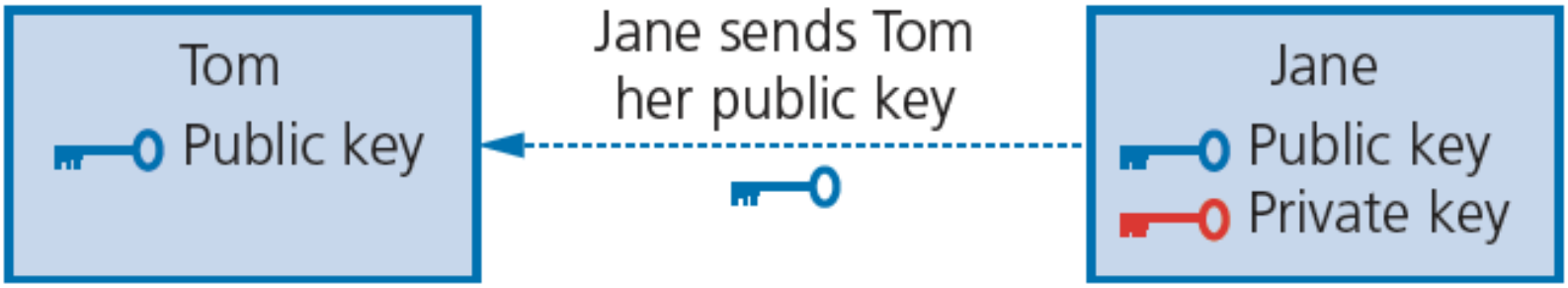


Asymmetric encryption

Example: Tom and Jane work for the same company. Tom wishes to send a confidential document to Jane.

1 Jane uses an algorithm to generate a matching pair of keys (private and public) that they must be keep stored on their computers. The keys are mathematically linked, but one cannot be worked out from the other.

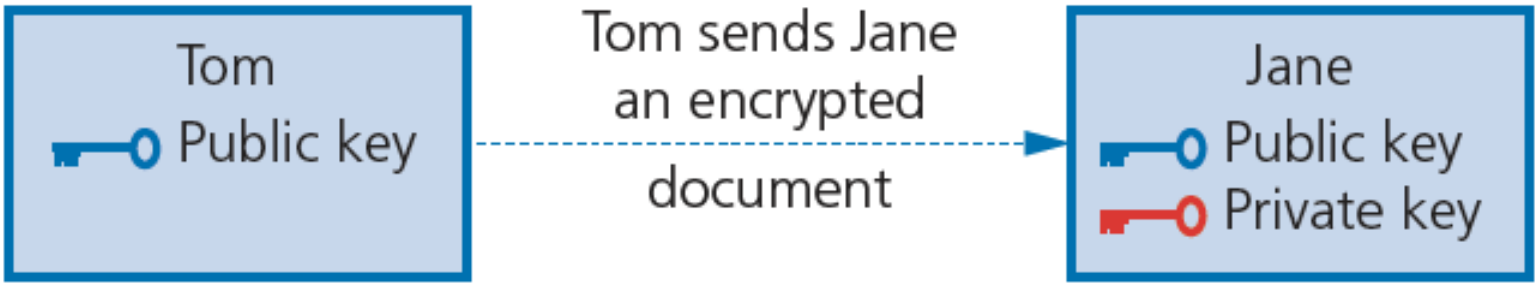
2 Jane sends Tom her public key.



Asymmetric encryption

Example: Tom and Jane work for the same company. Tom wishes to send a confidential document to Jane.

3. Tom now uses Jane's public key to encrypt the document he wishes to send her. He then sends his encrypted document (ciphertext) back to Jane.



Asymmetric encryption

Example: Tom and Jane work for the same company. Tom wishes to send a confidential document to Jane.

4. Jane uses her matching private key () to unlock Tom's document and decrypt it.

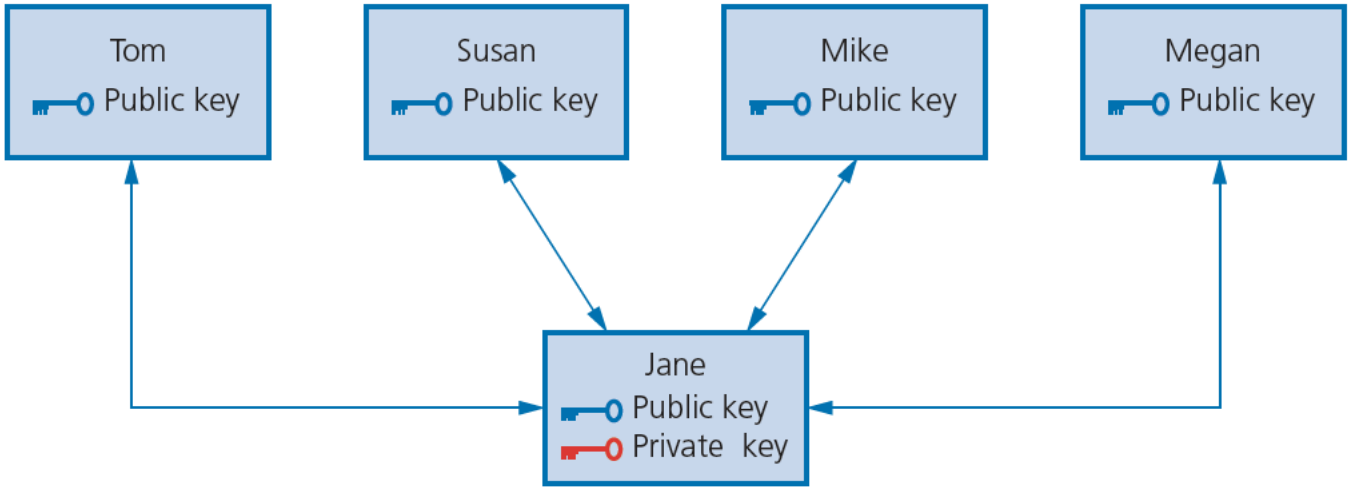
   Remember, the public key used to encrypt and the private key used to decrypt are a matching pair generated on Jane's computer.

   She cannot use the public key to decrypt the message.

   She can also send her public key to other people working in her company, so that she is able to receive encrypted message from them, and decrypt them using her private key.

Asymmetric encryption

Example: Tom and Jane work for the same company. Tom wishes to send a confidential document to Jane.



She can also send her public key to other people working in her company, so that she is able to receive encrypted message from them, and decrypt them using her private key.

What would need to happen for two-way communication between all five workers of encrypted documents?

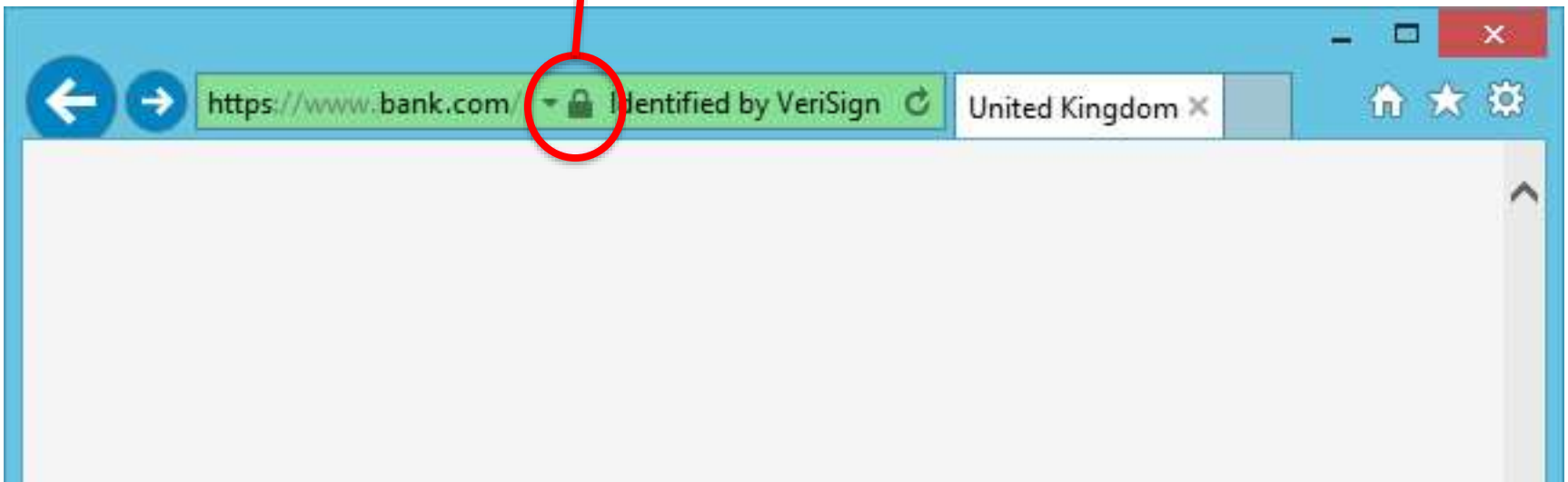
Brute-force attack

- Every possible key is tried
- On average, half of all possible keys need to be tried, so the longer the key, the more time it takes to find it
- The Caeser Cipher uses a shift of between 1 and 26 assuming only uppercase letters of the alphabet are used
- This means that there are 26 possible keys requiring an average of about 13 tries to crack the code

Key strength:
- Increasing the number of bits used for a key, increases the strength of the encryption

https:

- The '**s**' in **https** stands for '**s**ecure'
- This means that a security protocol is used to ensure a trusted, encrypted data connection
- Why do online banking and shopping websites have a padlock symbol?



Secure Sockets Layer (SSL)

- **SSL** is a protocol for transmitting private documents via the Internet
- It uses asymmetric (public key) encryption to encrypt data before transmission
  - Digital certificates are sent to the browser containing the website's public key
  - This is used to authenticate the website before any data or transactions are passed
- Many websites use SSL to receive confidential information, including credit card details
  - URLs that require an SSL connection start with **https:**

Transport Layer Security (TLS)

- **TLS** is a more recent upgrade to **SSL**, used with newer browsers, first defined in 1999
- The latest version of the protocol, TLS 1.3, was implemented in August 2018
- Like SSL, it uses asymmetric encryption
    - In July 2013, Google announced it would change the size of its public/private encryption keys from 1024 bit to 2048 bit keys to increase the security of the TLS encryption it provides to its users

SSL or TLS handshake

- When a user logs on to a secure site (identified by a URL beginning with **https:**)
  - A “handshake” procedure enables the SSL or TLS client and server to establish the secret keys with which they communicate
  - Asymmetric encryption techniques generate a secret shared key
  - SSL or TLS then uses the shared key for the symmetric encryption of messages, which is faster than asymmetric encryption

Plenary

- Encryption is the process of encoding a message such that only authorised users can understand the message.
- When the message is received by an unauthorised user, it is unintelligible as they cannot decode it.
- The Caesar Cipher algorithm is one such example in which the letters in a message are replaced by letters in the alphabet that are one or more positions away.
- For example: encrypting the word ‘computer’ with letters that are +3 positions away becomes ‘frpsxwhu’.

Plenary

- Symmetric encryption uses the same key to encrypt and decrypt a message
- Asymmetric encryption includes the use of public and private keys
- The longer the key, the stronger the encryption
- The complexity of the encryption algorithm also affects the strength of the encryption
- https is the secured form of the http webpage. Inputs from the user are encrypted to offer a secure online experience such as banking, shopping, etc.
- SSL and TLS protocols both use asymmetric encryption

Vocabulary

- **encryption**
- **symmetric**
- **asymmetric**
- **encoding**
- **cipher**
- **key**
- **intercept**
- **authorised**
- **network**
- **encrypt**
- **decrypt**

Past paper question examples ...

(b) The data stored by the library is archived at the end of each day. The archive is held on a server in the library office.

The data is encrypted with an 8-bit key. As some of the data is confidential, the library wants to make the encryption more secure.

(i) State how the library could make the encryption more secure.

.....

..... [1]

(ii) The term used to describe data before it is encrypted is plain text.

State the term used to describe encrypted data.

..... [1]

Past paper question examples ...

(b) The finance company is also worried about the security of the data stored on its servers.

The company has decided to encrypt the data to improve the security.

Describe how the data are encrypted.

.....

.....

.....

.....

.....

.....

.....

.....

..... [4]

Key strength

- 5 bits would enable 2^5 key combinations; one for each letter (plus a few spare) so it could be said that Caesar was using 5-bit encryption
- Increasing the number of bits used for a key, increases the strength of the encryption

Key Size (bits)	Number of Alternative Keys	Time required at 10^6 decryptions/microsecond
32	$2^{32} = 4.3 \times 10^9$	2.15 milliseconds
56	$2^{56} = 7.2 \times 10^{16}$	10.01 hours
128	$2^{128} = 3.4 \times 10^{38}$	5.4×10^{18} years

Cryptanalysis

- The objective of cryptanalysis is to decode the ciphertext - typically by finding out the secret key
- How did you find the key for "**RETREAT**"?
- There are two general approaches:
  - Brute-force attack
  - Non-brute-force attack (cryptanalytic attack)

Non brute-force attack

- At Bletchley Park in England, during World War II, a team of code breakers decoded German ciphers. The cipher key was changed every day.
- One important clue was that some messages always started with the words *"Heil Hitler"*
- Other messages often ended with *"Nieder mit die Englander"* (Down with the English)
  - This gave the code-breakers a starting point!

Frequency analysis

- Letters are not used equally often
- In English, **E** is by far the most common letter, followed by **T**, **A**, **O**, **I**, **N**, **S**, **R**, then **H**
- Other letters like **Z**, **J**, **K**, **Q**, **X** are fairly rare
- In Czech, the letter **Z** is only worth 4 points in Scrabble. It's worth 10 in the English version

Modern ciphers

- Modern ciphers are created using two very large prime numbers multiplied together
- The larger the prime number, the more difficult it is find the two numbers needed to break the code
- The largest known prime number is $2^{57,885,161} - 1$, with 17,425,170 digits!
- Calculate the two prime factors of the following:
  - 15?
  - 91?
  - 143?

Algorithmic security

- Ciphers are based on computational security
  - The keys are determined using a computer algorithm
  - A key derived from an algorithm, can also be unpicked
  - Given enough ciphertext, computer power and time, any key can be determined and the message cracked

Strong and weak encryption

- These terms are relative, but:
  - Encryption can be considered to be ‘strong’ when the useful lifetime of the encrypted data is less than the time taken to break the code
  - With weak encryption, the code may be broken in time to use the information, but it wouldn’t be worth the effort trying
- Some governments have banned encryption above a certain strength. Why?

Worksheet

- Try **Task 1**