Cryptography

A Short History

Steganography

- The secret message is hidden (on a messenger).
- Greeks
  - Message written on a shaved head.
  - Hair let grow.
  - And shaved when the messanger arrived.
  - Required a certain lack of urgency!

Steganography (cont)

- Invisible ink.
- Known to Greeks in first century AD.
- Milk of a plant.
- When the parchment was heated, the writing reappeared.

Steganography (cont)

- Second World War
  - The microdot.
  - Photographically shrink a page of text down to the size of a dot and hide it on top of a full stop.

Steganography (cont)

- The problem is that if the message is discovered, the message is revealed.

Cryptography

- The aim is to hide the meaning of a message (rather than hide the existence of the message.)
- Two types
  - Transposition
  - Substitution (the one used)

Transposition

- The letters are rearranged.
- Effectively generating an anagram.
- "For example consider this short sentence".
- 35 characters.
- Number of arrangements of 35 different characters is 35!
- 35! is approximately 1×10^{40}

Transposition

- In fact there are less than that as some letters are repeated, but the number of arrangements is still very large.
- The problem is the algorithm needed to recover the original text.
- Generally not used.

Substitution - Caesar Shift Key

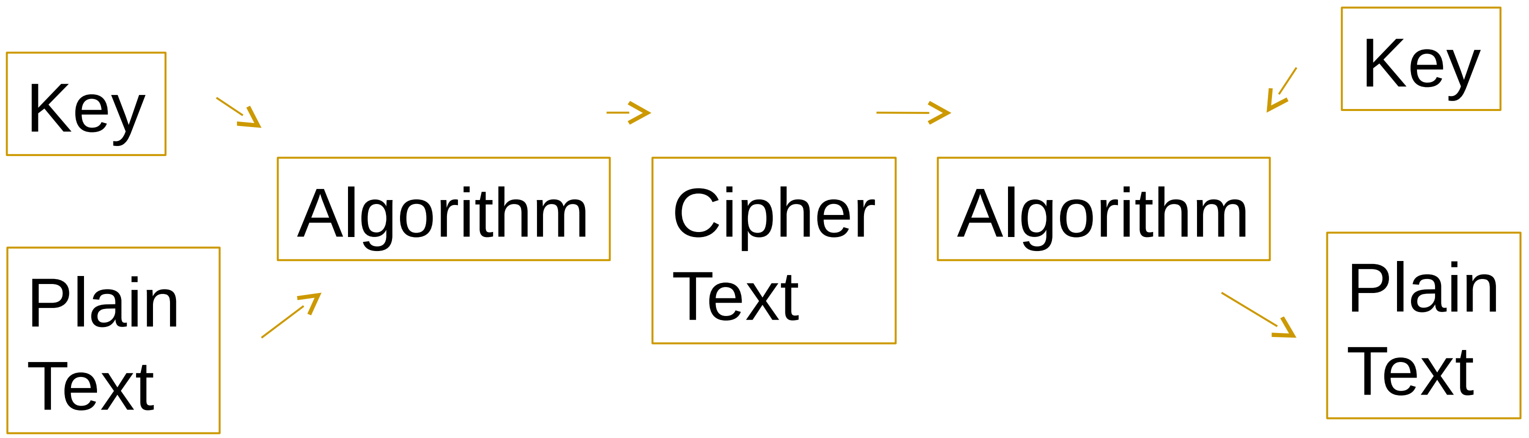
- Replace each letter by for example, the letter three places down in the alphabet.
- This allows for 25 different substitutions.

General Substitution

- First letter can be replaced by one of 25 other letters.
- The second by 24 etc. etc.
- $25! = 1.5 \times 10^{25}$.

Substitution

- Now see that in fact we have



The Key as a Number

- 120321…
- A replaced by L
- B replaced by C
- C replaced by U etc. etc.

Kerckhoff's Principle (1883)

- "The security of a crypto-system must not depend on keeping secret the crypto-algorithm. The security depends on keeping secret the key."
- Modern cryptography algorithms work like this.

Cryptography – Size of Key Space

- For a key to be useful, there must a large number of them.
- (This is a necessary, but not a sufficient condition for a good cryptographic algorithm.)
- For example, Caesars Shift Key has 25 different keys. Not good.
- A general substitution has $1.5x10^{25}$
- Size of key space is a necessary but not sufficient condition for a good algorithm.
- It means that the algorithm is not going to be broken using an exaustive search.

Frequency Analysis

- First carried out in Arabia around 750AD.
- Realized that the frequency of letters in text could be used to easily break a substitution cipher.
- Each letter occurs with a known (within limits) frequency.
- E, t, and a are the most frequently used letters.

Letter Frequency

- are the most frequently used letters.
- e – 7
- t – 4
- a – 1
- r – 3
- u – 2

- But of course not always.

Other Patterns

- E can appear before and after most letters.
- T rarely appears before a lot of letter, b, d, g, k, m, q, v.
- Some letters are a lot more friendly than others.
- Some letters repeat. Others don't.
- A word always has a vowel.
- Words also have frequencies.

Other Patterns

- "The" and "and" are the most frequent three letter words in English.
- Frequently used text are a real gift for cryptanalysts.

The Black Chambers in Vienna

- About 18th century.
- All diplomatic messages into and out of embassies in Vienna were intercepted and copied.
- A group of cryptanalysts worked round the clock on decrypting these.
- 100 a day.
- Results used in Austria and sold to other countries.

Polyalphabetic Ciper

Polyalphabetic Ciper

- Instead of a single (monoalphabetic) substitution, multiple substitutions are used depending on where in the message a letter occurs.
- The main example is the Vignere cipher.

Vigenere Cipher

- a bc de f g hi
- ABCDEFGHI
- BCDEFGHI J
- CDEFGHI
- DEFGHIJ
- EFGHIJK
- FGHIJKL

- BAD B ADB AD (key)
-
-
- Had a bad da (plaintext)
- IAG B BDE DD (ciphertext)

Vigenere Cipher

- Key is repeated above the plain text.
- The key letter determines which row in square is to be used to encipher the plain text letter.

Vigenere Cipher

- Not susceptible to frequency analysis.
- Know as a poly-alphabetic cipher.
- For a long time was though to be unbreakable.

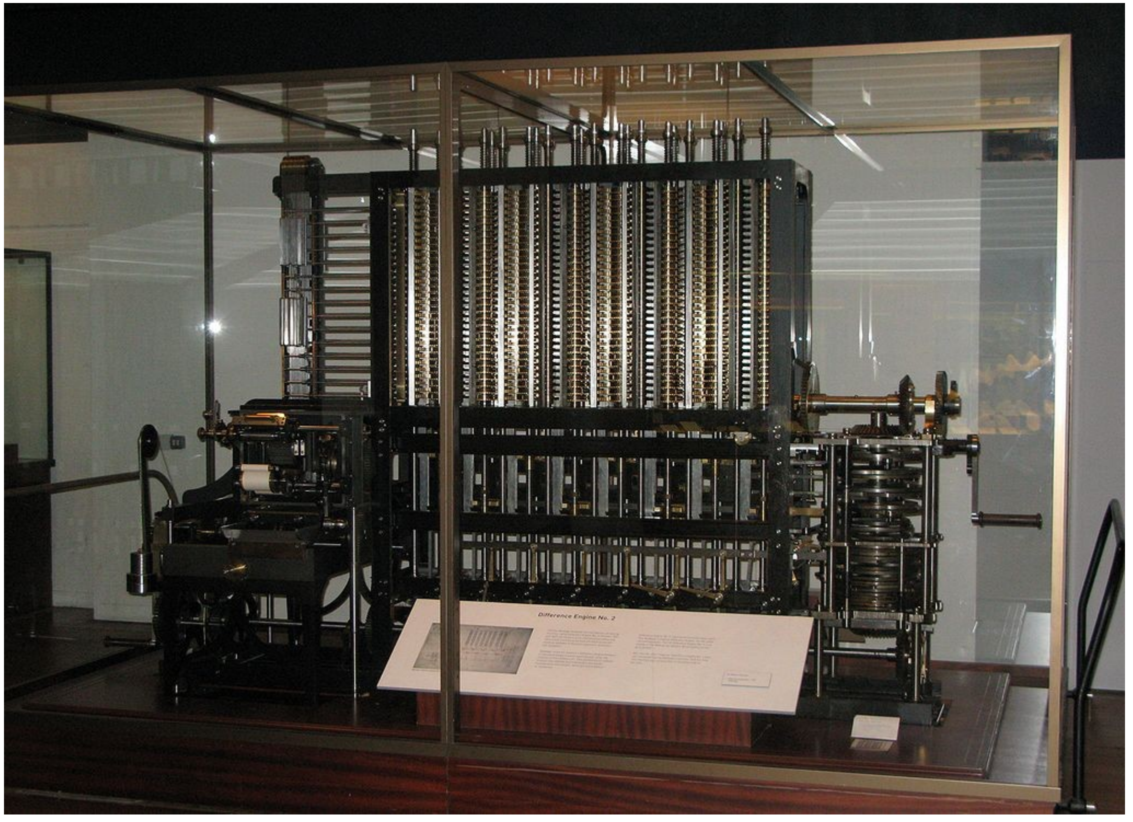
Babbage - Breaking of Vigenere Cipher

- 1854, Babbage broke the Vigenere Cipher.
- He used a weakness caused by the repetition of the keys.
- He looked for frequent words like "the" and using this analysis was able to regenerate the key.
- Working in intelligence for the government he never published solution.

Aside – Babbage's Difference Engine

- Mechanical Calculator
- Designed to calculate polynomial functions
- Polynomial functions can approximate most mathematical functions
- Because most mathematical tables at the time were calculated by hand and contained many errors.
- Caused many problems for engineering and science.

Babbage’s Difference Engine



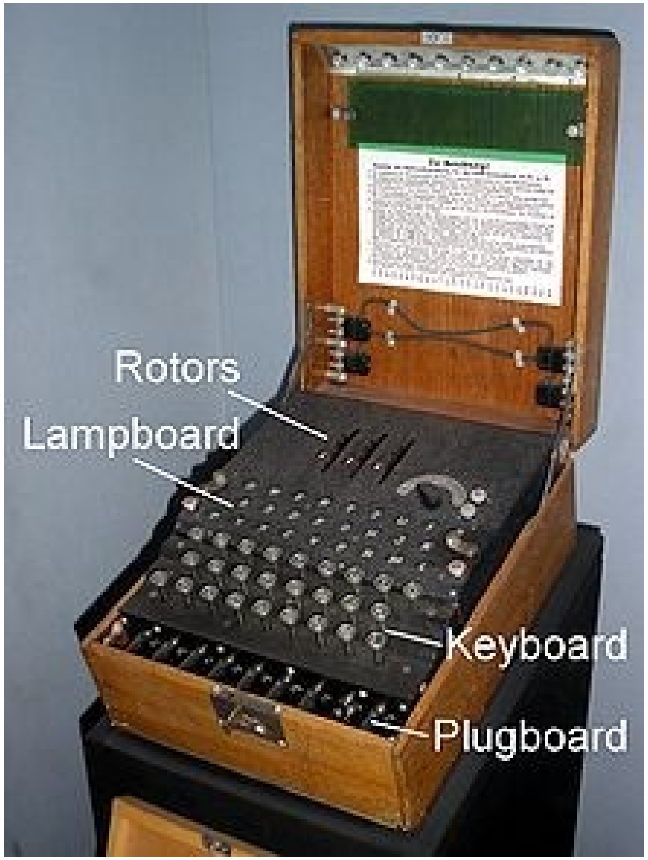
Kasiski - Breaking of Vigenere Cipher

- 1863, Kasiski independently broke the Vigenere Cipher.
- He published and the technique became known as the Kasiski test.

The Enigma Machine

Keyspace of the Enigma Machine

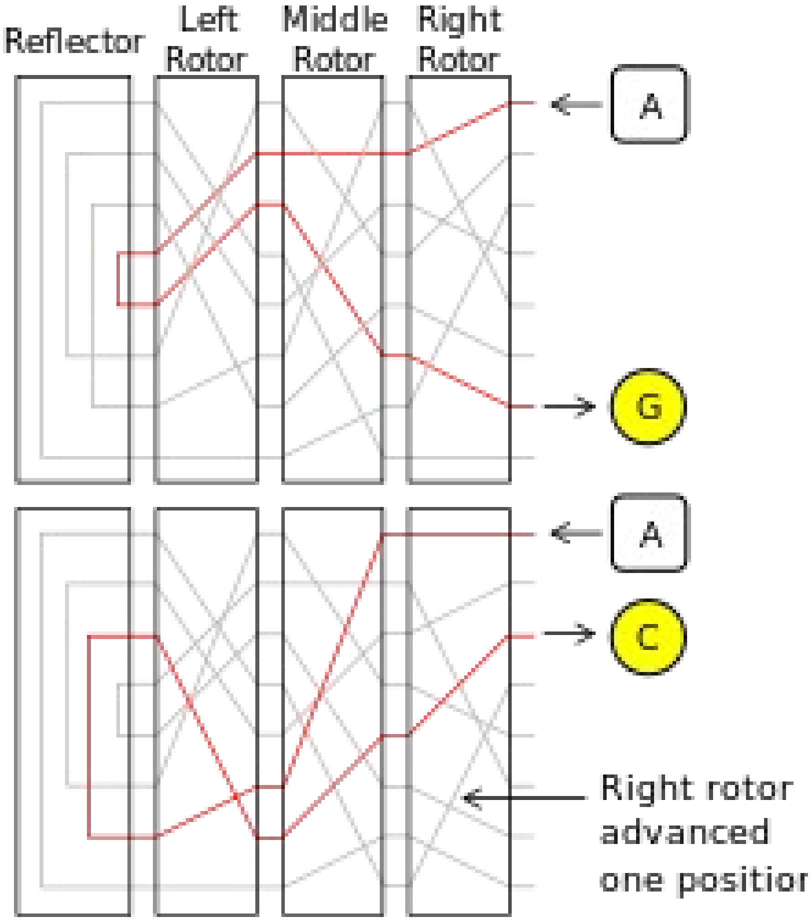
The Enigma Machine



The Enigma Machine

- Letters were keyed into the keyboard.
- Electrical signals passed through a number of rotors (scamblers), to a reflector and back through the rotors to the lampboard.
- This converted plaintext to ciphertext.
- On the receiving end the ciphertext was keyed in and the plaintext was recovered (lampboard).

Rotors


Reflector
Left Rotor
Middle Rotor
Right Rotor
A
G
A
C
Right rotor advanced one positior

Rotors

- After every letter entered, the first rotor advanced one.
- After 26 letters, the second also advanced one.
- Resulting in a poly-alphabetic cipher.

- The rotors (scramblers) were removable.

The PlugBoard

- Allowed the swapping of 6 letters (with another 6)

Example Key

- Plugboard
  - A/L P/R T/D K/F O/Y Z/B
- Scrambler (Rotor) arrangements
  - 2-3-1
- Scrambler (Rotor) orientations

Number of keys

- Scrambler arrangements
  - 3x2x1 = 6
- Scrambler orientations
  - 26x26x26 = 17,576
- Plugboard
  - 26x25x24x23x22x21x20x19x18x17x16x15/ 6
  - = 4,626,053,752,320,000 /6 = 7 x 10^{14}

Number of keys

- Large number of keys.
- Mainly because of plugboard settings.
- But plugboard settings couldn't be used on their own.
- That would give a mono-alphabetic substitution which could be broken using frequency analysis.
- Hence the need for the scrambler wheels.

Use of the Machine

- There was a day key for every day.
- The day key was in turn used to encrypt a three letter key for each message which gave the scrambler orientations.
- This was done twice and placed at the start of the message.
- For example
  - DWK-SHY

Use of the Machine (cont)

- Field operators choose these keys.
- They were transmitted twice to ensure there were no mistakes.
- In effect a different key was being used for each message sent.
- The scrambler arrangements and plugboard setting were not changed during the day, so only a part of the key was being changed.

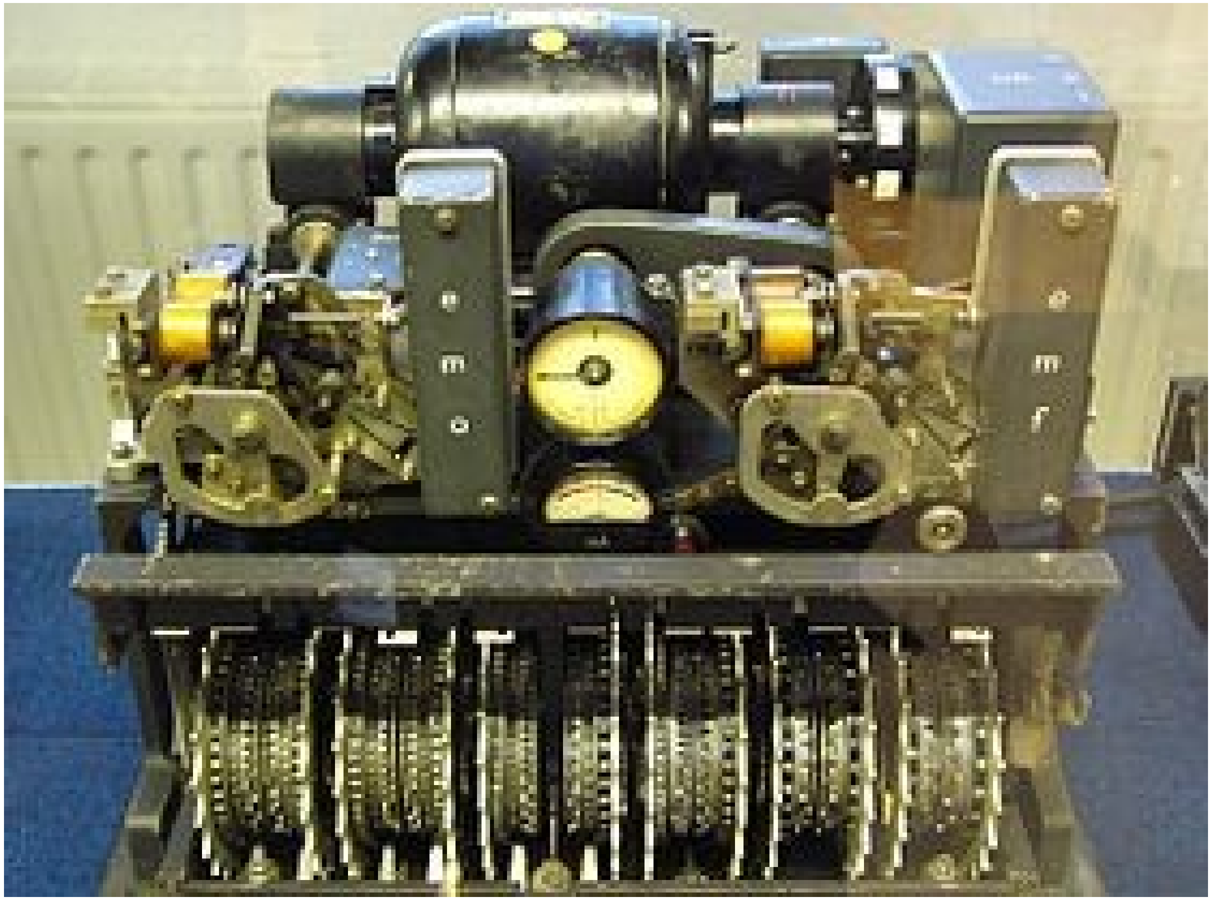
Cryptanalysis - Poland

- An early version of the enigma machine was broken by Marian Rejewski working in the Polish Cipher Bureau in the early 1930’s.
- This was achieved by seperating the scrambler settings from the plugboard settings.
- The scrambler settings were broken using an exhaustive search (6*17,576).
- And the plugboard setting were obtained using a form of frequency analysis.

Cryptanalysis - Blechley Park

- Later versions of the machine were eventually broken in Bletchly Park, England.
- Alan Turing was one of the main scientists involved.
- Bletchly Park had over 7,000 employees by the end of the war.
- Had the resources to break the new cipher.
- Eventually the machine had 12 scrambler wheels (known as the Lorentz machine).

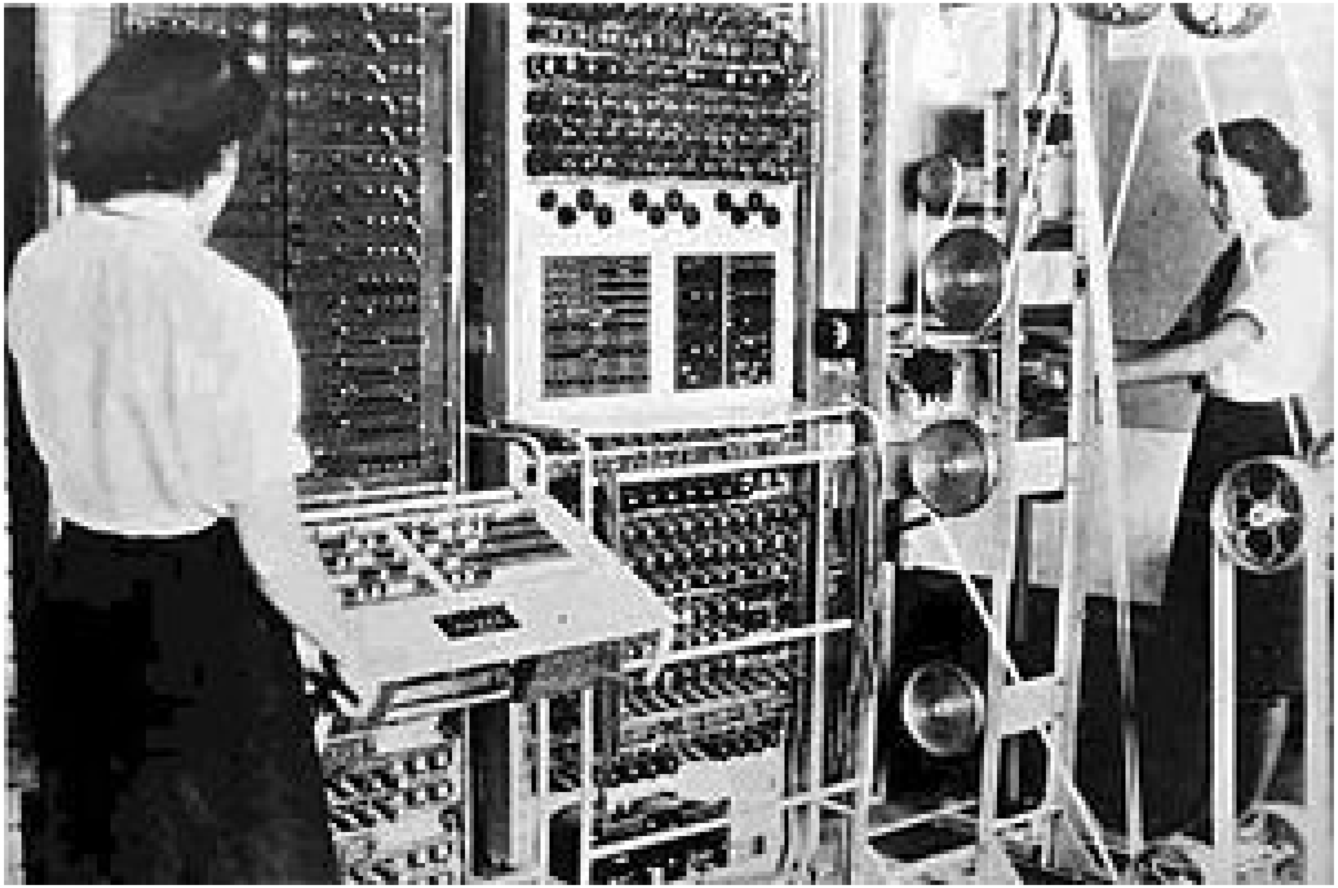
The Lorentz Machine



Aside - The Collossus

- In order to break the cipher, the world's first electronic, digital, programmable computer was built in Bletchley Park by Tommy Flowers, a telecoms engineer.
- 10 were in use by the end of the war.
- Machines were classified and destroyed after the war.
- Two went to GCHQ and were probably used there during the cold war.
- Blueprints were burned by Tommy Flowers personally.

Collossus (Mk II)



Modern Encryption

DES – Data Encryption Standard

- Modern encryption algorithms are built around results from mathematical number theory.
- DES algorithm submitted by IBM in response to a request for an encryption algorithm.
- DES was approved as a federal standard in November 1976.
- [Initially there were suspicions about a NSA back door.]
- Symmetric key encryption. 56 bit key.

The Key Distribution Problem

- To communicate securely with a symmetric key both parties need the same key.
- The key distribution problem was huge.
- Literally tons of material, (paper, floppy disks, punch cards, tape) were transported every day by the US government.
- All with very stringent security requirements.

Diffie/Hellman Key Exchange (1976)

- Each party generates a pair of keys.
- Known as a Diffie-Hellman key pair.
- Each party keeps one key private and sends the other (public key)
- Each takes their own private key and the others public key and generates a number.
- The magic is that the generated number is the same and can be used as a symmetric key.

Diffie/Hellman

- Bob Alice
- publicB publicA
- privateB privateA
- publicA publicB
- symmetric key == symmetric key

Diffie/Hellman

- A solution.
- Working in Stanford.
- Requires the active participation of both parties.
- (Used in SSL.)

RSA

RSA (1978)

- Rivest, Shamir, Aldeman working in MIT.
- Asymmetric encryption algorithm.
- Keys are generated in pairs. Encrypt with one key. Decrypt with the other.
- Is computationally expensive compared to symmetric key encryption.

Public Private Key Encryption

- Keep one key private.
- Make the other public.

- I send you data by encrypting with your public key. (Only you can decrypt it.)
- I can prove my identity (authentication) by encrypting with my private key.
- When you decrypt it and it make sense, only I could have encrypted it.

Public Private Key Encryption

- The final major piece of the jigsaw.
- Leads to Digital Signatures and Digital Certificates.