Cryptography

Cryptography

Crypto + graphy = secret + writing

- to make information secret, use a **cipher**, an algorithm that converts plain text to **ciphertext**, which is gibberish unless you have a **key** that undo the cipher.
- **encryption**: the process of making text secret
- **decryption**: the reverse process

Julius Caesar uses a **Caesar cipher**; he shifts all of the letters forward by three.

- A becomes D and “brutus” becomes “euxwxv”.
- to decipher, you need both the algorithm and the shift number, which acts as the key. In this case, the key is 3.
- The Caesar cipher uses 26 keys and is easy to break.

Substitution Cipher

Caesar cipher is an example of a larger class of ciphers called **substitution cipher**.

- every letter is replaced by a different letter by a translation.
- one drawback of basic substitution cipher is letter frequency is preserved.
- the letter e is the most common letter in English. If cipher translates e to x, then x would be the most common letter in ciphertext.
- it was the breaking of a substitution cipher that led to the execution of Mary, Queen of Scots, in 1587, for plotting to kill Queen Elizabeth.

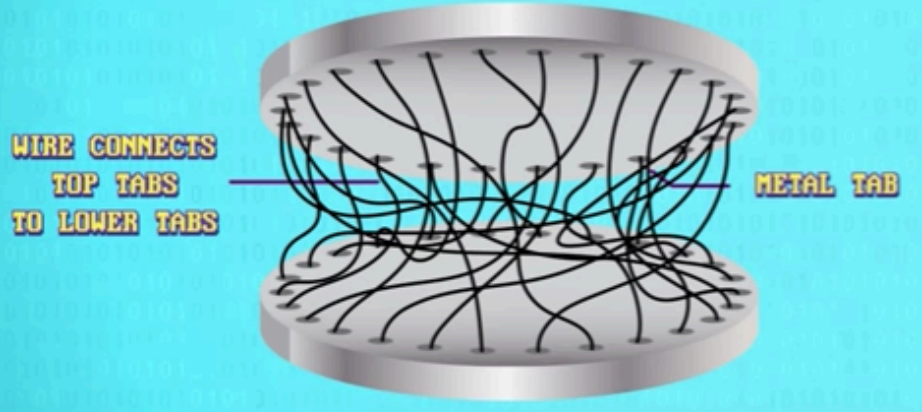
Enigma

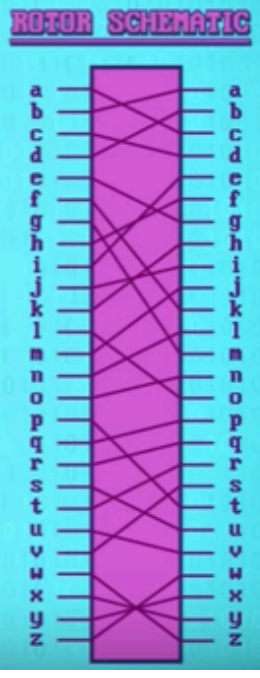
In the 1900s, cryptography was performed using machines.

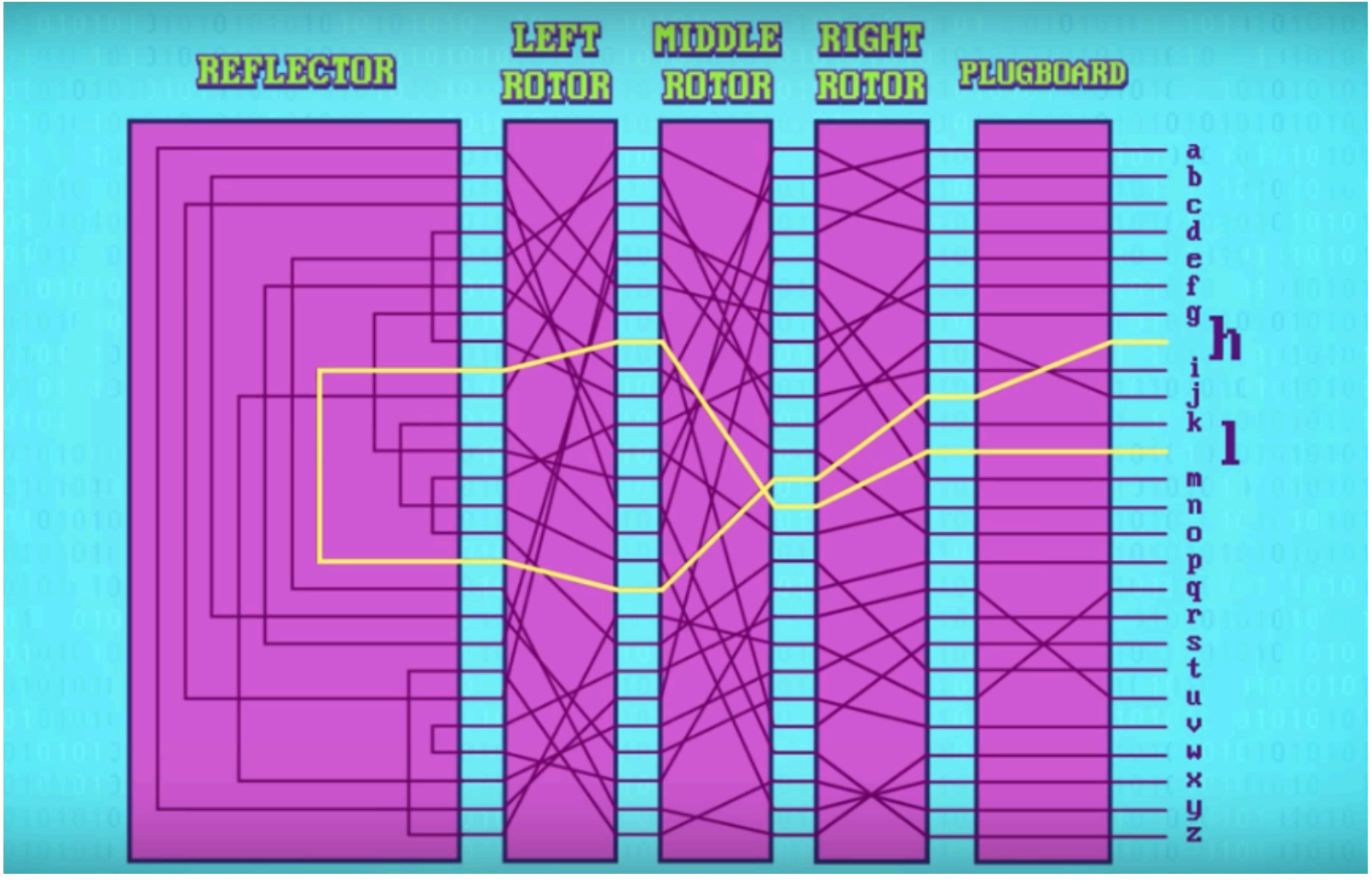
The German Enigma was an example of one such machine. It encrypt wartime communication.

The Enigma was a typewriter-like machine with a keyboard with the full alphabet and configurable rotors that were the key to the encryption.

ENIGMA ROTOR



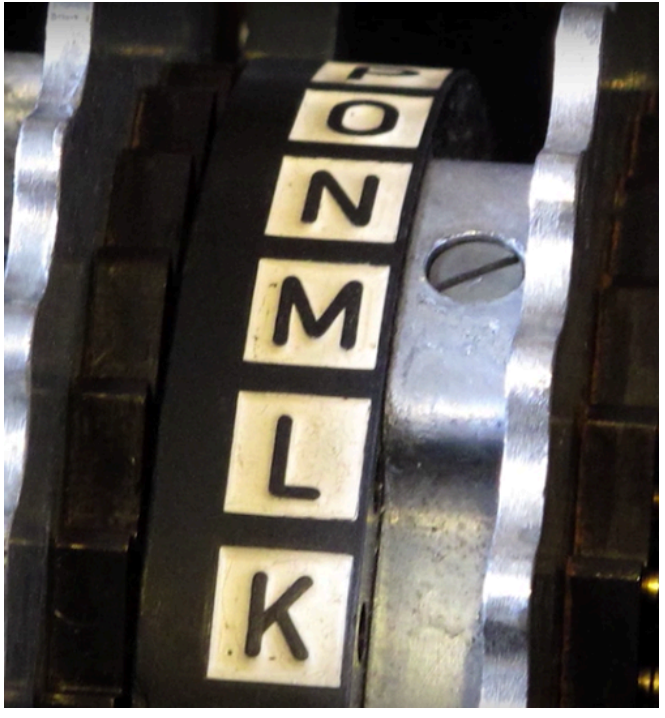



REFLECTOR
LEFT ROTOR
MIDDLE ROTOR
RIGHT ROTOR
PLUGBOARD
a b c d e f g h i j k l m n o p q r s t u v w x y z

The Enigma

To prevent it from being simple substitution cipher, after every letter, the rotor rotate by one, changing both the cipher and key.

- For example “AAA” might be encoded as “BDK”



The enigma was hard to crack but Alan Turing and his team at Bletchley Park were able to solve it and even automate the entire process!

One of Enigma’s flaw was every letter can’t be mapped into itself.

(for a movie version of this story, watch “The Imitation Game”, Benedict Cumberbatch.)

DES

One of the earliest widespread software ciphers was the **Data Encryption Standard(DES)** developed by NASA and IBM in 1977.

It was a 56-bit encryption and thus has 2^56 keys which is 72 quadrillion(thousand trillions).

Back in 1977, no one has the computing power to brute force that many keys.

In 1999, powerful computers can begin to crack DES.

AES

In 2001, **Advanced Encryption Standard(AES)** was published

- uses up to 256 bits, brute force is virtually impossible.
- For example, even with 128 bits, if you use every computer on the planet, it'll take trillion of years to try every combination of keys. With 256 bits, it would take up about the age of the universe to crack.

AES chops data into 16-byte chunks, perform a series of permutations and substitutions and other operations to obscure the message, then repeat 10 or more times.

- No fancy math, unlike the other popular alternative, RSA.

AES is used everywhere.

- encrypting files on the Iphone.
- encrypting data over WiFI with WPAS and HTTPS protocol.

In October 2017, "Krack Attack" was leveraged against WPA2 by researchers, particularly its handshake protocol.

- The vulnerability was not with the AES algorithm but its implementation.

Symmetric Encryption

Cryptographic techniques so far rely on keys known by both sender and receiver.

The sender encrypt message with a key and the receiver decrypt it with the same key.

Symmetric encryption: encryption technique that uses the same key for both encryption and decryption.

In the old days, keys are shared physically or by voice.

- Germans use codebooks with daily settings for Enigma machines.

Today, server needs to send a key over the internet.

- making a purchase on Amazon, login on to Gmail, etc.
- keys can be intercepted.

Key Exchange

Solution is to use **key exchange**.

key exchange: an algorithm that allows two computer to agree on a key without sending one!

This is done using one-way functions.

A **one-way function** is a function that is easy to compute but hard to invert.

- it's easy to mix colors but hard to unmix, i.e., figure out the component colors used to get a mixed color.

We'll use painting to illustrate how this is done conceptually.

Key Exchange

Alice and Bob want to establish a shared secret key for encryption/decryption.

Each starts with a secret color.

Eve is listening.

(Eve for "eavesdrops")

1. Alice's secret color

1. Bob's secret color

Key Exchange(colors)

Eve is listening.

1. Alice's secret color

2. Mix in public color

1. Bob's secret color

2. Mix in public color

Both mix the public color with their secret color.

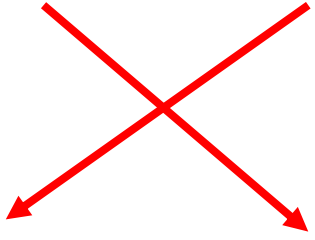
Key Exchange(colors)

Eve is listening.

1. Alice's secret color

2. Mix in public color

3. Send to Bob



1. Bob's secret color

2. Mix in public color

3. Send to Alice

Then send it to each other publicly. Then what?

Key Exchange(colors)

Eve is listening.

1. Alice's secret color

2. Mix in public color

3. Send to Bob

4. Mix in secret color

1. Bob's secret color

2. Mix in public color

3. Send to Alice

4. Mix in secret color

Each then adds in his/her secret color to obtain a shared secret color(key).
The key is then used for symmetric encryption.

Diffie-Helman

Diffie-Helman key exchange uses modular exponentiation as its one-way function.

Two positive integers x and y are **congruent modulo n**, denoted

x = y mod n if x % n = y % n.

For example, 7 = 2 mod 5 since both 2 and 7 when divided by 5 give a remainder of 2.

- 13 = 25 mod 2 = 1 mod 2
- 45 = 15 mod 10 = 5 mod 10

When performing math operations, we reduce modulo n so that the result is a remainder from 0 to n-1.

- 3^5 mod 31 = 243 mod 31 = 26 mod 31

Diffie-Helman

Modular exponentiation is a one-way function because it is easy to raise a power but hard to invert.

For example, it is easy to compute 3^5=26 mod 31.(One line of Python code.) But it's hard to find x such that 3^x=26 mod 31 if the base and modulo are very large integers(hundreds of digits long).

- This is called the **discrete logarithm problem**.

The base that is used for the modulo is a large prime(hundreds of digits long).

Solve this:

232413423432423424234234324^x = 234234234234230034382042 34 mod 2352354234324234234325345324234234234234534535235

Diffie-Helman is an algorithm that uses this one way function of modular exponentiation to do key exchange. The process is similar to the painting example earlier.

Diffie-Helman Key Exchange

Eve is listening.

		Public: Base B Modulo M		
1. Alice's secret exponent	Secret Exponent X		Secret Exponent Y	**1. Bob's secret exponent**
2. Compute	B^X mod M		B^Y mod M	**2. Compute**
3. Send to Bob	B^Y mod M		B^X mod M	**3. Send to Alice**
4. Compute	(B^Y)^X mod M) B^{XY} mod M		(B^X)^Y mod M) B^{XY} mod M	**4. Compute**

B^{XY} mod M is the shared key.

An Example

Eve is listening.

Public:
Base 3
Modulo 17

1. Alice's secret exponent	Secret Exponent 15		Secret Exponent 13	**1. Bob's secret exponent**
2. Compute	$3^{15} = 6 \mod 17$		$3^{13} = 12 \mod 17$	**2. Compute**
3. Send to Bob	$3^{13} = 12 \mod 17$		$3^{15} = 6 \mod 17$	**3. Send to Alice**
4. Compute	$12^{15} = 10 \mod 17$		$6^{13} = 10 \mod 17$	**4. Compute**

Thus the secred shared key is 10.

Diffie-Helman

One shortcoming of the Diffie-Helman key exchange is that it uses communication overhead to established a shared key.

- Alice and Bob sends back and forth different results of their calculations.

Another shortcoming is if Alice needs to messages to different people, she needs to exchange different keys.

- if she is a bank for example, she needs thousands of distinct keys to communicate with each person.
- she would then need to send thousands of messages just to establish these secret keys with each person.

Asymmetric Encryption

James Ellis, a British Engineer, was working on a non-secret encryption in 1977.

His idea is clever yet simple: lock and unlock are inverses. Use this as a one-way function!

- Easy to lock, but hard to unlock.

Ellis' idea is an example of **asymmetric encryption(public-key cryptography)**: encryption which uses a public and a private key. One for encryption and one for decryption.

Someone can encrypt a message with the public key but only the recipient with their private key can decrypt.

- the public key is use for encryption, the private key for decryption.

Alice has a private key and a lockbox(public key).

She wishes to receive an encrypted message from Bob.

Alice sends Bob the lockbox(public key).

Bob put his message inside the lockbox.

Lock it(one-way function, easy to lock)

Sends it back.

Alice can open it with her private key and receives the message. (even if the lockbox is intercepted, it cannot be opened without the private key)

If Alice is a bank, she can duplicates the lockbox(public key) and sends them to everyone she wishes to receive encrypted messages.

She only has to manage one private key.

Digital Signatures

Similar to the previous example, a public digital key can encrypt something that can only be decrypted by a private digital key.

The reverse is also possible: encrypting with the private key something that can be decrypted with a public key.

- Use for digital signatures
- Server encrypts data with private key.
- Anyone can decrypt it using the server's public key.
- This acts as an unforgeable signature that only the owner, using the private key, can encrypt.
- This proves that you're getting data from the right person, not an imposter.

The most famous example of **asymmetric encryption is RSA** named after their inventors Rivest, Shamir and Adleman.

The math behind RSA is beyond scope of the class. There's an optional video that you can watch at the end of the slides.

Modern Cryptography

The “key” parts of modern cryptography:

- symmetric encryption
- key exchange
- public-key cryptography

When you connect to a secure website, such as Facebook, that has the green padlock, you know that cryptography was used to encrypt data.

- Facebook receives a digital certificate from a **certification authority(CA)** which includes a public and private key.
- Your browser uses the public known key to validate facebook.com. If facebook is down, a hacker can’t create a clone of facebook without the correct private key to validate.(no green padlock on facebook clone’s site)

Specifically, it means

- your computer used public-key cryptography to verify the server.
- key exchange to establish a temporary secret shared key
- and symmetric encryption to send data back and forth.

Cryptocurrency

A **cryptocurrency** (or crypto currency) is a digital asset designed to work as a medium of exchange that uses cryptography to secure its transactions, to control the creation of additional units, and to verify the transfer of assets. (Wiki)

cryptocurrency:

- decentralized currency
- Bitcoin was the first, established in 2009. Since then other coins have been established, called altcoins.
- Bitcoin was invented by Satoshi Nakamoto.(Even today, no one knows who he is.)
- Uses public-key cryptography to make, secure transactions, verify ownerships, etc...
- Uses advanced math(elliptic curves) to do encryption.

Homework

1)Read and reread these lecture notes.

2)Watch(Required):

a)How bitcoin works(non technical)
https://www.youtube.com/watch?v=l9jOJk30eQs

3)Optional:

a)How RSA works. This is technical with some math but still accessible.
https://www.youtube.com/watch?v=wXB-V_Keiu8

b)How bitcoin works under the hood. Technical but still accessible. May need to rewatch, rewind several times.
https://www.youtube.com/watch?v=Lx9zgZCMqXE

c)Most of this lecture's material including some images come from PBS Digital Video on Cryptography.
https://www.youtube.com/watch?v=jhXCTbFnK8o

References

Part of this lecture is a recap of the following an episode from PBS Crash Course in Computer Science series.

PBS Crash Course in Computer Science. Cryptography. Retrieved from https://www.youtube.com/watch?v=jhXCTbFnK8o