ANU hack and UNSW cyber email

- ANU hack - what would you advise UNSW Vice Chancellor?
    - Don't suggest hiring consultant
    - Security engineering students to try and hack the system à penetration testing
        - Normal answer - make it harder to break into uni with firewall
- Consider and analyse UNSW cyber email - what is the likely objective? how successful is it in achieving that objective?
    - Email essentially called its users being idiots and criticising their online habits that should be fixed

Also mentioned the raid on journalist office

Security by Design

- 4 primary colours for a security engineer
    - Trust
    - Secrets
    - Humans
    - Engineering - risk, complexity
- Trust
    - Start with a healthy dose of scepticism when you begin
    - Trust no one, and don't build security on someone else's perceptions of trust
    - Want to make sensible choices with the aim of perfection
    - Never believe anyone who says something is perfect
    - Engineering is really about how can we achieve almost perfection with limited resource and time
- Defence in depth
    - System should be able to work if something fails e.g. submarine division into compartments
    - Want to avoid a single point of failure – once barrier is broken, everything fails
    - Concentric rings of walls for castles - if one wall breaks, limits areas of access while defending from other barriers
    - Segmentation of confidential data
    - Uni doesn’t really do this
- Security by design
    - Build security in from the beginning, instead of adding security when the breaches occurs (responsive security)
- Apes vs ants – represent 2 different paths of complexity
    - Apes – complexity between interactions between individuals
        - Have the capability from breaking from role within society and have freedom
        - Brain develops to allow us to work in larger groups and perform more complex tasks
    - Ants – themselves are not complex, dumb

        - Their society + colony is complex
        - No ant by itself is the one making the decisions, but the colony works as a whole
        - Colony is rigid + unwavering following of roles à predictable
        - e.g. human wearing a castle suit
- Bell LaPadua – levels of confidentiality
    - Not allowing people to have the choice to access specific parts of info
    - Relies on complete human obedience
    - Can only increase controls + penalty to deter
    - Could have an idiot on top = single point of failure
- Physical security
    - Everything we deal with is in a virtual world – within it, can't really tell anything
    - If someone can access the hardware they can do anything
    - Everything we run is running on the machine – if someone interferes with the machine à game over
    - Physical security - the thing we trust all the time as programmers
    - Consider hardware security measures before the software security measures
    - Hardware attacks - have become increasingly more common e.g. Huawei
    - Passwords - can retrieve it by setting up a camera around the space they type it, check the keyboard for imprints
    - Side channels - for every cyber action, there is a corresponding physical trace. “Every contact leaves a trace in the physical world”.
    - Side channel attack - based on information gained from the implementation of a system rather than weakness in the implemented algorithm itself

Vigenere (Caesar + Password)

- Take a password and shift the encrypted message by the alphabetical order of the password.

    e.g. ABBA would give the shift (0,1,1,0) on repeat, so “Hello” would become

    ```
    H+0(A) = H
    E+1(B) = F
    L+1(B) = M
    L+0(A) = L
    O+0(A) = O
    ```
- The Kasiski test
    - Used to defeat ciphers where the password/offset repeats, such as the Vignere cypher
    - Find the length of the key by looking for repetitions (longer repetitions are better)
    - Line up each section, then solve the lined up letters as single substitution cipher
- The Index of Coincidence

- The coincidence index is an indicator used in cryptanalysis which makes it possible to evaluate the global distribution of letters in encrypted message for a given alphabet. **If the index is high,** (similar to plain text) then the message has probably been encrypted using **a transposition cipher** or a monoalp habetic substitution. I**f the index is low** (similar to random text), then the message has probably been encrypted using a **polyalphabetic cipher** (letter can be replaced by multiple other ones)
- The easiest ways to determine if you have a monoalphabetic substitution cipher. The encoded message should have the same coincidence index as the language it was encoded in. (E.g. English texts typically have a coincidence index of 1.73)
- Formula

$$\mathbf{IC} = c \times \left(\left(\frac{n_\mathrm{a}}{N} \times \frac{n_\mathrm{a} - 1}{N - 1} \right) + \left(\frac{n_\mathrm{b}}{N} \times \frac{n_\mathrm{b} - 1}{N - 1} \right) + \ldots + \left(\frac{n_\mathrm{z}}{N} \times \frac{n_\mathrm{z} - 1}{N - 1} \right) \right)$$

$$\mathbf{IC} = \frac{\sum_{i=1}^{c} n_i (n_i - 1)}{N(N-1)/c}$$

  - Where *c* is the normalizing coefficient (26 for English), n_a is the number of times the letter "a" appears in the text, and *N* is the length of the text.
- Manual Calculation
  - Line the text up with a copy of itself shifted slightly.
  - Count the number of times the letter in the first copy matches the second copy.
  - Divide this by the number of aligned pairs of letters.
  - Multiply that by the size of the alphabet (26 for english)
  - The result is an estimate of the Coincidence Index

Enigma Machine

- Enigma Machine was an encryption device used by Nazi Germany in WW2 to protect commercial, diplomatic and military communication. Enigma has an electromechanical rotor mechanism that scrambles the 26 letters of the alphabet
- caesar plus almost infinite password (password extender)
- Look it up and make sure you can use it
  - Each key press causes one or more rotors to step by a letter, changing the substitution
- Understand the weakness and how it was used that allowed it to be cracked
  - Repeatedly used stereotypical expressions
  - Repetition of message key
  - Easily guessed keys
  - Re-transmitting a message on different cipher networks (solving a bad cipher of same message -> solved that enigma key)

- Not allowing repetitions (e.g. a letter cannot map to itself), reduced possibilities

One Time Pad

- In cryptography, the one-time pad (OTP) is an encryption technique that cannot be cracked, but requires the use of a one-time pre-shared key the same size as, or longer than, the message being sent.
- In this technique, a plain text is paired with a random secret key (also referred to as a one-time pad). Then, each bit or character of the plaintext is encrypted by combining it with the corresponding bit or character from the pad using modular addition.
- **Like Vignere** but the password is the length of the message and completely random, once used the cipher should be discarded

Type 1 & Type 2 Errors

- Type 1: False Positive - reject a true null hypothesis
  - E.g. a medical test saying you **do** have a disease when in reality you **do not**
- Type 2: False Negative - fail to reject a false null hypothesis
  - E.g. a medical test saying you **do not** have a disease when in reality you **do**
- There is a tradeoff between the two

Null hypothesis: statement that there is no relationship between two measured phenomena
Richards example: Sydney Airport Passport Gates sometimes letting people through with the wrong passport, and sometimes not letting people through with the correct passport.

Approximate English letter frequency table

Letter	Frequency	Letter	Frequency
E	12.02	M	2.61
T	9.10	F	2.30
A	8.12	Y	2.11
O	7.68	W	2.09
I	7.31	G	2.03
N	6.95	P	1.82
S	6.28	B	1.49
R	6.02	V	1.11
H	5.92	K	0.69

D	4.32	X	0.17
L	3.98	Q	0.11
U	2.88	J	0.10
C	2.71	Z	0.07

Insiders

Often securities can be broken from insiders because of self-interest, corruption, etc. There are many examples in espionage between Americans and Russians.

Other

- **Entropy**: randomness. If there's order in the data, it's hard to get rid of it even if we encrypt it. Better codes hide the data/information better
- **Edit distance**: distance between valid sentences (very large for English)
- Work ratio: ratio between the work a legitimate recipient has to do to decrypt the message, and an interceptor trying to decrypt the message. We want the ratio to be as big as possible
- **Repeat/playback attack**: exploit system by repeatedly transmitting an authenticated message
- **Tempest attack**: Analysing electromagnetic radiation from a distance. Dutch elections
- **Number stations**: radio shortwave station that just repeats numbers. Countries sent messages to their spies which they decoded with one-time-pads. Side channel was the amount of activity on the number stations

How to differentiate between ciphers

1. Check letter frequencies - this will tell us whether the cipher is a transposition cipher or not. If the letter frequencies obey english standards then likely the cipher is transposition.

2. Check coincidence index - if the coincidence index is around 1.73, then we believe the cipher is likely mono-alphabetic substitution.

3. Check the periodic CI - if periodically the CI spikes, the cipher is likely poly-alphabetic substitution.