Shannon's Information Theory

David Malone

18 October 2017

Cryptography

Cryptography is how to write messages and keep them secret.

1. We encrypt the message (plaintext) using a secret key.
2. Now the message (ciphertext) should be secret.
3. To decrypt the message you need the key.

Encode and decode are technically something different - they are how information is transmitted or stored.

Encrypting data

Caesar Cipher:

$$\texttt{HelloWorld} \rightarrow \texttt{LippsAsvph}$$

Key is d.

In modern world, ciphers should be secure if you know everything but the key (Kerckhoffs's principle).

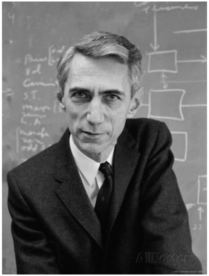
Brute Force Guessing

If you don't know the key, you can always guess. How could you decrypt `Wklv lv d vhfuhw phvvdjh`?

0	`Wklv lv d vhfuhw phvvdjh`	13	`Jxyi yi q iushuj cuiiqwu`
1	`Xlmw mw e wigvix qiwweki`	14	`Kyzj zj r jvtivk dvjjrxv`
2	`Ymnx nx f xjhwjy rjxxflj`	15	`Lzak ak s kwujwl ewkksyw`
3	`Znoy oy g ykixkz skyygmk`	16	`Mabl bl t lxvkxm fxlltzx`
4	`Aopz pz h zljyla tlzzhnl`	17	`Nbcm cm u mywlyn gymmuay`
5	`Bpqa qa i amkzmb umaaiom`	18	`Ocdn dn v nzxmzo hznnvbz`
6	`Cqrb rb j bnlanc vnbbjpn`	19	`Pdeo eo w oaynap iaoowca`
7	`Drsc sc k combod wocckqo`	20	`Qefp fp x pbzobq jbppxdb`
8	`Estd td l dpncpe xpddlrp`	21	`Rfgq gq y qcapcr kcqqyec`
9	`Ftue ue m eqodqf yqeemsq`	22	`Sghr hr z rdbqds ldrrzfd`
10	`Guvf vf n frperg zrffntr`	**23**	`This is a secret message`
11	`Hvwg wg o gsqfsh asggous`	24	`Uijt jt b tfdsfu nfttbhf`
12	`Iwxh xh p htrgti bthhpvt`	25	`Vjku ku c ugetgv oguucig`

Why did this work?

1940s: Claude Shannon



$$\texttt{HelloWorld} \rightarrow \texttt{PhymsQpvhv}$$

Key is idnbeubewsb. . . , but with different key `JumpyJoyce`

- Defined perfect secrecy.
- Seeing the cipher text tells you nothing about the plain text.
- Since guessing was involved, this is about probabilities.

Decoy Messages

- With more keys the message was safer.
- Some keys lead to decoy messages.
- Counting the number of messages.
- Counting the number of sensible messages.
- Look at chance of hitting a decoy.

Counting

Counting things in mathematics is called combinatorics.

For example, you want to buy a car. They offer three possible extras: a sun roof, seat DVD player or parking camera. How many combinations are possible?

- Sun roof, or not, gives 2.
- For each, DVD player, or not, gives 2. So,

$$2 + 2 = 2 \times 2 = 4.$$

- For each of those 4, could add a parking camera or not.

$$2 + 2 + 2 + 2 = 4 \times 2 = 8.$$

You end up multiplying the possibilities: $8 = 2 \times 2 \times 2 = 2^3$.

Counting in Bits

- Two choices N times give 2^N options.
- This is how computers store data.
- Each choice is a bit.

How many bits to store a letter?

$$2^4 = 16 < 26 < 32 = 2^5.$$

Or:

$$4 = \log_2 16 < \log_2 26 < log_2 32 = 5.$$

$\log_2 26 = 4.7004397181\ldots.$

Counting Messages

How many messages with 10 characters?

$26\times26\times26\times26\times26\times26\times26\times26\times26\times26 = 26^{10} = 141167095653376$

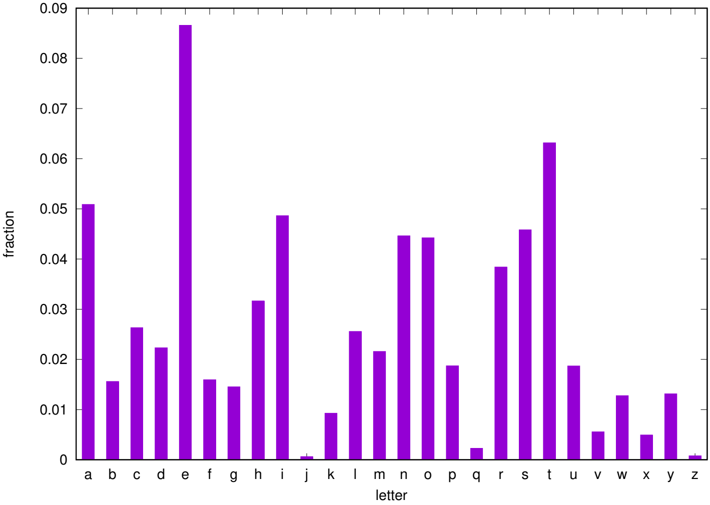
How many messages with N characters?

$$26 \times 26 \ldots 26 \times 26 = 26^N \approx 2^{N4.7}$$

Counting possible messages is relatively easy. What about sensible messages?

Counting Sensible Messages

Shannon realised taking the frequencies into account is important.



Have to *choose* the letters carefully.

Counting Ways to Choose

How may ways can we choose two extras for our car?

$$\binom{3}{2} = \frac{3 \times 2}{2}.$$

There is actually a formula:

$$\binom{N}{M} = \frac{N!}{M!(N-M)!}.$$

There is even a multi-choose formula for choosing k groups of things from N.

$$\binom{N}{n_1\ n_2\ \ldots n_k} = \frac{N!}{n_1!n_2!\ldots n_k!}.$$

where $n_1 + n_2 + \ldots n_k = N$.

Counting Sensible Messages

Shannon realised that if you want a sensible message of length N:

- You need $n_a = Np_a$ of letter a,
- You need $n_b = Np_b$ of letter b,
- ...
- You need $n_z = Np_z$ of letter z,

He used the multichoose formula to show this was about

$$2^{NH(p)}.$$

Counting Sensible Messages

$$H(p) = -\left(p_a \log_2 p_a + p_b \log_2 p_b + \ldots p_z \log_2 p_z\right),$$

is now called the *Shannon Entropy*.

- Shannon calculated $H(p)$ for English and got ≈ 4.
- He knew it was important to take into account pairs, ...
- For English is actually about 1.5 bits per character.

For Example . . .

Roughly how many 1-letter messages does English have?

$$2^{NH(p)} = 2^{1\times 1.5} = 2.8.$$

How many 10 letter messages?

$$2^{NH(p)} = 2^{10\times 1.5} = 32768.$$

How many 140 character English tweets?

$$2^{NH(p)} = 2^{140\times 1.5}$$

That's about
1645504557321206042154969182557350504982735865633579863348609024.

Imperfect Secrecy

- We have $26^N \approx 2^{N4.7}$ messages of length N.
- We have $2^{NH(p)} \approx 2^{N1.5}$ messages of length N.

Chance of hitting a decoy message

$$\frac{2^{N1.5}}{2^{N4.7}} = 2^{N1.5}2^{-N4.7} = 2^{N(1.4-4.7)} = 2^{-3.2N}.$$

This means that the chance of hitting a decoy gets lower and lower as the message gets longer, unless you add more keys to compensate.

Need to make sure brute force attacks aren't practical.

Summary

- Shannon counted messages and probabilities.
- Gave a theory for secret messages.
- Was able quantify how secret you were being.
- Shannon Entropy important not just for secrecy ...
- ... also for transmission and compression of data too.
- Maths can be applied in unexpected areas!

How to Share a Secret

There is a way to agree a secret with someone in public, so anyone listening can't figure out the secret in practice.

1976: Diffie and Hellman

Won the 2015 Turing Award.

Diffie-Hellman key exchange

Basic idea: Both people do the following:

1. Agree on a number as a *generator*, say 2.
2. Each Pick a secret number.
3. Multiply two by itself that many times.
4. Tell the other person the answer.
5. Multiply the other person's answer by itself the secret number of times.

Why does this work?

Rory	David
$2 \times 2 \times 2 = 8$	$2 \times 2 \times 2 \times 2 = 16$
$16 \times 16 \times 16 =$ $(2 \times 2 \times 2 \times 2)\times$ $(2 \times 2 \times 2 \times 2)\times$ $(2 \times 2 \times 2 \times 2) =$ 4096	$8 \times 8 \times 8 \times 8 =$ $(2 \times 2 \times 2)\times$ $(2 \times 2 \times 2)\times$ $(2 \times 2 \times 2)\times$ $(2 \times 2 \times 2) =$ 4096

It doesn't matter what order we do the multiplication in.

To make secure:

- Need big numbers and,
- work with remainder on division by big prime number.