Secret Codes, I: Cryptography

Peter J. Cameron

How the Light Gets In
Hay-on-Wye, June 2016

The how and why of information security

It is a truism that more and more information flies around the world and through our lives every day.

If I know a secret and have no intention of ever telling anyone, there is no problem. If I don't care who knows it, again there is no problem. A conflict only arises when (to use the traditional characters in this story) Alice has some information which she wants Bob to know but must keep out of the hands of the eavesdropper Eve.

The issue is exacerbated by modern communications, but is far from a new problem. I'll start by looking back at how we got to where we are.

Some distinctions

How can Alice get information to Bob without Eve's interference? This is more subtle than first appears.

If Alice writes a letter to Bob and entrusts it to the post, Eve might intercept the letter, steam it open and read it, and then send it on its way, taking care that the envelope shows no sign of her tampering.

Or she might replace the message by a completely different one, if she is skilled enough to fake Alice's handwriting.

Either of these would be a disaster.

So let's set out a few terms.

Some terms

- Alice might hide the message somehow and hope that Eve doesn't find it. This is steganography.
- The most common scenario: Alice will probably encode the message in some way so that, even if Eve intercepts it, she cannot read it. But she has to make sure Bob is able to decode the message! This is cryptography. (Strictly, the art of sending encoded messages is "cryptography" and the art of breaking the code is "cryptanalysis", but I will use the same term for both.)
- The communication channel may be noisy; messages may become distorted. Alice must ensure that her message can overcome the distortion. This is error correction.
- Also part of information security is the need to tell whether a message has been intercepted or tampered with.
- Finally, Alice has to be able to sign her message in such a way that Bob can be certain that it comes from Alice (and Alice cannot deny sending it).

Secrecy and error correction

In Hay-on-Wye, it is appropriate to tell an old story linking cryptography with error correction, which occurs in the tale of Lludd and Llevelys in *The Mabinogion*.

> When Lludd told his brother the purpose of his errand Llevelys said that he already knew why Lludd had come. Then they sought some different way to discuss the problem, so that the wind would not carry it off and the Corannyeid learn of their conversation. Llevelys ordered a long horn of bronze to be made, and they spoke through that, but whatever one said to the other came out as hateful and contrary. When Llevelys perceived there was a devil frustrating them and causing trouble he ordered wine to be poured through the horn to wash it out, and the power of the wine drove the devil out.

The purpose of the horn is to keep the communication secret from the Coannyeid. But the devil in the horn introduces errors, and washing it out with wine is the error correction.

Steganography

I will begin with steganography. This has a long history.

- Herodotus relates that Histiaeus (the tyrant of Miletus) shaved the head of his most trusted slave, tattooed a message on his head, and then waited for his hair to grow back. The slave was then sent to Aristagoras, who was instructed to shave the slave's head again and read the message, which told him to revolt against the Persians. [Not useful if your message is urgent!]
- The Galician writer Alvaro Cunqueiro describes another classical technique, "skitale":

  > The code often used in the Arthurian Chancellery … involved a rod of olive-wood about a span and a half in length, around which was obliquely wrapped a bit of skin; on this the message was written, from top to bottom, in such a way that when the skin was unrolled only detached letters appeared, and to read the message the recipient had to roll the skin again around a rod of the same dimensions.

Modern steganography

- Invisible ink has always been popular with children setting up secret communications.
- Spy stories from the twentieth century often use *microdots*, where the message is photographically reduced to the size of a dot and is stuck over a full stop in some innocent text.
- A big file such as a picture or music file can carry a secret message as a watermark; this is done by altering a few bytes of the file, not enough to be noticed. Copyright information is often included in this way.
- A more up-to-date version uses the genetic code. A molecule of DNA is a long message in an alphabet of four letters A, C, G, T (adenine, cytosine, guanine and thymine). Using genetic manipulation techniques a few molecules of DNA can be imprinted with a secret message, which can be recovered by amplifying with PCR (polymerase chain reaction) techniques.

Catching a liar

We now turn to error correction.

I need a volunteer. Your job is to think of a number, a whole number between 0 and 15 inclusive. I am going to ask you some questions about your number. You are allowed to lie in your answer to **at most one** of my questions. I will tell you what your number is and which question you lied to (if any).

1. Is the number 8 or greater?
2. Is it in the set $\{4, 5, 6, 7, 12, 13, 14, 15\}$?
3. Is it in the set $\{2, 3, 6, 7, 10, 11, 14, 15\}$?
4. Is it odd?
5. Is it in the set $\{1, 2, 4, 7, 9, 10, 12, 15\}$?
6. Is it in the set $\{1, 2, 5, 6, 8, 11, 12, 15\}$?
7. Is it in the set $\{1, 3, 4, 6, 8, 10, 13, 15\}$?

Error-correcting codes

Here is a very brief explanation of how it works.

If you write out the correct answers to the seven questions for each of the numbers from 0 to 15, you find that any two numbers give different answers to at least three questions.

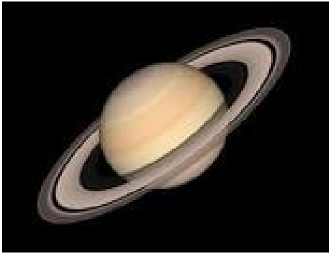
So the answers you give me are only “one step away” from the correct answers for your number, but at least “two steps away” from the answers for any other number. So I can figure out the correct number.

(If it is 3 kilometres from Hay-on-Wye to Clyro, and I travel 1 kilometre from Hay, I can be sure that I am closer to Hay than to Clyro.)

This example is based on an error-correcting code called a binary Hamming code.

Random errors or deliberate alterations?

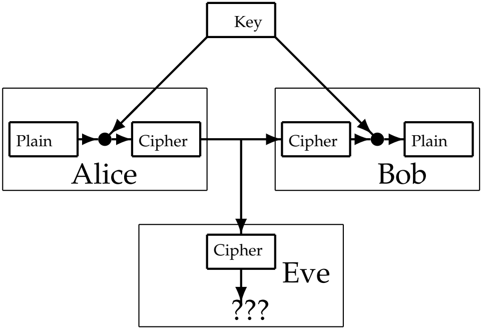
Similar codes are used for error-free information transmission in contexts ranging from backing up a computer onto a remote hard disk to sending information and pictures from distant planets.



This technique is for correcting random errors. We assume that Nature is not maliciously changing the message into a different one as Eve might do. Nevertheless, the ideas of error correction are related to those of preserving the authenticity of a message.

General principles of cryptography

We now come to cryptography. First, some general principles. In order for Alice and Bob to communicate, they must share some information called the key. Knowledge of the key is necessary for encryption and decryption. The process can be represented schematically like this:



Alice encrypts the plaintext using the key to produce the ciphertext, and Bob uses the key to decrypt the ciphertext to recover the plaintext. Eve, not knowing the key, cannot decrypt.

Kerckhoffs' Principle

A very important principle is Kerckhoffs' principle, formulated by Auguste Kerckhoffs in the 19th century. We must assume that Eve knows the type of cipher being used, and can intercept the ciphertext. All that she doesn't know is the key.

Even if Alice and Bob have invented a clever new cipher, they have to expect that knowledge of how it works will rapidly spread through the intelligence community.

In more sophisticated cryptanalysis, we may also assume that Eve has part of the key, and would like to know that this gives no information about the remainder of the key.

In this analysis, concepts from information theory such as entropy arise naturally.

Caesar cipher

The simplest technique is the substitution cipher, where we systematically substitute a letter in the plaintext by another letter or symbol. If you do Codeword puzzles in the newspapers, you will understand the method.

Julius Caesar reportedly used a simple substitution cipher in the Gallic wars. He simply shifted each letter forward or backward in the alphabet by a fixed number of places, as in the following table.

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

Thus “Send a hundred slaves as tribute to Rome” would be enciphered as `Xjsi f mzsiwji xqfajx fx ywngzyj yt Wtrj`.

A puzzle for you: What is the longest English word which is transformed into another English word by some Caesar shift?

Substitution ciphers

The key to a Caesar cipher is just the size of the shift: in the above case, 5.

The number of keys is so small (only 26) that it is a simple matter to check them all. Almost certainly only one shift will lead to a sensible message. (We see that a necessary condition for a secure cipher is having a large number of keys.)

We can improve things by using arbitrary substitutions. Here the number of keys is the number of substitutions, which is 26 factorial:

$$26! = 403291461126605635584000000.$$

However, even with this many keys, as you probably know from doing Codeword puzzles, the cipher is not secure! So having a large number of keys is not sufficient for security.

Frequency analysis

More than a millennium ago, Arab scholars engaged in textual analysis of the Qu'ran observed that, like any language, Arabic has patterns: certain letters or combinations of letters occur more often than others.

In English, the most common letters (in order) are

`E T A I O N S H R D L U`

There are also statistics on letter pairs or triples. For example, `TH` is the commonest digram, while `Q` is always followed by `U` except in foreign words.

Knowledge of these patters allows tentative identification of letters in the ciphertext; as more letters are found, the subsequent choices become easier.

The end of substitution ciphers

Half a millennium after the work of the Arab cryptanalysts, Mary Queen of Scots, who was held prisoner by the English, used a substitution cipher with a few embellishments (special letters, a codebook, nulls) to communicate secretly with her supporters. They plotted to assassinate Queen Elizabeth. However, Mary's circle of associates had been infiltrated by agents of Elizabeth's spymaster, Sir Francis Walsingham. His team were able to decrypt the intercepted messages and reveal the plot. This information led to Mary's execution at Fotheringhay in 1587.

The message was clear: cryptanalysis was a powerful tool; and substitution ciphers were not up to the job.

Two interesting fictional examples occur in Edgar Allen Poe's story "The Gold Bug", and the Sherlock Holmes story "The Adventure of the Dancing Men", by Arthur Conan Doyle.

Stream ciphers

However, a modification of the Caesar cipher paradoxically gives us the only known completely secure cipher!

We remarked that the key to a Caesar cipher was the size of the shift. This can be given as a single letter: in our example, the shift of 5 places turns `A` into `F`, and so can be described by the single letter `F`.

Now we can strengthen the cipher as follows. Instead of using the same shift throughout, we take the key to be a string of letters at least as long as the message being encrypted. The first letter of the message is shifted by the amount specified by the first letter of the key; the second, by the second; and so on. If Alice and Bob both have the key, they have a workable cipher system.

Vigenère cipher

In this form, the key is a single word, which is repeated as often as necessary. For example, if the key is `FOXES`, then we shift the first letter by 5 places, the second by 14, the third by 23, the fourth by 4, the fifth by 18, the sixth by 5 again, and so on.

Two refinements are possible:

To speed the process, we can use a substitution table, a 26×26 table in which looking up the row of the plaintext letter and the column of the key will give the ciphertext letter. This is just obtained by writing `A...Z` in the first row and then shifting each row one place from the one above. But all we need is that each letter occurs once in each row and each column. This is a structure called a Latin square.

The effect is that instead of a different *Caesar shift*, we are using a different *substitution cipher* for each letter.

In the Second World War, the Japanese navy used this system with a 10-symbol alphabet (the digits $0 \ldots 9$), and changed the substitution table regularly to make decrytion harder.

Vigenère cipher, 2

The second modification addresses a weakness in the Vigenère cipher. As described above, with a keyword such as `FOXES` of length 5, the first, sixth, eleventh, … letters of the message are all encrypted by the same Caesar shift. So, if the keyword is not too long and its length can be guessed, then we only have to break a small number of Caesar ciphers.

Now suppose that having encrypted with `FOXES`, we then re-encrypt with `WOLVES`. Since 5 and 6 have no common factor except 1, the result is that the repeated Caesar shifts will only occur in every 30th place, rather than every 5th. Re-encrypting again with, say, `JAGUARS` would increase the period to 210. This principle was used in the second world war by the German FISH cipher.

FISH

The Second World War saw the introduction of mechanical encryption: messages were too long, and encryption methods had become too complicated, for the process to be done by hand in the field. The machines included the German Enigma (whose story has been often told) and Japanese Purple machines.

The Fish cipher is much less well known than the Enigma cipher, but arguably much more important. While Enigma was used for communicating with U-boat commanders, Fish handled high-level communications between the German High Command and commanders in the field. Breaking this cipher, which has been described as one of the greatest intellectual achievements of the war, gave vital strategic information to the Allies.

Colossus

The Sägefisch machine used eight wheels with 23, 37, 43, 47, 51, 53, 59 and 61 teeth. These numbers are coprime, and so the number of keys is their product, which is 16033955073056318658. Probably the Germans trusted that this large number would keep communications secure. Yet the cipher was broken at Bletchley Park, without even having a copy of the machine (unlike Enigma).

The cryptanalysts could have built a replica machine, but ordering the parts might have given away the information that they had broken the cipher. So instead they built the world's first general-purpose computer, the Colossus, from readily available parts (telephone switchgear).

After the war, the machine was smashed. But undoubtedly this introduced a new age of electronic cryptography.

Stream ciphers

We saw that, for a stream cipher, the key must be as long as the message. This poses a difficult problem: key distribution. If two corporations want to share megabytes of data with one another, they must first share the key; and this must be done securely! Within living memory, couriers criss-crossed the City of London carrying keys from one institution to another.

There are two ways to proceed:

- Share a genuinely random key, and face the difficult problem of key distribution.
- Mechanical and electronic equipment allows another solution: the key is pseudo-random, that is, it has the unpredictability associated with a random sequence of symbols, but can be generated from a much shorter amount of information.

We conclude with a brief look at the two approaches.

The one-time pad

The great pioneer of information, Claude Shannon, among many other discoveries, showed that a one-time pad is completely secure.

This is a stream cipher where the key is a random string and the substitution table is a Latin square. Security means that the ciphertext is also a random string and gives Eve no information. Indeed, even if part of the plaintext is known, the remaining ciphertext is still a random string.

Peter Wright, in *Spycatcher*, tells of searching the apartments of suspected Soviet spies in London during the Cold War, finding and copying their one-time pads, so that their messages could be read. According to urban legend, there was a room deep inside the Pentagon where people sat tossing coins all day to generate random strings for one-time pads.

Of course, nothing is really secure. If Alice and Bob use a one-time pad, Eve can search their dustbins or ransack their hard disks in the hope of finding the key.

Binary addition

The simplest Latin square is

0	1
1	0

This corresponds to binary addition, the “native tongue” of computers:

$$0+0=0 \quad 0+1=1 \quad 1+0=1 \quad 1+1=0$$

Using this Latin square, if p is the plaintext written in binary, as a string of 0s and 1s (as any computer file is), and k is a random sequence of 0s and 1s (produced, for example, by tossing coins), we can write the ciphertext as $p+k$, meaning “add the first terms of p and k by binary addition, then add the second pair, and so on”.

By Shannon’s Theorem, this is a secure cipher, provided the key really is random and is kept secret.

Other stream ciphers

Taking the other approach means using a pseudo-random number generator which can produce a random-looking key from a short input. Many commercial ciphers are of this kind.

According to Kerckhoffs' Principle, we must assume that Eve has the program for the pseudo-random number generator, and only the short input is secret; so it must be long enough to prevent exhaustive search!

From time to time, governments try to restrict the strength of such encryption systems. The American Data Encryption Standard (DES) was restricted to 56-bit keys because the government could break these ciphers. The same battle is now being re-fought with longer keys, as we see in the recent cases involving Apple iPhones.

In the next lecture, we will look at modern developments, and see that although much has changed, much still remains the same …