Ciphers

Ælfred se leof

1 Introduction

Cryptography is the science of scrambling a message (called the *plaintext*) into a *ciphertext* that is unreadable to any party without access to special information called the *key*. These days it's an important topic in computer security but its history dates back almost as far as writing itself.

Cryptography has traditionally been a secretive business by virtue of its applications in espionage and the occult, and a belief (now held to be false, at least in academic circles) that secret methods are better than open ones. Hence, finding practitioners willing to divulge their art hasn't always been easy. When I first looked at the history of cryptography, the only substantial work on the subject was Kahn's *The Codebreakers*. Since then, Singh has also published *The Code Book* (which also spawned a television series).

There are many entertaining stories to be found in the history of cryptography, which you can find in Kahn's book. In this short article, I'll only give an overview of the development of ciphers over the SCA period.

2 Substitution Ciphers

The earliest and simplest form of cryptography is the *substitution cipher*, in which each letter is replaced by another symbol, such as Q for A, W for B, and so forth. Only someone who knows the mapping between the symbols is able to reverse the substitution process and recover the plaintext. Messages enciphered this way seem to be nearly as old as writing itself though the technique itself isn't discussed in anything that has come down to us until the time of the Greeks and Romans.

The technique for breaking these ciphers is first known to be described by the Arab philosopher al-Kindi (c.800-c.870)[1], and is now well-known. Arab cryptographers noted that natural languages have structure that can be exploited to guess the mapping between the plaintext and ciphertext symbols. In English, for example, a ciphertext symbol that appears very often is likely to be E or another common letter. A pair of ciphertext symbols that appears commonly is likely to be TH or EA, and so on. Using these sorts of observations, it is easy to make some likely guesses and once a few symbols are guessed correctly it is easy to work out the rest by guessing words and grammatical constructions.

2.1 Nomenclators

This technique became known in Europe during the fifteenth century. In response, the basic substitution system was extended by having a list of several ciphertext symbols to which a single plaintext letter can map. To encipher a letter, we make a random choice from the list. This way, the frequent appearances of letters like E and A are divided amongst several different ciphertext symbols and it is no longer possible to identify symbols by the frequency of their appearance.

A *nomenclator* is a cipher made by using the aforementioned technique, in combination with code words representing common names or phrases. This method seems to have been the favoured method in Europe from the fifteenth century right up until the advent of the telegraph in the nineteenth century.

Nomenclators, properly implemented, are harder to break than simple subsitution ciphers, but they can still be broken using the same kinds of techniques. Furthermore, lazy clerks charged with encrypting messages using a nomenclator weren't always very good random number generators and might always choose the same symbol

[1] Kahn, writing in 1967, has ibn ad-Duraihim (1312-1361) for this; al-Kindi's manuscript was only discovered in 1987

from the list of possible symbols, effectively reducing the nomenclator to a simple substitution cipher.

3 Polyalphabetic Ciphers

The basic substitution cipher and the nomenclator are examples of *monoalphabetic substitution* in which a given ciphertext symbol always represents the same plaintext symbol. We have seen how these kinds of ciphers can be broken by making educated guesses at the identity of ciphertext symbols.

3.1 Vigenère Ciphers

The first known *polyalphabetic* cipher is due to Leon Battista Alberti (1404-72), an Italian best known for his architecture. In an essay on cryptography written in 1466, he described a method of encipherment which used two disks placed on top of each other and fastened together at the centre. The edge of the outer disk was labelled with the letters of the alphabet in their usual order, and the inner disk was labelled with a scrambled version. (Some readers may recognise this device as a "decoder ring"). This is the forerunner to rotor-based cipher machines such as the Enigma that were used in the Second World War.

To encrypt a message, the sender finds the first letter to be encrypted on the outer ring, and writes down the adjacent letter from the inner ring. The sender then turns the inner ring by one letter, and encrypts the second letter in the same way as the first. The sender turns the ring again for the third letter, and so on. To decrypt the message, the receiver must have a copy of the inner disk and know the starting position of the disks. He or she can then find the first letter of the message by going from the inner ring to the outer ring, turning the disks by one, and so on.

This is one way of implementing what is now known as a *Vigenère cipher*, after the French cryptographer Blaise de Vigenère (1523-96) even though he, apparently, had nothing to do with it. In this kind of cipher, we choose a different ciphertext alphabet for each position in the plaintext. Using Alberti's cipher disk, the current ciphertext alphabet is chosen by the relative position of the two disks, and will (for English) repeat after twenty-six letters.

Another, better-known, implementation is to choose a key word and encrypt by "adding" the letters of the key word to the letters of the plaintext. Think of A = 1, B = 2, etc. with 27 = 1, 28 = 2, etc., so that, for example, A + C = D and Y + B = A. To encrypt the first letter of plaintext, we add it to the first letter of the key word. The second letter is encrypted by adding it to the second letter of the key word, and so on, starting again with the first letter of key word when we come to the end of the key. This way we don't need to make any disks, though the method is weaker because the key is shorter and easier to guess.

In either implementation, each ciphertext symbol may map to many different plaintext symbols, and the kinds of educated guesses used for solving nomenclators are no longer helpful. During the SCA period, cryptanalysts could only break these kinds of ciphers by guessing parts of the plaintext (e.g. guessing a message began with 'Dear John') and working backwards to calculate a guess at the secret key. A general method for breaking these kinds of ciphers was not discovered until the nineteenth century.

3.2 Autokey Ciphers

Another Italian, Girolamo Cardano (1501-1576) described a different kind of polyalphabtic cipher in *De Subtilitate* (1550), in which each letter is encrypted using the letter before it. This kind of cipher is called an *autokey* cipher and the classic implementation is due to Vigenère's *Traicté des Chiffres* (1585).

In Vigenére's version, the secret key is a single letter of the alphabet. The first plaintext letter is encrypted by adding it to the secret key as for the Vigenère cipher above. The second plaintext letter is encrypted by adding it to the first plaintext letter, and the third by adding it to the second plaintext letter, and so on.

Kahn seems to think this a better cipher than the one that bears Vigenère's name; in fact it has an obvious problem in that (for English) there are only twenty-six possible secret keys and it is feasible to find the correct one by trying them all. Nonetheless, the idea of using the plaintext itself to control encryption is a significant advance over simple substitution.

3.3 Polyalphabetics vs Nomenclators

Though polyalphabetic ciphers could not be broken using techniques known during the Rennaissance, they did not displace nomenclators in diplomatic practice until the nineteenth century. Kahn quotes several authors reporting polyalphabetic ciphers to be laborious and error-prone: if the encipherer makes a mistake in a monoalphabetic cipher, only the letter in question becomes garbled; but in a polyalphabetic cipher, the whole ciphertext after that point becomes indecipherable (as far the average Rennaissance decipherer was concerned).

4 Key Management

The reader might wonder how the secret key was shared between the sender and the recipient in the first place. Up until the invention of Diffie-Hellman key exchange in the 1970s, this had to be done in person – spies and diplomats had to be given their secret keys before leaving home. If the spy was intercepted, or otherwise compromised, the key was revealed.

Similarly if the key was easy to guess, the cipher was quick to follow. The few Vigenère ciphers that were broken in period used keys that were easy for the attacker to guess.

Today, network administrators are still tearing their hair out over users who choose easily-guessed passwords, or who write them down for attackers to find, or give them away to their friends.

5 Conclusion

What was state-of-the-art communications security in the Rennaissance is now largely relegated to the domain of childrens' spy books. However, these basic techniques, and lessons learnt from using and abusing them, form the basis of much of modern cryptography.

References

[1] Kahn, D., *The Codebreakers*, Scribner, New York, Revised Ed. 1996. ISBN: 0684831309

[2] Singh, S., *The Code Book*, Fourth Estate, London, 1999. ISBN: 1857028791