13 Variants of Cryptographic Procedures

Some Definitions

Substitution: Letters or groups of letters are replaced by other ones.

Monoalphabetic substitution: Each letter is replaced by another letter that is always the same.

Polyalphabetic substitution: Each letter is replaced—depending on its position in the text—by another letter. (The most important method of classical cryptography in the 20th century up to the sixties)

Monographic substitution: Letters are replaced by symbols one at a time.

Polygraphic substitution: In each step one or more letters are replaced by several symbols.

Homophonic substitution: For some plaintext letters or groups there are several choices of ciphertext symbols.

A mathematical model uses a probability space Ω and considers encryption functions of the type

$$f_k : M \times \Omega \longrightarrow \Sigma^*.$$

This is called **probabilistic encryption**.

Transposition: The letters of the plaintext are permuted.

Codebook: Letter groups of various lengths (for example entire words) are replaced by other ones according to a list. Since the Renaissance this was in use under the denomination **Nomenclator**. It was the most used encryption method even in the 20th Century, especially by diplomats.

Source coding (superencrypted code): The plaintext is transformed with a codebook, and the resulting "intermediate text" is encrypted by some kind of substitution.

Book cipher: Plaintext words or letters are looked up in a certain book. As ciphertext one takes the position of the word or letter in the book, for example page number, line number, number of the word (or number of the letter).

Block cipher: In each step a fixed number of letters is substituted at once.

Stream cipher: In each step a single letter is substituted, each time in another way, depending on its position in the plaintext.

Product cipher: A sequence of several transpositions and block substitutions is applied one after the other (also called cipher cascade).

Polygraphic Substitution

For a fixed l in each step an l-gram (block of l letters) is encrypted at once.

As simplest nontrivial example we consider **bigraphic substitution**. Here pairs of letters are encrypted together. The easiest description of the cipher is by a large square of sidelength $n = \#\Sigma$. An example for the standard alphabet:

	a	**b**	**c**	**d**	**...**
a	CA	FN	BL	...	...
b	SK	WM	...	...	...
c	HP	...	...	...	...
d	...	...	...	...	...
...	...	...	...	...	...

With this table BA is encrypted as SK .

The earliest historical example was given by PORTA in 1563. His bigram table however contained strange symbols meeting the spirit of the time. A picture is on the web page http://www.staff.uni-mainz.de/pommeren/Cryptology/Classic/1_Monoalph/PortaBi.gif

Properties of the Polygraphic Substitution

1. The key space of a bigraphic substitution is the set $\mathcal{S}(\Sigma^2)$ of all permutations of the Cartesian product $\Sigma \times \Sigma$. It contains the huge number of $n^2!$ keys. (Of course one also could restrict the keys to a subspace.) The effective keylength is

$$d(F) = \log_2(n^2!) \approx n^2 \cdot \log_2(n^2) = 2 \cdot n^2 \cdot \log_2(n).$$

   For $n = 26$ this amounts to about 4500. Exhaustion surpasses all present or future computer capacity.

2. Compared with a monoalphabetic (and monographic) substitution the frequency distribution of single letters is flattened down. A statistical analysis therefore must resort to bigram frequencies and is a lot harder. Pattern recognition and search for probable words also is harder, but not so much. Also more general attacks with known plaintext are feasible.

3. We may interpret a polygraphic substitution of l-grams as a monographic substitution over the alphabet $\tilde{\Sigma} = \Sigma^l$ of l-grams. The larger

l, the more complicated is the cryptanalysis. However for the *general* polygraphic substitution also the complexity of specifying the key grows with n^l, that is exponentially with l. Therefore this encryption method is useful only with a restricted keyspace. That means we need to fix a class of substitutions $\Sigma^l \longrightarrow \Sigma^l$ whose description is much shorter than the complete value table of n^l entries.

A bigraphic example from history is the PLAYFAIR cipher, invented by WHEATSTONE.

4. Polygraphic substitutions are the predecessors of modern block ciphers.

Codebooks

See the web page `http://www.staff.uni-mainz.de/pommeren/Cryptology/Classic/1_Monoalph/Codebook.html`