6. Block Ciphers

In many of the aforementioned cryptosystems, changing one letter in the plaintext changes exactly ne letter in the ciphertext. In the shift, affine, and substitution chiphers, a given letter in the ciphertext always comes from exactly one letter in the plaintext. This greatly facilitates finding the key using frequency analysis. In the Vigenere system, the use of blocks of letters, corresponding to the length of the key, made the frequency analysis more difficult, but still possible, since there was no interaction among the various letters in each block. The modern study of symmetric-key ciphers relates mainly to the study of block ciphers and stream ciphers and to their applications. A block ciphers take as input a block of plaintext and a key, and output a block of ciphertext of the same size. Since messages are almost always longer than a single block, some method of knitting together successive blocks is required. Block chiphers avoid these problems by encrypting blocks of several letters or numbers simultaneously. A change of one chacter in a plaintext block should change potentially all the characters in a plaintext block should change potentially all the characters in the corresponding ciphertext block.

The Playfair cipher is a simple example of a block cipher, since it takes two-letter blocks and encrypts them to two-letter blocks. A change of one letter of a plaintext pair will always change at least one letter, and usually both letters, of the ciphertext pair. However, blocks of two letters are too small to be secure, and frequency analysis, for example, is usually successful.

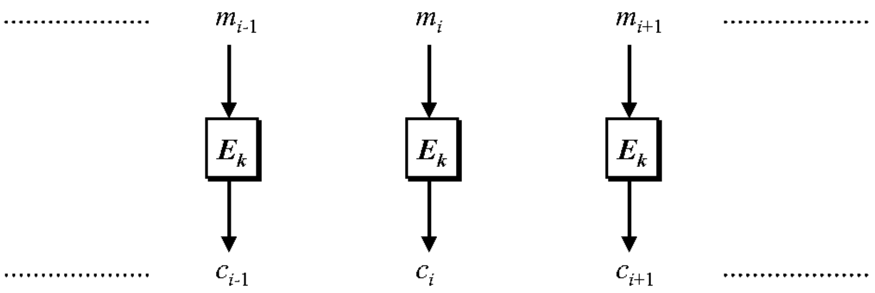
Stream ciphers, in contrast to the 'block' type, create an arbitrarily long stream of key material, which is combined with the plaintext bit-by-bit or character-by-character, somewhat like the one-time pad. In a stream cipher, the output stream is created based on a hidden internal state which changes as the cipher operates. That internal state is initially set up using the secret key material. RC4 is a widely used stream cipher.

Cryptographic hash functions are a third type of cryptographic algorithm. They take a message of any length as input, and output a short, fixed length hash which can be used in (for example) a digital signature. For good hash functions, an attacker cannot find two messages that produce the same hash. MD4 is a long-used hash function which is now broken; MD5, a strengthened variant of MD4, is also widely used but broken in practice.

Many of the modern cryptosystems that will be treated later in this book are block chiphers. For example, Data Encryption Standard DES operates on block of 64bits. Advanced Encryption Standard AES uses blocks of 128 bits. RSA uses blocks several hundred bits long, depending on the modulus used. All of these block lengths are long enough to be secure against attacks such as frequency analysis.

The standard way of using a block cipher is to convert blocks of plaintext to blocks of chiphertext, independently and one at a time. This is called the electronic codebook (ECB) mode. However, there are ways to use feedback from the blocks of chiphertext, in the encryption of subsequent blocks of plaintext. This leads to the cipher block chaining (CBC) mode and cipher feedback (CFB) mode of operation.

In ECB mode (see Figure 6.1), each plaintext block is encrypted independently with the block cipher.

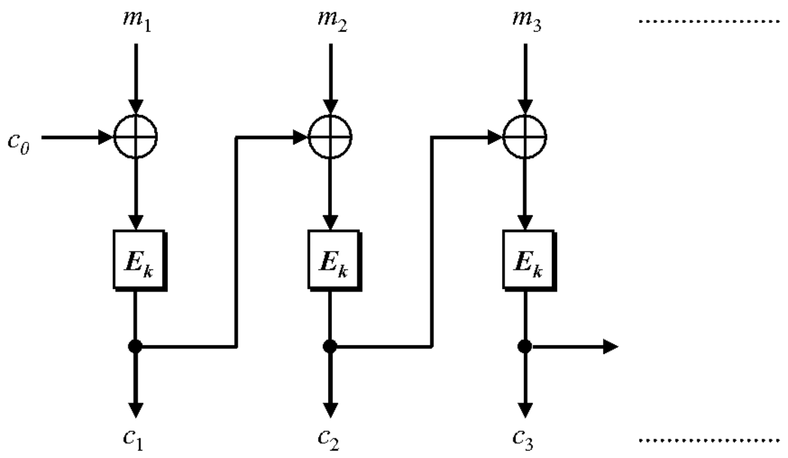


$$c_i = E_k(m_i) \quad m_i = D_k(c_i)$$

Figure 6.1 Electronic Code Book mode

ECB mode is as secure as the underlying block cipher. However, plaintext patterns are not concealed. Each identical block of plaintext gives an identical block of ciphertext. The plaintext can be easily manipulated by removing, repeating, or interchanging blocks. The speed of each encryption operation is identical to that of the block cipher. ECB allows easy parallelization to yield higher performance.

In CBC mode (see Figure 6.2), each plaintext block is XORed with the previous ciphertext block and then encrypted. An initialization vector c_0 is used as a "seed" for the process.



$$c_i = E_k(c_{i-1}\,\text{Å}m_i) \quad m_i = c_{i-1}\,\text{Å}D_k(c_i)$$

Figure 6.2 Cipher Block Chaining mode

CBC mode is as secure as the underlying block cipher against standard attacks. In addition, any patterns in the plaintext are concealed by the XORing of the previous ciphertext block with the plaintext block. Note also that the plaintext cannot be directly manipulated except by removal of blocks from the beginning or the end of the ciphertext. The initialization vector should be different for any two messages encrypted with the same key and is preferably randomly chosen. The speed of encryption is identical to that of the block cipher, but the encryption process cannot be easily parallelized, although the decryption process can be.

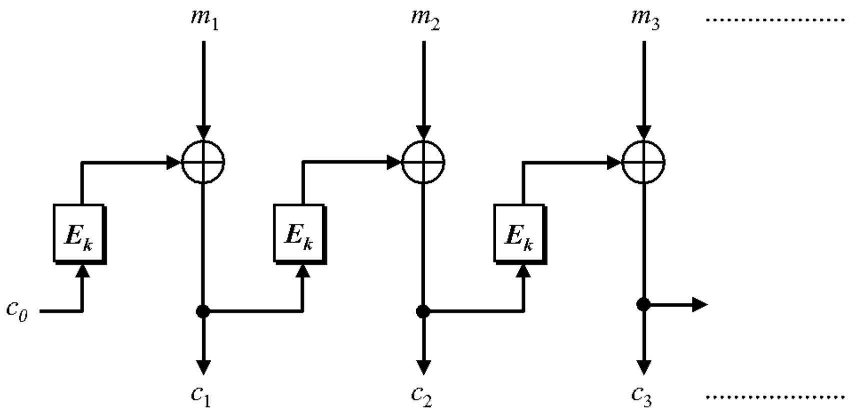
PCBC (Propagating Cipher Block Chaining) mode is a variation on the CBC mode of operation and is designed to extend or propagate a single bit error in the ciphertext. This allows errors in transmission to be captured and the resultant plaintext to be rejected. The method of encryption is given by

$c_i = E_k(c_{i\text{-}1}\ \text{Å}m_{i\text{-}1}\ \text{Å}m_i)$

and decryption is achieved by computing

$m_i = c_{i\text{-}1}\ \text{Å}m_{i\text{-}1}\ \text{Å}D_k(c_i)$.

In CFB mode (see Figure 6.3), the previous ciphertext block is encrypted and the output produced is combined with the plaintext block using XOR to produce the current ciphertext block. It is possible to define CFB mode so it uses feedback that is less than one full data block. An initialization vector c_0 is used as a "seed" for the process.



$c_i = E_k(c_{i\text{-}1})\ \text{Å}m_i \quad m_i = E_k(c_{i\text{-}1})\ \text{Å}c_i$

Figure 6.3 Cipher Feedback mode

CFB mode is as secure as the underlying cipher and plaintext patterns are concealed in the ciphertext by the use of the XOR operation. Plaintext cannot be manipulated directly except by the removal of blocks from the beginning or the end of the ciphertext; see next question for some additional comments. With CFB mode and full feedback, when two ciphertext blocks are identical, the outputs from the block cipher operation at the next step are also identical. This allows information about plaintext blocks to leak. The security considerations for the initialization vector are the same as in CBC mode. Instead, the last ciphertext block can be attacked.

When using full feedback, the speed of encryption is identical to that of the block cipher, but the encryption process cannot be easily parallelized.