CS 161 Computer Security
Fall 2005 Joseph/Tygar/Vazirani/Wagner **Notes 9**

1 Block Ciphers:

In symmetric encryption schemes, Alice and Bob share a random key and use this single key to repeatedly exchange information securely despite the existence of an eavesdropping adversary, Eve. The block cipher is a fundamental building block in implementing a symmetric encryption scheme.

In a block cipher, Alice and Bob share a k bit random key K, and use this to encrypt an n bit message into an n bit ciphertext. In mathematical notation this can be said as follows. There is an encryption function $E : \{0,1\}^k \times \{0,1\}^n \to \{0,1\}^n$. Once we fix the key K, we get a function mapping n bits to n bits: $E_K : \{0,1\}^n \to \{0,1\}^n$ defined by $E_K(M) = E(K,M)$. E_K is required to be a permutation on the n bit strings. The inverse mapping of this permutation D_K is the decryption algorithm. $D_K(E_K(M)) = M$.

The Data Encryption Standard (DES) is an example of a block cipher. It was designed by IBM in 1974 in response to a request from NIST for an encryption algorithm that could be standardized.

DES uses a key length of $k = 56$ and a block length of $n = 64$. It was designed to have extremely fast VLSI implementations. In terms of security, DES has proved to be an impressively strong algorithm. After all these years, the best practical attack known is exhaustive key search (a symmetry in the key structure can be used to halve the search space) which requires 2^{55} computations. Thus DES behaves very differently than the one-time pad. Even given a very large number of plaintext, ciphertext pairs, there appears to be no effective way to decrypt any new ciphertexts. The best such technique for DES is through linear cryptanalysis and requires 2^{42} such chosen plaintext-ciphertext pairs. We will formalize these security properties of DES by saying that the function E_K for a randomly chosen key K "behaves like" a random permutation on the n bit strings.

Formally, we shall measure the security of the block cipher by performing the following experiment: suppose that the adversary, Eve, is given a box which contains either (I) the block cipher with a random key or (II) a uniformly random permutation on n bits. Eve is allowed T steps in which to play with the box to guess whether it type I or type II. Define the advantage of Eve to be Adv(Eve) = Pr[Eve guesses I when the box is of type I] $-$ Pr[Eve guesses I when the box is of type II]. If the maximum advantage over all guessing strategies for the adversary is ε, then we say that the block cipher is (T,ε) secure. For DES, the above discussion says that if we want $\varepsilon = 1$ we need $T/T_{DES} \geq 2^{55}$, where T_{DES} is the time required to run DES on a single input. In general there is a tradeoff between T and ε, and so we could say that $\frac{T}{T_{DES}\varepsilon} \geq 2^{55}$. In some sense $log\ T/\varepsilon$ is the effective key length of the block code. We could also write a more detailed tradeoff in terms of the number q of plaintext-ciphertext pairs available. In this case we would write: $\varepsilon \leq \frac{T/T_{DES}}{2^{55}} + \frac{q}{2^{42}}$.

In 1998 NIST announced a competition for a new block cipher. One motivation is that the block length of $n = 64$ is just too short to guarantee security. Another motivation is speed - DES was really designed for hardware implementation, and is quite slow in software. In the summer of 2001 NIST announced their choice - the Advanced Encryption Standard (AES) which was designed by Joan Daemen and Vincent Rijmen, both from Belgium. AES has a block length of $n = 128$ bits and a key length k that may be either 128, 192 or 256 bits.

2 Symmetric Encryption Schemes:

A symmetric encryption scheme allows Alice and Bob to privately exchange a sequence of messages in the presence of an eavesdropper Eve. We will assume that Alice and Bob share a random secret key K. How Alice and Bob managed to share a key without the adversary's knowledge is not going to be our concern here. The encryption scheme consists of an encryption algorithm $\mathscr{E}$ that takes as input the key K and the plaintext message $M \in \{0,1\}^*$, and outputs the ciphertext. The decryption algorithm $\mathscr{D}$ takes as input the key and the ciphertext and reconstructs the plaintext message M. In general the encryption algorithm builds upon a block cipher to accomplish two goals: one is to show how to encrypt arbitrarily long messages using a fixed length block cipher. The other is to make sure that if the same message is sent twice, the ciphertext in the two transmissions is not the same. The encryption algorithm to achieve these goals can either be randomized or stateful - it either flips coins during its execution, or its operation depends upon some state information. The decryption algorithm is neither randomized nor stateful.

ECB Mode (Electronic Code Book): In this mode the plaintext M is simply broken into n bit blocks $M_1, \dots M_l$, and each block is encoded using the block cipher: $C_i = E_K(M_i)$. The ciphertext is just a concatenation of these individual blocks: $C = C_1 \cdot C_2 \cdots C_l$. This scheme is adequate for simple tasks such as encrypting PINs for cash machine systems. However any redundancy in the blocks will show through and allow the eavesdropper to deduce information about the plaintext. We will discuss this in more detail after formalizing the notion of the security of a symmetric encryption scheme below.

CBC Mode (Cipher Block Chaining): This is a popular mode for commercial applications. A random n bit string, the initial vector or IV is selected. Define $C(0) = E_K(IV)$. The i^{th} encrypted block $C_i = E_K(C_{i-1} \oplus M_i)$. The ciphertext is the concatenation of the initial vector and these individual blocks: $C = IV \cdot C_1 \cdot C_2 \cdots C_l$. The CBC mode does provide strong security guarantees on the privacy of the plaintext message. This is formalized later in this lecture and explored further in the homework.

OFB Mode (Output Feedback Mode): In this mode, the initial vector IV is repeatedly encrypted to obtain a set of keys K_i as follows: $K_0 = IV$ and $K_i = E_K(K_{i-1})$. These keys K_i are now used as keys for a one-time pad, so that $C_i = K_i \oplus M_i$. The ciphertext is the concatenation of the initial vector and these individual blocks: $C = IV \cdot C_1 \cdot C_2 \cdots C_l$. This scheme suffers from an important weakness — the message is malleable. i.e. suppose that the adversary happens to know that the j^{th} block of the message specifies the amount of money being transferred to his account from the bank and is 100. Since he knows both M_j and C_j, he can determine K_j. He can then substitute any n bit block in place of M_j and get a new ciphertext where the 100 is replaced by any amount of his choice.

Counter Encryption: One drawback of the CBC mode is that successive blocks must be encrypted sequentially. For high speed applications it is useful to parallelize these computations. This is easily achieved by encrypting a counter initialized to IV to obrain a set of keys $K_i = E_K(IV + i)$. As before the block M_i is then encrypted simply as $K_i \oplus M_i$.

3 Notions of Security:

The CBC mode can be shown to be secure under a chosen plaintext attack. Let us first define what it means for a symmetric encryption scheme to be secure. We consider the following experiment: the adversary, Eve, has access to a box which takes as input a pair of plaintext messages (M, M') and if the box is of type I it outputs the CBC encryption of M and if it is of type II it outputs the CBC encryption of M'. The adversary is allowed to play with the box for the available time and must guess which type it is. As before, the adversary's advantage is defined to be: Adv(Eve) = Pr[Eve guesses I — box I] - Pr[Eve guesses I — box

II]. If we can show that the adversary's advantage is negligibly small in the available time, then we will have shown that the encryption scheme is secure.

To show that ECB mode encryption is not secure in this sense, let x and y be any two distinct n bit strings. Let $M = x \cdot x$ and $M' = x \cdot y$. Then we claim that the output of the box is of the form $c \cdot c$ iff the box is of type I.

The homework outlines a proof of the fact that the CBC mode encryption is secure under chosen plaintext attack. If the total number of times the block cipher is invoked j times by the box during the T units of time that the adversary plays with it, and if the block cipher is (T, ε)-secure then $Adv(Eve) \leq \varepsilon + 2j^2/(2^n - j)$