INAP
INSTITUTO NACIONAL DE ADMINISTRACIÓN PÚBLICA

Curso *online*: **Seguridad en Redes WAN e Internet**

Módulo2 CRIPTOGRAFÍA

Autores: Daniel Díaz y Andrés Marín

Índice de contenidos

Capítulo 1 INTRODUCCIÓN A LA CRIPTOGRAFÍA

El término criptología deriva del griego "kriptos", que significa oculto. Se puede definir como la disciplina que estudia los principios, métodos y medios de ocultar la información contenida en un mensaje. Dentro de esta ciencia se pueden diferenciar dos partes: criptografía, protección de la información a través de su codificación mediante claves y criptoanálisis, supresión de esa protección sin el conocimiento de las claves.

Desde muy antiguo, todas aquellas personas que poseían informaciones privilegiadas (reyes, militares, sacerdotes, etc.) utilizaron métodos y mecanismos para salvaguardarlas de los posibles ataques. Se pretendía mantener la información en secreto o, por así decirlo, darle un carácter confidencial. El razonamiento para ello es sencillo: puesto que la información puede proporcionar beneficios, siempre existirá alguien que ponga todos los medios a su alcance para obtenerla. Todo bien tangible lleva aparejado un conjunto de riesgos por parte de su poseedor y un conjunto de amenazas por parte de quienes quieren poseerlo.

Como en muchas áreas científicas, el mayor desarrollo de la criptología tuvo lugar durante las dos guerras mundiales, debido a la necesidad de establecer comunicaciones militares y diplomáticas secretas utilizando las nuevas tecnologías del momento, como la telegrafía y la radiotecnia. La Segunda Guerra Mundial marcó un hito en el desarrollo de las tecnologías de la información, a lo largo de ella se impulsó el desarrollo de grandes máquinas computadoras, como es el caso de la máquina ULTRA, desarrollada por los británicos para destruir el cifrado de la máquina ENIGMA alemana.

Ahora bien, es en la segunda mitad del siglo XX, con el desarrollo de la informática, cuando han surgido nuevas aplicaciones de la criptografía, debido fundamentalmente al manejo de gran cantidad de información. En algunos casos, como en las redes informáticas, dicha información está a disposición de muchos usuarios, lo cual plantea la necesidad de que los datos estén protegidos durante su transmisión y almacenamiento.

Al mismo tiempo, el desarrollo de la informática produjo un cambio radical en el concepto de seguridad de los sistemas criptográficos, pues aquellos que eran supuestamente seguros frente a procedimientos manuales sucumbieron ante la eficacia de los ordenadores. De esta forma, la supuesta seguridad de los sistemas clásicos ha tenido que ser sustituida por una seguridad matemática y computacionalmente demostrable con los sistemas modernos.

Aún más, la criptografía es la base de los mecanismos de autenticación y no repudio, de capital importancia en el comercio electrónico y, en general, en la seguridad en redes. Así mismo, si tradicionalmente esta disciplina se aplicó exclusivamente a información textual, hoy en día se usa para preservar informaciones de cualquier naturaleza: voz, imágenes fijas, vídeo, etc.

1.1. CRIPTOSISTEMAS

Se llama cifrado (o transformación criptográfica) a una transformación del texto original (llamado también texto en claro) que lo convierte en el llamado texto cifrado o criptograma. Análogamente, se llama descifrado a la transformación que permite recuperar el texto original a partir del texto cifrado.

Para cifrar se pueden utilizar dos métodos:

1. Un algoritmo secreto
2. Un algoritmo público y clave secreta, esta clave es imprescindible para cifrar y descifrar

Todos los sistemas actuales, excepto algunos militares, utilizan algoritmos públicos y claves secretas, debido a que:

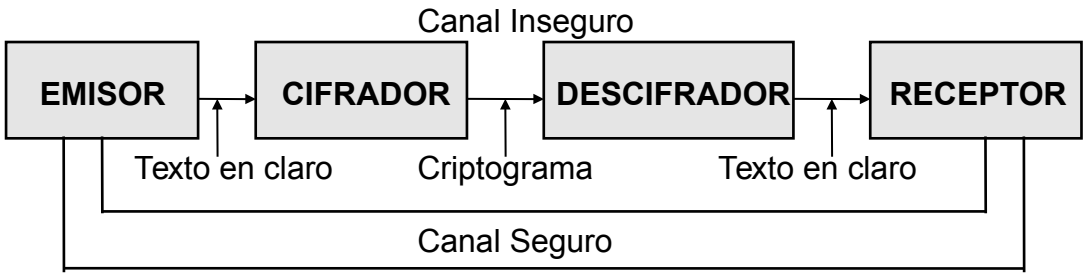
- Los algoritmos públicos están más probados, ya que toda la comunidad científica puede trabajar sobre ellos buscando fallos o agujeros.
- Es más fácil y más seguro transmitir una clave que todo el funcionamiento de un algoritmo.
- Existen soluciones hardware y software para los algoritmos públicos, de está manera su aplicación resulta más barata.

En las comunicaciones civiles, se emplean criptosistemas conocidos y perfectamente estandarizados. Estos criptosistemas han sido sometidos al escrutinio público, ampliamente estudiados y criticados por la comunidad criptográfica. Por tanto, los algoritmos conocidos gozan de ciertas garantías de robustez de las que no gozan los algoritmos secretos, por no haber pasado el filtro de la crítica pública. En este tema nos referiremos exclusivamente a los criptosistemas con algoritmos conocidos.

Un criptosistema es un conjunto de sistemas de cifrado y descifrado y los correspondientes protocolos de transmisión de claves. Un criptosistema está constituido por los siguientes elementos funcionales:

- Emisor. Genera un mensaje, texto en claro
- Cifrador. Transforma el texto en claro en un mensaje ininteligible, texto cifrado o criptograma
- Descifrador. Realiza la función inversa del cifrador, transforma el criptograma en texto en claro
- Receptor
- Protocolo de intercambio de claves

Un esquema simplificado de un criptosistema puede ser el siguiente:



El elemento fundamental de un criptosistema es el cifrador, que está compuesto por un dispositivo físico o por un programa, que implementa el algoritmo de cifrado, que suele ser el mismo que el de descifrado. Usualmente, este algoritmo está constituido por una función matemática que depende de un parámetro, la clave de cifrado o de descifrado según corresponda.

Formalmente todo criptosistema consta de cinco componentes:

- Espacio de mensajes: $M = \{m_1, m_2, ...\}$
  Todos los textos en claro, formables a partir de un alfabeto y unas reglas sintácticas y semánticas

- Espacio de textos cifrados: $C = \{c_1, c_2, ...\}$
  Donde el alfabeto puede ser el mismo o distinto que el del Espacio de mensajes

- Espacio de las claves: $K = \{k_1, k_2, ...\}$
- Familia de transformaciones de cifrado: $E_k : M \rightarrow C$
  Los distintos algoritmos que se pueden utilizar para cifrar los mensajes en claro

  Donde $k \in K$ es el parámetro que define cada transformación de la familia

- Familia de transformaciones de descifrado: $D_k : C \rightarrow M$

Se debe cumplir que $E_k (D_k (c)) = c$, si se cumple que $m = D_k (E_k (m))$, entonces D_k es la inversa de E_k, pero no siempre se cumple la viceversa.

O sea: $E (k,m) = c$; $D (k',c) = m$ Donde k y k' no tienen por qué ser iguales

Todo criptosistema debe cumplir los siguientes condicionantes:

- Las transformaciones E_k y D_k deben ser computacionalmente eficientes (y no sólo eficaces) para todas las claves.
- La seguridad del sistema debe depender exclusivamente del secreto de las claves y no de las funciones E y D, que se supone son públicas

1.2. TÉCNICAS BÁSICAS

La principal amenaza de los criptosistemas clásicos proviene de la alta redundancia de los lenguajes naturales, y de sus normas sintácticas y gramaticales, que permite diversos tipos de ataques. Por ejemplo, en castellano: la "q" siempre va seguida de la "u"; la "e" es la letra más frecuente; no aparecen tres consonantes seguidas; se conoce la frecuencia de aparición de cada letra en diferentes contextos: coloquial, culto,... En el diseño de cifradores hay dos principios básicos:

- CONFUSIÓN

  Consiste en establecer una relación lo más compleja posible entre la clave y el cifrado, de forma que el texto cifrado no conserve gran parte de las normas que rigen el lenguaje del texto original. Por ejemplo, alterando la posición de los caracteres, técnica de permutación, se destruyen los diagramas, triagramas, etc.

- DIFUSIÓN

  Consiste en distribuir las propiedades estadísticas del mensaje en claro sobre todo el texto cifrado, mediante la transposición, que evita los criptoanálisis basados en las frecuencias de las n-palabras o haciendo que cada carácter del texto cifrado dependa de un gran número de caracteres del texto original.

  Si los textos en claro y cifrado estuvieran en binario, lo ideal sería que cada bit de la salida dependiera de todos los bits de la entrada, a estos algoritmos se les llama completos.

Para dar mayor fortaleza al criptograma la mayoría de los cifradores utilizan conjuntamente la confusión y la difusión, como por ejemplo ocurre con el DES (*Data Encryption Standard*)

1.3. TIPOS DE CRIPTOSISTEMAS

Atendiendo a diferentes criterios se pueden establecer las siguientes clasificaciones:

1.3.1. CRIPTOSISTEMAS SIMÉTRICOS Y ASIMÉTRICOS

Según la relación existente entre las claves de cifrado y de descifrado, los criptosistemas se pueden dividir en dos grandes grupos:

a) Los criptosistemas simétricos, también llamados de clave única o clave secreta, utilizan la misma clave para cifrar que para descifrar, u obtenible una de la otra.

$c_i = E(k_i, m_i)$ (en el emisor) $m_i = D(k_i, c_i)$ (en el receptor)

La consistencia comporta que: $m_i = D(k_i, E(k_i, m_i))$

Hasta 1976 eran los únicos criptosistemas conocidos.

- Desventajas:
  - La distribución de claves exige un canal seguro, el canal del criptosistema es por hipótesis vulnerable
  - El número de claves implicadas es muy grande: (n (n-1)) / 2, siendo n el número de usuarios
- Ventajas:
  - Simetría. Los papeles de emisor y receptor son intercambiables
  - Mantiene la confidencialidad y confiere autenticidad al emisor (requiere conocer la clave)

b) Los criptosistemas asimétricos o de clave pública, constan por cada usuario, de una clave pública K_u de general conocimiento y una clave privada K_v secreta. Si suponemos que K_u y K_v son las claves pública y privada del emisor, para cifrar/descifrar un mensaje se podría:

EMISOR		RECEPTOR
$c_i = E(K_u, m_i)$	$\rightarrow$	$m_i = D(K_v, c_i)$

La consistencia comporta que: $m_i = D(K_v, E(K_u, m_i))$

Ambas claves son intercambiables se puede cifrar con una y descifrar con la otra y viceversa.

Una de las características principales es que de una clave no se puede deducir la otra, a no ser que se tenga algún dato adicional que también tiene que ser secreto. Estos datos adicionales se suelen destruir en el momento en que se genera el par clave pública -clave privada

Desventaja. No hay autenticación del emisor

Ventajas:

- No exige un canal seguro para la distribución de la clave pública
- El número de claves a gestionar es menor que con los sistemas simétricos (2n claves, n públicas y n privadas)

1.3.2. CIFRADORES DE FLUJO Y CIFRADORES DE BLOQUE

Según la fuente que genera el texto, los cifradores simétricos se pueden clasificar en dos grandes grupos:

a) Los cifradores de flujo, cifran cada carácter del texto en claro con el correspondiente símbolo de la clave, cifran carácter a carácter. Son muy rápidos pero no tienen difusión. El criptosistema se expresa:
$E_k (m) = E_{k1} (m_{1)}, E_{k2} (m_2), ... E_{ki} (m_i)$

b) Los cifradores de bloque, descomponen el texto en claro en bloques de caracteres o bits (n-palabras), de igual longitud cifrando cada bloque con la misma clave. Son más lentos que los de flujo pero tienen difusión, un carácter en claro se difunde en varios cifrados
$c = E_k (m_1) E_k (m_2) ..... E_k (m_n)$ donde k es la clave del bloque

El DES, adoptado en 1976 como estándar por el gobierno de EEUU, es el cifrador de bloques mas conocido que utiliza un cifrado producto de transposiciones y sustituciones.

1.3.3. CRIPTOSISTEMAS REVERSIBLES E IRREVERSIBLES

a) Los reversibles, son los más comunes y parten de que conociendo el algoritmo correspondiente y la clave utilizada, se puede obtener del mensaje cifrado el mensaje original.

b) Los irreversibles, se basan en algoritmos no invertibles, sólo trabajan en un sentido. Dado el criptograma no es posible obtener el mensaje en claro en un tiempo razonable, ni aún con los ordenadores más potentes de hoy en día. Se usan en aplicaciones que no requieren descifrar los criptogramas. Por ejemplo sirven para determinar si un cierto mensaje coincide con otro almacenado en memoria, mecanismo utilizado para las contraseñas de acceso.

1.3.4. CRIPTOANÁLISIS (ATAQUES)

El criptoanálisis, como ya se ha dicho, es la parte opuesta a la criptografía, donde se trata de obtener, sin conocer las claves, los mensajes en claro de los mensajes cifrados.

Las posibilidades de éxito de un ataque a un criptosistema dependen en gran medida de las circunstancias que lo rodean y de la información disponible. El criptoanálisis abarca diversas técnicas, algunas de ellas no dependen del conocimiento del algoritmo sino que mediante sistemas de aproximación matemática se puede descubrir el texto en claro o la clave. Existen dos mecanismos básicos de ataque:

- Realizar un análisis estadístico del mensaje
- Búsqueda exhaustiva, también conocido como ataque de fuerza bruta. Cuando se conoce el algoritmo de cifrado y descifrado, consiste en hacer una prueba exhaustiva con todas las claves del espacio de claves

Un criptosistema puede implementar dos tipos de secreto:

- Teórico o incondicional, se basa en que la información disponible por el enemigo no es suficiente para romper el sistema. Es seguro contra cualquier enemigo aunque tenga recursos y tiempo ilimitados.
- Práctico, se mide de acuerdo con la complejidad computacional del criptoanálisis si es seguro contra aquellos enemigos que tengan menos de una cantidad de tiempo y/o recursos.

A modo de resumen podemos establecer que:

- El secreto perfecto es teóricamente posible. Se puede conseguir por dos vías: una con una clave cuya longitud sea igual a la del mensaje y otra con un mensaje aleatorio.
- No existen sistemas prácticos completamente seguros, siempre se pueden violar probando todas las claves posibles. Por lo tanto, en criptografía se buscan sistemas que cumplan una de siguientes condiciones:
  - El precio para romperlo es más caro que el valor de la información.
  - El tiempo necesario para romperlo es más largo que el tiempo de vida de la información.

Capítulo 2 MÉTODOS CRIPTOGRÁFICOS CLÁSICOS

Los métodos denominados clásicos, usados hasta 1949, constituyen el fundamento de los métodos actuales, por lo que se considera muy conveniente su conocimiento, aunque, salvo una excepción, no son prácticamente usados hoy en día. Los métodos criptográficos clásicos se basan en:

- La sustitución, consiste en sustituir las unidades de texto original por otras
  La sustitución definida mediante una clave k= (γ_0, γ_1, γ_2,..γ_n....) es la transformación criptográfica E_k, que cifra la n-palabra original

  (x_0, x_2,...,x_{n-1},....) como la n-palabra (y_0,y_1,y_2,....,y_{n-1}) con $y_i=\gamma_i(x_i)$, $0 \leq i \leq n$, $\forall$ n= 1,2, ...

  Se llama sustitución monoalfabética si γ_i es la misma $\forall$ i=1,2,....., mientras que en caso contrario se llama sustitución polialfabética

- La transposición o permutación, consiste en crear el texto cifrado simplemente desordenando las unidades que forman el texto original.

2.1. MÉTODOS DE SUSTITUCIÓN

2.1.1. SUSTITUCIÓN SIMPLE O MONOALFABETO

Es la técnica más sencilla y consiste en sustituir cada carácter (letra) por otro del mismo alfabeto, de acuerdo con:

$$E(m) = (am \pm b) \bmod (n)$$

Donde:

m: valor asociado a cada letra A B C D E F G H I J K L .....

1 2 3 4 5 6 7 8 9 10 11 12.....

b: desplazamiento

a: intervalo de separación (para eliminar la aparición de letras consecutivas)

n: número de letras del alfabeto

a y b constituyen la clave

Para que el algoritmo sea biunívoco se debe cumplir que el m.c.d.(a,n) = 1, en caso contrario diferentes letras del alfabeto en claro darían lugar a la misma letra en el alfabeto equivalente.

Ejemplo: Con el alfabeto castellano, para a = 5 y b = 15:

PLANTA NUCLEAR → RTSDLS DPCTMSB

Cuando a = 1 y b = 3 se le denomina tipo Cesar, se utilizó en la época de Julio Cesar

2.1.2. SUSTITUCIÓN POLIALFABÉTICA

La sustitución polialfebética se define mediante una clave k = (γ_0, γ_1, ...), que contiene al menos dos sustituciones γi distintas. Utiliza varias sustituciones simples en el cifrado de un mensaje. Las letras de una palabra clave definen los desplazamientos de los alfabetos en sustituciones tipo "Cesar", de acuerdo con la siguiente transformación:

$$E = E(m_j) = (m_j + K_i) \bmod 27$$

Donde:

K_i es el desplazamiento de cada letra de la clave

$i \rightarrow 1, d$ (d es la longitud de la clave)

$j \rightarrow 1, r$ (r es la longitud del mensaje)

Destruye la frecuencia de las letras del texto en claro, sustituye una misma letra del texto en claro por letras diferentes del alfabeto utilizado. Ejemplo:

Mensaje:	PLANTA	ATÓMICA
Clave:	SOL SOL	SOLSOLS
Criptograma	IZLFIL	SIZEWNS

2.2. MÉTODOS DE TRANSPOSICIÓN O PERMUTACIÓN

No se sustituyen las letras, se cambia su posición dentro del mensaje, al no romper la frecuencia del lenguaje son fácilmente destructibles, entre ellos cabe citar: la Escitala (bastón de mando de los generales lacedomios, siglo V a. C.) que consiste en una cinta de pergamino enrollada en un cilindro cuyo diámetro determina la clave; el Posicionamiento en "zigzag" o "valla de rieles"; la Distribución en figuras geométricas y Funciones matemáticas de permutación

2.3. MÉTODOS BASADOS EN CÁLCULOS NUMÉRICOS O LÓGICOS

2.3.1. MÉTODOS ARITMÉTICOS

Utilizan cálculos aritméticos sencillos para realizar las conversiones, dan mucha seguridad pues destruyen la frecuencia del lenguaje y son fáciles de implantar:

a) Operaciones de suma y multiplicación

Los caracteres del texto en claro y de la clave se codifican en base n, siendo n = nº de caracteres del alfabeto. Para cifrar se suman o se multiplican los números resultantes y luego para descifrar se resta o divide el criptograma por el número de la clave. Se utilizan la suma o la multiplicación porque son operaciones que tienen inversa, la resta y la división respectivamente.

b) Cambio de base

El mensaje en claro se divide en bloques, se transforma a números y se cambia de base para codificarlo, para decodificarlo se convierten los números a la base inicial y se pasan a caracteres. Por ejemplo:

SOS $\rightarrow$ 191519 $\rightarrow$ 22112034 $_{(5}$

2.3.2. TRANSFORMACIONES LÓGICAS BOOLEANAS

Son transformaciones muy apropiadas para implantar en los sistemas informáticos. Las operaciones a utilizar deben tener inversa. De las 16 operaciones booleanas sólo 3 tienen inversa:

"-" negación $E(M) = \overline{M} \rightarrow M = \overline{E(M)}$

"⊕" O exclusivo $E_K(M) = M \oplus K \rightarrow M = E_K(M) \oplus K$

"≡" equivalencia $E_K(M) = M \equiv K \rightarrow M = E_K(M) \equiv K$

2.3.3. TRANSFORMACIONES MATRICIALES

Son transformaciones muy seguras, porque rompen la frecuencia del lenguaje, pero muy laboriosas. El mensaje original se transforma en una serie de ceros y unos que se colocan en una matriz de r filas y s columnas, la matriz resultante se suma o multiplica por otra matriz clave:

$(C) = (M) + (K)$ o $(C) = (M) * (K)$

$(M) = (C) - (K)$ $(M) = (C) * (K)^{-1}$

La operación de descifrado es posible porque la suma y multiplicación de matrices poseen operaciones inversas bajo ciertas condiciones:

- En la suma las matrices deben tener las mismas dimensiones
- En la multiplicación, que K posea una inversa única, por lo que la matriz debe ser cuadrada y no singular.

2.4. CIFRADO PRODUCTO

Es el método más usual de crear confusión y difusión. Consiste en la aplicación sucesiva de diferentes cifradores, cada uno de los cuales actúa sobre el criptograma obtenido de la aplicación del cifrado anterior (texto en claro para él), normalmente lo componen un número alto de cifradores. Hay que tener presente que cifrar dos veces con una misma función, con clave distinta, se obtiene el mismo resultado que cifrando una sola vez con una determinada clave.

El cifrado producto se puede expresar:

$C = E(M) = E_n(\ldots\ldots E_2((E_1(M))\ldots\ldots)$

Donde E_i son los distintos métodos criptográficos empleados: transposiciones sustituciones, agrupaciones, etc.. Descifrándose el texto cifrado mediante:

$$M = D_1(D_2 \ldots.(D_{n-1}(D_n(C)\ldots.)$$

Donde D_1 representa el método de descifrado correspondiente al método de cifrado E_1

Un ejemplo clásico de cifrado producto es el LUCIFER, que posteriormente evolucionó al DES, que está compuesto de sucesivos cifrados de sustitución y transposición.

Capítulo 3 CRIPTOSISTEMAS SIMÉTRICOS O DE "CLAVE SECRETA"

Las principales características de este tipo de criptosistemas son:

- Utilizar para el cifrado y descifrado la misma clave: “simétricos”
- La clave se debe mantener secreta. Toda la seguridad se basa en la privacidad de esta clave secreta, por ello la clave se debe transmitir por un canal seguro.
- Proporcionan secreto y autenticidad a la información cifrada
- Requieren gran cantidad de claves (el cuadrado del nº corresponsales). El mayor inconveniente que plantea su uso es la distribución de claves.

Los algoritmos simétricos, actualmente utilizados en sistemas de información, cifran bloques de texto de tamaño constante o variable según el tipo de algoritmo. Existen 4 formas básicas de funcionamiento:

- *Electronic CodeBook (ECB).* (Libro de Código Electrónico). Cada bloque de texto se cifra por separado, usando la misma clave. Trabaja bien en canales de transmisión con interferencia, la alteración de algún bit sólo corromperá el bloque en que se encuentre. No permite detectar cuando se han insertado o eliminado porciones de un mensaje
- *Cipher Block Chainning (CBC).* (Encadenamiento de Bloques de Cifrado). Los bloques del criptograma se relacionan entre ellos mediante funciones or-exclusiva (XOR). El texto en claro pasa primero un XOR con el valor cifrado del bloque previo, el resultado se cifra posteriormente usando la clave. Para el primer bloque se utilizan valores conocidos, vector de inicialización. La mayoría de las implantaciones del DES utilizan el modo CBC
- *Cipher FeedBack (CFB).* (Retroalimentación de Cifras). La salida se retroalimenta al mecanismo. Después de cada bloque se cifra, parte de él se recorre (con shift) dentro de un registro. El contenido de este registro se cifra con el valor de la clave usando el modo ECB, y esta salida se transforma con un XOR usando la cadena de datos para producir el resultado del cifrado.
- *Output FeedBack (OFB).* (Retroalimentación de Salidas). Igual que el CFB, se realiza una or-exclusiva entre caracteres o bits aislados del texto y las salidas del algoritmo. Pero éste

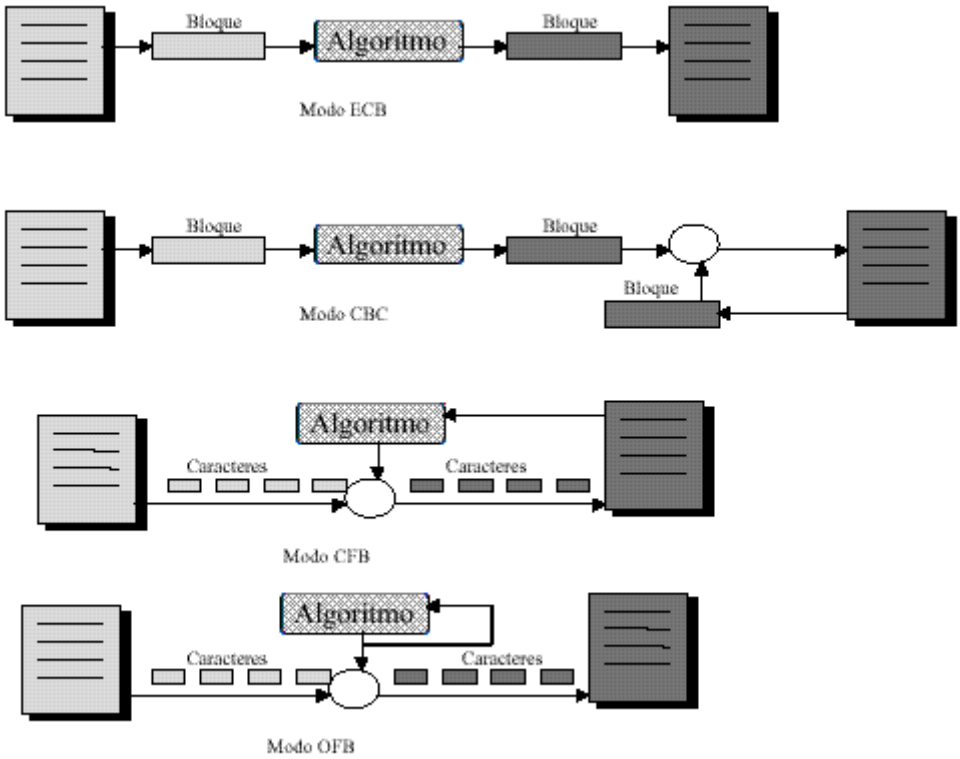
utiliza como entradas sus propias salidas, por lo tanto no depende del texto, es un generador de números aleatorios.

Los algoritmos simétricos más utilizados son AES, DES e IDEA

DES. El más antiguo. Actualmente se utiliza 3DES para poder utilizar claves de tamaño superior

IDEA (*International Data Encryption Algorithm*). Se utiliza mucho en sistemas europeos. No está sujeto a las leyes de ningún país.

AES. El más utilizado hoy en día.



3.1. ALGORITMO DES

El DES (*Data Encryption Standard*) es un algoritmo desarrollado originalmente por IBM a requerimiento del NBS (*National Bureau of Standard)* de los E.E.U.U, en la actualidad NIST, y posteriormente modificado y adoptado por el gobierno de EE.UU. en 1977 como estándar de cifrado de todas las informaciones sensibles no clasificadas. En 1981, ANSI (*American National Standar Institute*), adoptó el DES, con el nombre DEA (*Data Encryption Algorithm*), como norma en el sector privado (ANSI X3. 106) y para el cifrado en redes (ANSI X3. 105). Las normas ISO

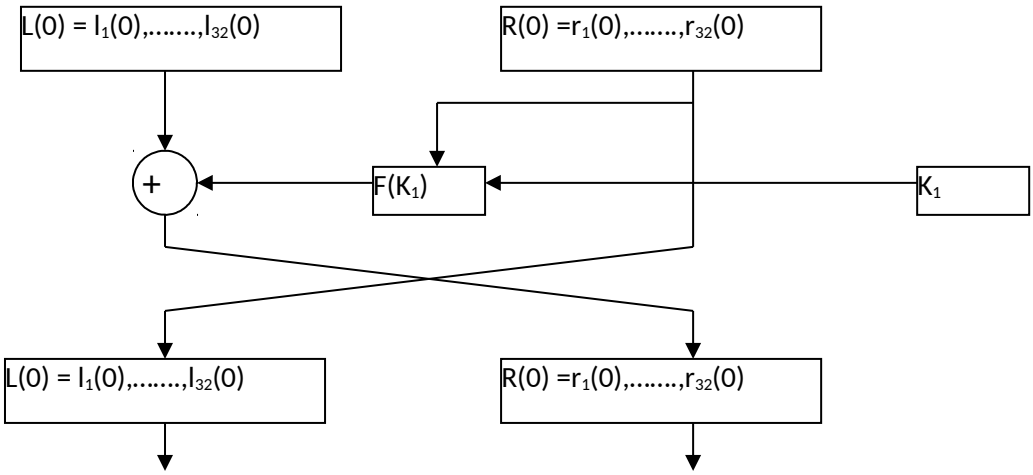
relativas al DES son: ISO 8382 (equivalente a ANSI X3.106-1981), ISO 9797, ISO 9798 e ISO 10118.

El algoritmo está formalmente descrito en FIPS-PUB 46-2 (*Federal Information Processing Systems Publication*). La norma exige que el DES se implemente mediante un circuito integrado electrónico, el chip del DES es un producto estratégico USA, no está permitida su exportación sin un permiso especial, y no se permite comercializar en USA chips fabricados en el exterior. Sin embargo la norma de la versión (DEA) estandarizada por ANSI no exige implementación en un circuito integrado y puede ser implementado por software.

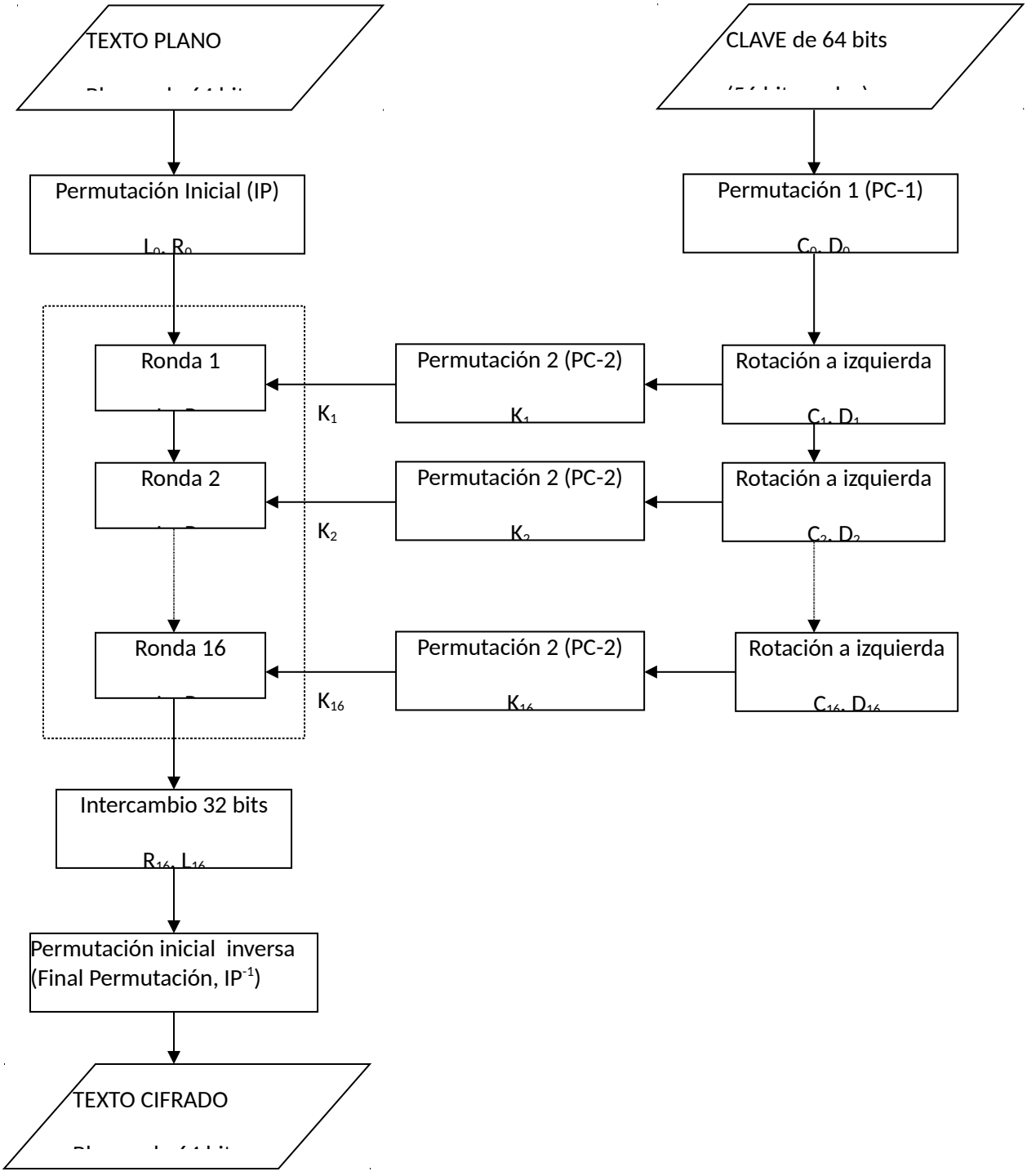
DES cifra, mediante permutaciones y sustituciones, bloques de datos de 64 bits (8 caracteres ASCII), con una clave de 64 bits, (de ellos 8 bits son de paridad, en realidad usa 56 bits, lo que da un total de 2^{56}, o sea 72.057.594.037.927.936 claves posibles), produciendo 64 bits cifrados.

El cifrado de DES consta de 19 etapas diferentes. La primera etapa es una transposición, una permutación inicial (IP) del texto plano de 64 bits, independientemente de la clave. La última etapa es otra transposición (IP-1), exactamente la inversa de la primera. La penúltima etapa intercambia los 32 bits de la izquierda con los 32 de la derecha. Las 16 etapas restantes son una Red de Fiestel de 16 rondas. Se denominan sistemas de Fiestel a los que dividen el bloque de datos en dos mitades y en cada vuelta de cifrado se trabaja con una de las dos mitades.

La operación de cada ronda consiste en sumar modulo 2 a la parte izquierda con una función $F(K_i)$ de la parte derecha, gobernada por una clave K_i, y después se intercambian las partes derecha e izquierda

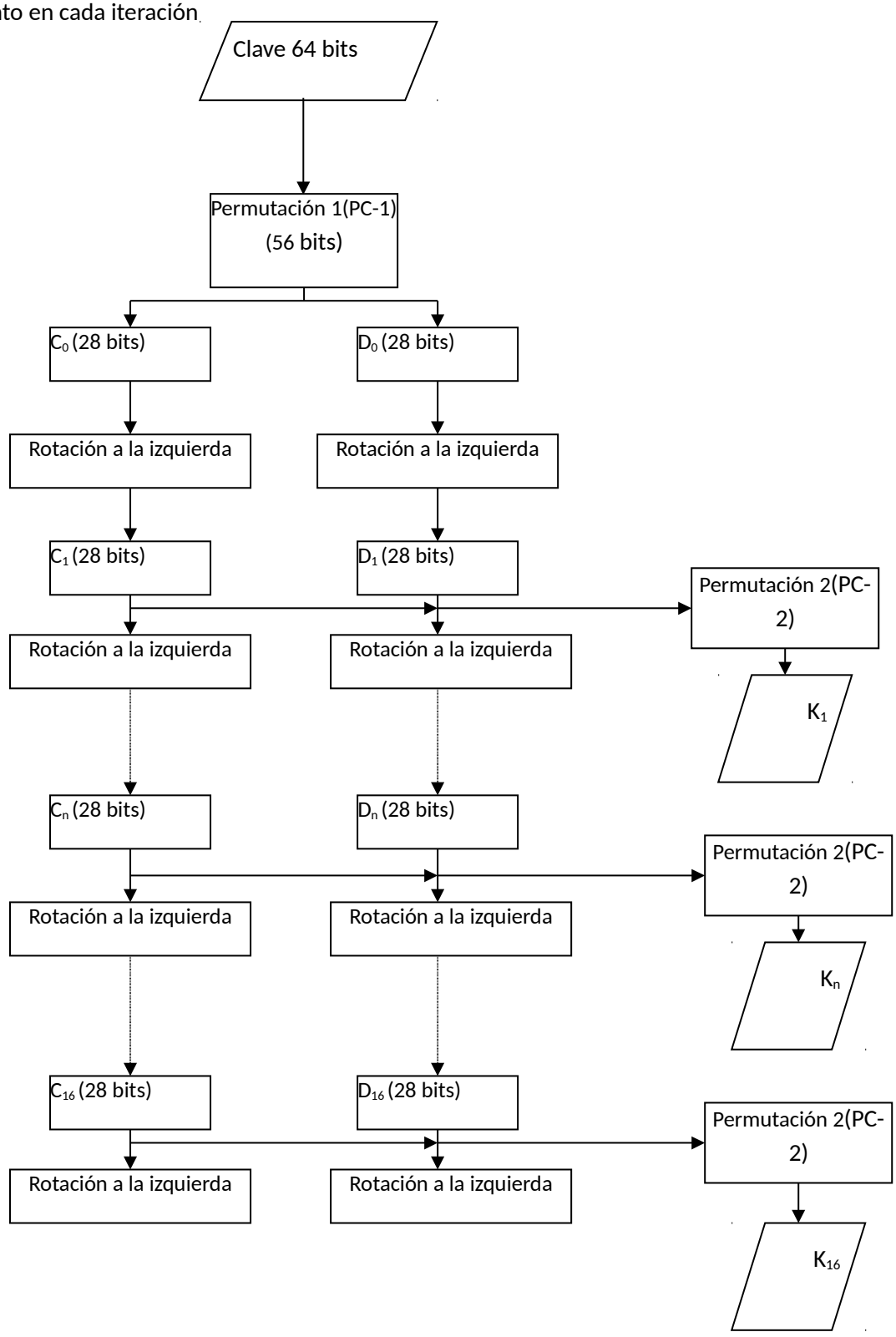


Primera ronda de la estructura del DES



Esquema general del algoritmo DES

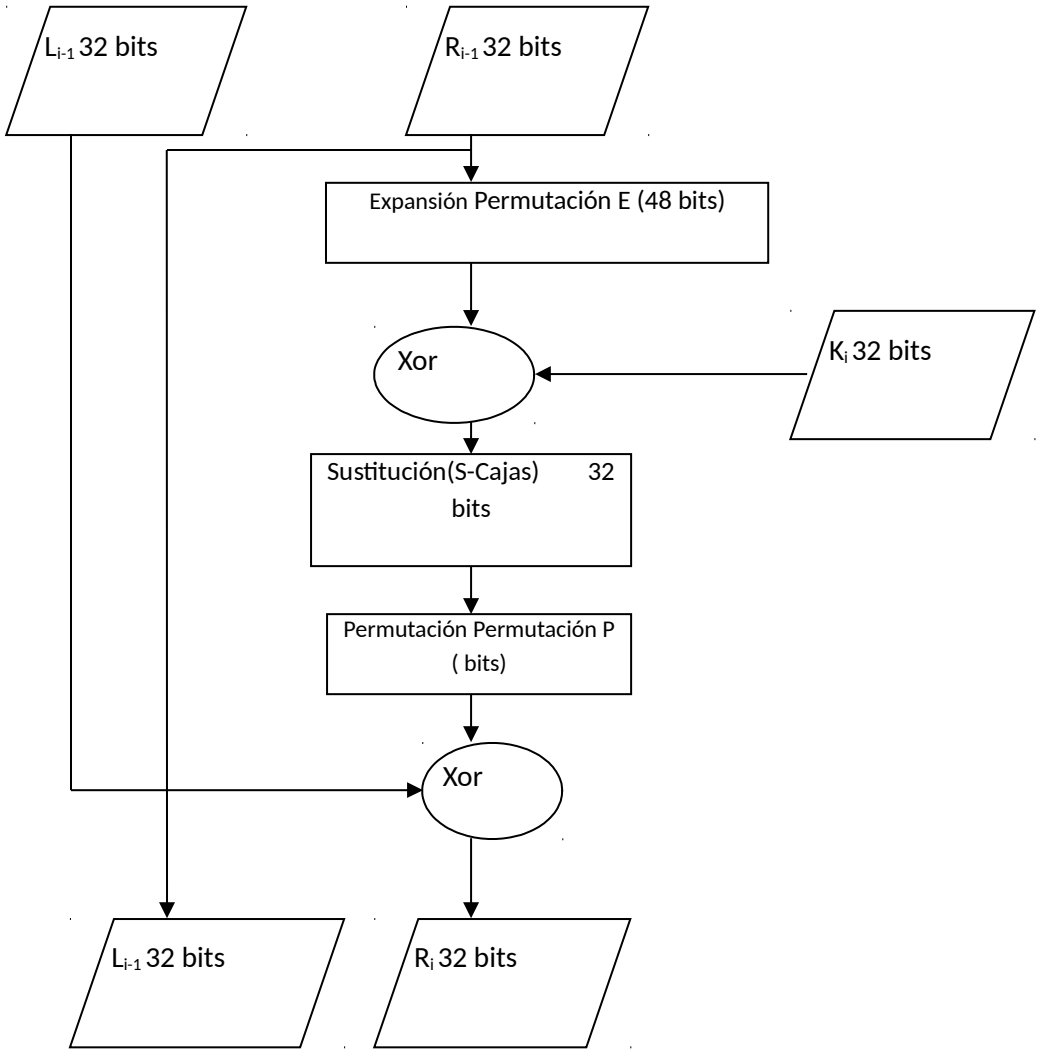
En cada una de las 16 iteraciones se emplea un valor de, K_i, obtenido a partir de la clave de 56 bits y distinto en cada iteración.



Cálculo de las subclaves, K_i

Se realiza una permutación inicial (PC-1) sobre la clave, y luego la clave obtenida se divide en dos mitades de 28 bits, cada una de las cuales se rota a izquierda un número de bits determinado que no siempre es el mismo. K_i se deriva de la elección permutada (PC-2) de 48 bits de los 56 bits de estas dos mitades rotadas.

La función f de la red de Feistel se compone de una permutación de expansión (E), que convierte el bloque correspondiente de 32 bits en uno de 48, que aporta las propiedades de difusión al algoritmo. Después realiza una or-exclusiva con el valor K_i, también de 48 bits, aplica ocho (sustituciones no lineales $f(x)+f(y) \neq f(x+y)$) S-Cajas de 6*4 bits, y efectúa una nueva permutación (P). Es en las cajas S donde reside la potencia, y la relativa invulnerabilidad computacional del cifrado, pudiéndoselas considerar como el núcleo principal del método.



Ronda del algoritmo DES

El sistema de descifrado es muy similar, así se facilita la implementación tanto en hardware como por software. Para descifrar basta con usar el mismo algoritmo empleando las Ki en orden inverso, no es necesario invertir la función sino repetirla. Esto es así porque la operación "o exclusivo" es un involución su aplicación repetida dos veces conduce a los valores originales.

Propiedades fundamentales del DES:

- Dependencia entre símbolos. Cada bit del texto cifrado es una función compleja de todos los bits de la clave y de todos los bits del texto original (por bloques).
- Cambio de los bits de entrada. Un cambio de un bit en el mensaje original produce el cambio del 50%, aproximadamente, de los bits del bloque cifrado.
- Cambio de los bits de la clave. Un cambio en un bit de la clave produce, aproximadamente, el cambio de la mitad de los bits del bloque cifrado.
- Claves débiles: existen cuatro claves "débiles" que producen un mensaje cifrado fácil de descifrar, porque todas las claves parciales K_1 a K_{16} son iguales. Existen 28 claves "semidébiles" que producen un mensaje cifrado fácil de descifrar, porque producen sólo dos o cuatro subclaves parciales diferentes. Cuando se elige una clave al azar, es preciso asegurarse de que no es una de esas claves.
- Un error en la transmisión de un texto cifrado se propaga a todo el bloque del que forma parte.

El inconveniente del algoritmo es que la clave de 64 bits es relativamente corta, hasta hace unos años era suficiente para las máquinas existentes. Pero hoy en día se puede romper con máquinas potentes trabajando en paralelo a través de una red, por este motivo ya no es el estándar de seguridad de USA. Una máquina con 10 millones de chips podría probar 10 millones de claves/seg. consiguiendo la vulneración del cifrado en unas 24 horas.

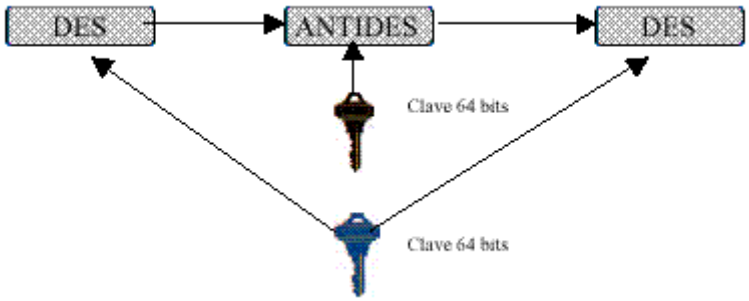
Ventajas del algoritmo:

- Es el más extendido en el mundo, más barato y más probado (sobre todo en máquinas UNIX). Hay módulos hardware de aceleración.
- Es muy rápido y fácil de implementar

3.1.1. Triple DES (TDES)

En un intento de hacer más invulnerable al DES o mejorar su rendimiento, a lo largo de los últimos años, se han propuesto numerosas variantes. Una solución para incrementar la fortaleza del algoritmo es emplear el cifrado producto con varios DES y otras tantas claves distintas. Supuesto que el DES no sea un grupo, como así sucede (lo que se sospechó desde siempre, aunque no se llegó a demostrar hasta el 1992), el resultado sería un algoritmo notoriamente más robusto. Además, esta forma de proceder presenta la ventaja de amortizar las enormes inversiones en equipos DES en todo el mundo.

Con este planteamiento se ha diseñado un sistema basado en tres iteraciones del algoritmo DES, llamado triple DES o TDES:



El TDES utiliza una clave de 128 bits (16 de paridad y 112 de clave), se aplican 64 bits a los dos DES y los otros 64 bits al DES inverso (ANTIDES) que se realiza entre los otros dos. La expresión del cifrado sería:

$$C = E(k_1, (D(k_2, E(k_1, M)))$$

Con tres algoritmos se podría aplicar tres claves distintas pero no se hace así para que sea compatible con el DES. Si la clave de 128 está formada por dos claves iguales de 64 el sistema se comporta como un DES simple:

$$EK\,[DK[EK[Texto]]] = EK[Texto]$$

Para darnos una idea de la seguridad del Triple DES, supongamos que tenemos un circuito integrado que realiza un millón de cifrados triple DES por segundo y que se construye un ordenador paralelo que contiene un millón de estos circuitos, este ordenador sería capaz de probar 10^{12} cifrados por segundo, pero para forzar todas la claves de los Triple DES requeriría :

$2^{112} = 5.19 * 10^{33}$ operaciones de cifrado

$1.19 * 10^{33} / 10^{12} = 5.19 * 10^{21}$ seg = $1.65 * 10^{14}$ años

Esto es más de 16453 veces la edad actual estimada del universo (aproximadamente 10^{10} años).

El triple DES, es una alternativa utilizada por numerosos estándares y productos comerciales. Entre los primeros están la norma de gestión de claves ANSI X9. 17 e ISO 8732, y entre los segundos el PEM (*Privacy Enanced Mail)* popular producto de correo electrónico.

3.2. IDEA

En 1990 Lay, X y Massey J. del Instituto Federal de Tecnología Suizo inventaron un algoritmo denominado PEES *(Proposed European Encryption Standard*). En 1992 se publicó la segunda versión bajo el nombre IDEA, resistente a ataques de criptología diferencial. Este algoritmo está libre de restricciones y permisos nacionales y es de libre distribución por Internet. Esto ha hecho que sea un algoritmo muy popular, sobre todo fuera de los EE.UU.

En IDEA (*International Data Encryption Algorithm)* tanto los datos en claro como los datos cifrados están compuestos por bloques de 64 bits, mientras que la clave consta de 128 bits. El cifrado se basa en el concepto de mezclar operaciones aritméticas de grupos algebraicos diferentes. El algoritmo consiste en ocho vueltas de cifrado idénticas seguidas de una transformación de salida.

Durante el proceso de cifrado se utilizan operaciones de tres grupos aritméticos diferentes: multiplicativo, aditivo y aditivo bit a bit, sobre pares de sub-bloques de 16 bits. Estas operaciones son invertibles en si mismas, pero incompatibles entre si en el sentido de que no gozan de ley distributiva ni asociativa, no forman un grupo y la asociación de ellas no puede dar lugar a cancelación de operaciones.

IDEA presenta diferencias notables con respecto a DES que le hacen más atractivo:

- El espacio de claves es mucho más grande: $2^{128} = 3{,}4 * 10^{38}$
- Todas las operaciones son algebraicas, sin cajas S de oscura justificación
- Es más eficiente que los algoritmos de tipo Fiestel, porque a cada vuelta se modifican todos los bits del bloque y no solamente la mitad
- Se pueden utilizar todos los modos de operación definidos para el DES: ECB, CBC, CFB y OFB

Hasta ahora nunca ha sido roto, aunque no tiene la antigüedad del DES. Además su longitud de clave lo hace muy difícil de romper mediante ataques de fuerza bruta. Sería necesario probar 10^{38} claves, cantidad imposible de manejar con los medios informáticos actuales.

Numerosos paquetes de programas de seguridad para ordenadores personales o redes incluyen IDEA como algoritmo de cifrado. Por ejemplo, PGP, conocido producto de seguridad para redes.

3.3. AES

El cifrador sucesor de DES es el denominado Advanced Encryption Standard (AES), resultado de una competición entre muchas propuestas. La propuesta, denominada originalmente Rijndael por los apellidos de sus creadores los belgas Joan Daemen y Vincent Rijmen, fue elegida en el 2000.

AES no se basa en redes de Feistel, es una red de sustitución y permutación y se puede implementar de forma eficiente y rápida tanto en hardware como en software. Opera con bloques de 128 bits y tamaños de clave de 128, 192, y 256 bits en un campo finito de Galois denominado GF(28), que utiliza como polinomio reductor x8 + x4 + x3 + x + 1.

AES opera con un número especificado de repeticiones de varias rondas de transformaciones que se aplican al estado, una representación matricial de 4x4 obtenido a partir del texto en claro relleno con 0s. Concretamente se hace una fase inicial, 9 rondas principales y una ronda final. En la ronda inicial se derivan las claves de cada fase a partir de la clave dada. En las rondas principales, se realizan los procesos: subBytes, shiftRows, mixColumns, y addRoundKey.

subBytes es una sustitución no lineal de cada elemento del estado con una tabla, la S-BOX de Rijndael. shiftRows es un desplazamiento circular de las columnas del estado (ShiftRows). mixColumns multiplica cada columna por una matriz especial en el GF(28). addRoundKey es una suma de cada columna del estado con la clave de la ronda. La ronda final no incluye la operación mixColumns.

Se puede encontrar un vídeo ilustrativo realizado por Enrique Zabala para cryptool.org en la plataforma del curso en: Rijndael_Animation_v4_esp.swf. También hay una divertida presentación como viñetas en un pdf en: http://www.moserware.com/2009/09/stick-figure-guide-to-advanced.html.

Hoy en día AES es el estándar más utilizado de cifrado simétrico.

CAPÍTULO 4 CRIPTOSISTEMAS ASIMÉTRICOS O DE "CLAVE PUBLICA"

La posibilidad de ataques activos además de garantizar la confidencialidad, obliga a preservar la integridad, la actualidad de la información y la autenticidad de los actores.

Por ejemplo, si *B* recibe un mensaje de *A* se pueden plantear las siguientes preguntas:

- El mensaje lo envía *A* u otro usuario (autenticidad)
- Lo que se recibe es lo que envió *A* (integridad)
- Lo ha enviado *A* en este momento u otro usuario ha repetido algo que *A* envió con anterioridad
- *B* repudia los datos enviados por *A* o viceversa

La criptografía moderna trata de dar una solución a toda esta casuística, por ello la podemos definir, acorde al objeto de este tema, como "La disciplina que estudia los principios, métodos y medios de transformar los datos para ocultar la información contenida en ellos, garantizar su integridad, establecer su autenticidad y prevenir su repudio" .

En 1976 Diffie y Hellman publicaron el artículo *"New directions in cryptography"*. En el que proponían un nuevo tipo de criptografía basado en utilizar claves distintas para cifrar y descifrar. Estos criptosistemas no precisaban, previamente al establecimiento de una transmisión cifrada, transferir una clave secreta entre el emisor y el receptor, evitando así los problemas inherentes a la búsqueda de canales seguros para tal transferencia.

4.1. DEFINICIÓN DE CRIPTOSISTEMAS ASIMÉTRICOS

Antes de definir los sistemas de clave pública conviene clarificar el concepto de función unidireccional, $f : M \rightarrow C$ como una función invertible, de modo que es "fácil" calcular $f(m) = c$, mientras que es difícil calcular $f^{-1}(c) = m$. (No se sabe si hay funciones de este tipo, aunque se

supone su existencia). Una función unidireccional se dice que es una *función unidireccional tramposa* si puede ser invertida fácilmente cuando se conoce alguna información adicional extra. Tal información extra se conoce como *trampa*.

Se define un criptosistema de clave pública como una familia de funciones unidireccionales tramposas, $\{f_k\}$, para cada clave k de K, de modo que la trampa $t(k)$ sea fácil de obtener. Además para cada k de K se debe poder describir un algoritmo eficiente que permita calcular f_k , pero de modo que sea intratable la determinación de k y $t(k)$.

Para implementar un criptosistema de clave pública, dada una familia de funciones unidireccionales tramposas, cada usuario U elige una clave aleatoria u de K y pública E_u que permite calcular f_u ; E_u es su *clave pública*, mientras que la trampa $t(u)$, necesaria para invertir f_u, es su *clave privada*.

Si un usuario A desea enviar un mensaje m a otro B, mira la clave pública de B, E_b, y transmite $f_b(m) = c$ a B. Como B es el único capaz de invertir f_b, es el único que puede recuperar el mensaje m: $f_b^{-1}(f_b(m)) = m$.

En la actualidad se utilizan dos tipos de funciones unidireccionales con trampa. La primera de ellas es el producto de números enteros, cuya inversa es la factorización del número obtenido, y la segunda es la exponenciación discreta, cuya inversa es el logaritmo discreto. Las dos funciones son fáciles de calcular, mientras que no lo son sus inversas. Es decir, dado un número n, es difícil determinar su descomposición en factores primos y, por otra parte, dados a y b es difícil calcular x de modo que $a^x = b$.

4.2. CARACTERÍSTICAS DE LOS SISTEMAS DE CLAVE PÚBLICA

En estos criptosistemas se utilizan dos claves por participante: una denominada clave pública, que se hace de general conocimiento y la otra denominada clave privada, que se mantiene en secreto. Obviamente ambas claves no son independientes, pero del conocimiento de la pública no se infiere la privada, a no ser que se tenga algún dato adicional que también habrá de mantenerse en secreto o, mejor, destruirse una vez generado el par clave pública-clave privada.

Cada par de claves, pública y privada, cumple con las dos propiedades siguientes:

. Dada una clave, la pública, es computacionalmente imposible descubrir la otra, la privada.
. Cualquier información cifrada con una de las claves, únicamente puede ser descifrada por la otra clave. Sean EAB la clave pública y DAV la clave privada de un usuario, se cumple que:

M = DAV (EAB(M)) y

M = EAB (DAV(M))

Lo que significa que cualquier usuario puede cifrar un mensaje con su clave privada DAV y otro usuario recuperarlo con la clave publica EAB del primero, y viceversa cualquier usuario puede cifrar un mensaje con la clave pública de otro y únicamente este último podrá descifrar el mensaje

Estas ideas supusieron la revolución de la criptología, se podía utilizar para confidencialidad (como con los sistemas simétricos), autenticación y firma digital, además de facilitar de la distribución de claves.

Para cada tipo de servicio se cifra de manera diferente:

- Confidencialidad. El emisor cifra el texto con la clave pública del receptor y el receptor lo descifra con su clave privada. Sólo el receptor, que tiene la clave privada, y el emisor, que lo ha creado, pueden conocer el contenido (Figura II -8).
- Autenticación. Se cifra el mensaje, o un resumen de éste, mediante la clave privada y cualquier persona puede comprobar su procedencia utilizando la clave pública del emisor. El mensaje es auténtico porque sólo el emisor verdadero puede cifrar con su clave privada (Figura II-9).
- Firma digital. (propiedad privativa de un individuo o proceso que se utiliza para firmar mensajes). La información de la firma se añade al mensaje para garantizar la firma del remitente. El mecanismo es igual que en la autenticación pero siempre se cifra el resumen del mensaje, cuyo criptograma es la firma del emisor. Así el emisor no puede negar la procedencia ya que se ha cifrado con su clave privada. El receptor puede comprobar que el resumen coincide con la firma descifrada para ver si es auténtico (Figura II-10). La "firma digital" proporciona autenticación y garantía de integridad, pero no confidencialidad.

Se puede realizar sistemas completos con autenticación o firma y confidencialidad.

Este tipo de funciones ha dado lugar a los criptosistemas de clave pública, siendo los más conocidos:

. El RSA, el más utilizado, (desarrollado en el MIT, en 1977, por Rivest, Shamir y Adleman) basado en la factorización de números muy grandes.
. El algoritmo Diffie-Hellman, que se utiliza para distribuir claves simétricas, pero no sirve para confidencialidad, autenticación ni firma digital. Combinado con RSA se utiliza para conseguir protección en caso de que se desvele la clave privada RSA.
. El algoritmo DSA (*Digital Signature Algorithm*) que únicamente se suele utilizar para firma digital.
. ECDSA, qué utiliza curvas elípticas y permite obtener longitudes de clave menores y mejores rendimientos en dispositivos limitados (smartphones, tablets, etc.)

4.3. CIFRADO RSA

La primera realización del modelo fue publicada en 1978 con el nombre RSA (iniciales de los autores del artículo, Ronald **R**ivest, Adi **S**hamir y Leonard **A**dleman) en el artículo "*A method for obtaining digital signatures and public-key criptosystems*". RSA es el prototipo de los algoritmos de clave pública.
La clave pública (**e** y **n**) y la privada (**d**) están compuestas por un exponente y un módulo que es un producto de dos números primos grandes.

RSA se basa en las propiedades numérico-teóricas de la aritmética modular y los números enteros. Una propiedad hace uso de la función de Euler Totien de un número ∅ (n), que se

define como el conteo de enteros menores a ese número relativamente primos a ese número. (Dos números son relativamente primos si no tienen factores comunes; por ejemplo 9 y 8 son relativamente primos). La función $\varnothing(n)$ para un número primo es: $\varnothing(n)$ =n-1. Esto es porque todos los enteros positivos menores que el número son relativamente primos a él.

La propiedad que usa RSA fue descubierta por Euler y dice: cualquier entero *i* relativamente primo a *n* elevado a la potencia de $\varnothing(n)$ y tomando su mod *n* es igual a 1. esto es:

$$i^{\varnothing(n)} \bmod n \equiv 1$$

Supóngase que *e* y *d* son números enteros aleatorios que son inversos modulares $\varnothing(n)$, es decir:

$$ed \equiv 1 \bmod \varnothing(n)$$

Otra propiedad relacionada usada en RSA, también descubierta por Euler, dice que si *M* es cualquier número relativamente primo a *n*, entonces:

$$(M^e)^d \bmod n \equiv M \text{ y } (M^d)^e \bmod n \equiv M$$

Criptográficamente hablando, si M es parte de un mensaje, se tiene un medio sencillo de codificarlos usando una función:

$$s \equiv M^e \pmod n$$

y decodificarlo con otra función:

$$M \equiv s^d \pmod n$$

La seguridad del RSA se basa en el hecho de que no existe una forma eficiente de factorizar números que sean productos de dos números primos grandes.

Mediante “prueba y ensayo” es muy difícil calcular **d** ya que es un número de 2048 bits o más. Así el sistema de criptoanálisis utilizado es buscar la clave privada **d** a partir de las públicas **e** y **n**. Para esto hay que encontrar los números **p** y **q**, estos son la descomposición en factores primos de **n**, ya que **n** = **p** * **q.** No se ha descubierto aún ninguna forma analítica de descomponer números grandes en factores primos.

De este modo, no conociendo la factorización de **n,** la única posibilidad de ataque que tiene el criptoanalista es obtener **M** de la expresión:

$$M = \log_e C$$

lo que constituye un problema NP-completo y por tanto inabordable si *e* es un número elevado.

En la actualidad es frecuente tomar valores de *n* iguales a 2048 bits si se desea criptogramas seguros durante los próximos años, independientemente de la potencia de los ordenadores.

Factorizar un número de 665 bits usando uno de los algoritmos más veloces de factorización requeriría aproximadamente $1.2 * 10^{23}$ operaciones. Suponiendo que un ordenador pudiera hacer 10^{10} operaciones por segundo el tiempo de factorización de *n* con 200 dígitos sería de $1.2 * 10^{13}$ segundos, o 380267 años, por eso se recomienda esta longitud para proporcionar un alto grado de seguridad. Si se duplica el tamaño del número primo: un número de 400 dígitos requeriría nada menos que $8.6 * 10^{15}$ años para factorizarse.

Para dar otra perspectiva del tamaño de estos números, si se asume que de alguna manera se precalcularan los factores de todos los números de 200 dígitos decimales. Simplemente para almacenar los mismos números se requerirían aproximadamente $(9 * 10^{200}) * 665$ bits de almacenamiento (sin incluir ningún dato de control o de índice). Si suponemos que tuviéramos discos de 10 Terabytes, se necesitarían $6.12 * 10^{187}$ unidades. Si asumimos que cada uno de estos discos pesara una millonésima de gramo el peso de todo el almacenamiento sería $6.75 * 10^{177}$ toneladas. El planeta Tierra pesa $6588 * 10^{21}$ toneladas. No se tiene certeza sobre la cantidad de masa que tiene nuestra galaxia local, pero se sospecha que podría ser menor que la cantidad de masa necesaria para este proyecto. De todo ello se deduce que sin un descubrimiento real en la teoría de los números, el mecanismo RSA (y métodos similares) son completamente seguros de ataques por fuerza bruta, si se realiza una cuidadosa selección de los números primos para crear la clave

La implementación de los cifrados RSA se puede realizar por hardware o por software, aunque a partir de 512 bits la velocidad de la segunda opción suele ser desesperadamente lenta, por lo que usualmente es preferible la primera.

RSA es incomparablemente más lento que AES, no obstante, a pesar de esta diferencia de velocidad de ejecución, el criptosistema RSA (y otros de clave pública) se utilizan fundamentalmente para firmar digitalmente los mensajes que se envían y para cifrar una clave simétrica.

Normalmente se utilizan sistemas mixtos, simétricos para confidencialidad y asimétricos para distribución de claves simétricas, autenticación y firma digital. Por ejemplo si A quiere enviar un mensaje seguro a B, se podría utilizar el siguiente procedimiento:

- *A* cifra el mensaje *m*, mediante AES, con una clave aleatoria, y a su vez cifra la clave aleatoria utilizada con la clave pública de B, mediante RSA.
- *A* envía por el canal inseguro la pareja formada por el mensaje cifrado mediante AES y la clave de AES cifrada con RSA.

Para recuperar el mensaje recibido:

- *B* descifra la clave de AES mediante su clave privada del RSA y luego utiliza la clave obtenida para descifrar el mensaje *m*.

Este protocolo se conoce como *envoltura digital RSA o sobre digital*.

4.4. ALGORITMO DE INTERCAMBIO DE CLAVES DE DIFFIE-HELLMAN

El algoritmo se describía en el articulo "*New directions in Cryptography*" para ilustrar un ejemplo de la criptografía de clave pública que Diffie y Hellman acababan de descubrir.

Por medio de este protocolo dos personas pueden intercambiarse pequeñas informaciones secretas por un canal inseguro. La descripción del protocolo es la siguiente:

1. Los dos usuarios *A* y *B*, seleccionan públicamente un grupo multiplicativo finito, *G*, de orden *n* y un elemento $\alpha \in G$.
2. *A* genera un número aleatorio α., calcula α^a en *G* y lo transmite a *B*.
3. *B* genera un número aleatorio *b*, calcula α^b en *G* y lo transmite a *A*
4. *A* recibe α^b y calcula $(\alpha^b)^a$ en *G*
5. *B* con el α^a recibido calcula $(\alpha^a)^b$ en *G*

Ahora *A* y *B* poseen un elemento común y secreto del grupo *G:* α^{ab}. Un escucha (espía), S, podría conocer G, n, α^a y α^b y debería poder calcular el elemento α^{ab}, lo que hasta ahora es un problema intratable.

Este algoritmo sólo se suele utilizar para intercambiar claves simétricas, pero ésta es una de las principales funciones de los algoritmos asimétricos, así está muy extendido en sistemas de Internet con confidencialidad de clave simétrica (VPNs, SSL, etc...).
Tiene una gran aplicación en sistemas que garantizan la denominada PFS (Perfect Forward Secrecy), de forma que utilizando otro cifrado asimétrico (típicamente RSA) se intercambian parámetros Diffie-Hellman para generar un secreto común único para este intercambio. De esta forma aunque posteriormente alguien obtenga la clave privada RSA y haya grabado la sesión, no pueda romper la clave Diffie-Hellman negociada ad-hoc para dicha transferencia.