Sistemas de sustitución simple o monoalfabéticos monoliterales.

Denominamos así a todos aquellos sistemas en los que se sustituye un símbolo o letra por otro distinto (monoliteral), pero siempre el mismo, es decir, utilizando un solo alfabeto de sustitución (monoalfabético). Básicamente tenemos dos alfabetos que utilizaremos para cifrar o descifrar. La forma de cifrar (y descifrar) es sencilla, cogemos cada letra del mensaje que queremos cifrar, buscamos la posición en que está y la sustituimos por la que está en el otro alfabeto. Al alfabeto que cogemos como base lo denominaremos alfabeto en claro y al otro alfabeto cifrado. Por ejemplo, si tenemos los siguientes alfabetos:

Claro	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**	**N**	**Ñ**	**O**	**P**	**Q**	**R**	**S**	**T**	**U**	**V**	**W**	**X**	**Y**	**Z**
Cifrado	R	T	V	X	Z	B	D	F	H	J	L	N	O	Q	S	U	W	Y	A	C	E	G	I	K	M	Ñ	P

El mensaje en claro a cifrar sería: ATAQUE AL ALBA. Para cifrar buscamos la primera letra y la sustituimos por la que aparece debajo. En este caso la A quedaría convertida en la R. La segunda letra sería la T que sustituiríamos por la que está debajo, la E y segaríamos así hasta que acabáramos con todas la letras del mensaje en claro. Para hacer más difícil la labor al posible criptoanalista, eliminamos los espacios, con lo que el mensaje cifrado quedaría como:

Claro	A	T	A	Q	U	E	A	L	A	L	B	A
Cifrado	**R**	**E**	**R**	**Y**	**G**	**Z**	**R**	**N**	**R**	**N**	**T**	**R**

Como veremos, a pesar del elevado número de posibilidades teóricas (*n*!, con *n*=número de elementos del alfabeto), la realidad es que los cifrados de sustitución monoalfabéticos son muy fáciles de romper basándonos simplemente en las propiedades estadísticas del lenguaje.

Uno de los problemas que presenta la utilización de un alfabeto aleatorio es la dificultad de memorizar los alfabetos. Existe sin embargo una forma sencilla de generar un alfabeto de sustitución sin la necesidad de memorizar las equivalencias entre los mismos. Para ello basta con utilizar una palabra como clave. El alfabeto se genera poniendo la palabra clave al principio eliminando las letras duplicadas y poniendo a continuación el resto del alfabeto normal en un orden prefijado. Por ejemplo si utilizamos la palabra SEGURIDAD como clave, el alfabeto generado sería:

Claro	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**	**N**	**Ñ**	**O**	**P**	**Q**	**R**	**S**	**T**	**U**	**V**	**W**	**X**	**Y**	**Z**
Cifrado	S	E	G	U	R	I	D	A	B	C	F	H	J	K	L	M	N	Ñ	O	P	Q	T	V	W	X	Y	Z

Una variante más segura de este método es el de distribuir los caracteres del alfabeto en forma de tabla y coger los datos en columnas bien en orden posicional o en orden alfabético.

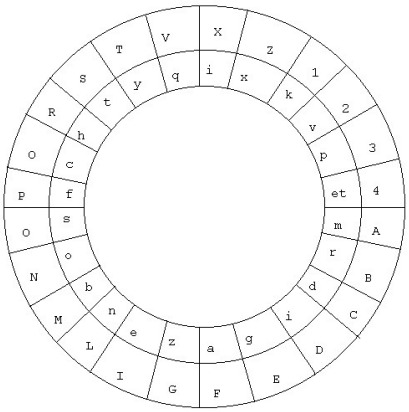
S	E	G	U	R	I
D	A	B	C	F	H
J	K	L	M	N	Ñ
O	P	Q	T	V	W

X	Y	Z				

Claro	A	B	C	D	E	F	G	H	I	J	K	L	M	N	Ñ	O	P	Q	R	S	T	U	V	W	X	Y	Z
Posic.	S	D	J	O	X	E	A	K	P	Y	G	B	L	Q	Z	U	C	M	T	R	F	N	V	I	H	Ñ	W
Alfab.	E	A	K	P	Y	G	B	L	Q	Z	I	H	Ñ	W	R	F	N	V	S	D	J	O	X	U	C	M	T

Un método muy utilizado en la historia es el de cambiar las letras por símbolos extraños o basados en formas geométricas. Este último caso es el del método pig pen o de los masones y el de los templarios. El primer caso es típico de la criptografía de los siglos XV y XVI. Simplemente decir que no aportan más seguridad que la utilización de un alfabeto normal aleatorio.

Finalmente podemos utilizar un esquema circular para representar la clave como en el cifrador de Alberti. ¡Ojo! Alberti lo utilizaba como un cifrado polialfabético, métodos de cifra que veremos en otro apartado.



Método de desplazamiento.

Así se denominan todos aquellos métodos que implican la sustitución de un caracter del alafbeto por otro desplazado *x* posiciones. De forma general, si *n* es el número de elementos del alfabeto y *k* el desplazamiento, cuyo valor tiene que estar entre 1 y *n-1* (si utilizamos *n* para el valor *k* obtendríamos el mismo valor, con lo que no estaríamos cifrando), tenemos que:

$$C_k = m + k (\mathrm{mod}\, n)$$

La operación mod o módulo, podemos definirla de forma sencilla como el resto de la división por n. Por ejemplo 35 mod 27 da como resultado 8.

Este método es muy fácil de romper con la simple aplicación de técnicas estadísticas, una tabla de las posibles permutaciones o un estudio exhaustivo de las claves hasta encontrar la correcta, ya que solo hay 26 posibles.

Curiosidades: Si utilizamos *k* = 3 obtenemos el método de cifrado de Julio César. Octavio Augusto era un poco más vago y utilizaba *k* =1. Para los amantes del cine, el nombre del ordenador de la película de Kubrick, 2001 una

odisea del espacio, estaba cifrado con *k* = -1, con lo que HAL se convertiría, al descifrarlo adelantando una posición cada letra, en IBM.

Cifrado Afín.

Se trata de una generalización del anterior en el que se utiliza una transformación lineal del tipo $C_k = am_k + b(\mathrm{mod}\, n)$. Tomando $a = 1$ tenemos el método de desplazamiento.

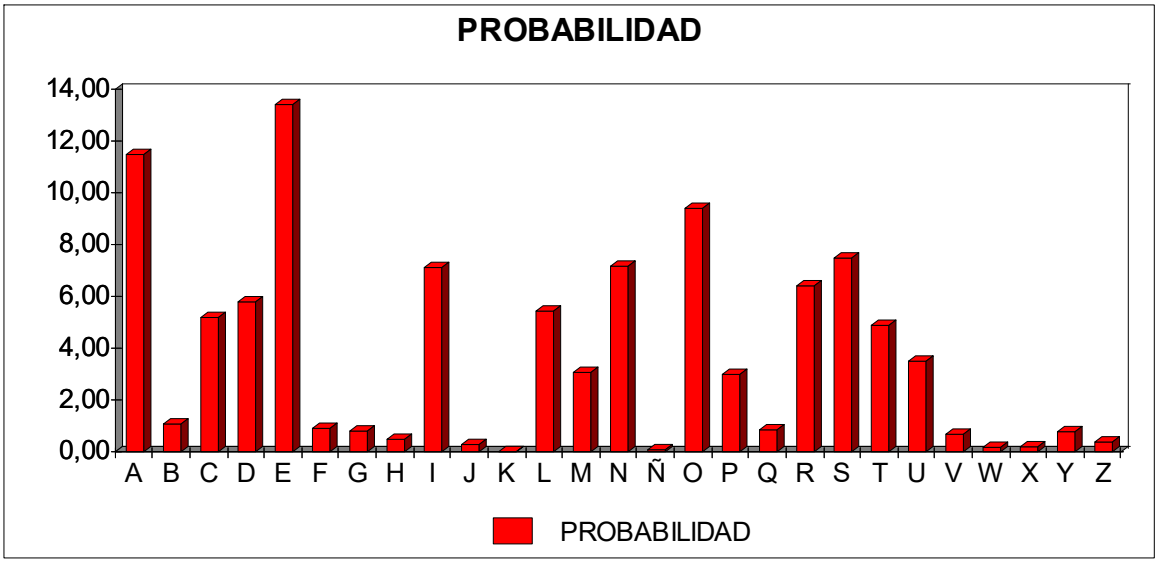
En este caso es necesario que *a* sea un entero primo con *n*, en caso contrario el descifrado podría dar lugar a ambigüedades. Sean $\left.\begin{array}{l} C_1 = am_1 + b(\mathrm{mod}\, n) \\ C_2 = am_2 + b(\mathrm{mod}\, n) \end{array}\right\}$, y supongamos que $C_1 = C_2$. En este caso tenemos que $a(m_1 - m_2) = 0(\mathrm{mod}\, n)$, o lo que es lo mismo $a(m_1 - m_2) = kn$ .Con lo que para que la solución sea única y $m_1 = m_2$ debe cumplirse que *a* y *n* no tengan ningún factor común, es decir, que sean relativamente primos. Se dice que dos números son relativamente primos si no tienen ningún factor común
Para descifrar tenemos que calcular el inverso de a y aplicar la siguiente fórmula:

$$m = (c_k - b).a^{-1} mod\ n$$

Criptoanálisis de los métodos de sustitución simples.

Para atacarlos se pueden utilizar varios métodos. El más rudimentario sería el ataque por fuerza bruta, que puede funcionar bien con el método de desplazamiento, ya que solo hay que probar 26 veces con un alfabeto de 27 letras, pero no sirve para los de alfabeto aleatorio.

Uno de los métodos más potentes es el de utilizar las características básicas del lenguaje como son la frecuencia de aparición de los caracteres que lo forman y las combinaciones más frecuentes de ellos. De esta manera reducimos el espacio de búsqueda de la solución. En cualquier lenguaje existen un conjunto de letras que se repiten frecuentemente y otras que se utilizan mucho menos. También es frecuente la aparición de pares de letras (digramas) o tripletes (trigramas), todas estas características son explotadas por el criptoanalista para deducir el mensaje original a partir del cifrado. Uno de los ejemplos más famosos de este tipo de ataque es el utilizado por el caballero Legrand en el escarabajo de oro de Edgar Allan Poe. A continuación vemos un esquema de la frecuencia de las letras en español. Si una letra aparece mucho es muy probable que sea una E o una A. Vamos haciendo pruebas hasta que obtenemos el alfabeto utilizado.



Otro ataque muy práctico es el de “palabra probable”. Si sabemos o sospechamos que una palabra aparece en el mensaje, podemos buscar el patrón de esa palabra en el texto cifrado y sustituir esa palabra por sus posibles apariciones. Por ejemplo, si sabemos que una de las palabras en el mensaje es “atacad”, buscaríamos un trozo del texto cifrado donde la primera, tercera y quinta letras fuesen iguales y supondríamos que esa es la posición donde aparece la palabra. Este método, combinado con el uso de las frecuencias del lenguaje da unos resultados muy buenos.

Trucos para hacer más difícil el criptoanálisis.

Lamentablemente si el sistema es débil de por sí, no vamos a conseguir convertirlo en un sistema fuerte. Podemos, sin embargo, darle un poco más de fortaleza utilizando la eliminación de espacios, para que el criptoanalista no sepa dónde empieza y acaba cada palabra; y la inserción de nulos en el mensaje a cifrar. A ser posible esos nulos deben ser de letras con poca probabilidad de aparición para que se disimule la frecuencia final resultante de las letras. Como inconveniente tenemos el hecho de que hace el mensaje a cifrar más largo. Por ejemplo, si queremos cifrar la frase “la inteligencia es como las pulgas, van saltando de una cabeza a otra, pero no todas se la quedan” empezaríamos quitando las separaciones y signos de puntuación, con lo que quedaría:

Lainteligenciaescomolaspulgasvansaltandodeunacabezaaotraperonotodasselaquedan

En un segundo paso lo que haremos será incluir una letra nula cada cinco en claro. Lo ideal sería que fuese aleatorio, pero nos servirá como ejemplo. Hemos marcado en rojo las letras nulas.

Laintzeligeñnciaefscomoxlaspuhlgasvpansalqtandowdeunaqcabezzaaotrvapero znotodyasseljaqueddan

El último paso sería cifrar el mensaje y enviarlo.