Foundations of Computer Security

Lecture 40: Substitution Ciphers

Dr. Bill Young
Department of Computer Sciences
University of Texas at Austin

Substitution Ciphers

A substitution cipher is one in which each symbol of the plaintext is exchanged for another symbol.

If this is done uniformly this is called a *monoalphabetic cipher* or *simple substitution cipher*.

If different substitutions are made depending on where in the plaintext the symbol occurs, this is called a *polyalphabetic substitution*.

Simple Substitution

A simple substitution cipher is an injection (1-1 mapping) of the alphabet into itself or another alphabet. *What is the key?*

A simple substitution is breakable; we could try all $k!$ mappings from the plaintext to ciphertext alphabets. *That's usually not necessary.*

Redundancies in the plaintext (letter frequencies, digrams, etc.) are reflected in the ciphertext.

Not all substitution ciphers are simple substitution ciphers.

Caesar Cipher

The Caesar Cipher is a monoalphabetic cipher in which each letter is replaced in the encryption by another letter a fixed "distance" away in the alphabet.

For example, A is replaced by C, B by D, ..., Y by A, Z by B, etc.
What is the key?

What is the size of the keyspace? Is the algorithm strong?

Vigenère Cipher

The Vigenère Cipher is an example of a polyalphabetic cipher, sometimes called a *running key cipher* because the key is another text.

Start with a key string: "monitors to go to the bathroom" and a plaintext to encrypt: "four score and seven years ago." Align the two texts, possibly removing spaces:

plaintext:	`fours corea ndsev enyea rsago`
key:	`monit orsto gotot hebat hroom`
ciphertext:	`rcizl qfkxo trlso lrzet yjoua`

Then use the letter pairs to look up an encryption in a table (called a *Vigenère Tableau* or *tabula recta*).

What is the corresponding decryption algorithm?

Vigenère Tableau

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

Cryptanalysis on Vigenère Cipher

The Vigenère Cipher selects one of twenty-six different Caesar Ciphers, depending upon the corresponding letter in the key.

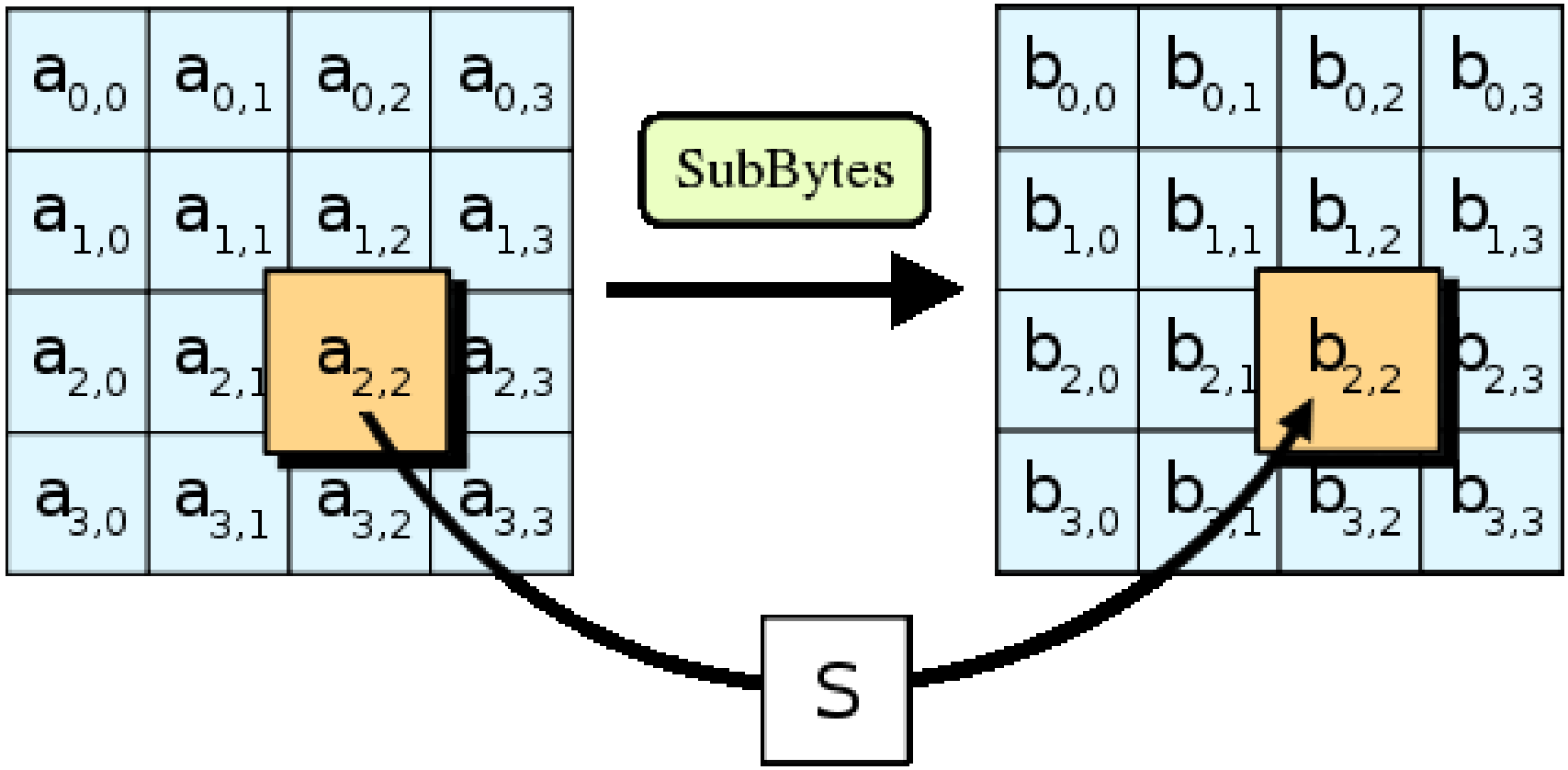
Running key ciphers are susceptible to statistical analysis. Both key and plaintext are English language strings and so have the entropy characteristics of English. In particular, the letters A, E, O, T, N, I make up approximately 50% of English text. Thus, at approximately 25% of indices, these can be expected to coincide.

This is an example of a *regularity* in the ciphertext that would not be expected merely from chance.

AES Substitution Step

Substitution need not only apply to symbols in a text.

The Advanced Encryption Standard (AES) contains a substitution step; each byte in a 16-byte array is replaced with a corresponding entry from a fixed 8-bit lookup table.



Lessons

- Substitution is one of the building blocks of encryption.
- Simple substitution means replacing symbols uniformly by others. The Caesar Cipher and our pirate example are instances.
- Polyalphabetic substitution means that the substitution varies according to the position in the text. The Vigenère Cipher is an example.

Next lecture: Using Information