MAD 4471: Cryptography and Coding Theory **Fall 2020**

Lecture 1: Introduction to Encryption

Lecturer: Jean-François Biasse *TA: William Youmans*

1.1 Caesar cipher

We describe the Caesar cipher because it is simple and to illustrate the concept of: plain text, cipher text, key, message space, etc...

Description:

- Caesar wants to encrypt m to M where m is plain text and M is message space.
- He chooses $k \in K = \{0, 1, \cdots, 25\}$ where k is *key* and K is *key space*.
- The encryption $Enc(m, k)$ is defined by $m_1 + k \mod 26, m_2 + k \mod 26, \cdots, m_l + k \mod 26$.
- In other words: each letter of m is shifted by the secret k.

Example 1 *$k = 3$, $m = ALEA\ JACTA\ EST^*$, once we shift by 3 we get $DOHD\ MDFWD\ HVW$.*
$*-$ "The die is cast". Phrase attributed to J. Caesar as he was crossing the Rubicon.

So the encryption is a function $Enc : \mathcal{M} \times \mathcal{K} \mapsto \mathcal{C}$ where $\mathcal{C}$ is the cipher text space. The decryption is a function $Dec : \mathcal{C} \times \mathcal{K} \mapsto \mathcal{M}$ which must satisfy,

$$Dec(Enc(m, k), k) = m$$

To encrypt the Caesar cipher, we use the key k to reverse the shift. $Dec(c, k)$ is defined by $c_1 - k \mod 26, c_2 - k \mod 26, \cdots, c_l - k \mod 26$.

Example 2 *With $k = 3$, $DOHD\ MDFWD\ HVW$ by shift of -3 we get $ALEA\ JACTA\ EST$.*

Remark 1 *The only secret is the key. By Kerckhoffs' principle: **the design is not secret**.*
Everyone knows how Caesar encrypts messages, but we hope that messages do not make sense without the key.

1.2 The security

- The security of the data doesn't depend on the design.

- Is it possible to make it impossible to recover the message?

Of course, we can always try all the keys. After $|\mathcal{K}|$ attempts, we will get the message.

In the case of the Caesar cipher, there are only 26 possible shifts. We can go over all of them in a reasonable time $\rightarrow$ the cipher is not very secure.

However, even when $|\mathcal{K}|$ is large, one can always try all the keys. Schemes must therefore allow the possibility to have a large key space. Besides that, we consider that an attack challenges the security of a scheme if it require less steps than trying all the keys.

Definition 1.1 (informal) *The scheme is broken if we recover the message in less steps than trying all the keys.*

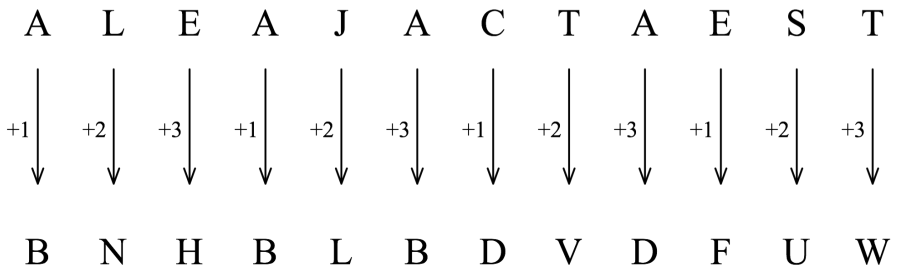
In fact, as we see later, recovering the message is not the only thing that an adversary may want to do.

1.3 The Vigenère cipher

The Caesar cipher can be modified to allow a larger key space. Thus avoiding the brute force attacks.

- The key is a sequence of k shift in $\{0, 1, \cdots, 25\}$ in a given k.
- Let $s_1, s_2, \cdots, s_h$ be the k (secret) shifts. $Enc(m, s_1, s_2, \cdots, s_h)$ is defined by $m_1 + s_1 \mod 26, \cdots, m_h + s_h \mod 26, m_{h+1} + s_1 \mod 26, \cdots$.

Example 3 *Assume the secret key is the sequence of shifts* $1, 2, 3$.



The decryption is just the application of these shifts backward. Here $\mathcal{K} = \{0, 1, \cdots, 25\}^k$, so $|\mathcal{K}| = 26^k$. If k is chose sufficiently large, $|\mathcal{K}|$ will be too big to brute force the search for a key.

1.4 Conclusion: What is security about?

Just because $|\mathcal{K}|$is large doen not mean that the scheme is secure. There is an obvious weakness to the substitution schemes (Caesar, Vigenère). Indeed, prior biases on the messages can be observed on the cipher text. For example, the most common letter in the English language is "e". So the most frequent letter is a cipher text from the Caesar cipher is the shift of "e", thus relatively the key $k = e - x \mod 26$.