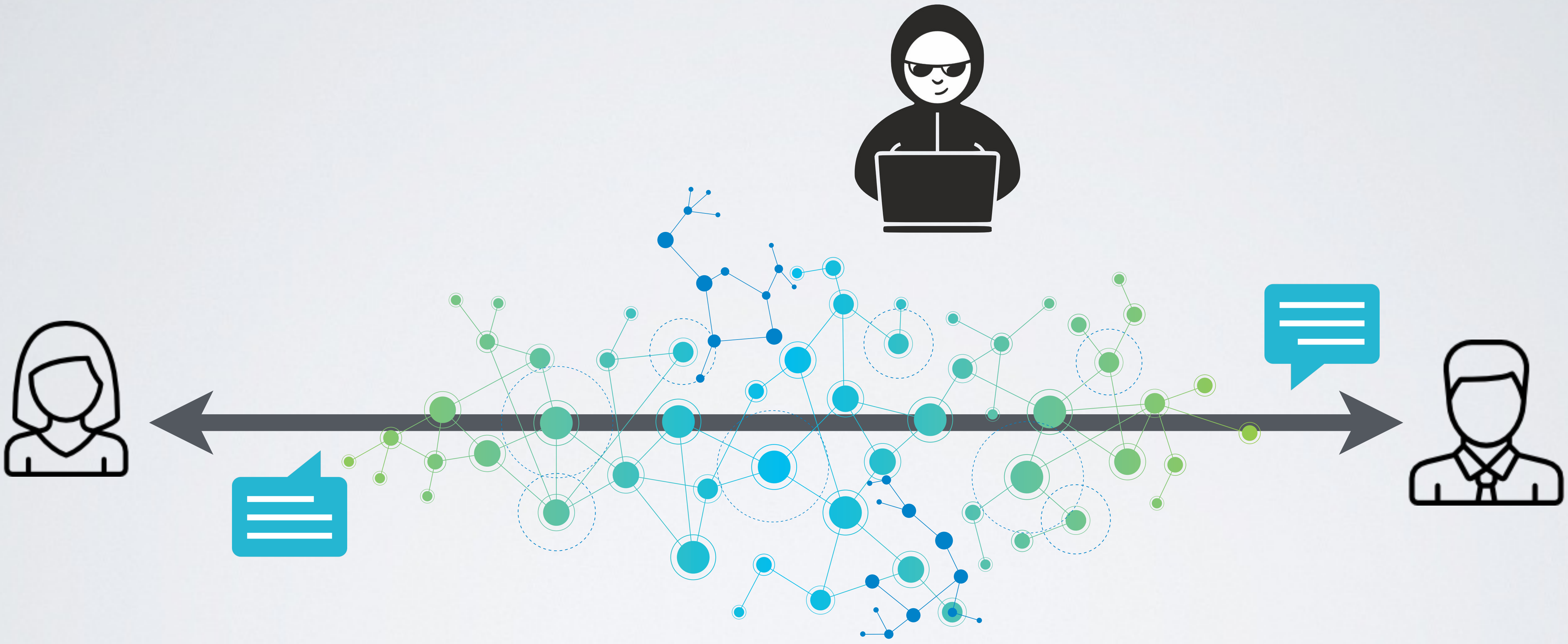
Introductory Cryptography

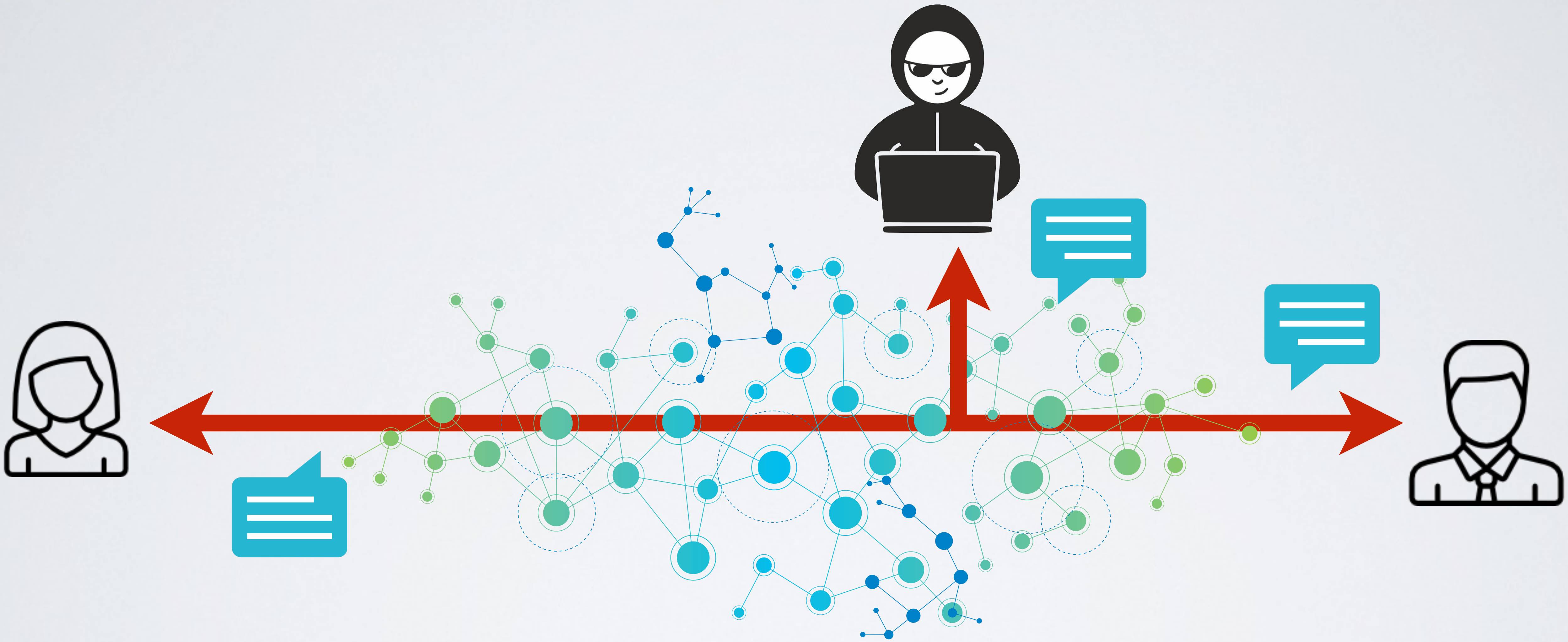
Classical Cryptography

Kc Udonsi

Communication over an **insecure** medium

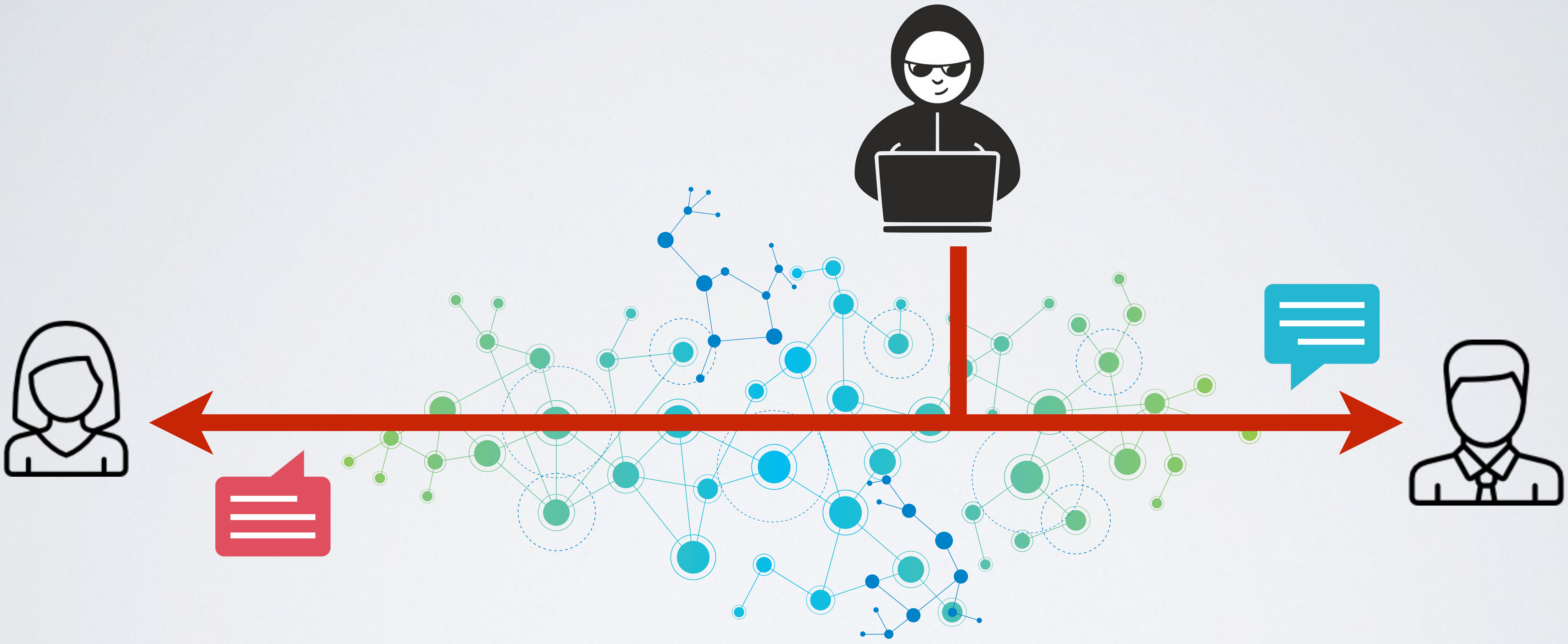


Threat 1 - **Interception**



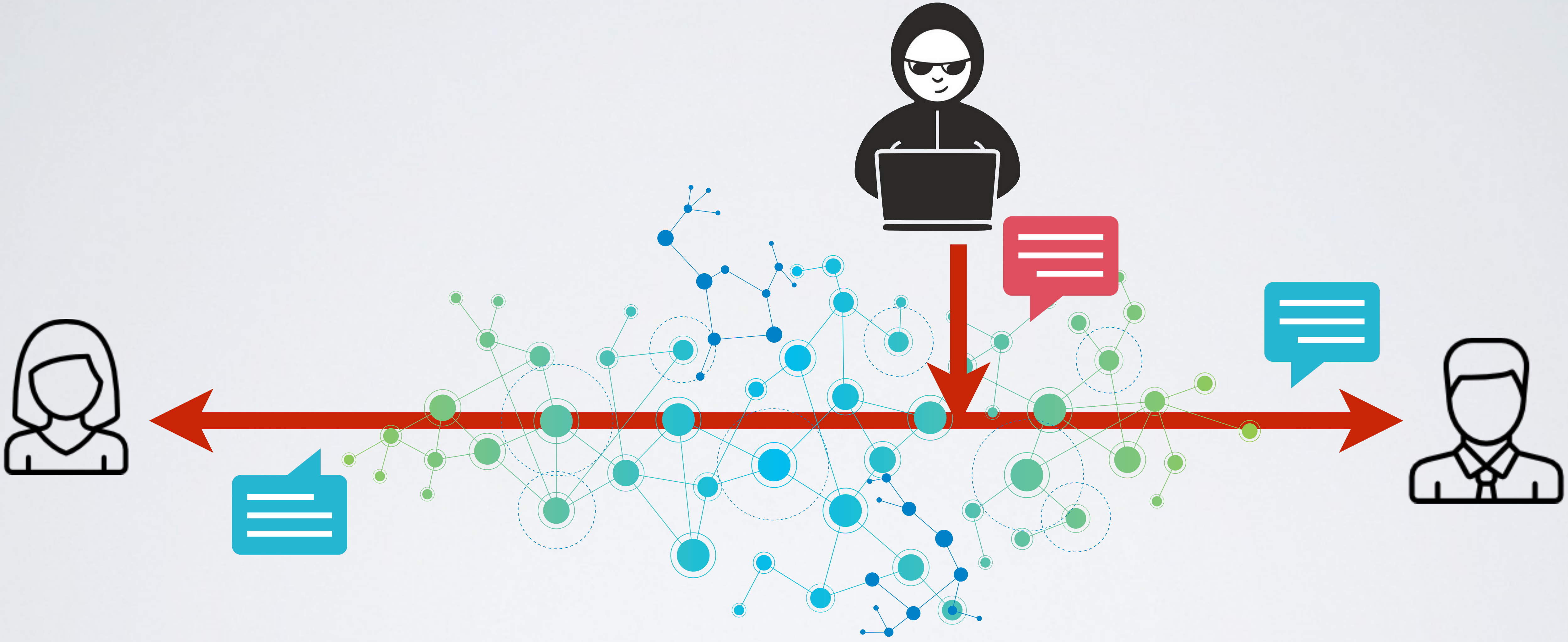
◉ **Interception** : an attacker can read messages

Threat 2 - **Modification**



◉ **Modification** : an attacker can <u>modify</u> messages

Threat 3 - **Fabrication**



- **Fabrication** : an attacker can <u>inject</u> messages

Threat 4 - **Interruption**

- **Interruption** : an attacker can block messages

Confidentiality and Integrity of communications

➡ Implement a **virtual trusted channel** over an insecure medium

Storage on an **insecure** medium

- Threat 1: **Loss**
  An attacker can corrupt or destroy data at rest
- Threat 2: **Disclosure**
  An attacker can disclose data at rest to other unauthorized parties
- Threat 3: **Theft**
  An attacker can obtain and store data at an arbitrary location
- Threat 4: **Modification**
  An attacker can compromise the integrity of data at rest

Storage on an **insecure** medium

- Threat 1: **Loss**
  Cryptography cannot be used to prevent loss. It may be used to yield loss. E.g ransomware
- Threat 2 & 3: **Disclosure & Theft**
  Encrypted data at rest cannot be meaningfully disclosed or utilized without decryption. E.g PGP Whole Disk Encryption
- Threat 4: **Modification**
  Cryptography can be used to verify the integrity of data at rest

Definitions of a cryptosystem

Definitions

Plaintext
The message in its clear form (the original message).

Ciphertext
The message in its ciphered form (the encrypted message).

Encryption
Transform a plaintext into ciphertext.

Decryption
Transform a ciphertext into a plaintext

Definitions

Cryptographic algorithm
The method to do encryption and decryption.

Cryptographic key
An input variable used by the algorithm for the transformation

N-bit security entropy (a.k.a. the key space)
The number of bits necessary to encode the number of possible keys (could be different than the key length)

Monoalphabetic cipher
A specific letter in the plaintext is consistently substituted with another letter in the cipher text

Definitions

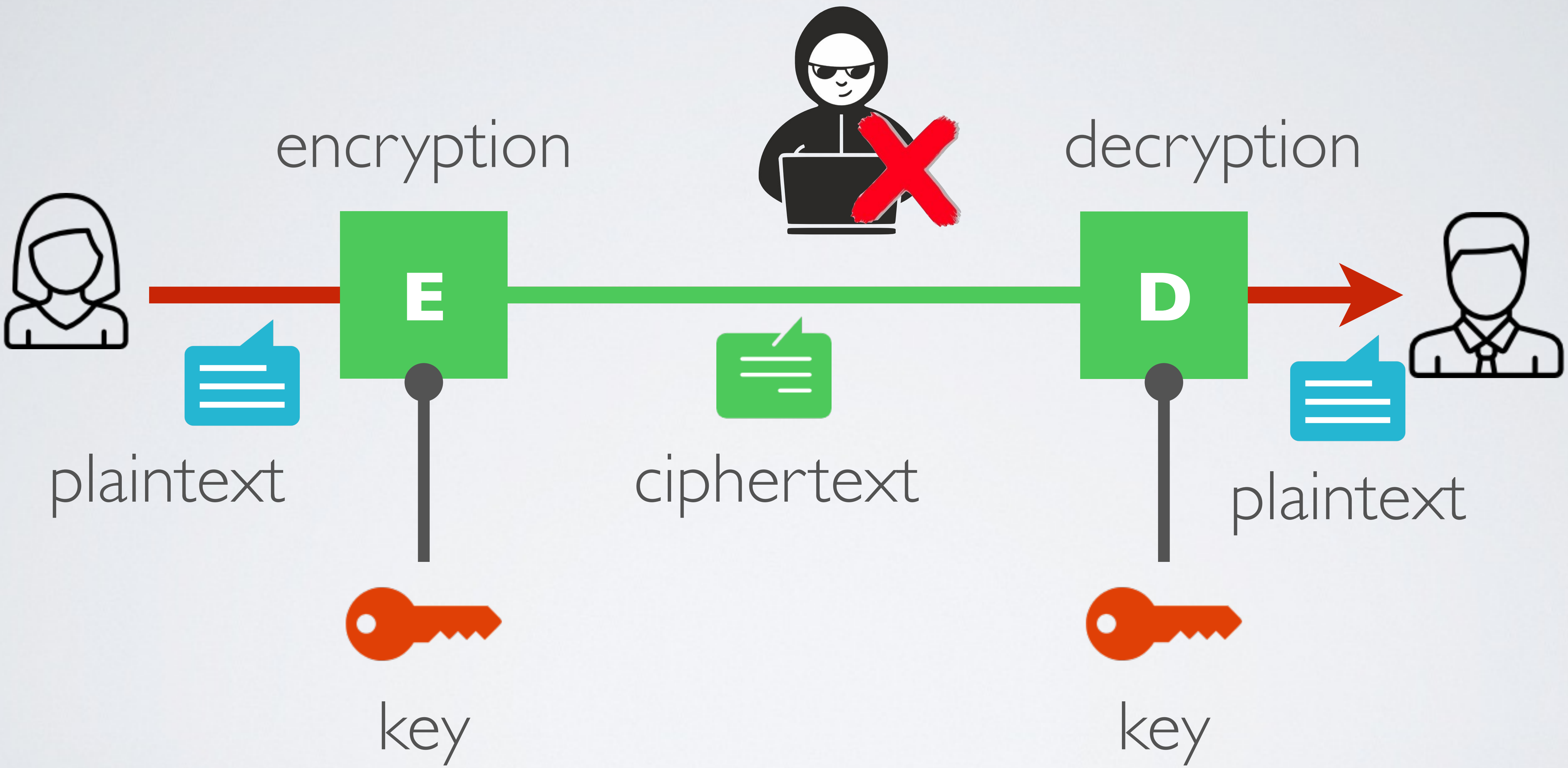
Polyalphabetic Cipher
A specific letter in the plaintext may be substituted with different letters in the cipher text

Cryptography
The art and science of securing messages

Cryptanalysis
The art and science of breaking secured messages to reveal the hidden message

Cryptography
The art and science of secure messaging which encompasses cryptography and cryptanalysis

The big picture



An early example…

Caesar Cipher - the oldest cryptosystem

A *shift* cipher – attributed to Julius Caesar (100-44 BC)

```
MEET ME AFTER THE TOGA PARTY
PHHW PH DIWHU WKH WRJD SDUWB
```

Shift the alphabet 23 places to the right and substitute letters

```
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
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
```

Representing data as numbers

Cryptographic algorithms are mathematical operations

➡ messages and keys must be represented as numbers
for instance : ASCII encoding

Back to Caesar Cipher

Algorithm : shift the alphabet of a certain number of positions

Key : the number of positions to shift

Key space : 25 possible rotations (~ 5 bits security)

Encoding :

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
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

Encrypting and decrypting one character is obtained as follows:

$$c = E(k,p) = (p + k) \bmod 26$$

$$p = D(k,c) = (c - k) \bmod 26$$

Cryptanalysis

Breaking the cipher

The Kerckhoffs' principle (1883)

"The enemy knows the system" - the security of a communication should not rely on the fact that the algorithms are secrets

➡ A cryptosystem should be secure even if everything about the system, except the key, is public knowledge

No security by obscurity

Breaking the cipher - the attacker's model

- **Exhaustive Search** (a.k.a brute force)
  Try all possible n keys (in average it takes n/2 tries)
- **Ciphertext only**
  You know one or several random ciphertexts
- **Known plaintext**
  You know one or several pairs of random plaintext and their corresponding ciphertexts
- **Chosen plaintext**
  You know one or several pairs of chosen plaintext and their corresponding ciphertexts
- **Chosen ciphertext**
  You know one or several pairs of plaintext and their corresponding chosen ciphertexts

➡ **A good crypto system resists all attacks**

Breaking Caesar cipher

Exhaustive search	Yes
ciphertext only	Statistical Analysis
known plaintext	Look at the first letter and get the shift
chosen plaintext	Choose "A" and get the shift
chosen ciphertext	Choose "A" and get the shift

Evolution of cryptosystems

A brief history of cryptography

~ 2000 years ago	Substitution ciphers (a.k.a mono alphabetic ciphers)
few centuries later	Transposition ciphers
Renaissance	Polyalphabetic ciphers
1844	Mechanization
1976	Public key cryptography

Substitution ciphers
(a.k.a mono alphabetic ciphers)

➡ Improvement over Caesar cipher

Algorithm : allow an arbitrary permutation of the alphabet

Key : set of substitutions

Key space : 26! possible substitutions (4×10^{26} ~ 89 bits)

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

```
if we wish to replace letters
WI RF RWAJ UH YFTSDVF SFUUFYA
```

Breaking substitution ciphers

Exhaustive search	Doable with a computer
ciphertext only	Statistical analysis
known plaintext	Match letters together
chosen plaintext	Choose ABCDE … and match letters
chosen ciphertext	Choose ABCDE … and match letters

Polyalphabetic ciphers (a.k.a Renaissance Cipher)

➡ Vigenere cipher

Algorithm : combine the message and the key

Key : a word

Key space : the length of the word

```
  wearediscoveredsaveyourself
+ deceptivedeceptivedeceptive   (mod 26)
  ---------------------------
  ZICVTWQNGRZGVTWAVZHCQYGLMGJ
```

Advantage : Encryption of a letter is context dependent

Breaking Polyalphabetic Ciphers

exhaustive search	Small key length only
ciphertext only	Statistical analysis for small key length and significant amount of ciphertext
known plaintext	Subtract plaintext from ciphertext
chosen plaintext	Choose AAAAA … and match letters
chosen ciphertext	Choose AAAAA … and match letters

OTP - One Time Pad

➡ Improvement over Vigenere cipher

Algorithm : combine the message and the key

Key : an infinite random string

Key space : infinite

```
  whatanicedaytoday
⊕ yksuftgoarfwpfwel
  -----------------
  ZZZJUCLUDTUNNWGQS
```

Advantage : this is the perfect cipher !

Disadvantage : hard to use in practice, how to transmit the key ?

XOR Cipher (a.k.a Vernham Cipher)

a modern version of Vigenere

Use $\oplus$ to combine the message and the key

$$E_k(m) = k \oplus m$$

$$D_k(c) = k \oplus c$$

$$D_k(E_k(m)) = k \oplus (k \oplus m) = m$$

Problem : known-plaintext attack

so $k = (k \oplus m) \oplus m$

$$x \oplus x = 0$$

$$x \oplus 0 = x$$

Mauborgne Cipher - a modern version of OTP

Use a random stream as encryption key

➡ Defeats the know-plaintext attack

Problem : Key-reused attack (a.k.a two-time pad)

$$C_1 = k \oplus m_1$$

$$C_2 = k \oplus m_2$$

so $C_1 \oplus C_2 = (k \oplus m_1) \oplus (k \oplus m_2)$

$$= (m_1 \oplus m_2) \oplus 0$$

$$= (m_1 \oplus m_2)$$

$$x \oplus x = 0$$

$$x \oplus 0 = x$$

The impossibility of breaking OTP

The ciphertext bears no statistical relationship to the plaintext

➡ No statistical analysis

For any plaintext and ciphertext, there exists a key mapping one to the other, and all keys are equally probable

➡ A ciphertext can be decrypted to any plaintext of the same length

The seeds of modern cryptography

1. **Diffusion**
   Mix-up symbols
   Transposition Cipher

2. **Confusion**
   Replace a symbol with another
   Polyaphabetic Cipher

3. **Randomization**
   Repeated encryption of the same text are different
   OTP

Types of Cryptographic Algorithms

Definitions

One way algorithms
Also known as message digests. No keys involved. Encryption cannot be reversed.

Symmetric Key algorithms
The keys used for encryption and decryption are the same OR two-way mathematically related (can be derived from the other reliably)

Public Key algorithms
Also known as asymmetric algorithms. The keys used for encryption and decryption are different but one-way mathematically related.