INTRODUCTION

Anwitaman DATTA
SCSE, NTU Singapore

⌘ Learning outcomes

⌘ Logistics and assessment

COURSE OUTLINE

Syllabus

⌘ Some fundamental (and basic) cryptography and network security concepts

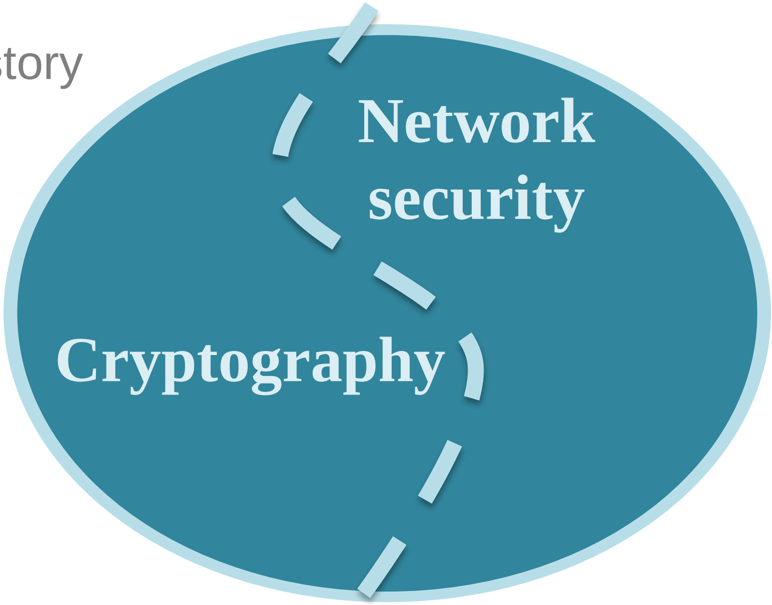
First half

⌘ Basic concepts & history

⌘ Foundational mathematics (number theory)

⌘ Symmetric & Public key cryptography

⌘ Hash functions (e-learning)



Second half

⌘ MAC

⌘ Key management

⌘ Authentication

⌘ Secure network architecture

Learning outcomes

- ⌘ Mathematical tools that form the basis of cryptographic algorithms
- ⌘ Design of cryptographic algorithms
- ⌘ Application of cryptography in real-world systems
- ⌘ Security issues in a Cyberspace environment
- ⌘ Secure network architecture
- ⌘ Basic secure network strategy based on a combination of cryptographic and network security control mechanisms

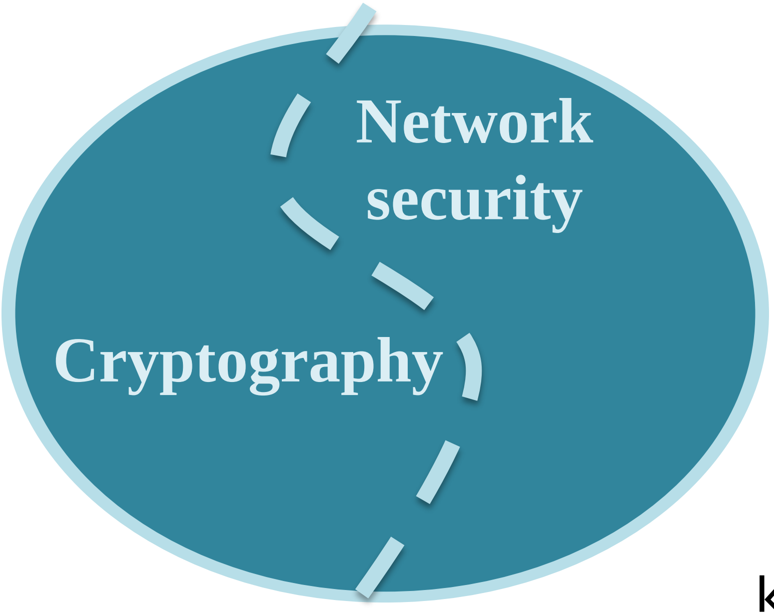
Detailed syllabus can be found on NTULearn course site.

Course delivery

⌘ Lectures (and tutorials): 2+1 hours a week
- Anwitaman DATTA
- Kwok Yan LAM

First half | **Second half**

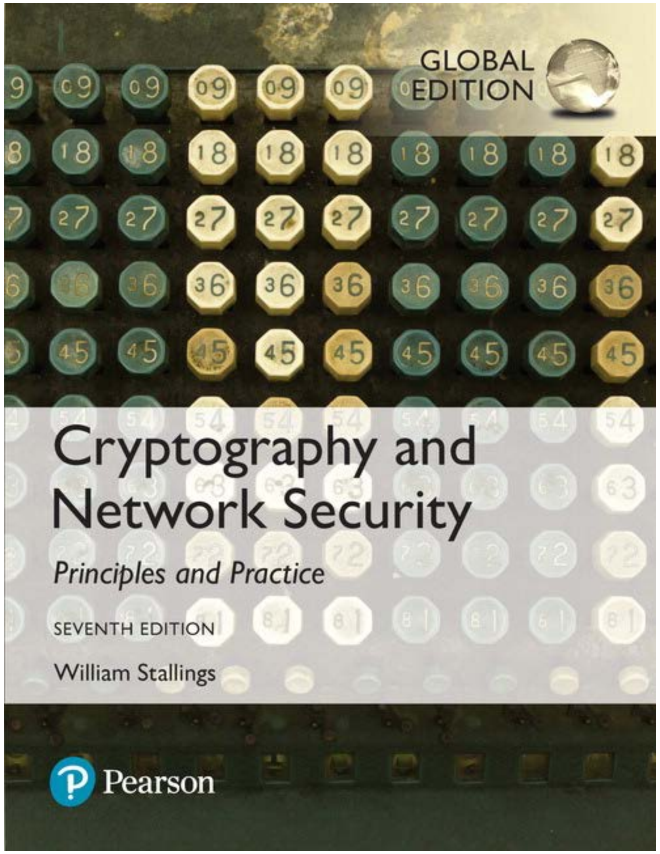
anwitaman@ntu.edu.sg

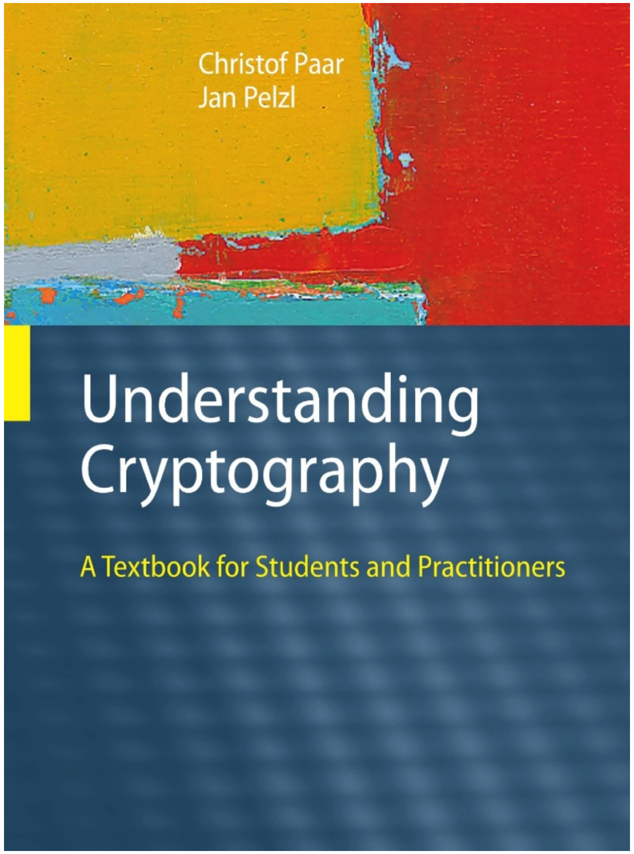


kwokyan.lam@ntu.edu.sg

Logistics

⌘ Text book/reference material





Assessment

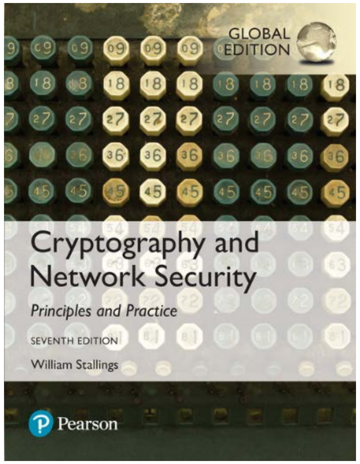
- ⌘ Final exam: 50%
- ⌘ Quizzes: 25%+25%
  - Week 6 [date TBA]
  - TBA for 2nd half

YMNX HTZWXJ BNQQ GJ KZS

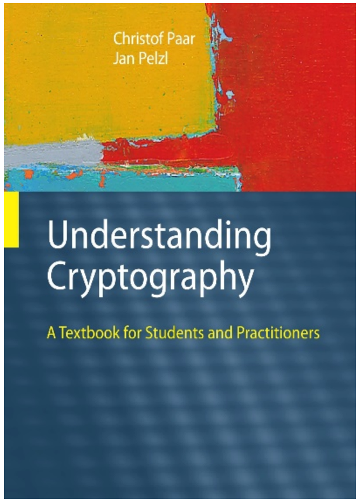
ANY QUESTIONS SO FAR?

Acknowledgement

⌘ The lecture slides have been created by extensively using material from the "Cryptography and Network Security" textbook in the reference, authored by W. Stallings. This includes not only the overall flow and examples used in the lecture materials, but also many images, tables and equations that have been directly derived from the book.



⌘ Likewise, some material from the reference book on "Understanding Cryptography" authored by Paar & Pelzl have also been used.



Disclaimer: I have used art works from third parties in these slides, but not for profit, and, (what I believe as) fair use. Nevertheless, if any such copyright owning party wishes their material to be removed or cited, kindly get in touch with me at anwitaman@ntu.edu.sg

⌘ Security incidences, threats and goals

⌘ Passive and active attacks

INTRODUCTION

The cyber security meltdown

- ⌘ Russian interference in US election
- ⌘ Bangladesh bank heist
- ⌘ Ukraine power-grid knocked out
- ⌘ Hollywood Presbyterian Hospital ransomware
- ⌘ Dyn (domain name service provider) DDoS

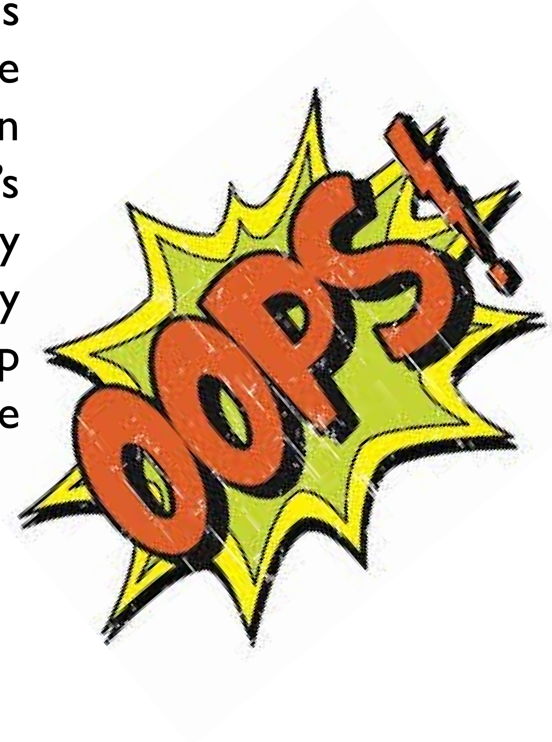
Attacks by mistake …

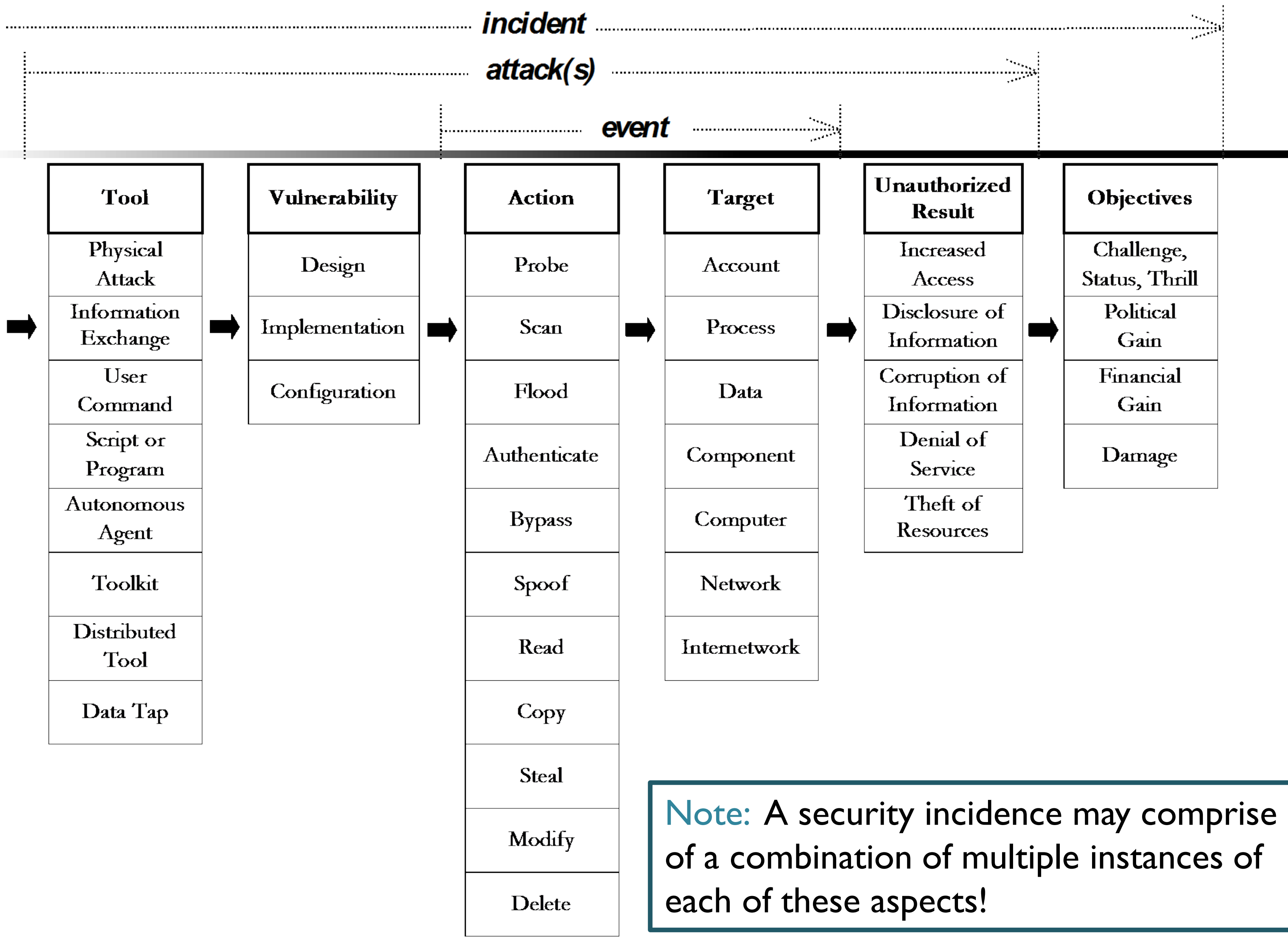
- February 2008: Pakistan censors YouTube globally

- April 2014: Indosat hijacks the world's internet

  Indosat, one of Indonesia's largest telecommunications providers, leaked large portions of the global routing table multiple times over a two-hour period. This means that, in effect, Indosat claimed that it "owned" many of the world's networks. Once someone makes such an assertion, typically via an honest mistake in their routing policy, the only question remaining is how much of the world ends up believing them and hence, what will be the scale of the damage they inflict?

  Source:
  http://research.dyn.com/2014/04/indonesia-hijacks-world/





Note: A security incidence may comprise of a combination of multiple instances of each of these aspects!

Security goals/objectives

- CIA triad: Confidentiality, Integrity, Availability
- Parkerian hexad: CIA + Possession, Authenticity, Utility
- McCumber's cube – multi-dimensional view of security objectives
  - CIA
  - of information under Transmission, Storage, Processing
  - taking into account technological, policy & practice, and human factors
- Violation of one security property may be a pathway to violate others

Note: None of these lists are holistic/exhaustive, and one can identify many other issues/security objectives.

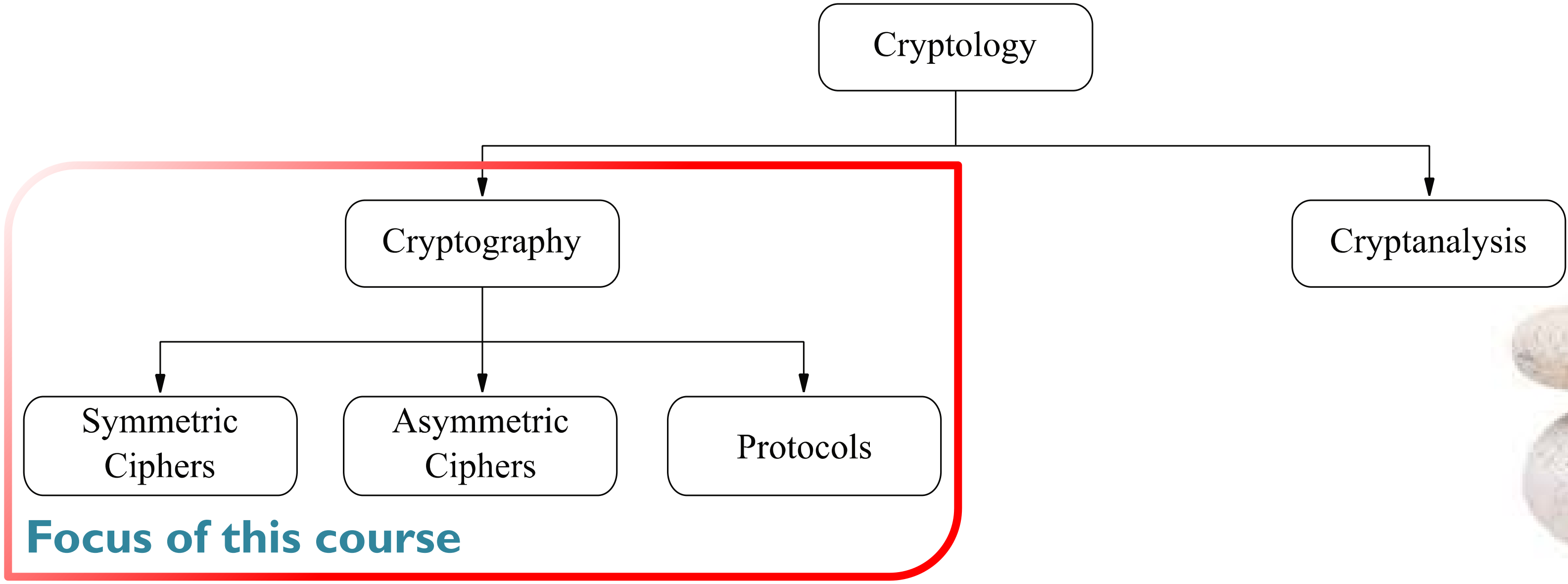
CIA triad

Confidentiality	• Preserving authorized restrictions on information access and disclosure
Integrity	• Guarding against improper information modification or destruction
Availability	• Ensuring timely and reliable access to and use of information

Note: Though these definitions originate from *information security* literature, the interpretation can be extrapolated to other domains/aspects. E.g., Availability of a *service*.

Achieving security

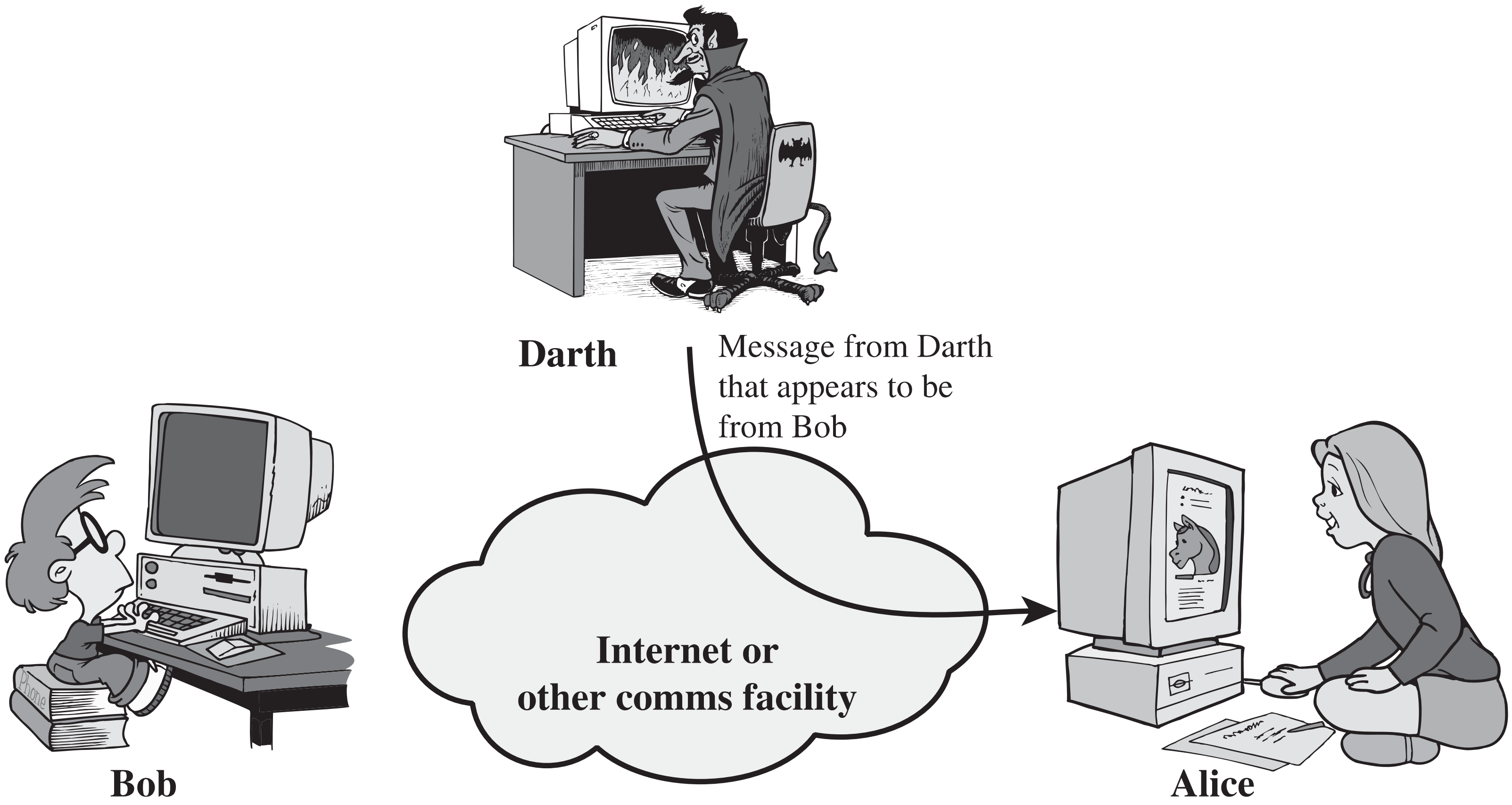
- Many aspects to realizing a proper security solution: cryptology is just one (but very *important and necessary*) part



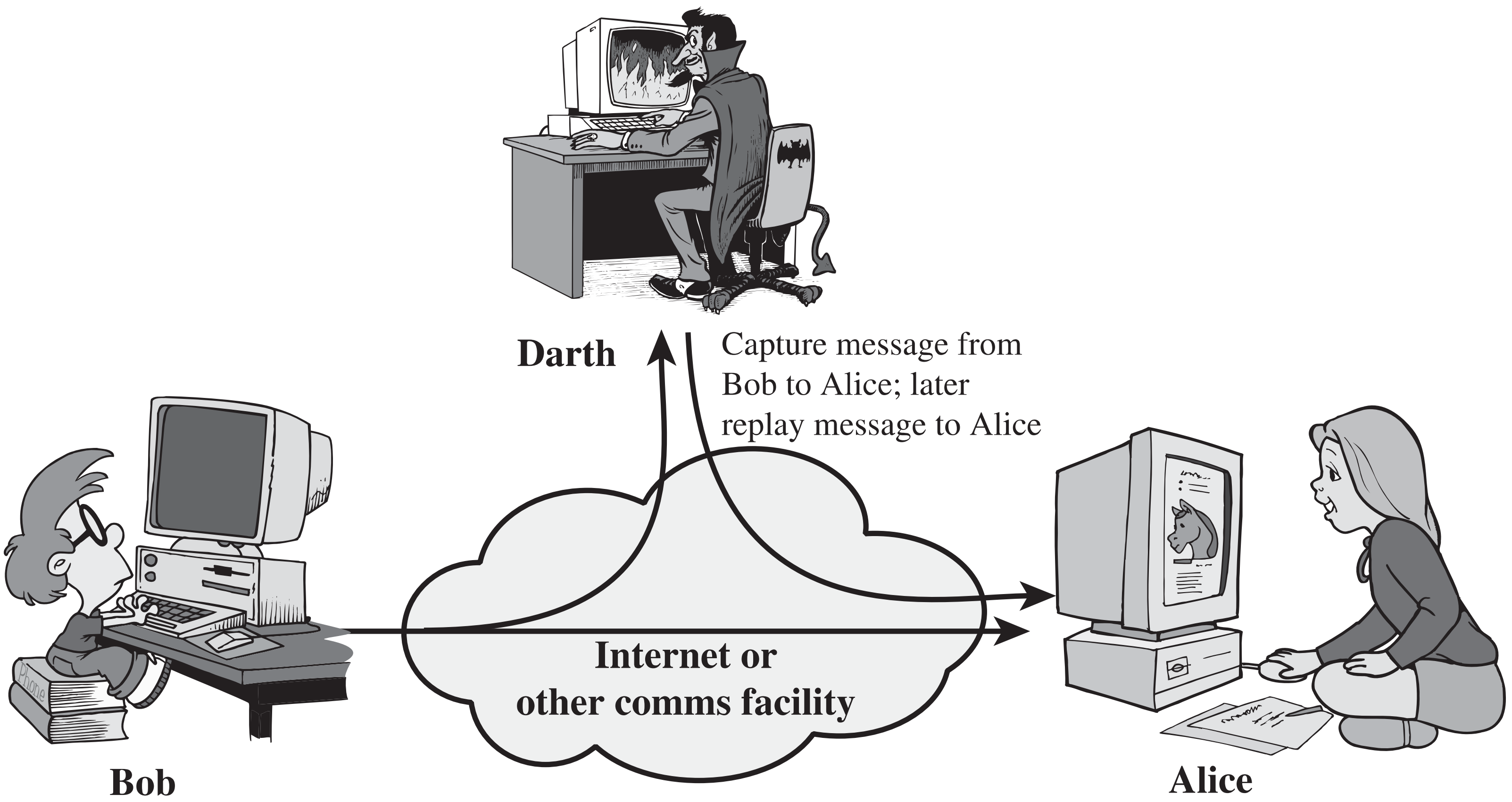
Types of attacks

Passive attacks	• Interception • Traffic analysis
Active attacks	• Impersonation/masquerading • Replay • Modification • DoS • …

Impersonation/masquerade

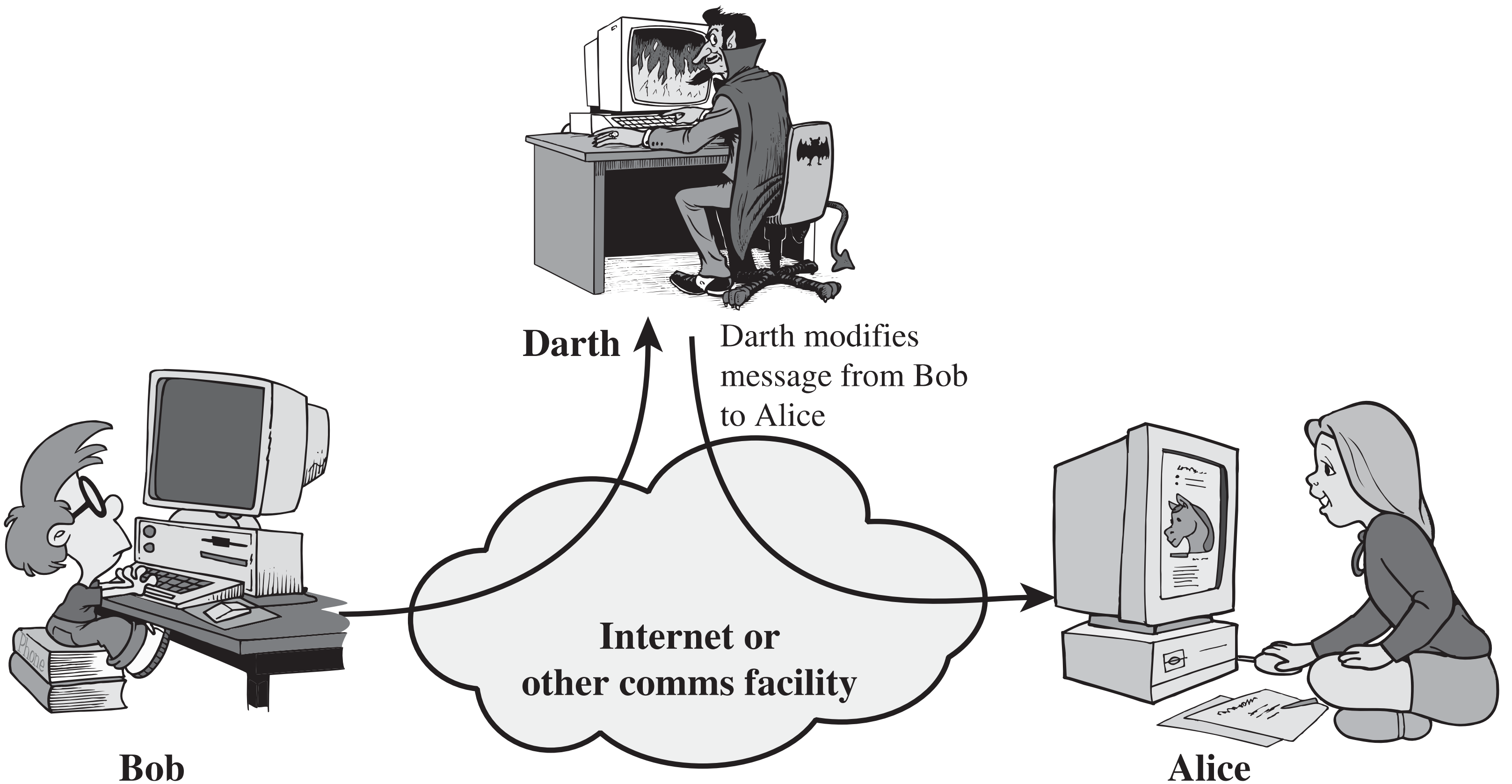


Replay

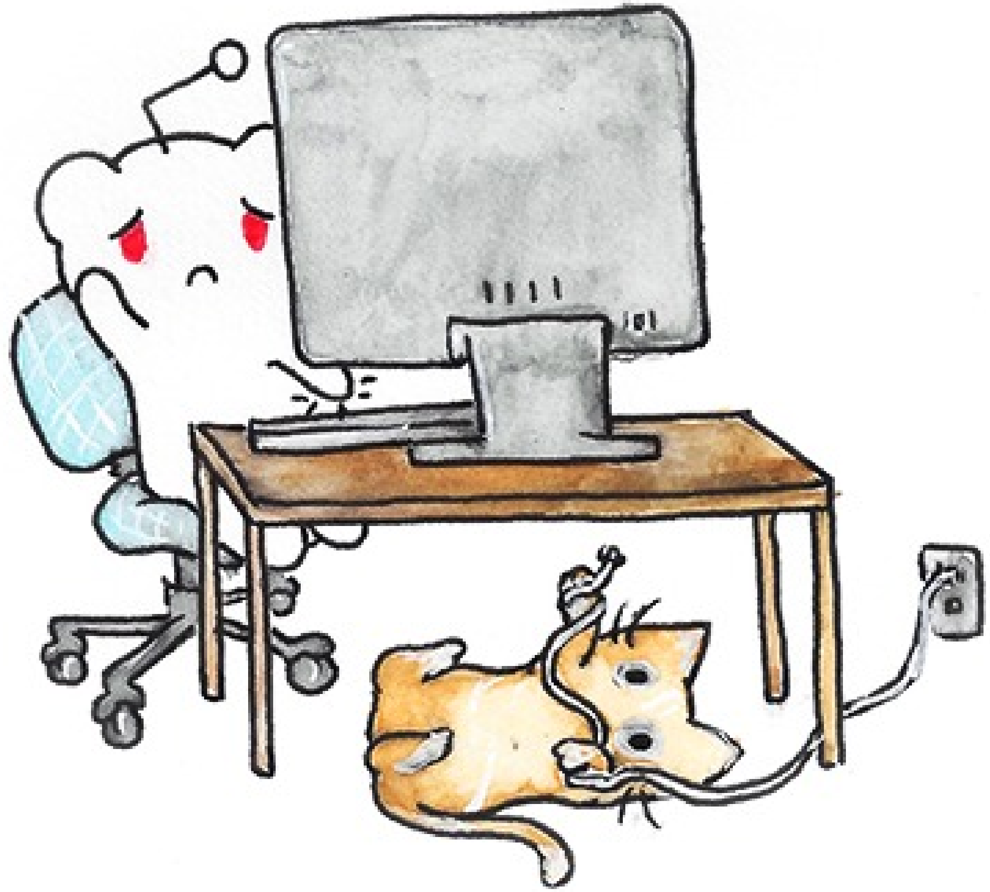


Darth
Capture message from
Bob to Alice; later
replay message to Alice
Internet or
other comms facility
Bob
Alice

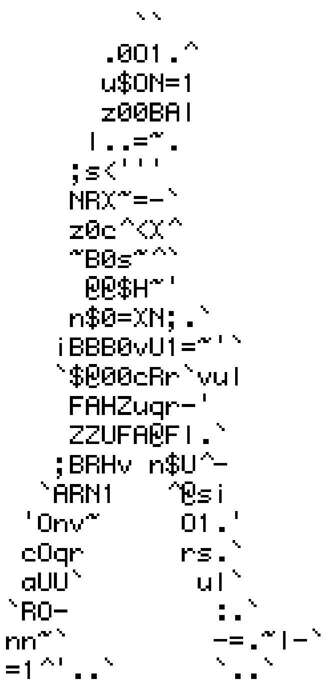
Modification



DoS



Rest of this course …



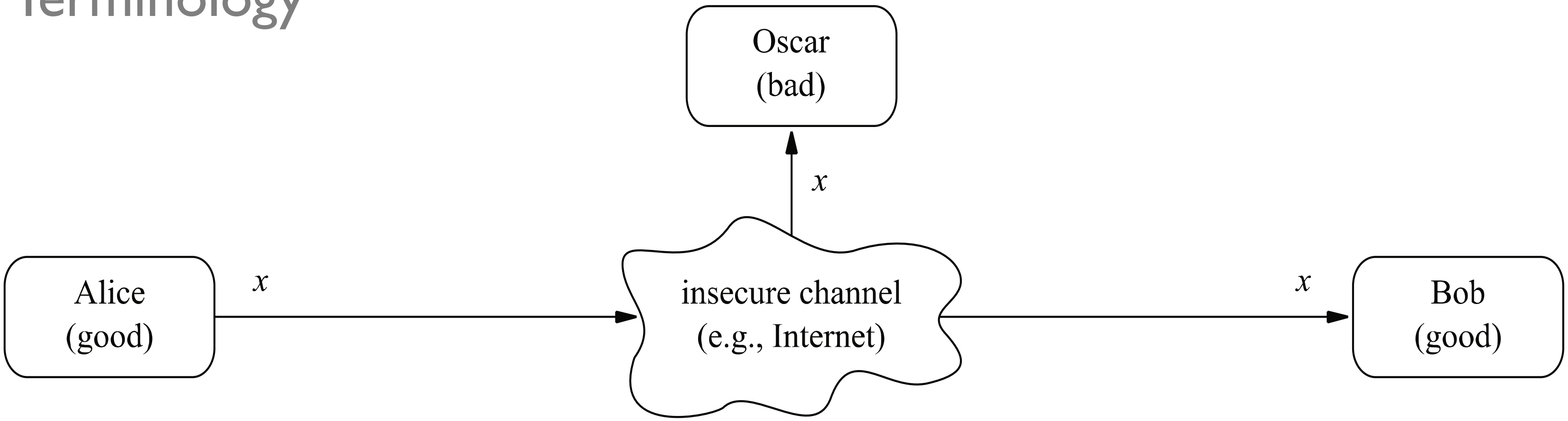
mainly about basic
cryptographic primitives & protocols for
confidentiality, integrity and authentication

- ⌘ Secret key cryptography
- ⌘ Classical ciphers

BASIC CONCEPTS

Private communication

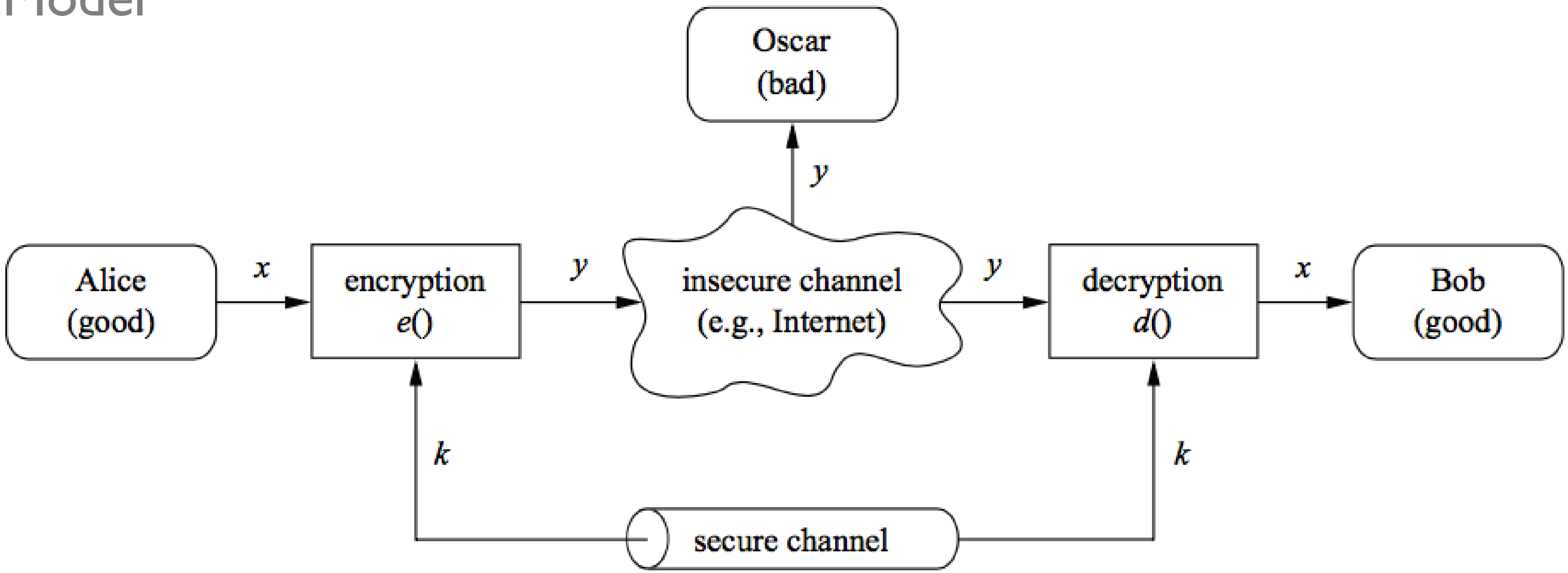
Terminology



⌘ **Alice** and **Bob** want to carry out **private communication** over an **insecure channel**
- **Oscar**, the adversary, trying to learn the content "x" of the private communication

Symmetric/secret key cryptography

Model



⌘ **Sender/Receiver** share a common secret key ***k***

- Encryption & Decryption both done with same key (hence, symmetric)

YMNX HTZWWXJ BNQQ GJ KZS

Remember?

YMNX HTZWXJ BNQQ GJ KZS

Remember?

The (cipher)text above is derived by substituting each letter of the alphabet with some other letter. Such a technique of "substituting" letters is called a **substitution cipher**.

YMNX HTZWXJ BNQQ GJ KZS

- One of the simplest form of substitution cipher: **k-shift cipher**
  - consider the following numerical equivalent assignment to each letter:

a	b	c	d	e	f	g	h	i	j	k	l	m
0	1	2	3	4	5	6	7	8	9	10	11	12

n	o	p	q	r	s	t	u	v	w	x	y	z
13	14	15	16	17	18	19	20	21	22	23	24	25

- The **k-shift cipher** uses the mappings:

Encryption: $C = E(k,p) = (p+k) \textbf{ mod } 26$

Decryption: $p = D(k,C) = (C-k) \textbf{ mod } 26$

LEGEND

p plain text
C cipher text
k secret key
E() encryption algorithm
D() decryption algorithm

YMNX HTZWXJ BNQQ GJ KZS

- Given that the above ciphertext (title) uses a k-shift cipher,
  decipher it without knowing the key

Try it out!

YMNX HTZWXJ BNQQ GJ KZS

⌘ Given that the above ciphertext (title) uses a k-shift cipher
decipher it without knowing the key

Since the "algorithm" is known, a brute-force attack* (exhaustive search for the "key"), i.e., checking 25 possibilities, in this case, would suffice. If one is lucky, the search can be terminated much earlier.

* Trivia: The term brute-force search has nothing to do with "Et tu, Brutus!", but a 3-shift cipher was used by Caesar (and the algorithm was not supposedly known to the adversaries). This specific instance (3-shift) cipher is thus known as **Caesar cipher**.

➡ k-shift cipher web demo

➡ modular arithmetic web demo

Kerckhoff's principle

⌘ A cryptosystem should be secure even if the attacker (Oscar) knows all details about the system, with the exception of the secret key.

In particular, the system should be secure even when the attacker knows the encryption and decryption algorithms (but not the secret key).

Security solely by obscurity is vulnerable to reverse engineering.

Use of known algorithms aide commoditization of cryptography.

Auguste Kerckhoffs (1835-1903)
Dutch linguist & cryptographer

Monoalphabetic cipher

⌘ Each plaintext symbol is substituted with a unique ciphertext symbol
- the interpretation of symbol can be flexible: single letters, n-grams, …

k-shift cipher: Assignment of substituting symbols is in a sequence
e.g., Caesar cipher: 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
Only 25 possible encryptions, easy to brute-force!

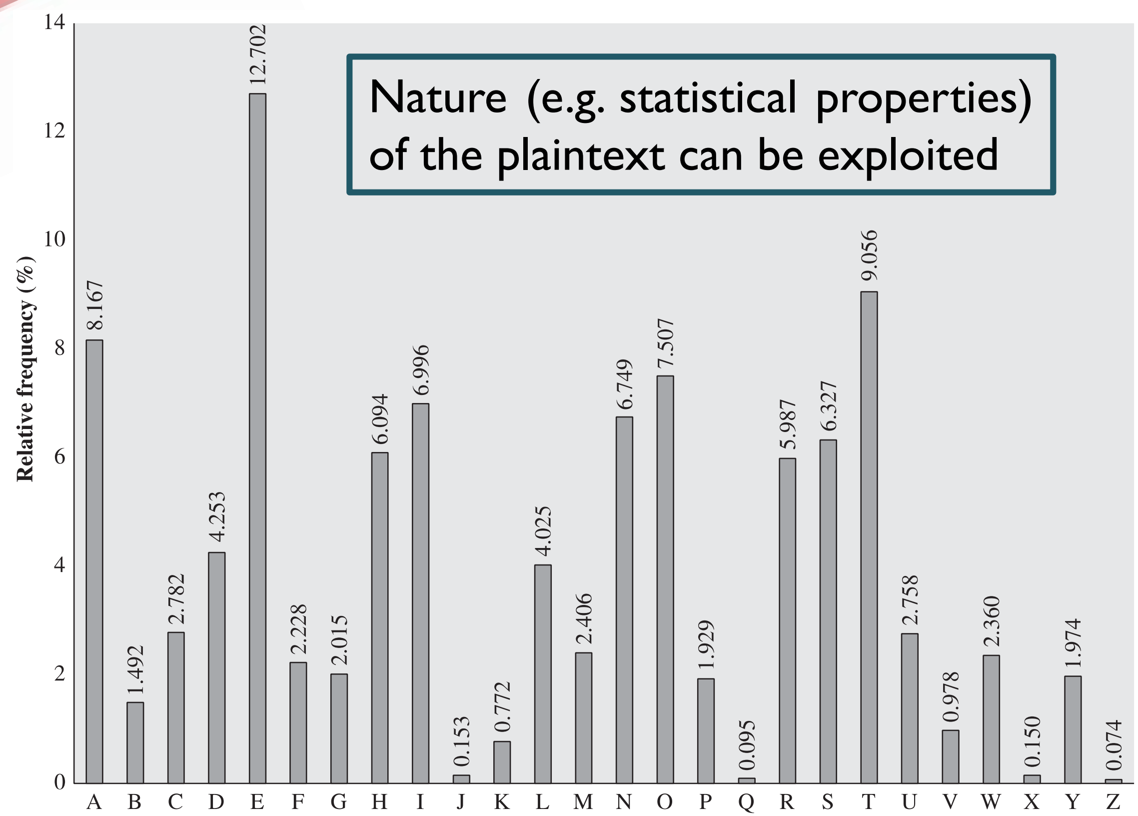
If any random permutation is used as a cipher:
Fragment of a possible cipher: X H R O Q U L …
How many possibilities?

Monoalphabetic cipher

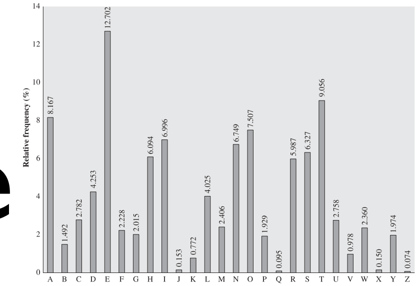
If any random permutation is used as a cipher:
Fragment of a possible cipher: X H R O Q U L …
Brute-force attack will take much longer than the age of the universe!

Not quite ಠ_ಠ

⌘ Cryptanalysis instead of brute-force



Frequency analysis example



- ⌘ Ciphertext

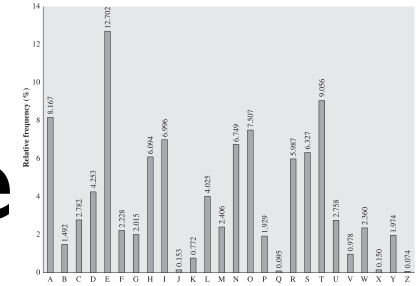
```
UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ
VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX
EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ
```

- ⌘ Symbol (relative) frequency in ciphertext

	13.33	H	5.83	F	3.33	B	1.67	C	0.00
Z	11.67	D	5.00	W	3.33	G	1.67	K	0.00
S	8.33	E	5.00	Q	2.50	Y	1.67	L	0.00
U	8.33	V	4.17	T	2.50	I	0.83	N	0.00
O	7.50	X	4.17	A	1.67	J	0.83	R	0.00
M	6.67								

- ⌘ Guess

Frequency analysis example



- Ciphertext

```
UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ
VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX
EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ
```

- Symbol (relative) frequency in ciphertext

	13.33	H	5.83	F	3.33	B	1.67	C	0.00
Z	11.67	D	5.00	W	3.33	G	1.67	K	0.00
S	8.33	E	5.00	Q	2.50	Y	1.67	L	0.00
U	8.33	V	4.17	T	2.50	I	0.83	N	0.00
O	7.50	X	4.17	A	1.67	J	0.83	R	0.00
M	6.67								

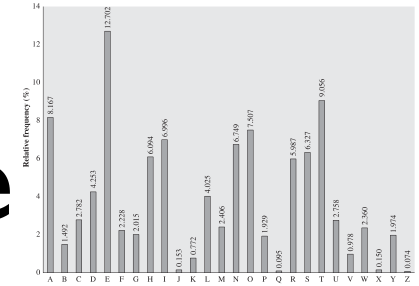
- Guess

$$\{P, Z\} \stackrel{?}{=} \{e, t\}$$

$$\{S, U, O, M, H\} \stackrel{?}{\subset} \{a, h, i, n, o, r, s\}$$

$$\{A, B, G, Y, I, J\} \stackrel{?}{\subset} \{b, j, k, q, v, x, z\}$$

Frequency analysis example



- Ciphertext

```
UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ
VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX
EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ
```

- Symbol (relative) frequency in ciphertext

	13.33	H	5.83	F	3.33	B	1.67	C	0.00
Z	11.67	D	5.00	W	3.33	G	1.67	K	0.00
S	8.33	E	5.00	Q	2.50	Y	1.67	L	0.00
U	8.33	V	4.17	T	2.50	I	0.83	N	0.00
O	7.50	X	4.17	A	1.67	J	0.83	R	0.00
M	6.67								

- Guess

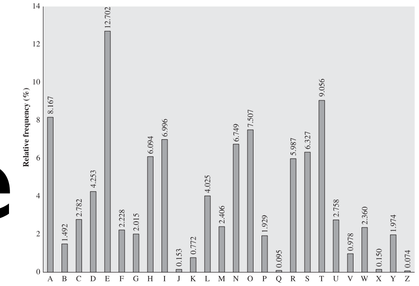
$$\{P, Z\} \stackrel{?}{=} \{e, t\}$$

$$\{S, U, O, M, H\} \stackrel{?}{\subset} \{a, h, i, n, o, r, s\}$$

$$\{A, B, G, Y, I, J\} \stackrel{?}{\subset} \{b, j, k, q, v, x, z\}$$

- Substitute and check
  May suffice for long text
- Otherwise, try n-grams
  e.g. most popular digram:
  th (= ZW?)

Frequency analysis example



- ⌘ Ciphertext

```
UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ
VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX
EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ
```

- ⌘ Symbol (relative) frequency in ciphertext

	13.33	H	5.83	F	3.33	B	1.67	C	0.00
Z	11.67	D	5.00	W	3.33	G	1.67	K	0.00
S	8.33	E	5.00	Q	2.50	Y	1.67	L	0.00
U	8.33	V	4.17	T	2.50	I	0.83	N	0.00
O	7.50	X	4.17	A	1.67	J	0.83	R	0.00
M	6.67								

- ⌘ Guess

$$\{P, Z\} \stackrel{?}{=} \{e, t\}$$

$$\{S, U, O, M, H\} \stackrel{?}{\subset} \{a, h, i, n, o, r, s\}$$

$$\{A, B, G, Y, I, J\} \stackrel{?}{\subset} \{b, j, k, q, v, x, z\}$$

- Substitute and check
  May suffice for long text
- Otherwise, try n-grams
  e.g. most popular digram:
  th (= ZW?)

Playfair cipher

⌘ Idea: multi-letter encryption to reduce structural information

e.g. aq → DM
av → GR
vq → XM

Note that these multi-letter n-grams (in fact, digrams) are each to be seen as single plaintext "symbol", and Playfair is thus still a monoalphabetic cipher.

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

➡ Playfair cipher web demo

Playfair cipher: Initialization

- Select a (secret) *keyword*, say CRYPTOCRYO
- Populate a 5*5 matrix, left-to-right, top-to-bottom with the keyword (omit duplicate letters)
- Complete the matrix alphabetically with unused letters

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

I/J are considered as *equivalent*

Playfair cipher: Preprocessing plaintext

key: CRYPTOCRYO

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

- ⌘ repeating "letter pairs" in the plaintext to be separated by a filler – say y

e.g. yummy → yu my my ~~(yu mm y)~~
google → go og le

Playfair cipher: Encryption

key: CRYPTOCRYO

- ⌘ If letters in a pair fall in same row, replace with letter on the right (warp)
- ⌘ If letters in a pair fall in same column, replace with letter beneath (warp)
- ⌘ Otherwise: Replace plaintext letter with letter in same row, but column of the paired letter

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

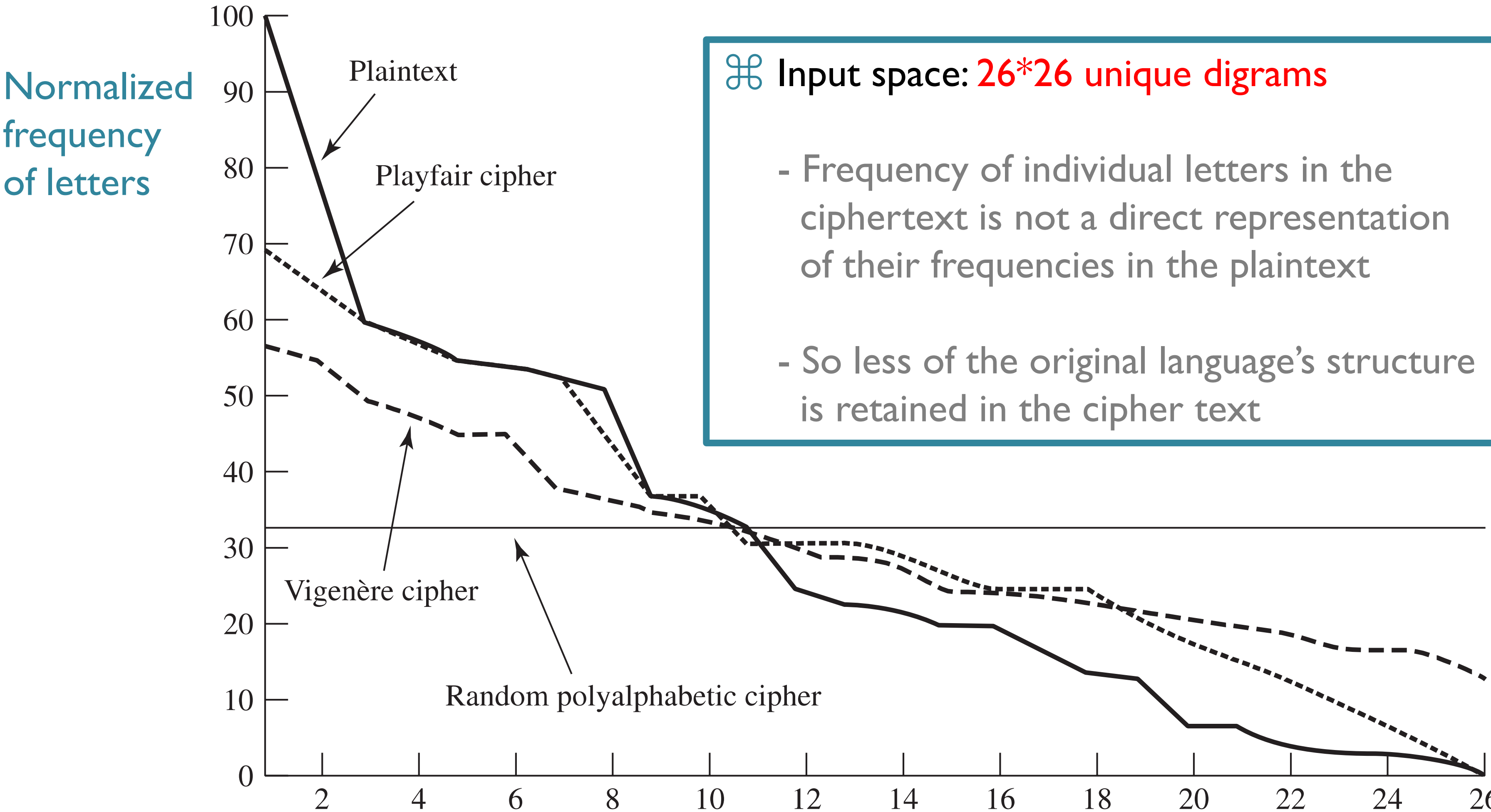
Example:

Plaintext: cool dude
Encryption input: co ol du de
Ciphertext: OF FU OX EO

Different plain text letters were mapped to same ciphertext letter

Same mapping is still possible depending on coincidental co-occurrences

Playfair cipher: Analysis



⌘ Input space: 26*26 unique digrams

- Frequency of individual letters in the ciphertext is not a direct representation of their frequencies in the plaintext
- So less of the original language's structure is retained in the cipher text

Relative frequency of letters is a good way to determine strength of a cipher against frequency analysis

Polyalphabetic substitution

⌘ A set of monoalphabetic ciphers used,
choice of cipher in each step determined by a key

e.g., Vigenère cipher

plaintext:	$p_0, p_1, p_2, \ldots p_{n-1}$
keyword:	$k_0, k_1, k_2, \ldots k_{m-1}$
encryption:	$C_i = (p_i + k_{i \bmod m}) \bmod 26$
decryption:	$p_i = (C_i - k_{i \bmod m}) \bmod 26$

➡ Vigenère cipher web demo

Vigenère cipher: example

```
key:          deceptivedeceptivedeceptive
plaintext:    wearediscoveredsaveyourself
ciphertext:   ZICVTWQNGRZGVTWAVZHCQYGLMGJ
```

key	3	4	2	4	15	19	8	21	4	3	4	2	4	15
plaintext	22	4	0	17	4	3	8	18	2	14	21	4	17	4
ciphertext	25	8	2	21	19	22	16	13	6	17	25	6	21	19

- Multiple substitutions for the same plaintext letter
  - However: there may be periodic repetitions
  - Once an attacker guesses the keyword length, he can attack individual monoalphabetic ciphers

One time pad (aka Vernam cipher)

- If the keyword is as long as the plaintext, and hence
  same substitution is never (systematically) repeated

  Mathematically (provably) impossible to break
  without knowing the key

> Alas, while providing perfect secrecy, one time pad is not practical!

TASOIINEHIUSRNPSTOTCNQE

- All the mechanisms discussed so far used substitution

A fundamentally different technique is to rearrange the plain text in some kind of permutation

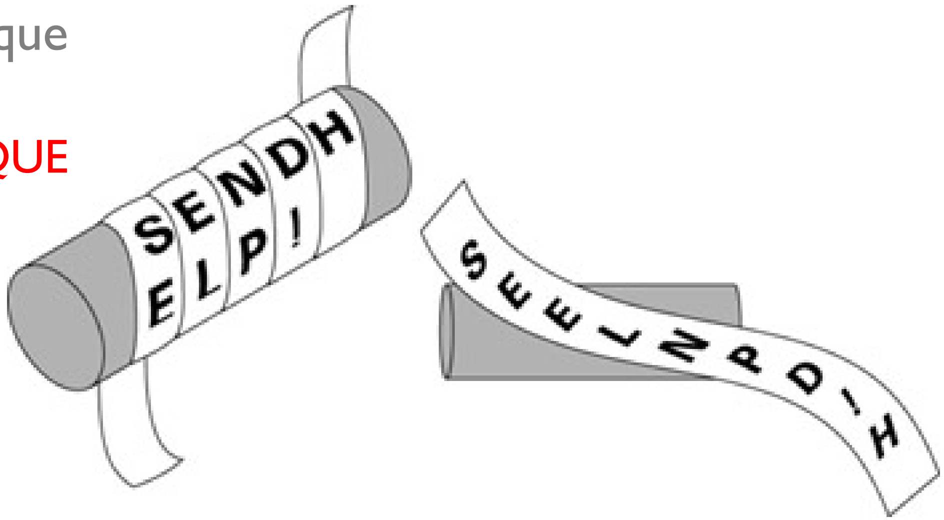
TASOIINEHIUSRNPSTOTCNQE

- All the mechanisms discussed so far used substitution

A fundamentally different technique is to rearrange the plain text in some kind of permutation

Simplest example: rail fence technique

Plaintext: TRANSPOSITION TECHNIQUE
Take odd letters: TASOIIN…
Take even letters: RNPSTOT…
Merge the two: ???



Transposition technique

- A slightly more sophisticated technique

Plaintext:	a	t	t	a	c	k	p
	o	s	t	p	o	n	e
	d	u	n	t	i	l	t
	w	o	a	m	x	y	z

Transposition technique

⌘ A slightly more sophisticated technique

Key:	4	3	1	2	5	6	7
Plaintext:	a	t	t	a	c	k	p
	o	s	t	p	o	n	e
	d	u	n	t	i	l	t
	w	o	a	m	x	y	z

Ciphertext: TTNAAPTMTSUOAODWCOIXKNLYPETZ

Easy to recognize: same letter frequencies as plain text.

- Arrange ciphertext in matrices of varying sizes, and play around with rearrangements
- Di/tri-grams help: in guessing matrix dimension, interpolating column permutation

Transposition technique

Key:	4	3	1	2	5	6	7
Plaintext:	a	t	t	a	c	k	p
	o	s	t	p	o	n	e
	d	u	n	t	i	l	t
	w	o	a	m	x	y	z

⌘ Reapply same transposition once more

Key:	4	3	1	2	5	6	7
Input:	t	t	n	a	a	p	t
	m	t	s	u	o	a	o
	d	w	c	o	i	x	k
	n	l	y	p	e	t	z

Output: NSCYAUOPTTWLTMDNAOIEPAXTTOKZ

Reapplication makes it harder to
- guess the matrix dimension
- interpolate the column permutation

Three ideas

Substitution

- Substitute plaintext symbols
- Poly-alphabetic substitution is better resilient to frequency analysis

Transposition

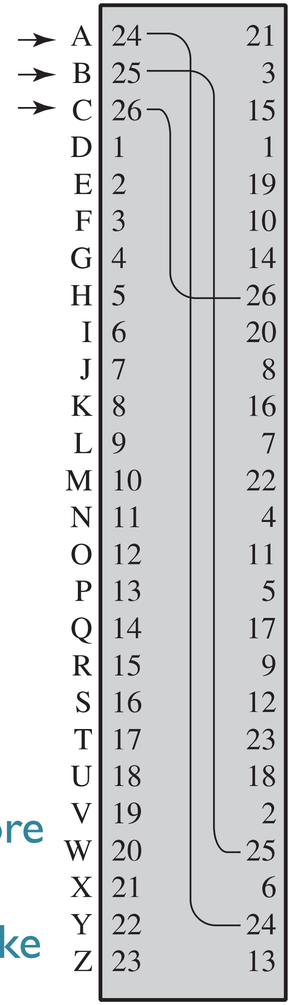
- Reorder (permute) the sequence of symbols

Cascade

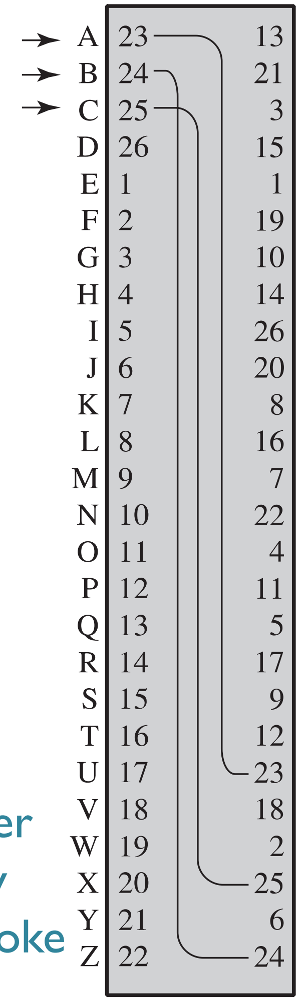
- (Re-)apply multiple times the smaller units of encryption, to realize a stronger encryption

Rotor machines

⌘ 1 rotor:
Polyalphabetic substitution of period 26
i.e. 26 different monoalphabetic ciphers



before key stroke

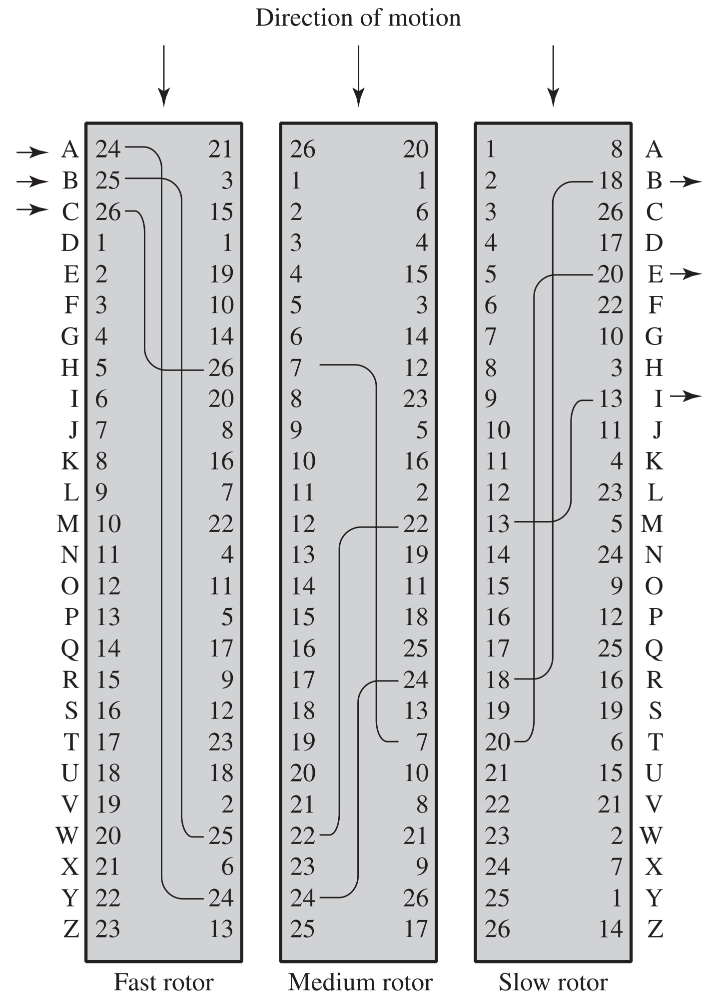


after key stroke

Rotor machines

⌘ Multiple rotors:

e.g. 3 rotors: 26*26*26 = 17,576 different monoalphabetic substitutions before repetitions



➡ Enigma machine web demo

Side note: Steganography

Yesterday I was thinking of how to teach this course meaningfully.
Over the years I have witnessed that students of varied mathematical skills take it.
Under the circumstance, I need to calibrate it to make things accessible for all.
Still, I also need to make sure that the sharpest of the students feel stimulated.
However, it then becomes difficult to find a meaningful balance.
Another thing to consider, is that, I must ensure that students do learn the skills.
Learning hard skills is however a difficult thing, and not everything about it is fun.
Lest you misunderstand me, I don't want to make it inaccessible for the sake of it.
Nonetheless, some difficult mathematical concepts will have to be mastered.
Otherwise, there is no way to explain the inner workings of crypto algorithms.
Therefore, finally I came to the conclusion that there is no easy way out of this.
Priority should be in making sure that the quality of learning is not compromised.
At the same time, we have to try to help the weaker students.
Still, as a student, ultimately you have to take the leadership in learning.
Security is difficult, yet ultimately crucial, and will be worth the effort, otherwise …

Side note: Steganography

Yesterday I was thinking of how to teach this cou
Over the years I have witnessed that students of
Under the circumstance, I need to calibrate it to
Still, I also need to make sure that the sharpest of
However, it then becomes difficult to find a meani
Another thing to consider, is that, I must ensure t
Learning hard skills is however a difficult thing, an
Lest you misunderstand me, I don't want to make
Nonetheless, some difficult mathematical concept
Otherwise, there is no way to explain the inner w
Therefore, finally I came to the conclusion that th
Priority should be in making sure that the quality
At the same time, we have to try to help the wea
Still, as a student, ultimately you have to take the
Security is difficult, yet ultimately crucial, and will

Steganography: Hiding in plainsight

Cryptanalyst models

Self-study
(examinable)

Type of Attack	Known to Cryptanalyst
Ciphertext Only	• Encryption algorithm • Ciphertext
Known Plaintext	• Encryption algorithm • Ciphertext • One or more plaintext–ciphertext pairs formed with the secret key
Chosen Plaintext	• Encryption algorithm • Ciphertext • Plaintext message chosen by cryptanalyst, together with its corresponding ciphertext generated with the secret key
Chosen Ciphertext	• Encryption algorithm • Ciphertext • Ciphertext chosen by cryptanalyst, together with its corresponding decrypted plaintext generated with the secret key
Chosen Text	• Encryption algorithm • Ciphertext • Plaintext message chosen by cryptanalyst, together with its corresponding ciphertext generated with the secret key • Ciphertext chosen by cryptanalyst, together with its corresponding decrypted plaintext generated with the secret key

Wrap up: Important concepts

- Security threats and goals
  - e.g., CIA triad
  - basic security concepts and definitions
- Types of attacks
  e.g., Active/Passive
- Cryptography vs Steganography
- Symmetric (secret) key cryptography
  - elaborated with classical cipher examples
  - three ideas: substitution, transposition, cascade
  - difference between brute-force vs cryptanalysis
  - different models of cryptanalysis

➡ Web demos

Self study (examinable)

- ⌘ Chapter 3, sections 3.1-3.5 from
  Cryptography & Network Security (7th ed) by W. Stallings
  including other specific ciphers (e.g., Hill cipher), and discussions (e.g., types of attacks based on what is known to the cryptanalyst) that have not been fully covered in the lectures