Introduction to Cryptography

January 20, 2022

Today

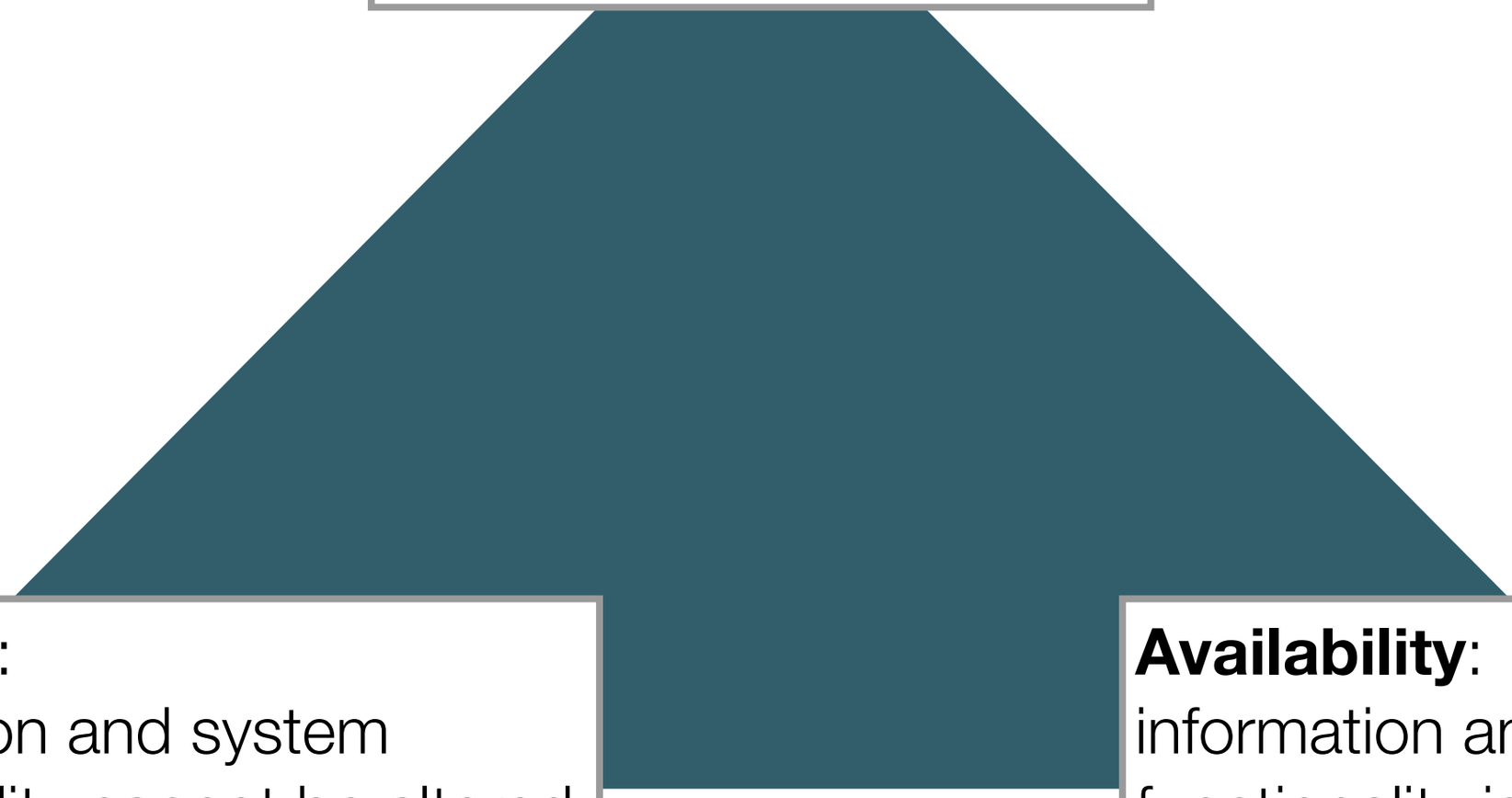
1. Attackers
   What should we assume about them?

2. Cryptography
   What is cryptography?

   - classic cryptography *(and why it doesn't work)*
   - towards perfect security *(but don't get your hopes up, it's not practical)*

Reminder:
Security Objectives

Confidentiality:
information is not available to unauthorized entities



Integrity:
information and system functionality cannot be altered by unauthorized entities

Availability:
information and system functionality is available to authorized entities

1. Attackers

“It is said that if you know your enemies and know yourself,
you will not be imperiled in a hundred battles;
if you do not know your enemies but do know yourself,
you will win one and lose one;
if you do not know your enemies nor yourself,
you will be imperiled in every single battle.”

— Sun Tzu, *Art of War*

1. Attackers

Is this secure?

Or is this?

Depends…

Attacker Model

- We can define security only **with respect to an attacker model**
  - what the attacker **can do**
  - what the attacker **knows**
  - what the attacker **wants to achieve**
  - ...
- *Example:* none of the commonly used cryptographic primitives can withstand an attack with unlimited computational power

Examples of Attacker Types in Practice

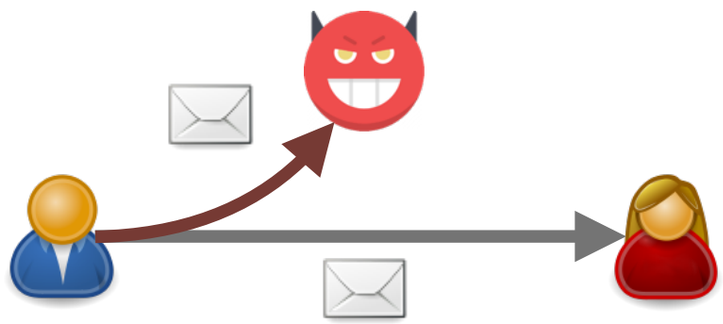
Type	Motivation	Capabilities	Objective
"Script-kiddies"	reputation gain	techniques and software developed by others	integrity (*e.g.*, website defacement) or availability
Cybercriminals	financial gain	expertise/investment in finding/purchasing vulnerabilities, infrastructure to support their operations	confidentiality (*e.g.*, stealing financial information) or integrity (*e.g.*, ransomware)
Industrial espionage	information gain	targeted attacks, ample resources	confidentiality (*e.g.*, learning trade secret)
Cyberwarfare	information gain or causing damage	heavy investments in multiple area of security, targeted attacks	confidentiality (*e.g.*, espionage) or integrity/ availability (*e.g.*, sabotage)

Attacker Model

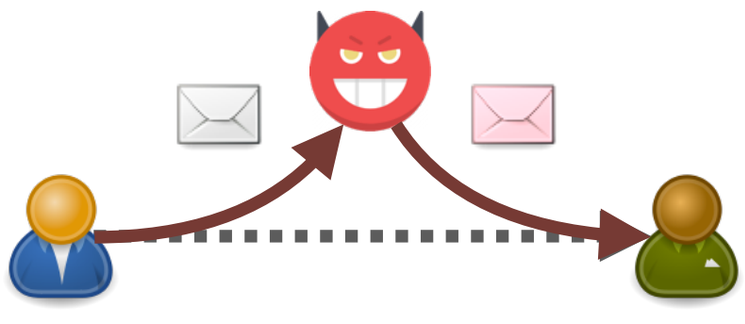
- We can define security only **with respect to an attacker model**
  - what the attacker **can do**
  - what the attacker **knows**
  - what the attacker **wants to achieve**
  - ...
- *Example:* none of the commonly used cryptographic primitives can withstand an attack with unlimited computational power
- It is generally **better to overestimate the attacker**'s capabilities, knowledge, and determination than to underestimate them

Examples of Attacker Capabilities

- *Communications security:*
  access to communication channel



passive (*e.g.*, eavesdropping, traffic analysis)



active (*e.g.*, tampering, replaying, interrupting, masquerading)

- *System security:*
  access to system
  - **remote**:
    without any prior authorization (*e.g.*, via Internet or other public network)
  - **local**: with some prior authorization (*e.g.*, using an unprivileged user account)
  - **physical access**

Safe Assumptions for Attacker's Knowledge

Attacker may know

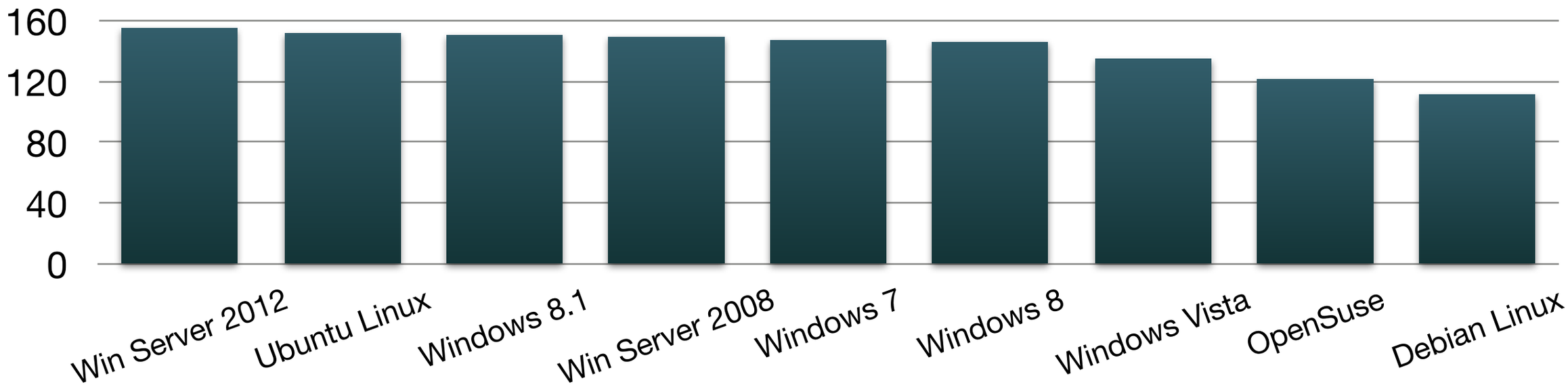
- algorithms
- system design
- implementation
- configuration
- …

Attacker cannot know

- truly random values

Security by Obscurity

- Providing security by keeping the design or implementation of a system secret
- Security experts, researchers, standards bodies, etc. generally **reject** this idea
- Obscurity may **slow down** an attack, but cannot stop it
  - if we thought of an idea, an attacker might also think of it
  - an attacker may try its attack for many possible design and implementation choices
- **False sense of security** may be very dangerous
- *Example*: security of closed-source vs. open-source software



number of vulnerabilities in the *National Vulnerability Database* for 2015

2. Cryptography

Secure communication in the presence of adversaries

Communications Security

`www.example.com`

web browser

web server

Communications Security

`www.example.com`

non-secure channel

web server

Communications Security

sender

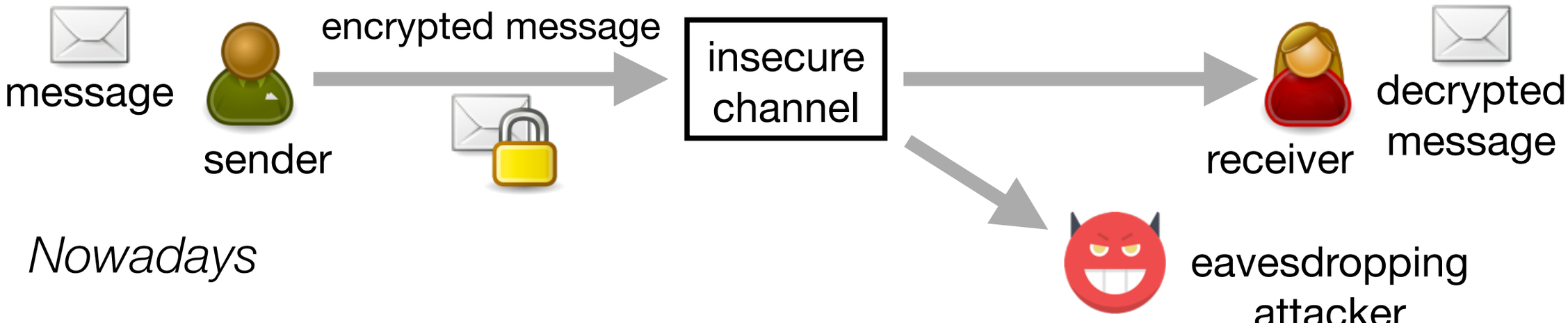
receiver

Cryptography

- Etymology

crypto + graphy
κρυπτός + γράφειν
“secret” + “writing”

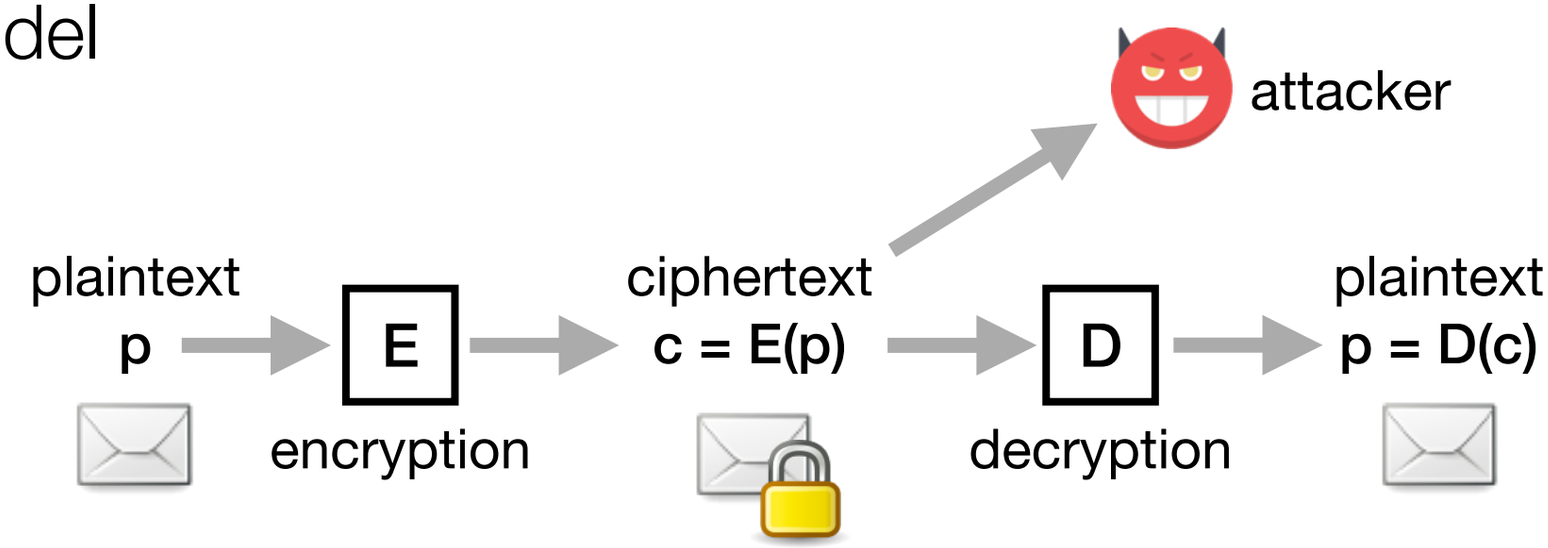
- *Traditionally*: **confidentiality → encryption**



- *Nowadays*
  - integrity → message authentication
  - non-repudiation → digital signatures
  - …

Encryption

- Basic model

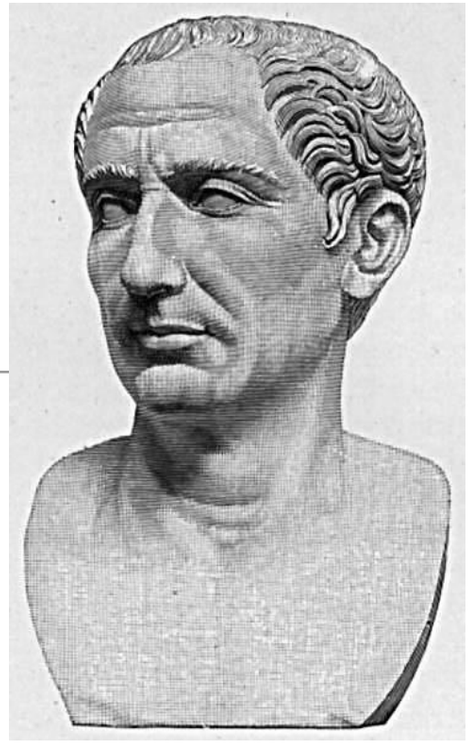


  - **p**: plaintext
  - **E**: encryption algorithm / cipher
  - **c**: ciphertext
  - **D = E^{-1}**: decryption algorithm / cipher

- Attacker's goal: recover the plaintext for a given ciphertext

CLASSICAL CRYPTOGRAPHY

Caesar Cipher



- Named after Julius Caesar, first recorded user of the scheme
- Shift each letter of the plaintext by three letters down the alphabet:

```
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
```

  - *plaintext:* “meet me after the toga party”
  - *ciphertext:* “phhw ph diwhu wkh wrjd sduwb”

- Formally
  - assign an integer to each letter:

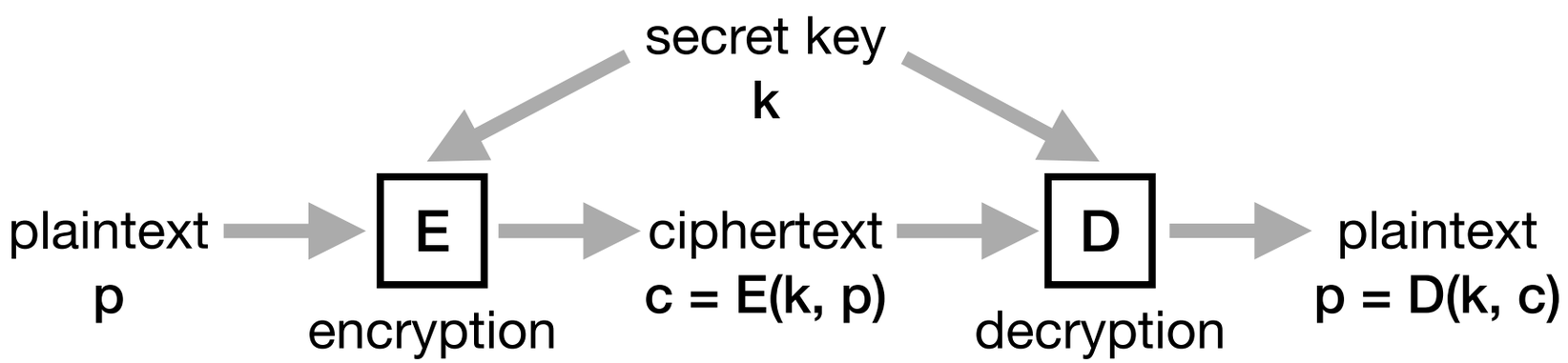
```
a  b  c  d  e  f  g  h  i  j  k  l  m  n  o  p  q  r  s  t  u  v  w  x  y  z
0  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
```

  - encryption: **c = E(p) = p + 3 mod 26**
  - decryption: **p = D(c) = c - 3 mod 26**

Security by obscurity:
confidentiality of the plaintext relies on the attacker not knowing the algorithm

Cryptographic Key

- Sender and receiver use a **random secret key k**
  - **secret key k**: chosen at random and known only by sender and receiver



- Attacker's goal: recover key or some plaintext
- Types of attacks
  - **ciphertext only**: attacker knows the algorithms and the given ciphertext
  - **known plaintext**: attacker also has one or more plaintext-ciphertext pairs
  - **chosen ciphertext**: attacker can also choose one or more ciphertexts (but not the given one) and obtain the corresponding plaintexts
  - **chosen plaintext**: attacker can also choose one or more plaintexts and obtain the corresponding ciphertexts

Generalized Caesar Cipher

- Secret key: **integer k randomly chosen from [1, 25]**
- Algorithms
  - encryption: **c = E(k, p) = p + k mod 26**
  - decryption: **p = D(k, c) = c - k mod 26**
- *Example*: with k = 5
  - plaintext: “`meet me after the toga party`”
  - encryption: c = p + 5 mod 26
  - decryption: c = p - 5 mod 26
  - ciphertext: “`rjjy rj fkyjw ymf ytlf ufwyd`”

Brute-Force Attack

- Brute-force attack:
  attacker tries **every possible key** on a given ciphertext until finding a correct translation into plaintext

- **Known plaintext attack**: search for **k** until a match **p = D(k, c)** is found

- **Ciphertext only attack**: attacker must be able to recognize the correct plaintext to find the key

  - *example*: attacker knows that plaintext is an HTTP request

```
GET / HTTP/1.1
Host: …
Connection: keep-alive
User-Agent: …
```

  - may be more difficult for fragments of compressed plaintexts

- On average, **half of all possible keys** must be tried to achieve success

Affine Cipher

- Modular arithmetic
  - operations: **addition +** and **multiplication** ·, which “wrap around”
  - **additive inverse -x**:
    0 = x + (-x) mod m
    (*example:* -5 = 21 modulo 26 since 5 + (-5) = 5 + 21 = 0 mod 26)
  - **multiplicative inverse x^{-1}**:
    $1 = x \cdot x^{-1} \bmod m$
    - exists only if x and m are coprime (i.e., their greatest common divisor is 1)
      (*example:* only 1, 3, 5, 7, 9, 11, 15, 17, 19, 21, 23, 25 are invertible modulo 26)

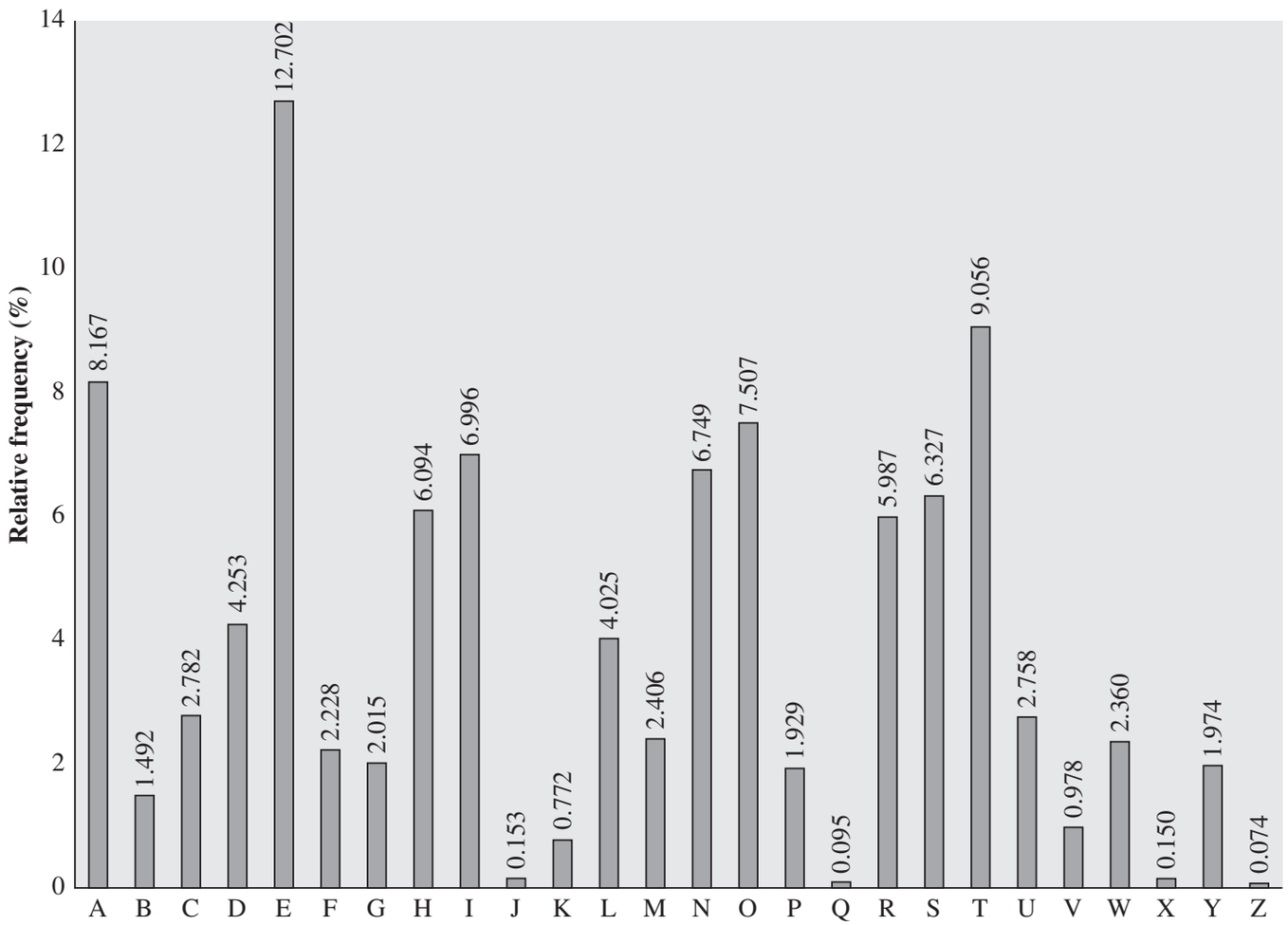
- Secret key: k_1 from {1, 3, …, 11, 15, …, 23, 25} and k_2 from [0, 25]
  - encryption: $\mathbf{c = p \cdot k_1 + k_2 \bmod 26}$
  - decryption: $\mathbf{p = (c - k_2) \cdot k_1^{-1} \bmod 26}$
  - number of possible keys = 12 · 26 = 312

Substitution Cipher

- Secret key: **permutation over the alphabet**
  - *example key:*
    ```
    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
    Q I V R A D X Z E S N C T U L J W F G H Y O P B M K
    ```
  - *plaintext:* “`meet me after the toga party`”
  - *ciphertext:* “`taah ta qdhaf hza hlxq wqfhm`”
- Number of possible keys = $26! \approx 4 \cdot 10^{26} \approx 2^{88}$
  - relatively strong against brute-force attacks
- Vulnerable to known plaintext attacks

Cryptanalysis

- Cryptanalytic attack:
  attacker relies on the **nature of the algorithm** and knowledge of the general **characteristics of the plaintext**
- Breaking a substitution cipher



- *example ciphertext:*
  TAAHTAQDHAFHZAHLXQWQFHM
- *frequency:*
  - A: 5 → E
  - H: 5 → T
  - Q: 3 → A
  - F: 2 → R
  - …
- *plaintext:*
  “meet me after the toga party”

Next lecture:

One-Time Pad and Stream Ciphers

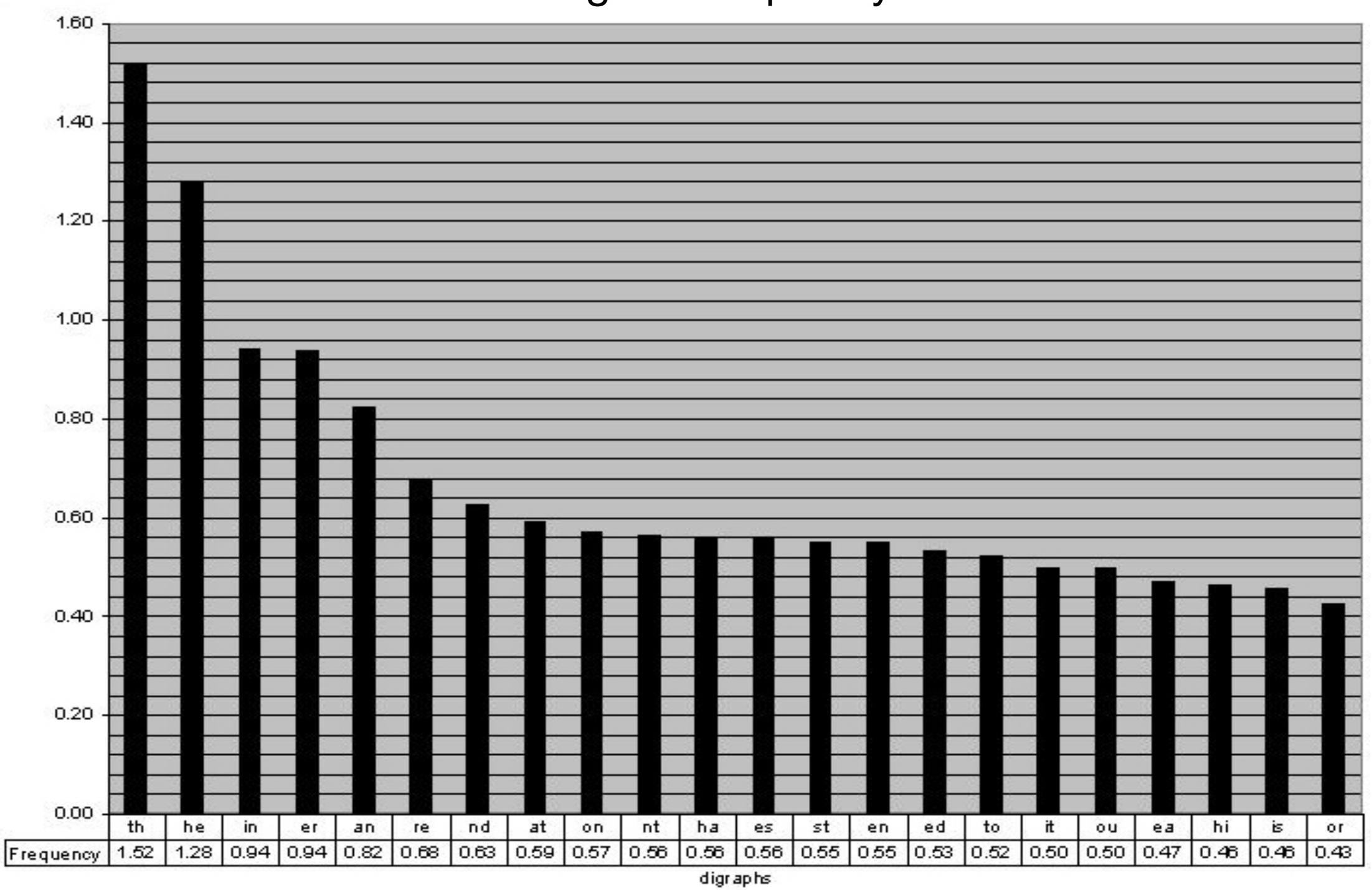
Playfair Cipher

- Invented in 1854 by Charles Wheatstone, but named after Lord Playfair, who promoted the use of the cipher
  - used by British forces in the Second Boer War and World War I
- Secret key: **5 x 5 table filled with letters**
  - keyword + remaining letters of the alphabet

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

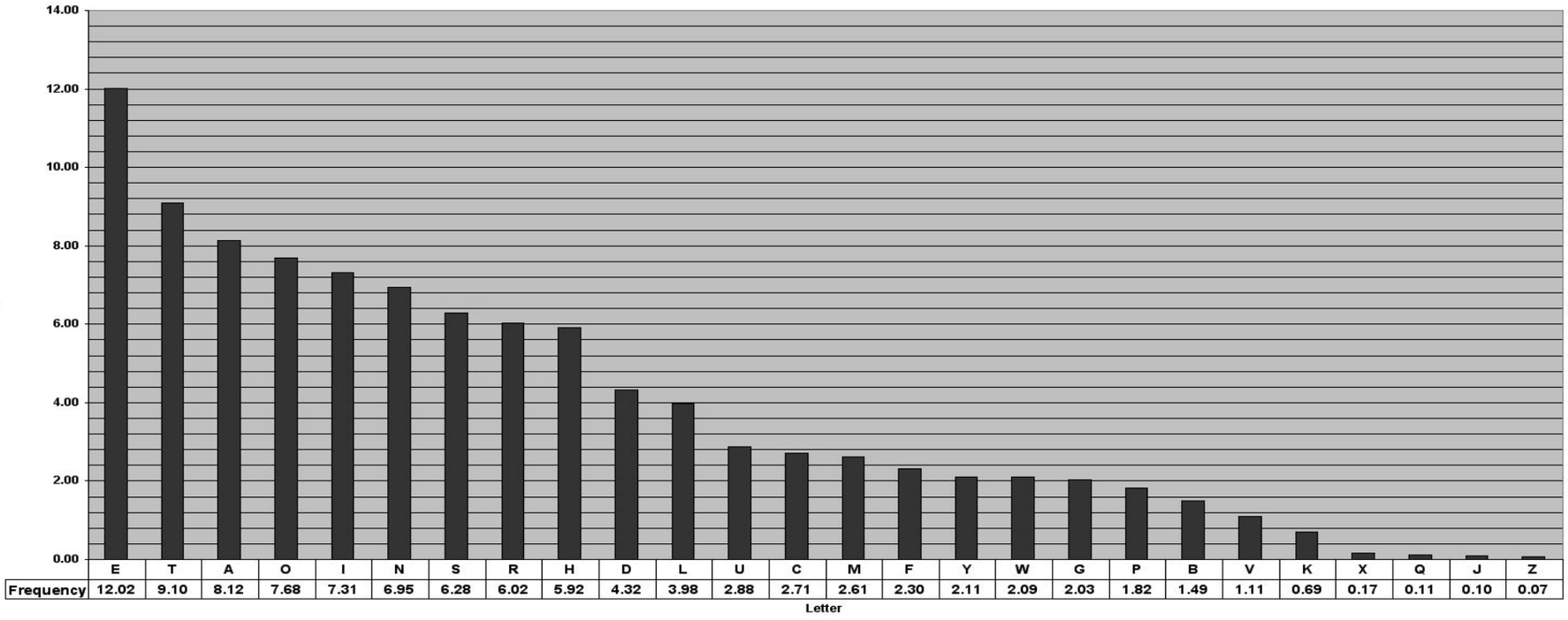
- Plaintext is encrypted **two letters at a time**
  - repeating plaintext letters that are in the same pair are separated with a filler (*e.g.*, X)
  - if the pair is on the same row/column, shift them right/down (with the rows/columns circularly following each other)
  - finally, replace each letter in a pair with the letter that lies in the same row but the column of the other letter (*e.g.*, plaintext pair “BP” becomes ciphertext pair “HS”)

Cryptanalysis for Multiple-Letter Ciphers

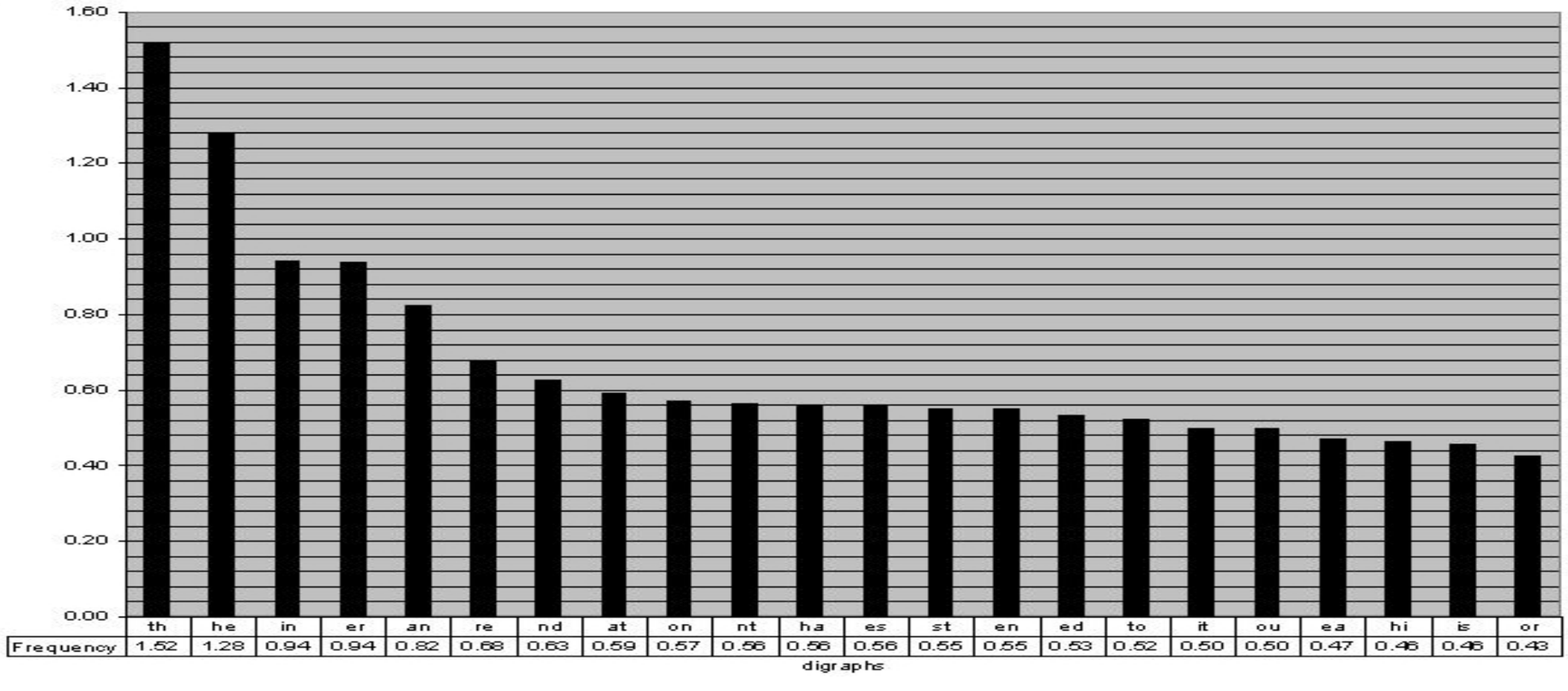


	th	he	in	er	an	re	nd	at	on	nt	ha	es	st	en	ed	to	it	ou	ea	hi	is	or
Frequency	1.52	1.28	0.94	0.94	0.82	0.68	0.63	0.59	0.57	0.56	0.56	0.56	0.55	0.55	0.53	0.52	0.50	0.50	0.47	0.46	0.46	0.43

Cryptanalysis for Multiple-Letter Ciphers



Single letter frequency



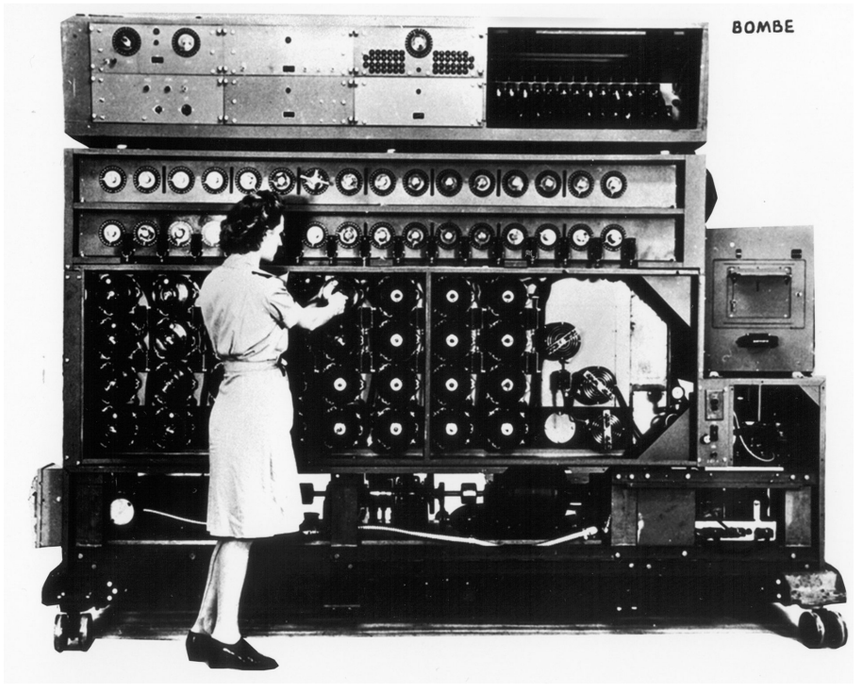
Bigram frequency

Vigenère Cipher

- Reinvented many times through history
  - originally described by Giovan Battista Bellaso in 1553, but named after Blaise de Vigenère
  - used by, for example, the Confederate States of America in the American Civil War
  - in 1917, *Scientific American* characterized this system as “impossible of translation”
- Key: letters k_1, …, k_N (each corresponds to a number from [0, 25])
- Encryption
  - repeat the **N** letters of the key so that it is as long as the plaintext
  - **i** th letter of ciphertext: $c_i = p_i + k_i \bmod 26$
  - *example:*
    - key: `DECEPTIVEDECEPTIVEDECEPTIVE`
    - plaintext: `wearediscoveredsaveyourself`
    - ciphertext: `ZICVTWQNGRZGVTWAVZHCQYGLMGJ`
- Vulnerable to cryptanalysis

Rotor Machines

- Machines based on the rotor principle were used by both Germany (Enigma) and Japan (Purple) in World War II
- British and Polish cryptologist developed electromechanical devices, called bombes, to break Enigma



US Navy Bombe

three-rotor Enigma