Cryptography I: Intro and Basics

CSE 565: Fall 2024
Computer Security

Xiangyu Guo (xiangyug@buffalo.edu)

University at Buffalo

Today's topic

- Announcement
- Review of Last Lecture
- Cryptography I
  - What is it?
  - Core application: Secure communication
  - Recap of discrete probability
  - Recap algorithm analysis
- Summary

Announcement

- Slides for last lecture has been uploaded.
- A quiz about Academic Integrity
  - **Due**: 23:59, Sep 19
  - **Where**: UBLearns
  - Count for 0 pts towards your final grade, but **You must complete this quiz in time, otherwise will receive an F in final grade**.

Review of last lecture

Security objectives

- “CIA Triad”
  - **C**onfidentiality: defend against improper *disclosure* of info
  - **I**ntegrity: defend against improper *modification* of info
  - **A**vailability: defend against improper *denial-of-service*
- Supplementary:
  - Authenticity: *verifiable* info source
  - Accountability: *attributable* activities

Review of last lecture

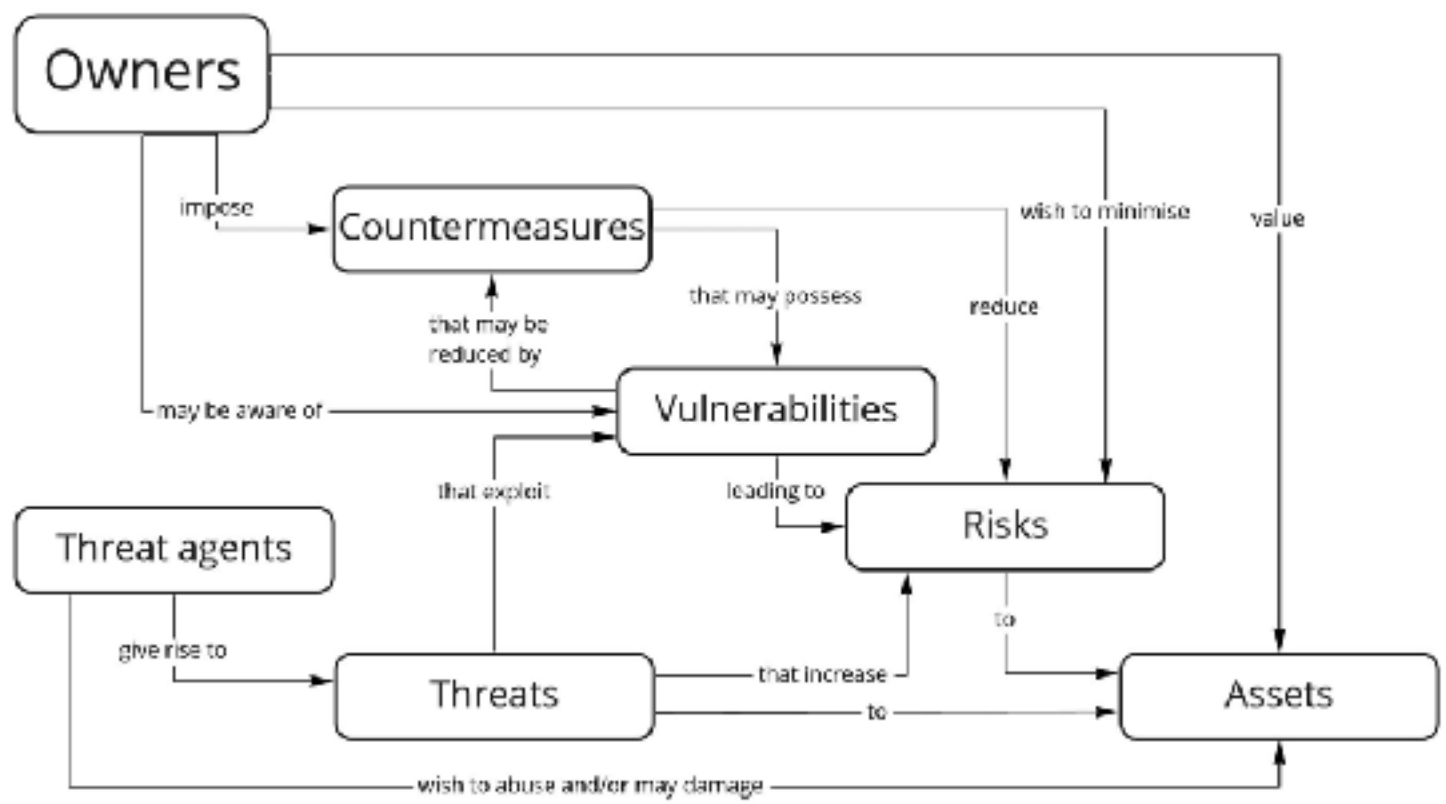
Tools for achieving security objectives

- Confidentiality: Encryption, Authentication -> Authorization
- Integrity: Backup, Checksum, Error Correcting Code
- Availability: Redundancy / Replicas, physical protection, ...
- Authenticity: Digital Signature (non-repudiation)
- Accountability: Logging, Monitoring, Version Control, ...

Review of last lecture

Terminologies

- Adversary (Active; Passive; Insider; Outsider), Attack, Countermeasure
- Risk, Policy, Asset
- Threat, Vulnerability.



Review of last lecture

(More) Terminologies

Attack Surface: the reachable and exploitable vulnerabilities in a system.

- Open ports on outward-facing Web and other servers, and code listening on those ports
- Services available on the inside of a firewall
- Code that processes incoming data, email, XML, office documents, and industry-specific custom data exchange formats
- Interfaces, SQL, and Web forms
- An *employee* with access to sensitive information that is vulnerable to a social engineering attack

Cryptography Intro

What is Cryptography

- Greek: “krypto” = hide
- **Cryptography** – secret writing. Originally, it is the study of encryption principles and methods
    - **Cryptanalysis** – analyzing and breaking secrets. Originally, the study of principles and methods of deciphering ciphertext without knowing key
- The *most basic* problem of cryptography is to ensure *security of communication over insecure media*

What is Cryptography

Cryptography is everywhere

- **Secure communication**:

  - web traffic: HTTPS
  - wireless traffic: 802.11i WPA2 (and WEP), GSM, Bluetooth

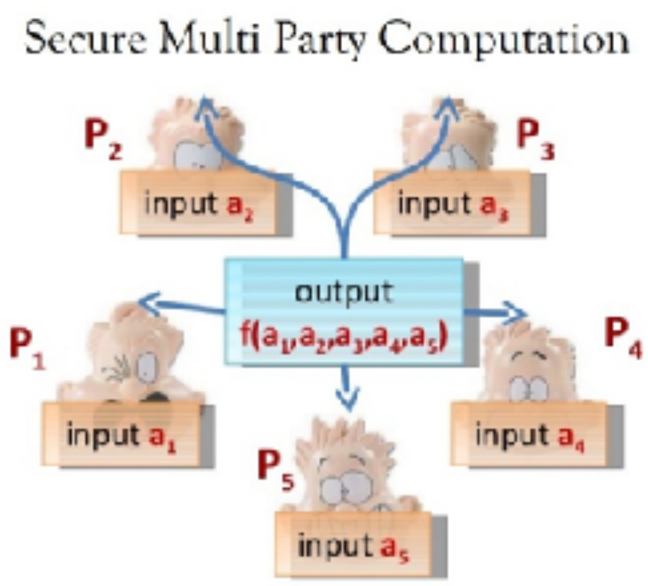
- **Encrypting files on disk**: (MS) BitLocker, TrueCrypt, (Apple) FileVault
- **Content protection** (e.g. DVD, Blu-ray): CSS, AACS, Denuvo
- **User authentication**

- ......

What is Cryptography

Cryptography is everywhere: even fancier apps

- **Cryptocurrency**:
    - Bitcoin, Dogecoin, Ethereum
    - Zero-Knowledge Proof
- **Anonymous voting**: MPC, FHE, ZKP
- **Private Search**: Duckduckgo (PIR),
- **Private Ads Recommendation**: Meta (MPC)
- **Anonymous social network**: Signals (Oblivious RAM)
- ……



Cryptography & Security

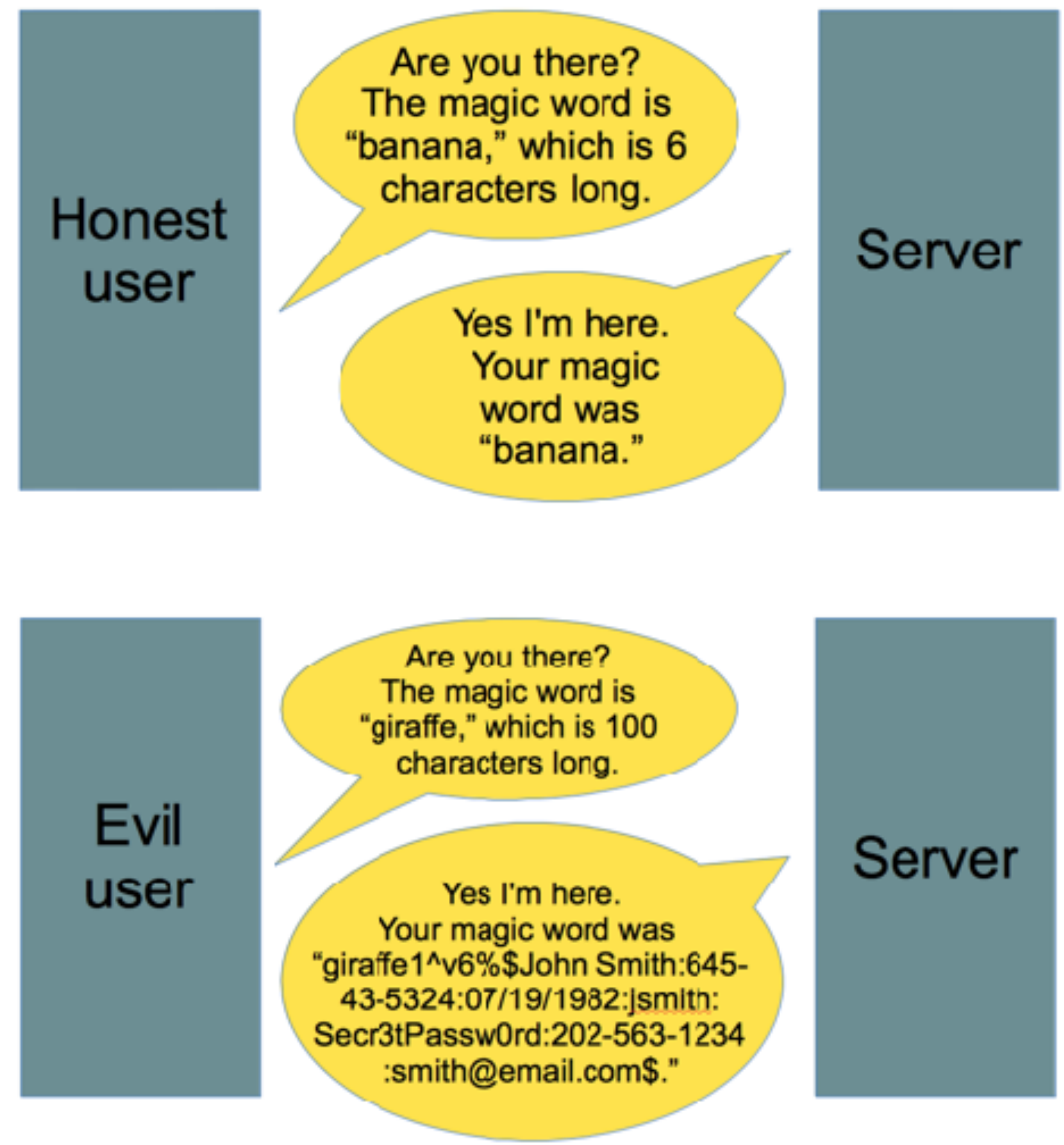
- Can help
  - **Confidentiality**
    - Obscure a message from eavesdroppers (Encryption)
  - **Integrity**
    - Assure recipient that the message was not altered (MAC)
  - **Authenticity**
    - Verify the identity of the source of a message (Digital Signature; MAC)
  - **Non-repudiation**
    - Convince a 3rd party that what was said is accurate (Digital Signature)

Cryptography & Security

- Most people argue cryptography is a branch of mathematics
- Security is about math, engineering, hardware, software, people, etc.
- Attackers try find the weakest link. In most cases, this is *not* the mathematics.

Cryptography & Security

- Example: The HeartBleed bug for OpenSSL
  - Nothing goes wrong with the crypto part
  - Mem leaked due to a forgotten bound check in the implementation of the protocol.

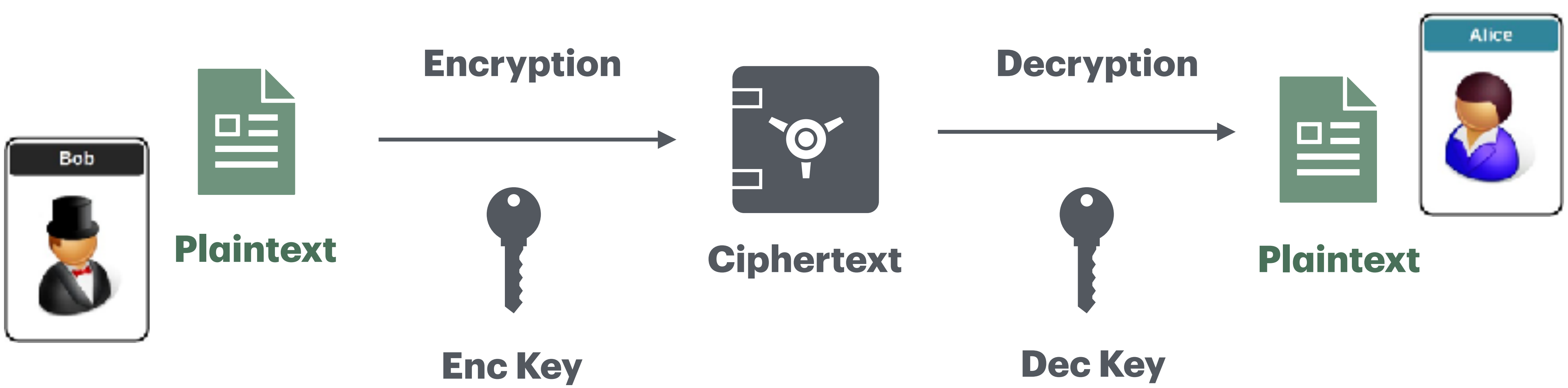


Cryptography & Security

- Cryptography is
  - A basic tool for achieving security objectives
- Cryptography is **NOT**:
  - The solution to all security problems
  - Reliable unless implemented and used properly
    - Failed examples: WEP, Heartbleed, etc
  - Something you should try to invent yourself
    - many many examples of broken ad-hoc designs

Terminologies

- Plaintext: the message to be transmitted or stored.
- Ciphertext: the disguised message.
- Key: Sequence that controls the operation and behavior of the cryptographic algorithm
- Encryption: the process of disguising a message so as to hide the information it contains
- Decryption: the *reverse* of Encryption



Terminologies

- Keyspace: The set of all possible values of keys in a crypto algorithm
- Protocol: an algorithm, defined by a sequence of steps, precisely specifying the actions of multiple parties in order to achieve an objective.
- Cryptosystem: The combination of algorithm, key, and key management functions used to perform cryptographic operations

Kerckhoff's Principle

- French handbook of military cryptography, 1883
- *A cryptosystem should be secure even if everything about the system, except the key, is public knowledge.*
- Don't assume enemy won't know algorithm
  - Can capture machines, find patents, etc.
  - Too expensive to invent new algorithm if it might have been compromised

JOURNAL

DES

SCIENCES MILITAIRES.

Janvier 1883.

LA CRYPTOGRAPHIE MILITAIRE.

« La cryptographie est un auxiliaire puissant de la tactique militaire. »
(Général Lewal, *Études de guerre.*)

I.

LA CRYPTOGRAPHIE DANS L'ARMÉE

Un grand nombre de combinaisons ingénieuses peuvent répondre au but qu'on veut atteindre dans le premier cas ; dans le second, il faut un système remplissant certaines conditions exceptionnelles, conditions que je résumerai sous les six chefs suivants :

1° Le système doit être matériellement, sinon mathématiquement, indéchiffrable ;

2° Il faut qu'il n'exige pas le secret, et qu'il puisse sans inconvénient tomber entre les mains de l'ennemi ;

3° La clef doit pouvoir en être communiquée et retenue sans le secours de notes écrites, et être changée ou modifiée au gré des correspondants ;

4° Il faut qu'il soit applicable à la correspondance télégraphique ;

5° Il faut qu'il soit portatif, et que son maniement ou son fonctionnement n'exige pas le concours de plusieurs personnes ;

6° Enfin, il est nécessaire, vu les circonstances qui en commandent l'application, que le système soit d'un usage facile, ne demandant ni tension d'esprit, ni la connaissance d'une longue série de règles à observer.

Secure communication

Secure Socket Layer (SSL) / TLS

No eavesdropping (*Confidentiality*) & No tampering (*Integrity*)

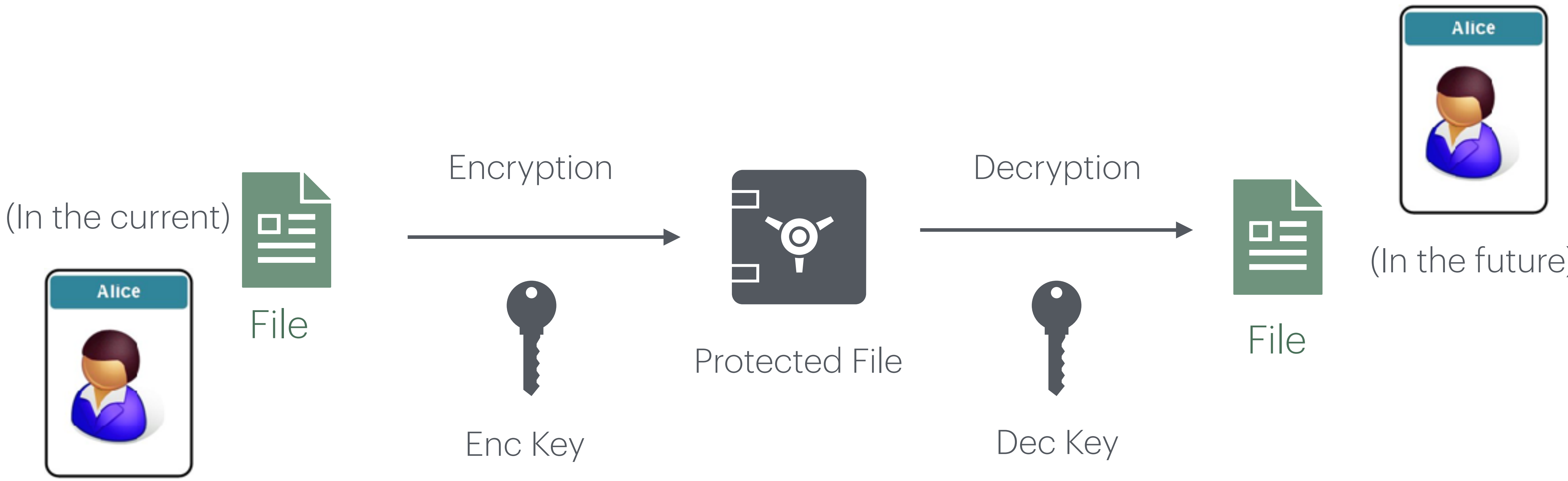
Secure communication

Secure Socket Layer (SSL) / TLS

- Two main parts
  - TLS Handshake: Establish *shared secret key* using **public-key cryptography**.
  - TLS Record: Transmit data using **symmetric encryption** based on the *shared secret key*. Ensure confidentiality and integrity.

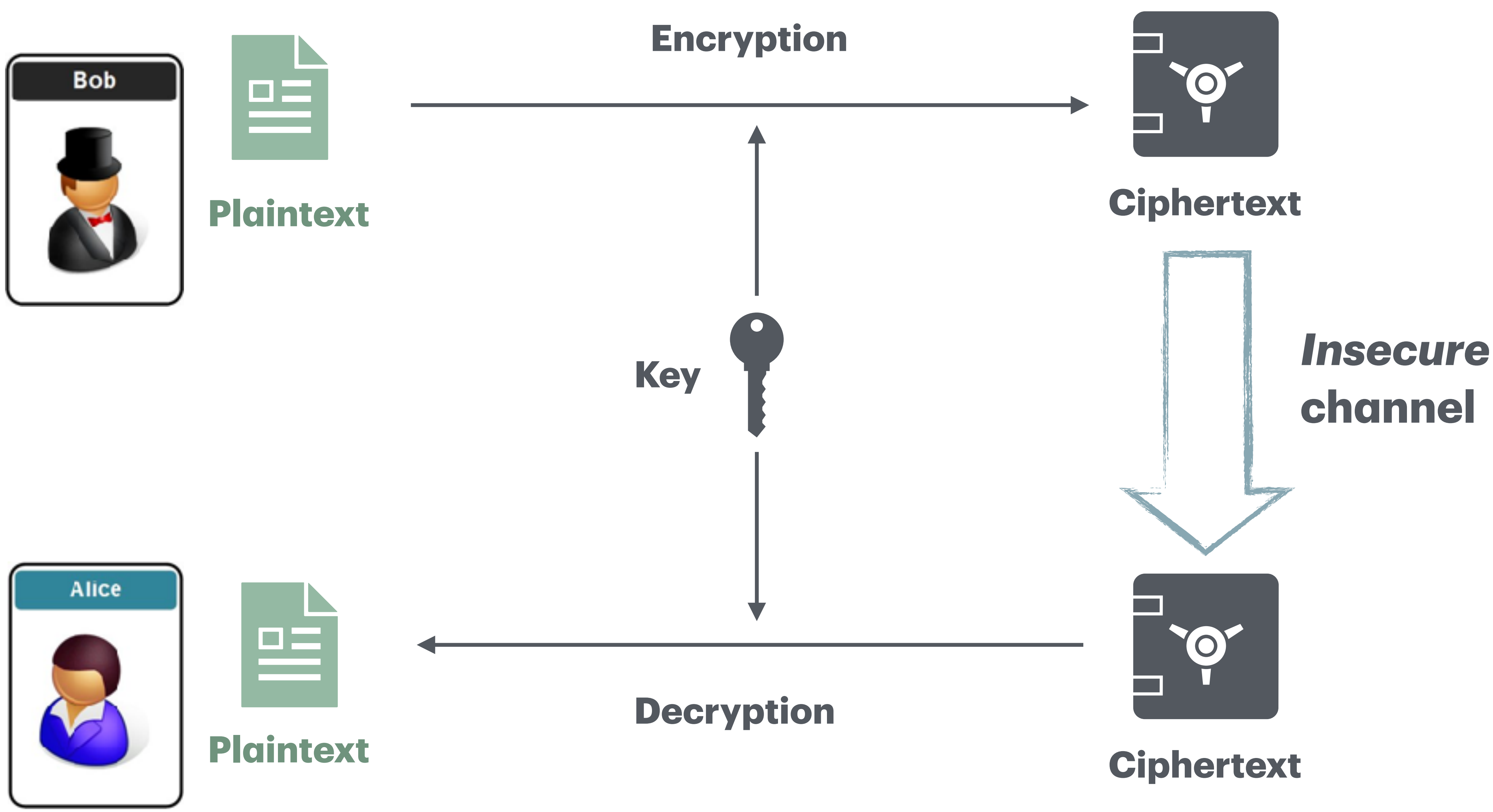
Secure communication

Protected files on disk

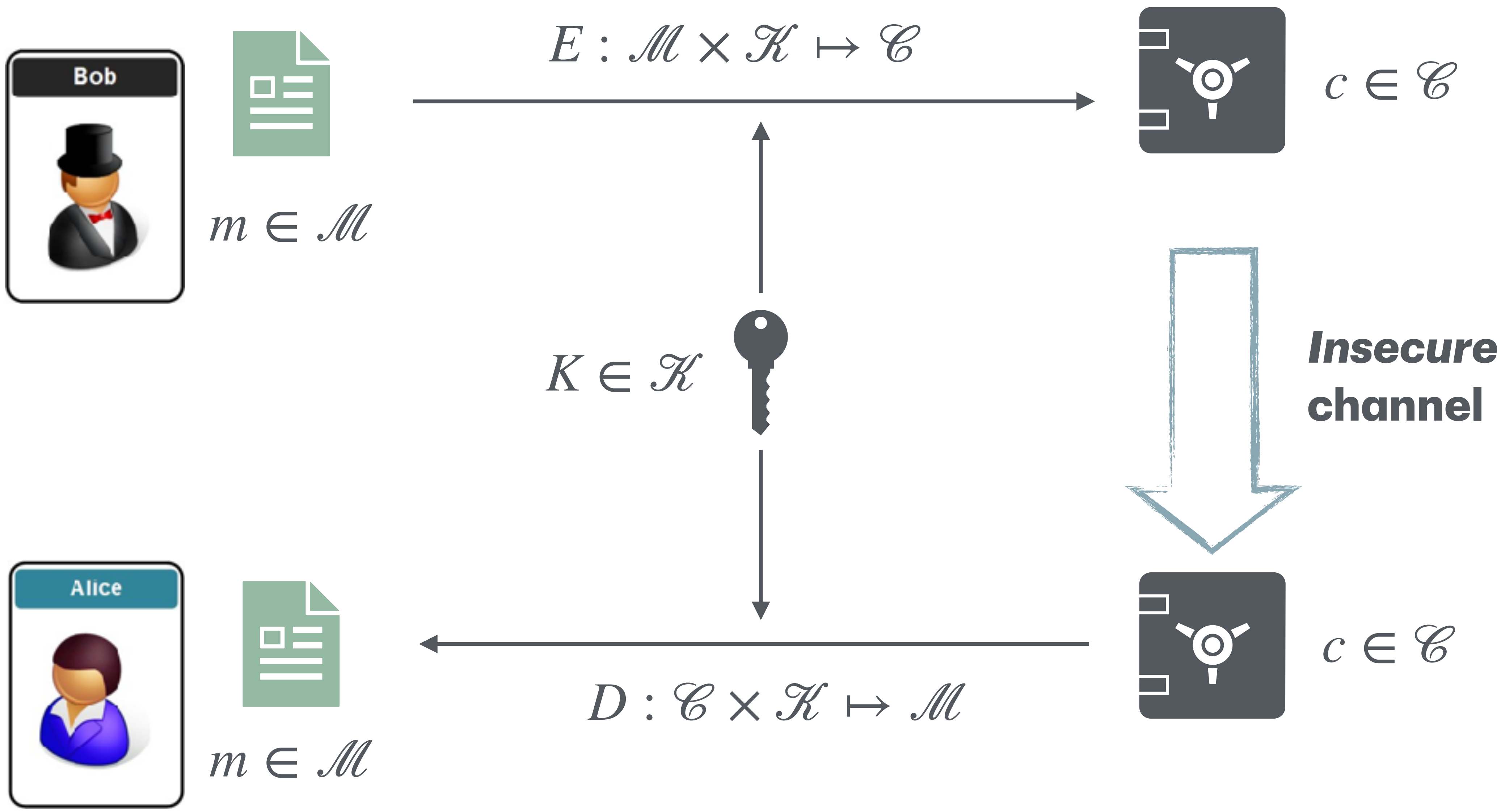


Analogous to communication: Alice *today* sends a message to Alice *tomorrow*.

Building Block: Symmetric Ciphers



Building Block: Symmetric Ciphers



$$D(E(\text{Plaintext}, K), K) = \text{Plaintext}$$

Building Block: Symmetric Ciphers

- **Def**: A symm. cipher defined over $(\mathscr{K}, \mathscr{M}, \mathscr{C})$ is a pair of ("efficient") algs (E, D) where
  - $E : \mathscr{M} \times \mathscr{K} \mapsto \mathscr{C}$: Enc(Ptext, Key)=Ctext
  - $D : \mathscr{C} \times \mathscr{K} \mapsto \mathscr{M}$: Dec(Ctext, Key)=Ptext
  - $D(E(\text{Ptext}, K), K) = \text{Ptext}$

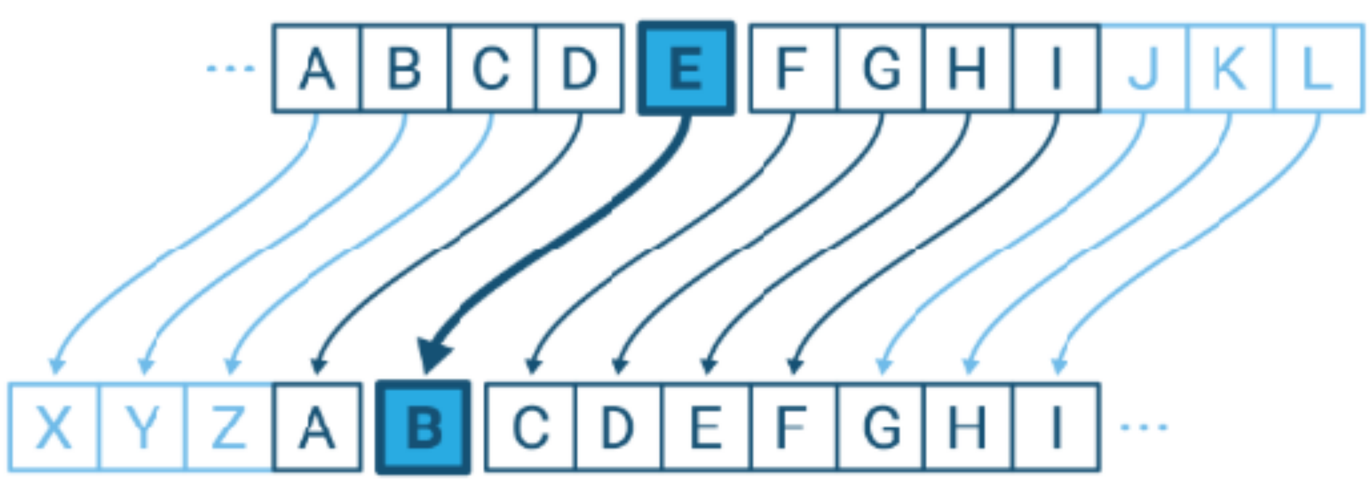
Building Block: Symmetric Ciphers

- Desired properties of cipher (E, D)
  - Kerckhoff's: secrecy depends only on key K
  - Without knowing K must be "<u>hard</u>" to invert E
  - "<u>Easy</u>" to compute E and D

Classical Symm. Ciphers (And how they broke)

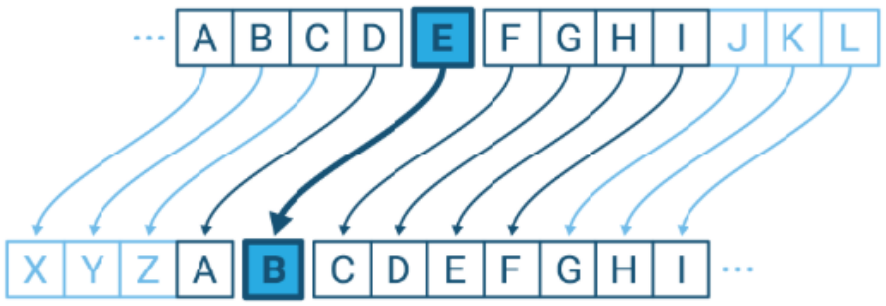
Ceasar Ciphers

- Simple cipher used by Caesar to encrypt messages
- Shift each letter in the alphabet by a fixed distance, *wrap around* at the end.
- Example: a shift of 23 (or -3 mod 26)



Ceasar Ciphers

- As a cipher, both the Enc E and Dec D are the shift table



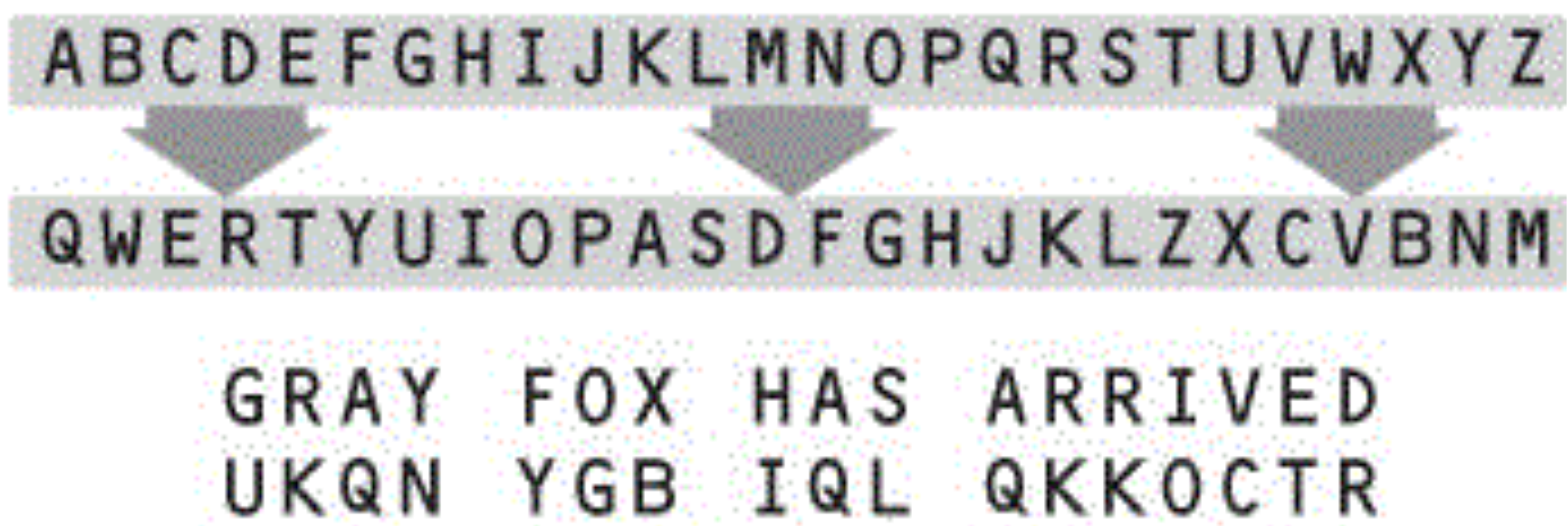
- Key: the shift distance (23).
- Keyspace: {0,...,25}
- Example:
  - Plaintext: ATTACK $\Longrightarrow$ Ciphertext: XQQXZH

Attack Ceasar Ciphers

- Attackers
  - Do not know the key or the shift table
  - Only knows ciphertext (a Ciphertext-only Attack)
- Make a list of all possible keys and corresponding plaintext. If you expected the unencrypted text to be in English, you could easily figure out which word was right
- Small key space (25)
- Not so safe

Substitution Ciphers

- Extend Caesar: instead of shifting, do a permutation



Substitution Ciphers

- Key: the substitution table
- Key space: 26 (ways to choose what **A** maps to) * 25 (**B** can map to anything else) * 24 (**C** can map to anything else) * ... * 1 (only one choice left for **Z**) = 26!
- $|\mathscr{K}| = 26! \approx 2^{88}$
- Not bad?

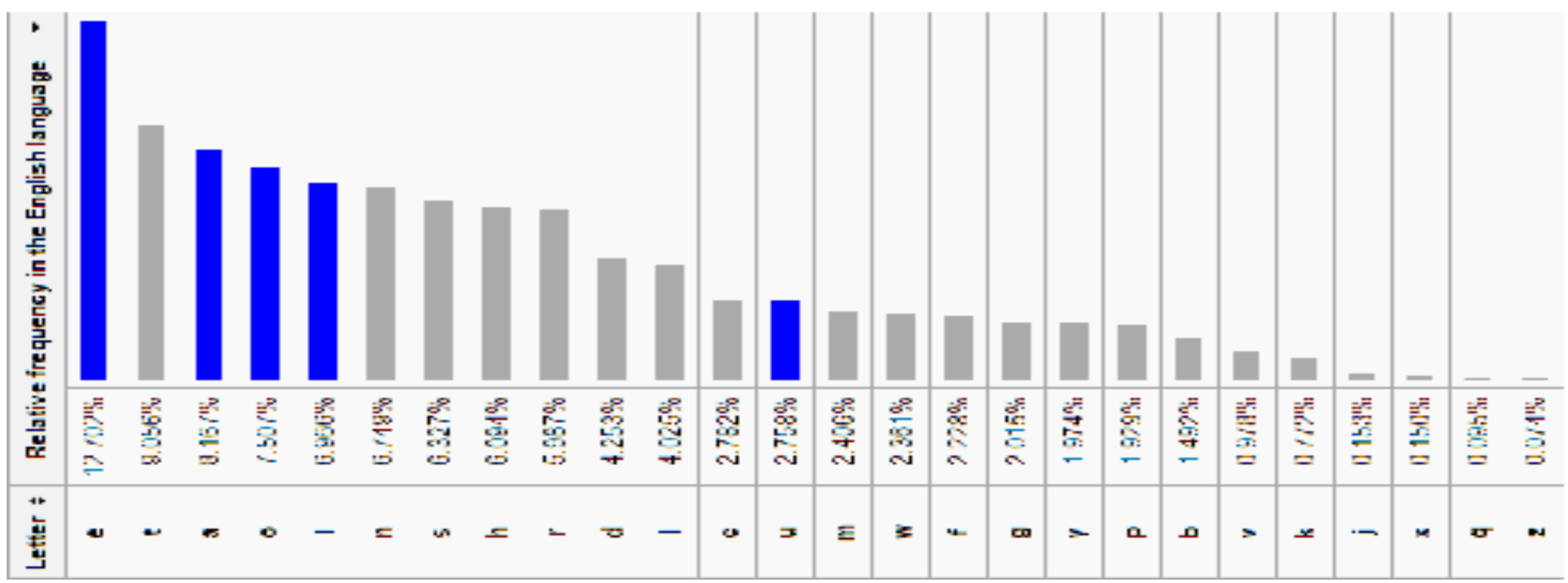
Attack the Substitution Ciphers

eohitkztbzl hkgrxetr wn q esqlloeqs eohitk (qfr lgdt dgrtkf eohitkl) voss ktctqs lzqzolzoeqs ofygkdqzogf qwgxz zit hsqofztbz, qfr ziqz ofygkdqzogf eqf gyztf wt xltr zg wktqa zit eohitk. qyztk zit rolegctkn gy yktjxtfen qfqsnlol, ftqksn qss lxei eohitkl egxsr wt wkgatf wn qf ofygkdtr qzzqeatk. lxei esqlloeqs eohitkl lzoss tfpgn hghxsqkozn zgrqn, zigxui dglzsn ql hxmmstl (ltt eknhzgukqd). zit qkqw dqzitdqzoeoqf qfr hgsndqzi qs-aofro vkgzt q wgga gf eknhzgukqhin tfzozstr kolqsqi yo olzoaikqp qs-dx'qddq (dqfxlekohz ygk zit rteohitkofu eknhzgukqhioe dtllqutl), vioei rtlekowtr zit yoklz afgvf xlt gy yktjxtfen qfqsnlol eknhzqfqsnlol zteifojxtl.

- <u>What is the most common letter in English text?</u>
  - Ans: **E**, with a frequency of 12.7%
  - So its substitution is likely to have the highest frequency.

Attack the Substitution Ciphers

eohitkztbzl hkgrxetr wn q esqlloeqs eohitk (qfr lgdt dgrtkf eohitkl) voss ktctqs lzqzolzoeqs ofygkdqzogf qwgxz zit hsqofztbz, qfr ziqz ofygkdqzogf eqf gyztf wt xltr zg wktqa zit eohitk. qyztk zit rolegctkn gy yktjxtfen qfqsnlol, ftqksn qss lxei eohitkl egxsr wt wkgatf wn qf ofygkdtr qzzqeatk. lxei esqlloeqs eohitkl lzoss tfpgn hghxsqkozn zgrqn, zigxui dglzsn ql hxmmstl (ltt eknhzgukqd). zit qkqw dqzitdqzoeoqf qfr hgsndqzi qs-aofro vkgzt q wgga gf eknhzgukqhin tfzozstr kolqsqi yo olzoaikqp qs-dx'qddq (dqfxlekohz ygk zit rteohitkofu eknhzgukqhioe dtllqutl), vioei rtlekowtr zit yoklz afgvf xlt gy yktjxtfen qfqsnlol eknhzqfqsnlol zteifojxtl.



Attack the Substitution Ciphers

eohitkztbzl hkgrxetr wn q esqlloeqs eohitk (qfr lgdt dgrtkf eohitkl) voss ktctqs lzqzolzoeqs ofygkdqzogf qwgxz zit hsqofztbz, qfr ziqz ofygkdqzogf eqf gyztf wt xltr zg wktqa zit eohitk. qyztk zit rolegctkn gy yktjxtfen qfqsnlol, ftqksn qss lxei eohitkl egxsr wt wkgatf wn qf ofygkdtr qzzqeatk. lxei esqlloeqs eohitkl lzoss tfpgn hghxsqkozn zgrqn, zigxui dglzsn ql hxmmstl (ltt eknhzgukqd). zit qkqw dqzitdqzoeoqf qfr hgsndqzi qs-aofro vkgzt q wgga gf eknhzgukqhin tfzozstr kolqsqi yo olzoaikqp qs-dx'qddq (dqfxlekohz ygk zit rteohitkofu eknhzgukqhioe dtllqutl), vioei rtlekowtr zit yoklz afgvf xlt gy yktjxtfen qfqsnlol eknhzqfqsnlol zteifojxtl.

Bigram Frequencies

TH	:	2.71	EN	:	1.13	NG	:	0.89
HE	:	2.33	AT	:	1.12	AL	:	0.88
IN	:	2.03	ED	:	1.08	IT	:	0.88
ER	:	1.78	ND	:	1.07	AS	:	0.87
AN	:	1.61	TO	:	1.07	IS	:	0.86
RE	:	1.41	OR	:	1.06	HA	:	0.83
ES	:	1.32	EA	:	1.00	ET	:	0.76
ON	:	1.32	TI	:	0.99	SE	:	0.73
ST	:	1.25	AR	:	0.98	OU	:	0.72
NT	:	1.17	TE	:	0.98	OF	:	0.71

Trigram Frequencies

THE	:	1.81	ERE	:	0.31	HES	:	0.24
AND	:	0.73	TIO	:	0.31	VER	:	0.24
ING	:	0.72	TER	:	0.30	HIS	:	0.24
ENT	:	0.42	EST	:	0.28	OFT	:	0.22
ION	:	0.42	ERS	:	0.28	ITH	:	0.21
HER	:	0.36	ATI	:	0.26	FTH	:	0.21
FOR	:	0.34	HAT	:	0.26	STH	:	0.21
THA	:	0.33	ATE	:	0.25	OTH	:	0.21
NTH	:	0.33	ALL	:	0.25	RES	:	0.21
INT	:	0.32	ETH	:	0.24	ONT	:	0.20

Attack the Substitution Ciphers

eohitkztbzl hkgrxetr wn q esqlloeqs eohitk (qfr lgdt dgrtkf eohitkl) voss ktctqs lzqzolzoeqs ofygkdqzogf qwgxz zit hsqofztbz, qfr ziqz ofygkdqzogf eqf gyztf wt xltr zg wktqa zit eohitk. qyztk zit rolegctkn gy yktjxtfen qfqsnlol, ftqksn qss lxei eohitkl egxsr wt wkgatf wn qf ofygkdtr qzzqeatk. lxei esqlloeqs eohitkl lzoss tfpgn hghxsqkozn zgrqn, zigxui dglzsn ql hxmmstl (ltt eknhzgukqd). zit qkqw dqzitdqzoeoqf qfr hgsndqzi qs-aofro vkgzt q wgga gf eknhzgukqhin tfzozstr kolqsqi yo olzoaikqp qs-dx'qddq (dqfxlekohz ygk zit rteohitkofu eknhzgukqhioe dtllqutl), vioei rtlekowtr zit yoklz afgvf xlt gy yktjxtfen qfqsnlol eknhzqfqsnlol zteifojxtl.

Ciphertexts produced by a classical cipher (and some modern ciphers) will reveal statistical information about the plaintext, and that information can often be used to break the cipher. After the discovery of frequency analysis, nearly all such ciphers could be broken by an informed attacker. Such classical ciphers still enjoy popularity today, though mostly as puzzles (see cryptogram). The Arab mathematician and polymath Al-Kindi wrote a book on cryptography entitled Risalah fi Istikhraj al-Mu'amma (Manuscript for the Deciphering Cryptographic Messages), which described the first known use of frequency analysis cryptanalysis techniques.

Simple Substitution Doesn’t Work

- A large space of keys is not enough
- Mono-alphabetic
  - The same plaintext letters are always replaced by the same ciphertext letters
- Doesn’t hide statistical properties of plaintext.
- Doesn’t hide relationships in plaintext
- Natural languages are very redundant

Make it Harder?

- Hide statistical properties
  - Encrypt “e” with 12 different symbols, “t” with 9 different symbols, etc.
- Poly-alphbetic cipher
  - Use different substitutions
- **Transposition** (permutation)
  - Scramble order of units; reorder units of plaintext

Transposition Cipher

- Scrambling the character order by row-column transposition

1. Tile the plaintext *"MY+COOL+CIPHER+IS+SIMPLE"* in row direction.

2. Read ciphertext in column direction. The columns are ordered based on the secret key.

Key:

3	1	4	2	5
M	*Y*	+	*C*	*O*
O	*L*	+	*C*	*I*
P	*H*	*E*	*R*	+
I	*S*	+	*S*	*I*
M	*P*	*L*	*E*	

Ciphertext: *YLHSPCCRSEMOPIM++E+LOI+I*

From Classical to Modern Cipher

- Many modern (symm.) ciphers are essentially combination of substitution (a.k.a. *"S-Box"*) and transposition (permutation, a.k.a. *"P-Box"*)
  - Combining multiple different "transformations" is more secure
    - *A Mathematical Theory of Cryptography,* Claude Shannon, 1945
  - [Shannon'45] two fundamental principles for statistical security
    - Confusion: produced by substitution
    - Diffusion: produced by transposition

Recap on Discrete Probability

Probability distribution

- Universe U: finite set (e.g. $U = \{0,1\}^n$, the set of all n-bit strings)
- Probability distribution P (over U): A function $P : U \mapsto [0,1]$ such that $\sum_{x \in U} P(x) = 1$
- Examples:
  - Uniform distribution: for all $x \in U$, $P(x) = 1/|U|$
  - Point distribution at x_0: $P(x_0) = 1$ and $P(x) = 0$ for any $x \neq x_0$

Probability distribution

Distribution on 2-bit strings

- Universe $U = \{00,01,10,11\}$
- An example distr $P(00) = 1/2$, $P(01) = 1/8$, $P(10) = 1/4$, $P(11) = 1/8$
- Uniform distribution: for all $x \in U$, $P(x) = 1/4$
- Point distribution at $x_0 = 01$: $P(01) = 1$ and $P(x) = 0$ for any $x \neq 01$

Events

- Events: Any subset A of U
  - Probability of an event: $\Pr(A) = \sum_{x \in A} P(x)$
- Example: $U = \{0,1\}^8$
  - $A = \{$all $x \in U$ such that the lowest two bits $= 11\} \subset U$
    - E.g. $(10111001)_2 \notin A, (00011011)_2 \in A$
  - For the uniform distr on U, $\Pr[A] = 1/4$

Random Variable

- Random Variable X is a function $X : U \mapsto V$ for some range V
- Example: $U = \{0,1\}^8$
  - $X : U \mapsto \{0,1,2,\ldots,8\}$
  - $X(y) =$ number of bit 1 in string y
  - For the uniform distr on U, $\Pr[X = 0] = 1/2^8$, $\Pr[X = 1] = 8/2^8 = 1/2^5$
- X also induces a distr on V:
  $P(v) := \Pr[X = v] = \Pr[X^{-1}(v)] = \sum_{y \in U} P(X(y) = v)$

Random Variable

- **Uniformly random variable**
  - Universe U, uniform prob distr $P(x) = 1/|U| \quad \forall x \in U$
  - $X : U \mapsto U$ is a uniform random var if $\Pr[X = x] = 1/|U|$
    - So essentially X is the identity function

Independence

- Events A and B are independent if $\Pr[A$ and $B] = \Pr[A] \cdot \Pr[B]$
- Random variables X, Y taking values in V are independent if $\forall a, b \in V : \Pr[X = a$ and $Y = b] = \Pr[X = a] \cdot \Pr[Y = b]$
- Example: $U = \{0,1\}^2 = \{00,01,10,11\}$ with uniform distribution
  - Define R.V.: $X(r) = r_0$, $Y(r) = r_1$, for every $r \in U$
  - $\Pr[X = 0$ and $Y = 0] = \Pr[r = 00] = 1/4 = \Pr[X = 0] \cdot \Pr[Y = 0]$

Example: XOR

- XOR ($\oplus$) of two strings from $\{0,1\}^n$ is there bitwise addition mod 2
- Example:
  - $(0110111) \oplus (1011010) = 1101101$

Symbol	A	B	Q
	0	0	0
	1	0	1
	0	1	1
	1	1	0
$Q = A \oplus B$			

Example: XOR

Hiding information

- **Theorem**: Let Y by an arbitrary random variable over $\{0,1\}^n$, and X be an independent uniform random variable on $\{0,1\}^n$, then $Z = X \oplus Y$ is an uniform random variable on $\{0,1\}^n$
- I.e., by XOR with an uniform random string X, we destroys all distribution information of Y
  - Y - plaintext, Z - ciphertext
  - Initial motivation for rigorously defining secrecy [C. Shannon, 1964]

The Birthday Paradox

- U: any finite universe
- Let $r_1, \ldots, r_n \in U$ be *independent identically* distributed random vars (but the distribution can be arbitrary)
- **Theorem**: When $n > 1.2 \times |U|^{1/2}$, then $\Pr[\exists i \neq j : r_i = r_j] \geq 1/2$
- Example: $U = \{0,1\}^{128}$
  - After sample 2^{64} random strings from U, two sampled strings will likely be the same.

The Birthday Paradox

- **Derivation**: suppose $|U| = N$, the probability of the n i.i.d. rv *all getting different* value (i.e., the prob. of *non*-collision) is:
  - $\dfrac{\binom{N}{n} \cdot n!}{N^n} \approx e^{-n(n-1)/2N} \approx e^{-n^2/2N}$ (by Stirling's approx)
  - When $n \approx 1.177\sqrt{N}$ we can reduce the above prob to 0.5

The Birthday Paradox

- **Application**:
  - The Birthday attack. Suppose we want to fabricate two different files that have the same checksum (hash) of 128 bit, then only $\sim 2^{64}$ trials are needed.
    - Informally speaking, the "effective strength" of some random-looking sequence (e.g. hash, key) is only half of its length.

Recap on Algorithm Analysis

Algorithm analysis 101

- **Computational problem**: any task that involves an *input*, some kind of *computing device*, and an *output*
  - **Computing device**: Random Access Machine (just like your laptop)
  - **Input**: *Reasonable* binary encodings
    - Input size: the length (# of bits) of the encoded input
  - **Output**: *Reasonable* binary encodings
    - Output size: the length (# of bits) of the encoded output
    - Usually not important in the analysis of alg.

Algorithm analysis 101

- **Example**: The Factoring Problem.
  - INPUT: An integer $N = p \cdot q$. (p and q are not given).
    - Input size = $\log_2 N$ (Reasonable?)
  - OUTPUT: p and q
- A concret instance:
  - INPUT: $N = 187 = (10111011)_2$
  - OUTPUT: $(p, q) = (17,11) = (10001,1011)_2$

Algorithm analysis 101

- **Example**: Ciphertext-only Attack.
  - INPUT: A n-bit ciphertext from some known encryption algorithm.
    - Input size = n
  - OUTPUT: The plaintext / The secret key

Algorithm analysis 101

- **Algorithm**: Procedures that *executed by* <u>*the computing device*</u> to solve a given computational problem.
  - Usually denoted as a mapping $A : \Sigma \mapsto \Pi$
    - $\Sigma \subset \{0,1\}*$: the set of input encodings
    - $\Pi \subset \{0,1\}*$: the set of output encodings

Algorithm analysis 101

- **Randomized Algorithm** $A(s, r)$: Aside from the input string s, A also accepts a (uniform) random string r
    - The length of r can be dependent on s (but usually smaller or on the same scale)
    - As a result, A's output is random
- Example:
    - Most ML algorithms: e.g., Stochastic Gradient Descent

Algorithm analysis 101

- **Why Randomized?**
  - Often makes algorithm simpler (and faster)
- Example:
  - Suppose an alg A for the ciphertext-only attack only succ with probability $1/1000$
  - Run A *independently* for 2000 times we can boost the succ prob to $1 - (1 - 1/1000)^{1000} \approx 86\,\%$
  - Good as long as each run is fast
- To learn more: CSE 632 (Sp 25)

Algorithm analysis 101

- **Running Time**
  - Given input s, the *number of steps* taken *for the underlying computing device* to execute A (in order to find an output)
    - E.g., The number of CPU instructions for your laptop to run a certain program.
    - Expressed as a function $T(n)$ of the input length $n = |s|$

Algorithm analysis 101

- **Asymptotic Running Time: Big-O**
  - Measures how fast $T(n)$ grows with the input length n
  - $T(n) = O(f(n)) \iff T(n) \leq C \cdot f(n)$ for some constant C
  - Helps simplify analysis and focus on the "essence" of alg. Performance.
    - Hardware independent.
    - Ignore the small edge cases.
    - Not the full picture, though.

Algorithm analysis 101

- **Example**: The Factoring Problem
  - INPUT: integer $N = p \cdot q$; length $n = \log_2 N$
  - Current state of the art:
    - GNFS (classical): $O\left(\exp\left(2(n)^{1/3}\right)\right)$
      - This is why RSA is safe
      - But not for small n: $n < 512$ can be factored on your laptop, while $n > 4096$ is currently out-of-reach for human (?)
    - Shor (quantum): $O(n^3)$
      - This is why RSA is unsafe (but not for now)

Summary

- Crypto basics
- Core application: Secure communication
    - Establish shared key: PKC
    - Transmitting msg with shared sec key: symm encryption
- Classical symm ciphers
    - Caesar; Substitution; Transposition; How they fail.
    - Modern ciphers: Combinations of the two. [C. Shannon]
- Recap on Probability
    - Uniform random var; Birthday Paradox.
- Recap on Algorithm
    - Big-O notation; Randomized Alg.

Questions?