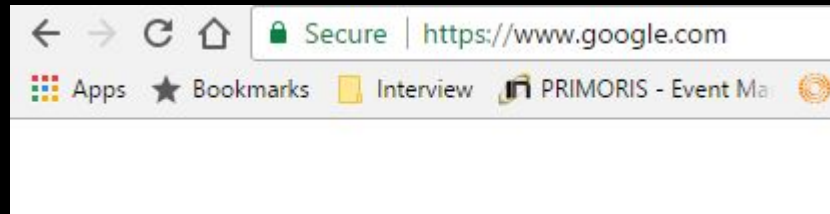


Introduction to Cryptography



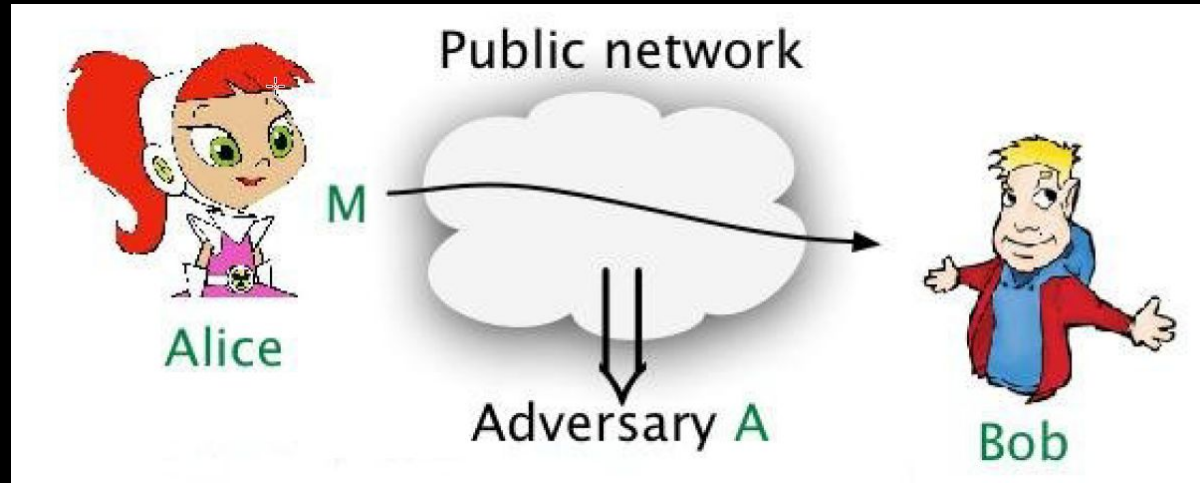
This research is supported by Award #1542465:
RET Site: Cyber Security Initiative for Nevada Teachers (CSINT)

Did you use any cryptography today?



- *https* invokes the *TLS* protocol
- *TLS* uses *cryptography*
- *TLS* is in ubiquitous use for secure communication: shopping, banking, Netflix, Gmail, Facebook, ...

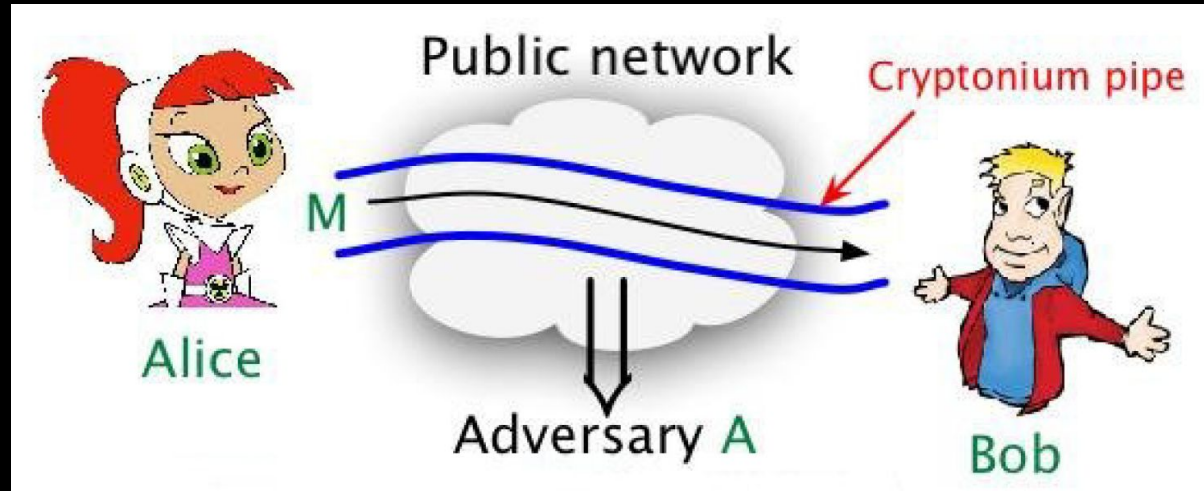
What is cryptography about?



Adversary: clever person with powerful computer

- Security goals:
 - **Data privacy**: Ensure adversary does not see or obtain the data (message) M .
 - **Data integrity and authenticity**: Ensure M really originates with Alice and has not been modified in transit.

Idea World



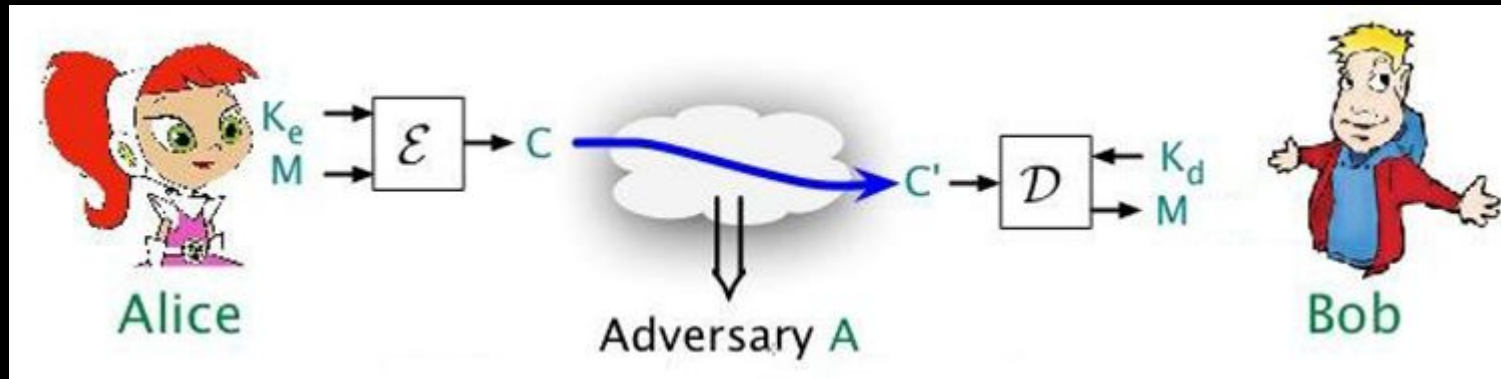
Adversary: clever person with powerful computer

Cryptonium pipe: Cannot see inside or alter content.

All our goals would be achieved!

But **cryptonium** is only available on **planet Crypton** and is in short supply.

Cryptographic schemes



\mathcal{E} : encryption algorithm

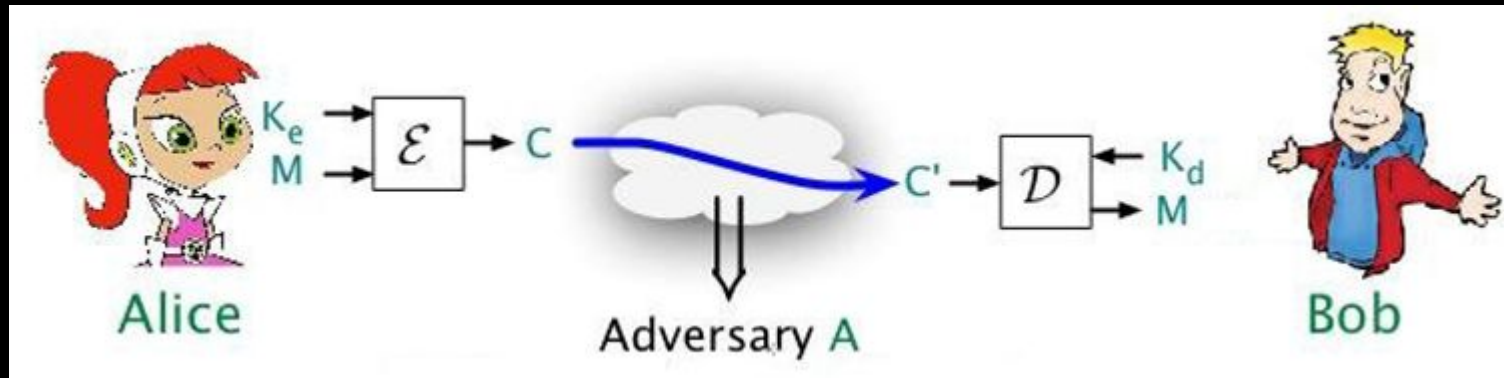
K_e : encryption key

\mathcal{D} : decryption algorithm

K_d : decryption key

Algorithms: standardized, implemented, public!

Cryptographic schemes



\mathcal{E} : encryption algorithm

K_e : encryption key

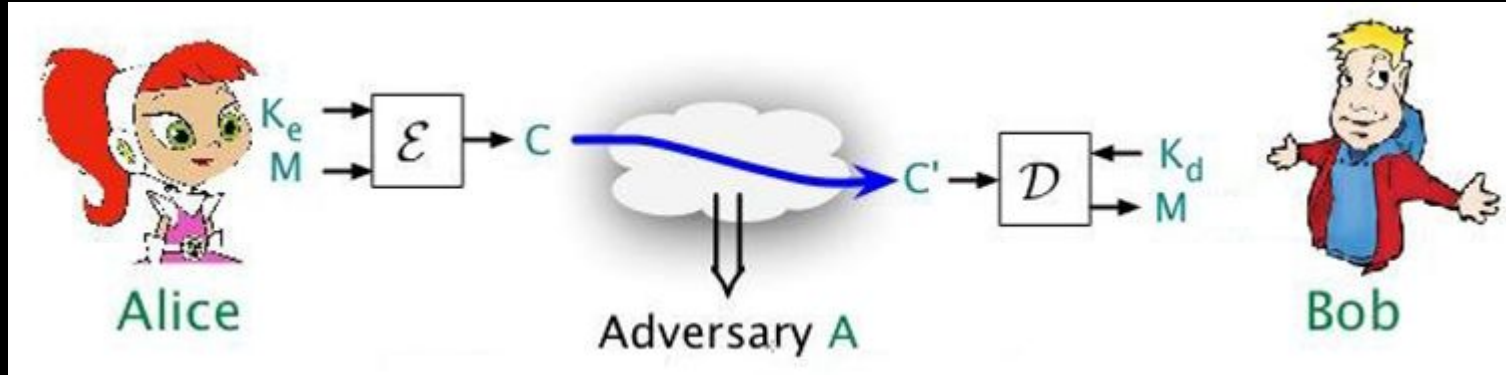
\mathcal{D} : decryption algorithm

K_d : decryption key

Settings:

- public-key (**asymmetric**): K_e public, K_d secret
- private-key (**symmetric**): $K_e = K_d$ secret

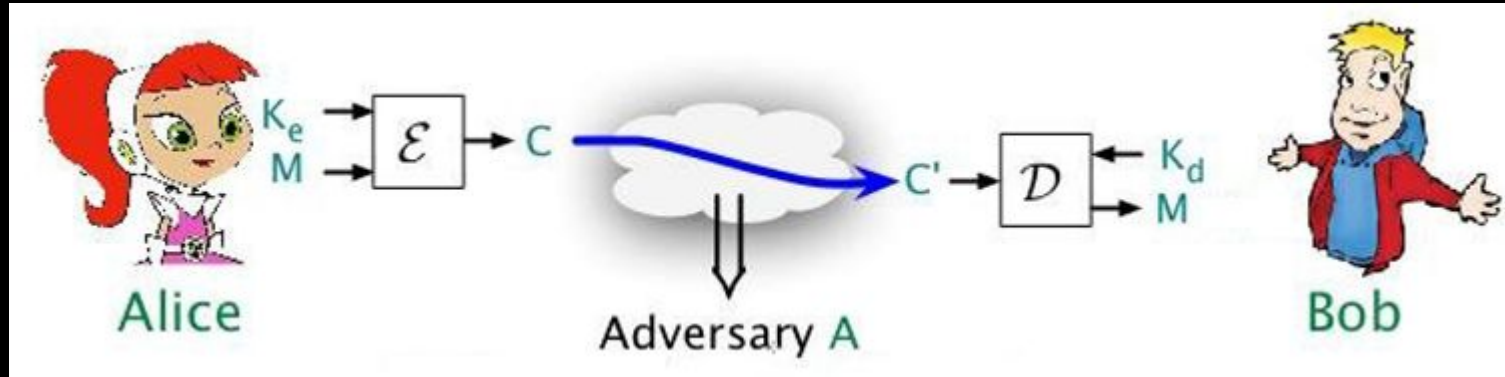
Cryptographic schemes



Our concerns:

- How to define security goals?
- How to design \mathcal{E}, \mathcal{D} ?
- How to gain confidence that \mathcal{E}, \mathcal{D} achieve our goals?

Cryptographic schemes



Computer Security: How does the computer/system protect K_e/K_d from break-in (viruses, worms, OS holes, . . .)?

Cryptography: How do we use K_e, K_d to ensure security of communication over an insecure network?

Cryptographic algorithms and protocols in 4 main areas:

Symmetric encryption

- Used to conceal the contents of blocks or streams of data of any size, including messages, files, encryption keys, and passwords

Asymmetric encryption

- Used to conceal small blocks of data, such as encryption keys and hash function values, which are used in digital signatures

Data integrity algorithms

- Used to protect blocks of data, such as messages, from alteration

Authentication protocols

- Schemes based on the use of cryptographic algorithms designed to authenticate the identity of entities

Definitions

Plaintext

- An original message

Ciphertext

- The coded message

Enciphering/encryption

- The process of converting from plaintext to ciphertext

Deciphering/decryption

- Restoring the plaintext from the ciphertext

Cryptography

- The area of study of the many schemes used for encryption

Cryptanalysis

- Techniques used for deciphering a message without any knowledge of the enciphering details

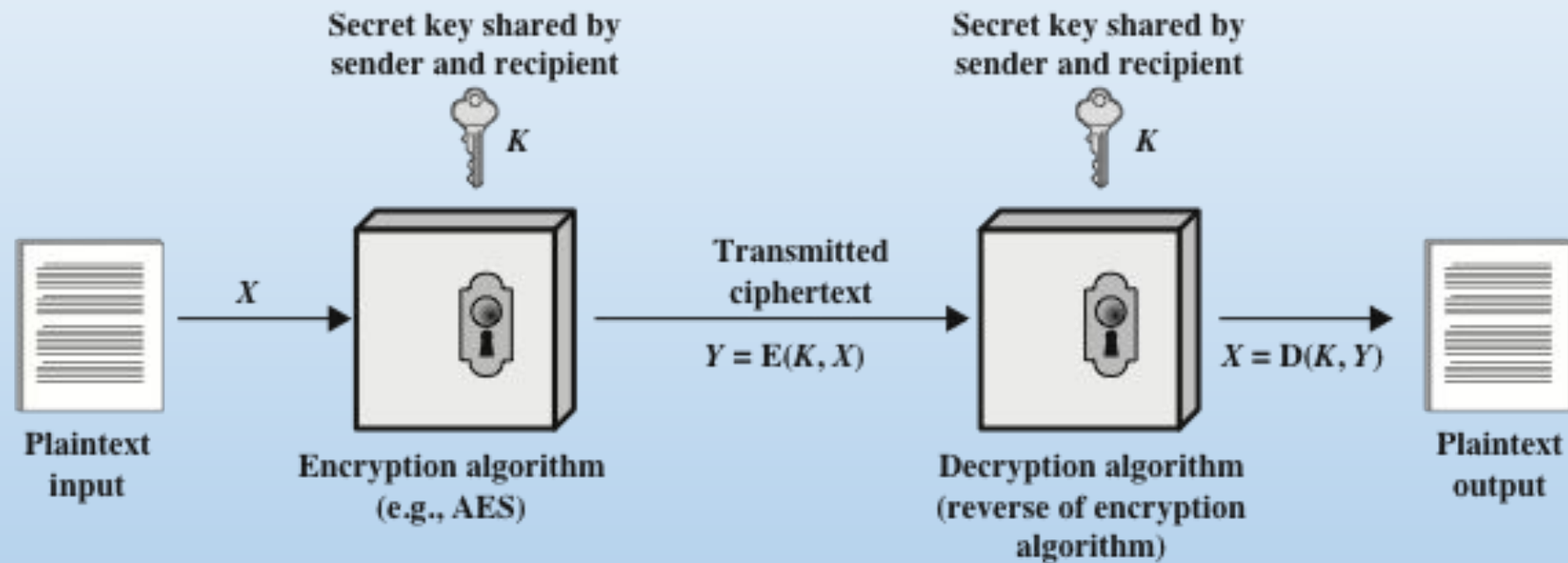
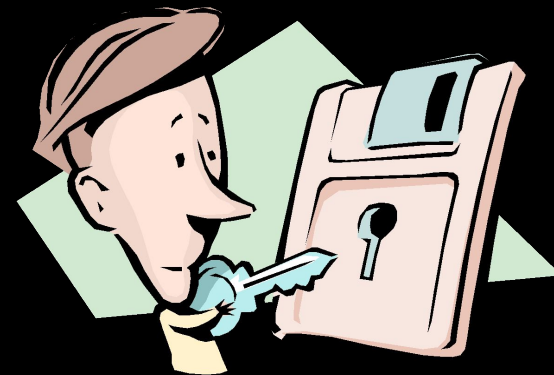


Figure 3.1 Simplified Model of Symmetric Encryption

Symmetric Cipher Model

- There are two requirements for secure use of conventional encryption:
 - A strong encryption algorithm
 - Sender and receiver must have obtained copies of the secret key in a secure fashion and must keep the key secure



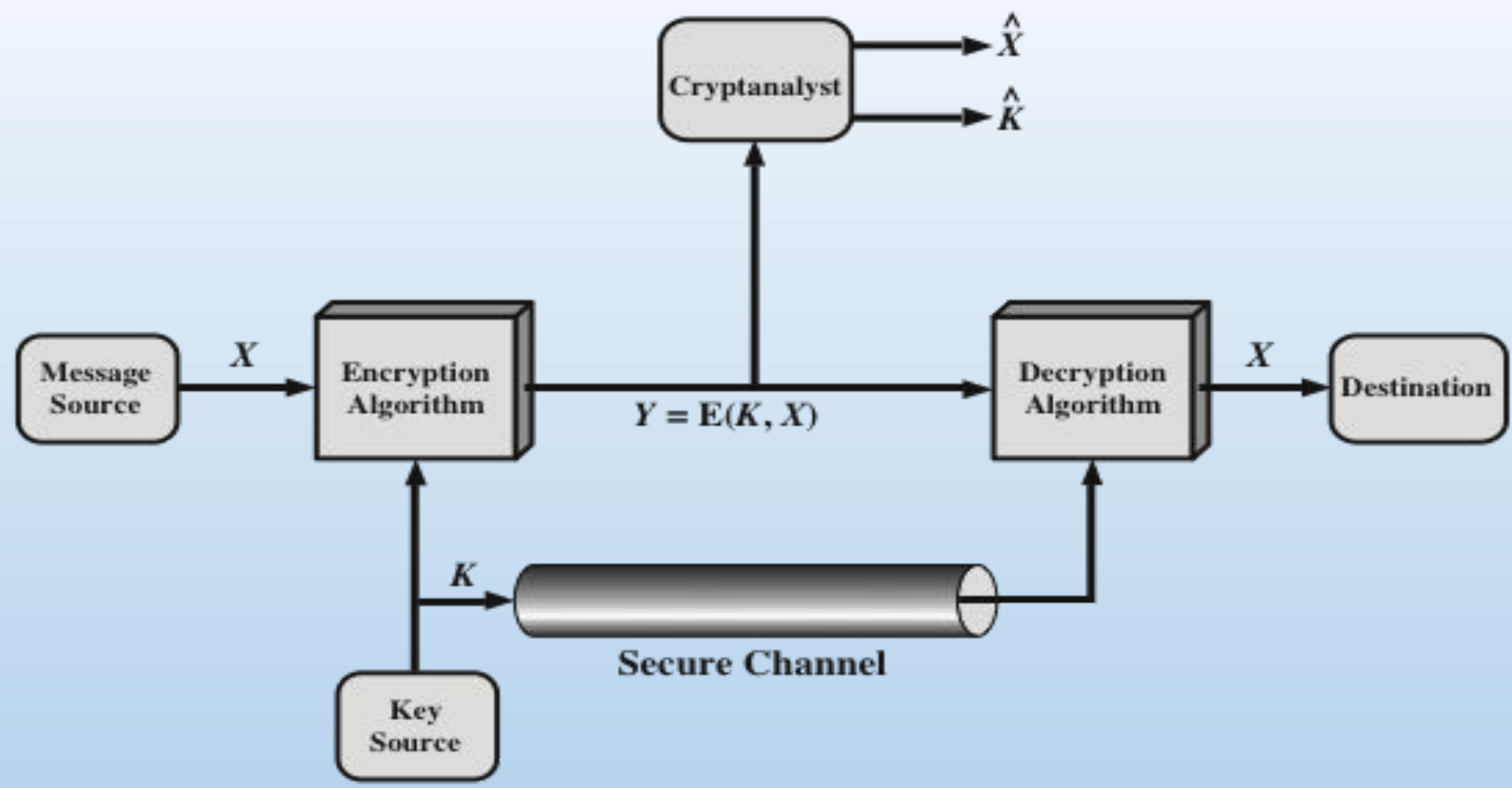
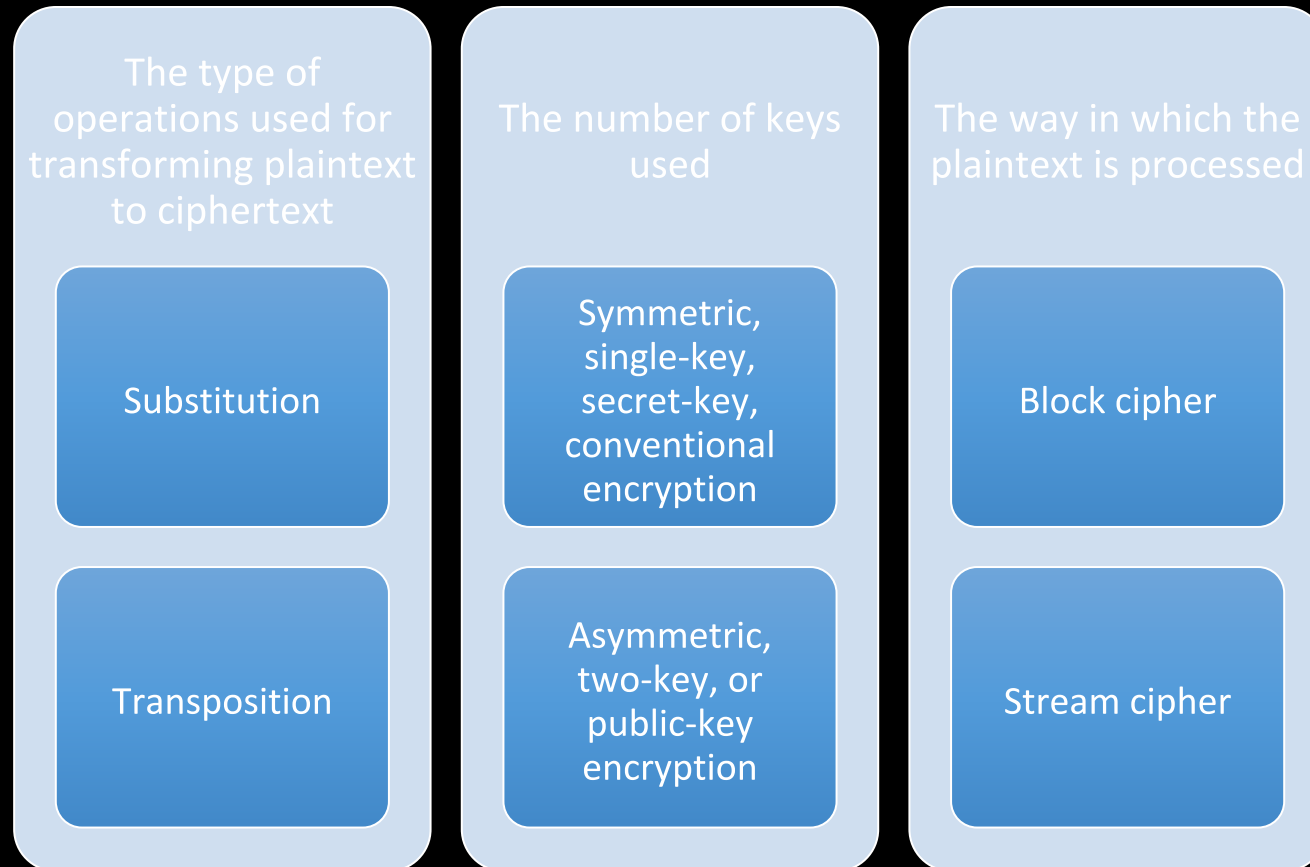


Figure 3.2 Model of Symmetric Cryptosystem

Cryptographic Systems

- Characterized along three independent dimensions:



Cryptanalysis and Brute-Force Attack

Cryptanalysis

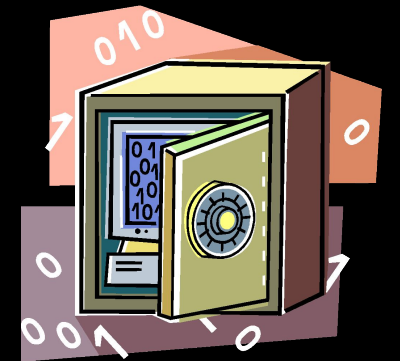
- Attack relies on the nature of the algorithm plus some knowledge of the general characteristics of the plaintext
- Attack exploits the characteristics of the algorithm to attempt to deduce a specific plaintext or to deduce the key being used

Brute-force attack

- Attacker tries every possible key on a piece of ciphertext until an intelligible translation into plaintext is obtained
- On average, half of all possible keys must be tried to achieve success


Encryption Scheme Security

- Unconditionally secure
 - No matter how much time an opponent has, it is impossible for him or her to decrypt the ciphertext simply because the required information is not there
- Computationally secure
 - The cost of breaking the cipher exceeds the value of the encrypted information
 - The time required to break the cipher exceeds the useful lifetime of the information



Brute-Force Attack

Involves trying every possible key until an intelligible translation of the ciphertext into plaintext is obtained



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



To supplement the brute-force approach, some degree of knowledge about the expected plaintext is needed, and some means of automatically distinguishing plaintext from garble is also needed



Caesar Cipher



- Simplest and earliest known use of a substitution cipher
- Used by Julius Caesar
- Involves replacing each letter of the alphabet with the letter standing 3 places further down the alphabet
- Alphabet is wrapped around so that the letter following Z is A

plain: meet me after the toga party

cipher: PHHW PH DIWHU WKH WRJD SDUWB

Caesar Cipher Algorithm

- Can define transformation as:

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

- Mathematically give each letter a number

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

- Algorithm can be expressed as:

$$c = E(3, p) = (p + 3) \bmod 26$$

- A shift may be of any amount, so that the general Caesar algorithm is:

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

- Where k takes on a value in the range 1 to 25; the decryption algorithm is simply:

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

Brute-Force Cryptanalysis of Caesar Cipher

KEY	PHHW	PH	DIWHU	WKH	WRJD	SDUWB
1	oggv	og	chvgt	vjg	vqic	rctva
2	nffu	nf	bgufs	uif	uphb	qbsuz
3	meet	me	after	the	toga	party
4	ldds	ld	zesdq	sgd	snfz	ozqsx
5	kccr	kc	ydrpc	rfc	rmey	nyprw
6	jbbq	jb	xcqbo	qeb	qldx	mxoqv
7	iaap	ia	wbpan	pda	pkcw	lwnpu
8	hzzo	hz	vaozm	ocz	ojbv	kvmot
9	gyyn	gy	uznyl	nby	niau	julns
10	fxxm	fx	tymxk	max	mhzt	itkmr
11	ewwl	ew	sxlwj	lzw	lgys	hsjlg
12	dvvk	dv	rwkvi	kyv	kfxr	grikp
13	cuuj	cu	qvjuh	jxu	jewq	fghjo
14	btti	bt	puitg	iwt	idvp	epgin
15	assh	as	othsf	hvs	hcuo	dofhm
16	zrrg	zr	nsgre	gur	gbtn	cnegl
17	yqqf	yq	mrfqd	ftq	fasm	bmdfk
18	xppe	xp	lqepc	esp	ezrl	alcej
19	wood	wo	kpdob	dro	dyqk	zkbdi
20	vnnv	vn	jocna	cqn	cxpj	yjach
21	ummb	um	inbmz	bpm	bwoi	xizbg
22	tlla	tl	hmaly	aol	avnh	whyaf
23	skkz	sk	glzkx	znk	zumg	vgxze
24	rjyy	rj	fkyjw	ymj	ytlf	ufwyd
25	qiix	qi	ejxiv	xli	xske	tevxc

Substitution cipher

-

$K_e = K_d = \pi: \Sigma \rightarrow \Sigma$, a secret permutation

e.g., $\Sigma \{ A, B, C, \dots \}$ and π is as follows:

	A	B	C	D	...
	E	A	Z	U	...

$$\mathcal{E}_\pi(\text{CAB}) = \pi(\text{C}) \pi(\text{A}) \pi(\text{B}) = \text{Z E A}$$

$$\mathcal{D}_\pi(\text{ZEA}) = \pi^{-1}(\text{Z}) \pi^{-1}(\text{E}) \pi^{-1}(\text{A}) = \text{C A B}$$

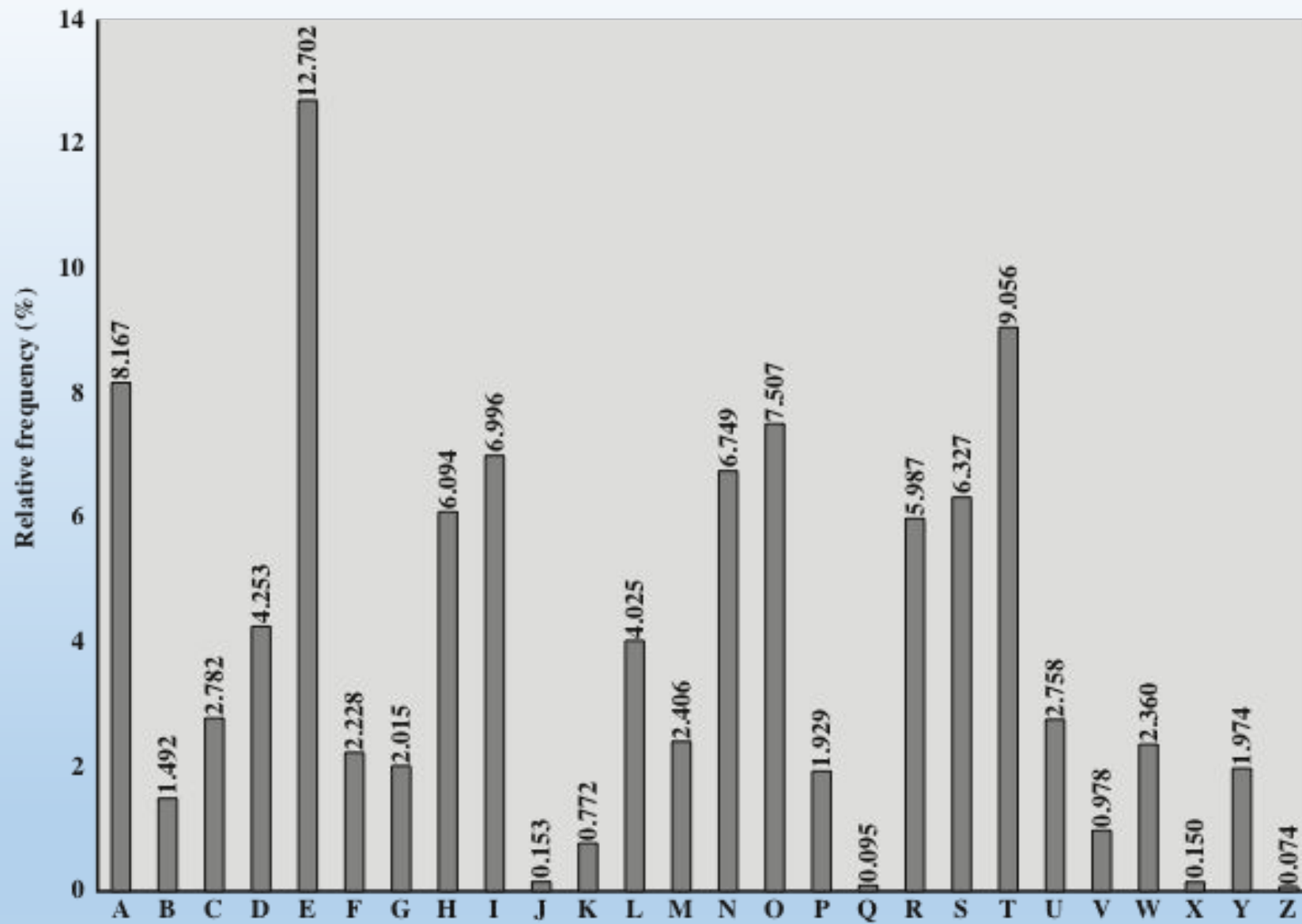


Figure 3.5 Relative Frequency of Letters in English Text

Vigenère Cipher

- To encrypt a message, a key is needed that is as long as the message
- Usually, the key is a repeating keyword
- For example, if the keyword is *deceptive*, the message “we are discovered save yourself” is encrypted as:

```
key:      deceptive  
plaintext: wearediscoveredsaveyourself  
ciphertext: ZICVTWQNGRZGVTWAVZHCQYGLMGJ
```

One-Time Pad

- Use a random key that is as long as the message so that the key need not be repeated
- Key is used to encrypt and decrypt a single message and then is discarded
- Each new message requires a new key of the same length as the new message
- Scheme is unbreakable
 - Produces random output that bears no statistical relationship to the plaintext
 - Because the ciphertext contains no information whatsoever about the plaintext, there is simply no way to break the code



Shannon and One-Time-Pad (OTP) Encryption

Theorem (Shannon): OTP is **perfectly** secure as long as only one message encrypted.

“**Perfect**” secrecy, a notion Shannon defines, captures **mathematical impossibility** of breaking an encryption scheme.

Fact: if $|M| > |K|$, then no scheme is perfectly secure.



Modern Cryptography: A Computational Science

Security of a “practical” system must rely not on the impossibility but on the computational difficulty of breaking the system.

“Practical” = more message bits than key bits

Modern Cryptography: A Computational Science

Rather than:

“It is impossible to break the scheme”

We might be able to say:

“No attack using $\leq 2^{160}$ time succeeds with probability $\geq 2^{-20}$ ”

i.e., Attacks can exist as long as **cost to mount them** is **prohibitive**, where **cost** = computing time/memory, \$\$\$

Modern Cryptography: A Computational Science

Security of a “practical” system must rely not on the impossibility but on the computational difficulty of breaking the system.

Cryptography is now not just mathematics; it needs to **draw on computer science**

- Computational complexity theory
- Algorithm design

The factoring problem

Input: Composite integer N

Desired output: prime factors of N

Example:

Input: 85

Output: 17, 5

$2,173 = 41 * 53?$

Can we write a factoring program? Easy!

Alg Factor(N)

For $i = 2, 3, \dots, \sqrt{N}$; do

If $N \bmod i = 0$ then return i

But this is very slow ...



Prohibitive if N is large (e.g., 400 digits)

44 digits prime: 20,988,936,657,440,586,486,151,264,256,610,222,593,863,921

Largest prime number $2^{77,232,917}-1$ (23,249,425 digits)

Can we factor fast?

- Gauss couldn't figure out how
- Nor does anyone know now



Nobody today knows how to factor a 400 digit number in a practical amount of time.

Factoring is an example of a **computationally hard**

DES - Symmetric Key

DES – Data Encryption Standard

- 1972 - NBS (now NIST) asked for a block cipher for standardization
- 1974 - IBM designs Lucifer
- Lucifer eventually evolved into [DES](#).

Widely adopted as a standard including by ANSI and American Bankers association

Used in ATM machines

Replaced (by AES) in 2001.

Problem with Symmetric Key Encryption

- Before Alice and Bob can communicate securely, they need to have a common secret key K_{AB} .
- If Alice wishes to also communicate with Charlie then she and Charlie must also have another common secret key K_{AC} .
- If Alice generates $K_{AB}; K_{AC}$, they must be communicated to her partners over private and authenticated channels.

RSA – Asymmetric Key

R - Rivest

S - Shamir

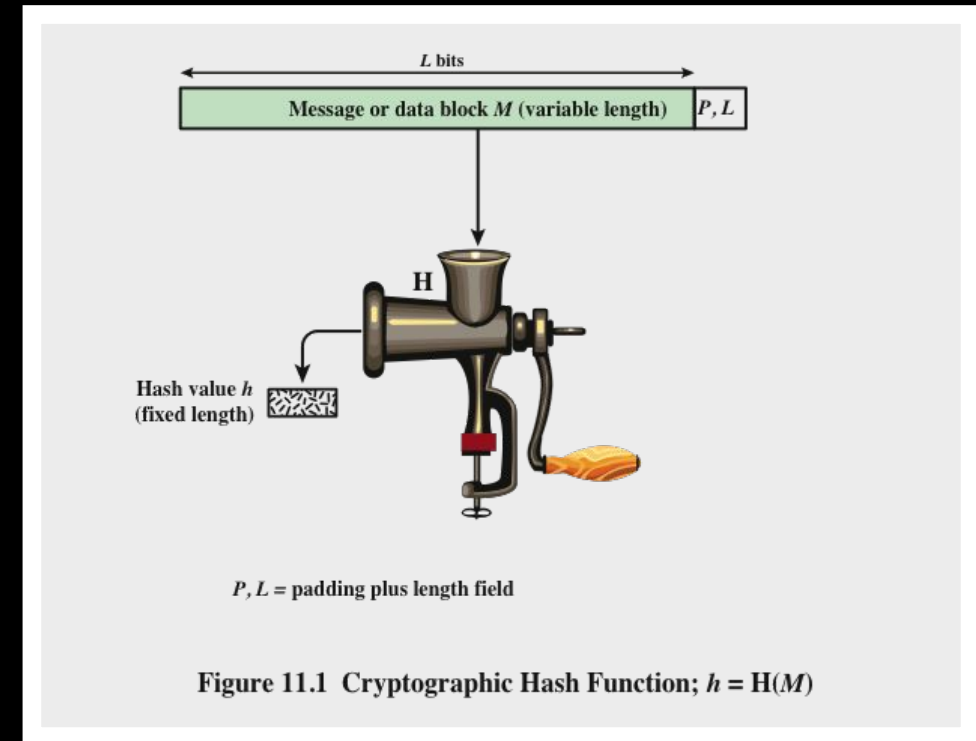
A - Adleman



- Alice has a **secret key** that is shared with nobody, and an associated **public key** that is known to everybody.
- Anyone (Bob, Charlie, ...) can use Alice's **public key** to send her an encrypted message which only she can decrypt.

Hashing

- Hash functions like **MD5, SHA1, SHA256**, ... are amongst the most widely-used cryptographic primitives.
- Their primary purpose is **collision-resistant** data compression, but they have many other purposes and properties as well.
- A good hash function is often treated like a magic wand ...



Password Verification

- Client A has a password PW that is also held by server B
- A authenticates itself by sending PW to B over a secure channel (SSL)



- **Problem:** The password will be found by an attacker who compromises the server.

Password Verification

- Client A has a password PW and server B stores $\overline{PW} = H(PW)$
- A send PW to B and B checks that $H(PW) = \overline{PW}$



Server compromise results in attacker getting \overline{PW} which should not reveal PW as long as H is one-way.