Cryptography Basics

Network Security

Instructor: Haojin Zhu

Cryptography

- What is cryptography?
- Related fields:
  - Cryptography ("secret writing"): Making secret messages
    - Turning plaintext (an ordinary readable message) into Ciphertext (secret messages that are “hard” to read)
  - Cryptanalysis: Breaking secret messages
    - Recovering the plaintext from the ciphertext
- Cryptology is the science that studies these both
- The point of cryptography is to send secure messages over an insecure medium (like the Internet)

Building blocks

- Cryptography contains three major types of components
  - Confidentiality components
    - Preventing Eve from reading Alice's messages
  - Integrity components
    - Preventing Mallory from modifying Alice's messages without being detected
  - Authenticity components
    - Preventing Mallory from impersonating Alice

Dramatis Personae

- When talking about cryptography, we often use a standard cast of characters
- Alice, Bob, Carol, Dave
  - People (usually honest) who wish to communicate
- Eve
  - A passive eavesdropper, who can listen to any transmitted messages
- Mallory
  - An active Man-In-The-Middle, who can listen to, and modify, insert, or delete, transmitted messages
- Trent
  - A Trusted Third Party

Why use Alice, Bob to represent attacker?

The reader is encouraged to read Diffie and Hellman's excellent article [1] for further background, for elaboration of the concept of a public-key cryptosystem, and for a discussion of other problems in the area of cryptography. The ways in which a public-key cryptosystem can ensure privacy and enable "signatures" (described in Sections III and IV below) are also due to Diffie and Hellman.

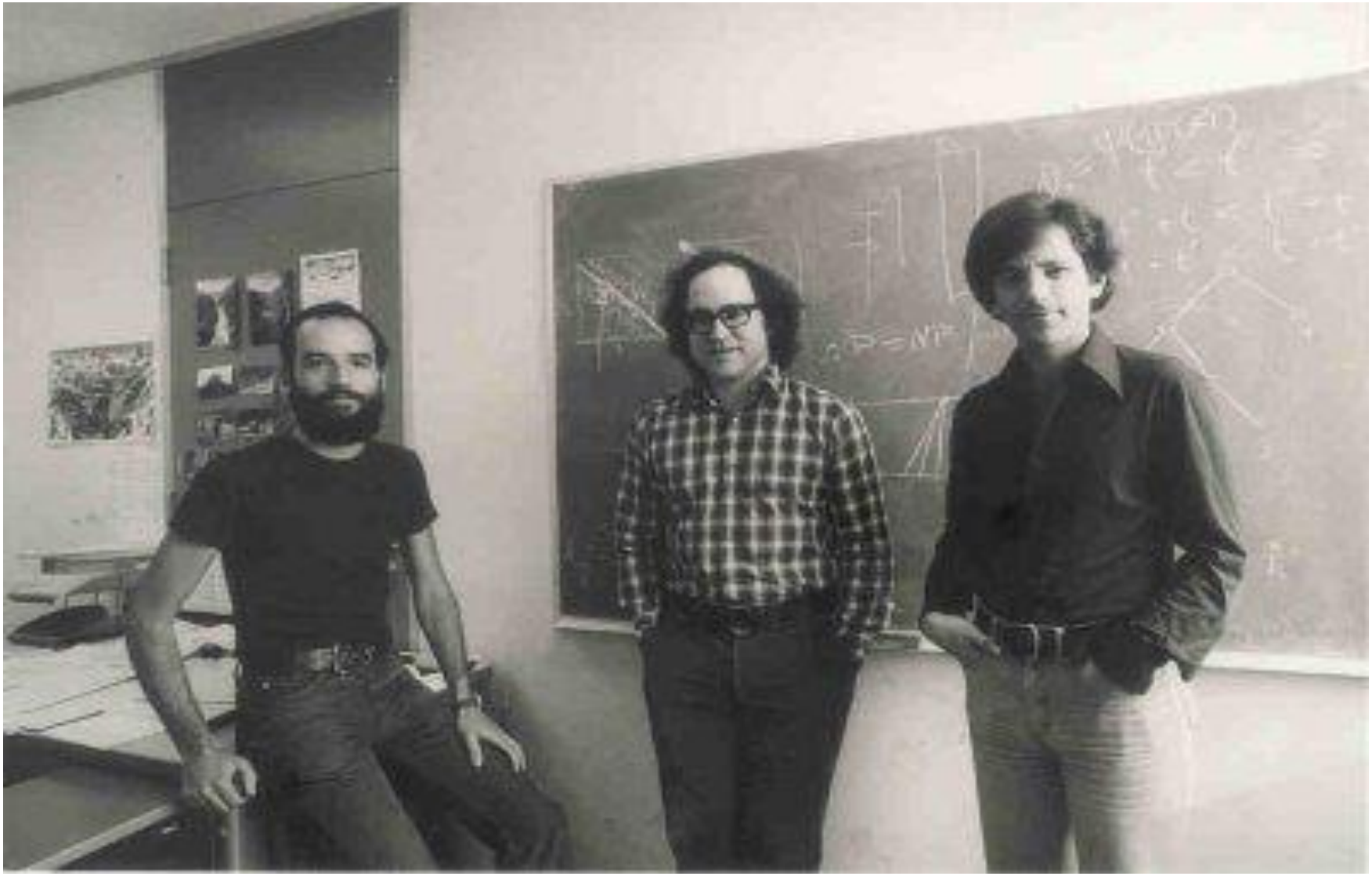
For our scenarios we suppose that A and B (also known as Alice and Bob) are two users of a public-key cryptosystem. We will distinguish their encryption and decryption procedures with subscripts: E_A, D_A, E_B, D_B.

III. Privacy

Two users can also establish private communication over an insecure communications channel without consulting a public file. Each user sends his encryption key to the other. Afterwards all messages are enciphered with the encryption key of the recipient, as in the public-key system. An intruder listening in on the channel cannot decipher any messages, since it is not possible to derive the decryption keys from the encryption keys. (We assume that the intruder cannot modify or insert messages into the channel.) Ralph Merkle has developed another solution [5] to this problem.

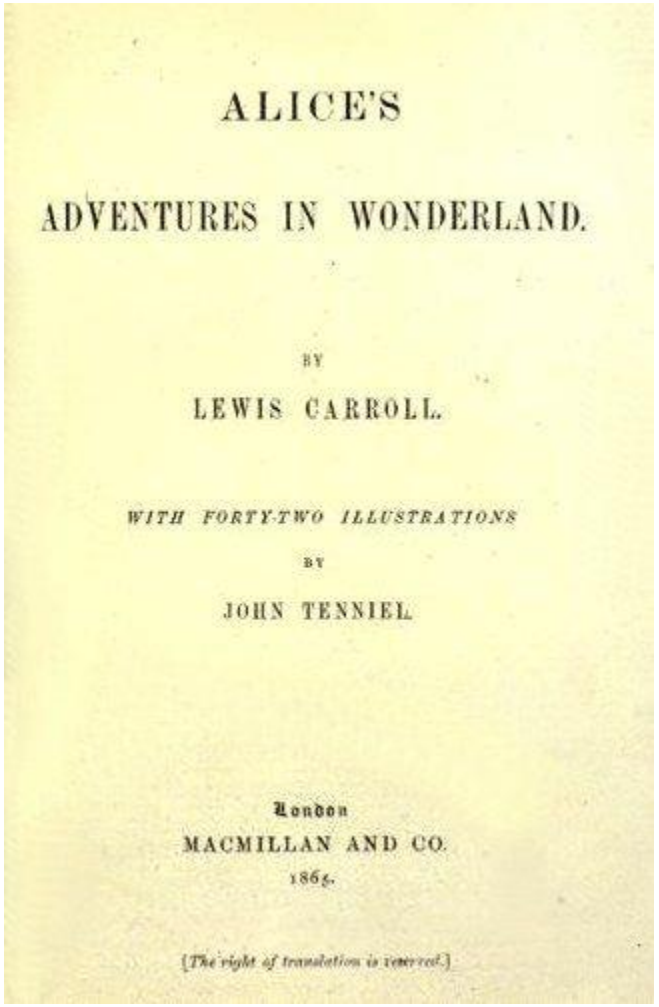
A public-key cryptosystem can be used to "bootstrap" into a standard encryption scheme such as the NBS method. Once secure communications have been established, the first message transmitted can be a key

Rivest、Shamir、Adleman, A Method of Obtaining Digital Signatures and Public-Key Cryptosystems, Communications of the ACM, 1978. (ACM Turing Award in 2002)



Shamir、Rivest、Adleman。https://cryptologicfoundation.org/

Rivest loves the movie "Alices adventures in wonder land"



ALICE'S

ADVENTURES IN WONDERLAND.

BY

LEWIS CARROLL.

WITH FORTY-TWO ILLUSTRATIONS

BY

JOHN TENNIEL.

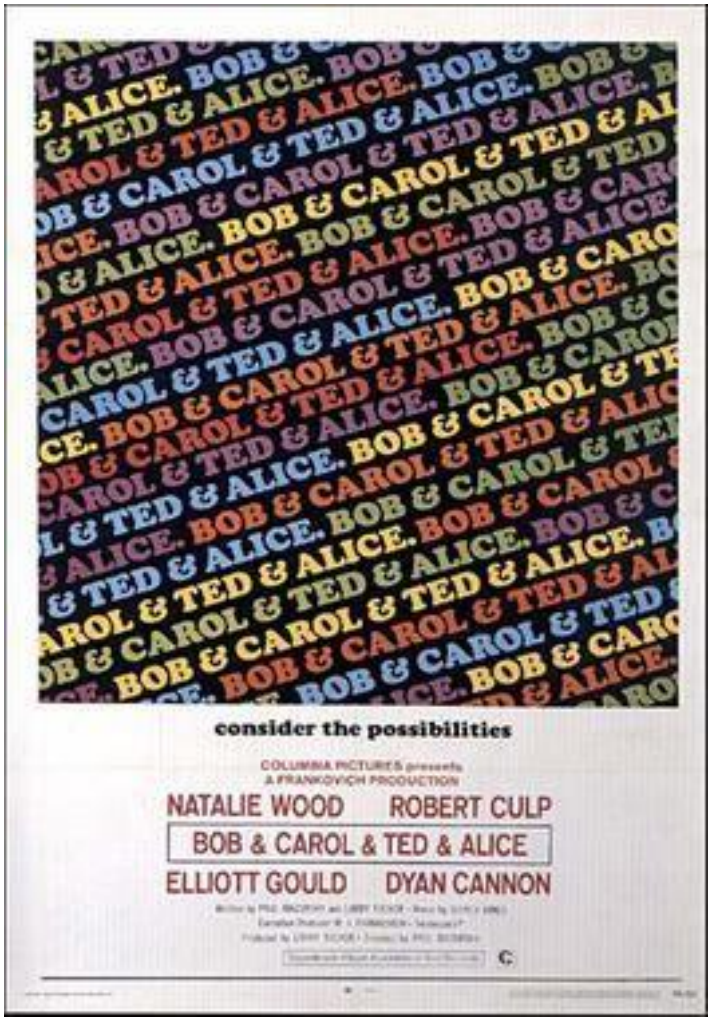
London

MACMILLAN AND CO.

1865.

[*The right of translation is reserved.*]

Another movie “Bob & Carol & Ted & Alice”



Kerckhoffs' Principle (19th c.)

- The security of a cryptosystem should not rely on a secret that's hard (or expensive) to change
- So don't have secret encryption methods
  - Then what do we do?
  - Have a large class of encryption methods, instead
    - Hopefully, they're all equally strong
  - Make the class public information
  - Use a secret key to specify which one you're using
  - It's easy to change the key; it's usually just a smallish number

Kerckhoffs' Principle (19th c.)

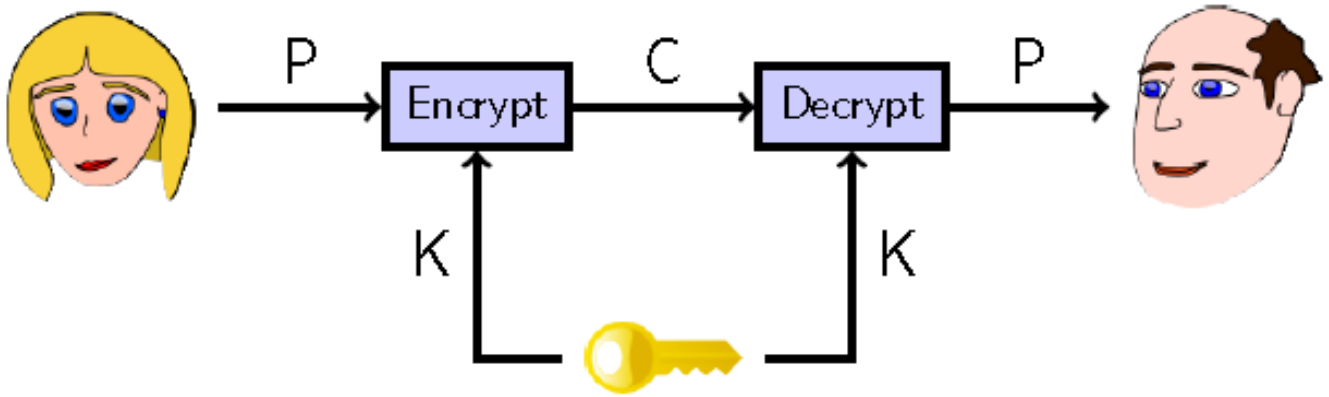
- This has a number of implications:
  - The system is at most as secure as the number of keys
  - Eve can just try them all, until she finds the right one
  - A strong cryptosystem is one where that's the best Eve can do
    - With weaker systems, there are shortcuts to finding the key
  - Example: newspaper cryptogram has 403,291,461,126,605,635,584,000,000 possible keys
  - But you don't try them all; it's way easier than that!

Strong cryptosystems

- What information do we assume the attacker (Eve) has when she's trying to break our system?
- She may:
  - Know the algorithm (the public class of encryption methods)
  - Know some part of the plaintext
  - Know a number (maybe a large number) of corresponding plaintext/ciphertext pairs
  - Have access to an encryption and/or decryption oracle
- And we still want to prevent Eve from learning the key!

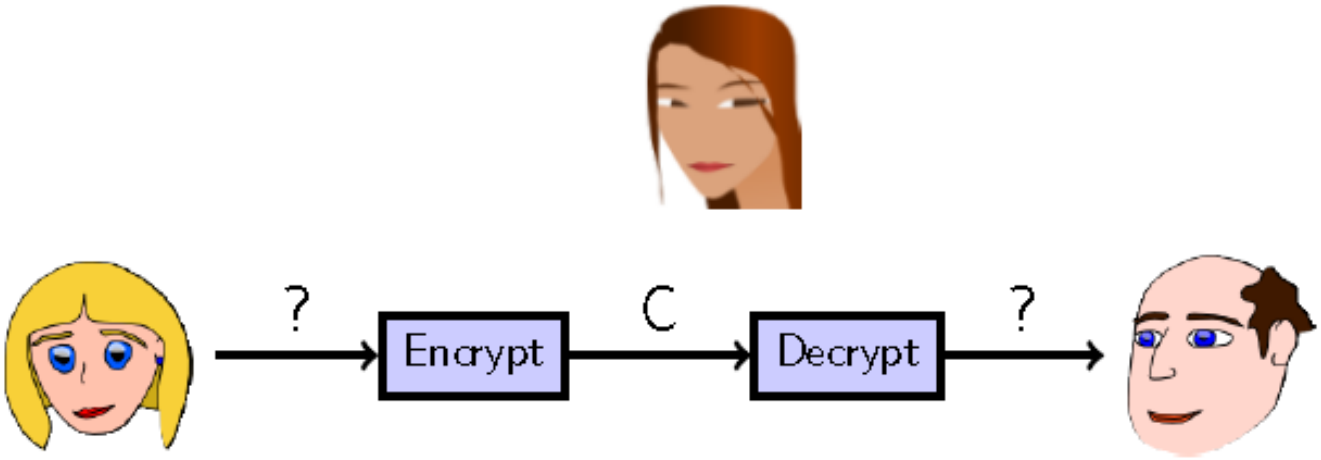
Secret-key encryption

- Secret-key encryption is the simplest form of cryptography
- Also called symmetric encryption
- Used for thousands of years
- The key Alice uses to encrypt the message is the same as the key Bob uses to decrypt it



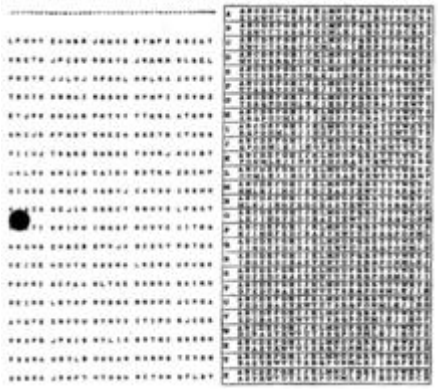
Secret-key encryption

- Eve, not knowing the key, should not be able to recover the plaintext



Perfect secret-key encryption

- Is it possible to make a completely unbreakable cryptosystem?



- Yes: the One-Time Pad

- It's also very simple:
  - The key is a truly random bitstring of the same length as the message
  - The “Encrypt" and “Decrypt" functions are each just XOR

One-time pad

- Q: Why does "try every key" not work here?

- It's very hard to use correctly
  - The key must be truly random, not pseudorandom
  - The key must never be used more than once!
    - A “two-time pad" is insecure!

- Q: How do you share that much secret key?

- Used in the Washington / Moscow hotline for many years

Key Randomness in One-Time Pad

- One-Time Pad uses a very long key, what if the key is not chosen randomly, instead, texts from, e.g., a book are used as keys.
  - this is not One-Time Pad anymore
  - this can be broken
  - How?
- Corrolary: The key in One-Time Pad should never be reused.
  - If it is reused, it is Two-Time Pad, and is insecure!
  - Why?

Usage of One-Time Pad

- To use one-time pad, one must have keys as long as the messages.
- To send messages totaling certain size, sender and receiver must agree on a shared secret key of that size.
  - typically by sending the key over a secure channel
- Key agreement is difficult to do in practice.
- Can't one use the channel for sending the key to send the messages instead?
- Why is OTP still useful, even though difficult to use?

Usage of One-Time Pad

- The channel for distributing keys may exist at a different time from when one has messages to send.

- The channel for distributing keys may have the property that keys can be leaked, but such leakage will be detected
  - Such as in Quantum cryptography

“墨子号”首次量子保密洲际通信细节公布

新华社华盛顿
１９日以封面论文

按照两国科学院２０１１年底在北京签署的洲际量子通信合作协议，中奥联合团队利用“墨子号”卫星于去年年中开展了距离达７６００公里的洲际量子密钥分发实验。在实验中，“墨子号”分别与河北兴隆、奥地利格拉茨地面站进行了星地量子密钥分发，然后以卫星作为中继，建立了兴隆地面站与格拉茨地面站之间的共享密钥，实验中获取共享密钥数据量约８００千比特。

基于共享密钥，采用一次一密的加密方式，中奥联合团队在北京到维也纳之间演示了图片加密传输。量子力学的开创者之一、曾获诺奖的奥地利物理学家埃尔温·薛定谔和中国古代哲学家墨子的图像加密后被分别传送到北京和维也纳。

结合高级加密标准ＡＥＳ－１２８协议，每秒更新一次种子密钥，中奥联合团队建立了一个北京到维也纳的加密视频通信系统，并利用该系统成功举行了两国科学院之间的７５分钟洲际视频会议，两院院长白春礼与安东·蔡林格进行了世界首次量子保密视频通话。

“洲际量子通信中，数据是使用量子密钥加密后传输的，监听者无法窃取中间的信息，这是目前人类唯一已知的无条件安全的通信手段，”论文第一作者、中国科技大学高级工程师廖胜凯告诉新华社记者，“除了使用量子密钥进行加密，其他环节与正常的照片传输和视频一样。”

http://www.xinhuanet.com/science/2018-01/21/c_136912037.htm

China launches quantum-enabled satellite Micius

16 August 2016

https://www.youtube.com/watch?v=qj22gj6vNX4

Computational security

- In contrast to OTP's "perfect" or "info-theoretic“ security, most cryptosystems have "computational" security
  - This means that it's certain they can be broken, given enough work by Eve
- How much is "enough"?
- At worst, Eve tries every key
  - How long that takes depends on how long the keys are
  - But it only takes this long if there are no "shortcuts"!

Some data points

- One computer can try about 17 million keys per second
- A medium-sized corporate or research lab may have 100 computers
- The BOINC project has 13 million computers

Berkeley Open Infrastructure for Network Computing

- Remember that most computers are idle most of the time (they're waiting for you to type something); getting them to crack keys in their spare time doesn't actually cost anything extra

40-bit crypto

- This was the US legal export limit for a long time
- 2^{40} = 1,099,511,627,776 possible keys
- • One computer: 18 hours
- • One lab: 11 minutes
- • BOINC: 5 ms

56-bit crypto

- This was the US government standard (DES) for a long time
- 2^{56} = 72,057,594,037,927,936 possible keys
- One computer: 134 years
- One lab: 16 months
- BOINC: 5 minutes

Cracking DES

"DES cracker" machine of Electronic Frontier Foundation

128-bit crypto

- This is the modern standard
- 2^{128} = 340,282,366,920,938,463,463,374,607, 431,768,211,456 possible keys

- One computer: 635 thousand million million million years

- One lab: 6 thousand million million million years

- BOINC: 49 thousand million million years

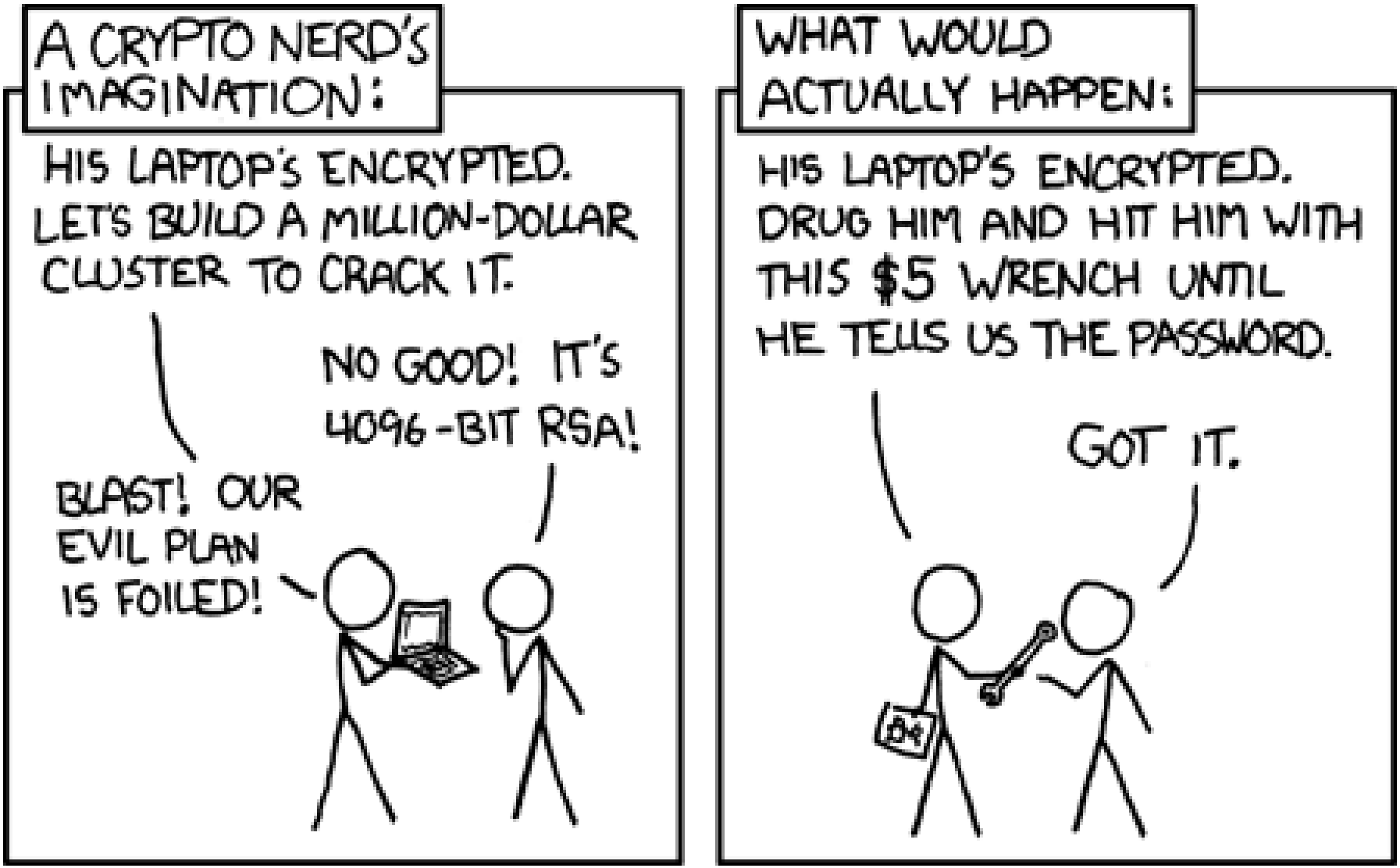
Well, we cheated a bit

- This isn't really true, since computers get faster over time
  - A better strategy for breaking 128-bit crypto is just to wait until computers get 2^{88} times faster, then break it on one computer in 18 hours.
  - How long do we wait? Moore's law says 132 years.
  - If we believe Moore's law will keep on working, we'll be able to break 128-bit crypto in 132 years (and 18 hours) :-)
    - Q: Do we believe this?

An even better strategy

- Don’t break the crypto at all!
- There are always weaker parts of the system to attack
  - Remember the Principle of Easiest Penetration
- The point of cryptography is to make sure the information transfer is not the weakest link

Rubber hose cryptanalysis



Encryption/Decryption

- Plaintext: a message in its original form
- Ciphertext: a message in the transformed, unrecognized form
- Encryption: the process that transforms a plaintext into a ciphertext
- Decryption: the process that transforms a ciphertext to the corresponding plaintext
- Key: the value used to control encryption/decryption.

Cryptanalysis

- “code breaking”, “attacking the cipher”
- Difficulty depends on
  - sophistication of the cipher
  - amount of information available to the code breaker
- Any cipher can be broken by exhaustive trials, but rarely practical

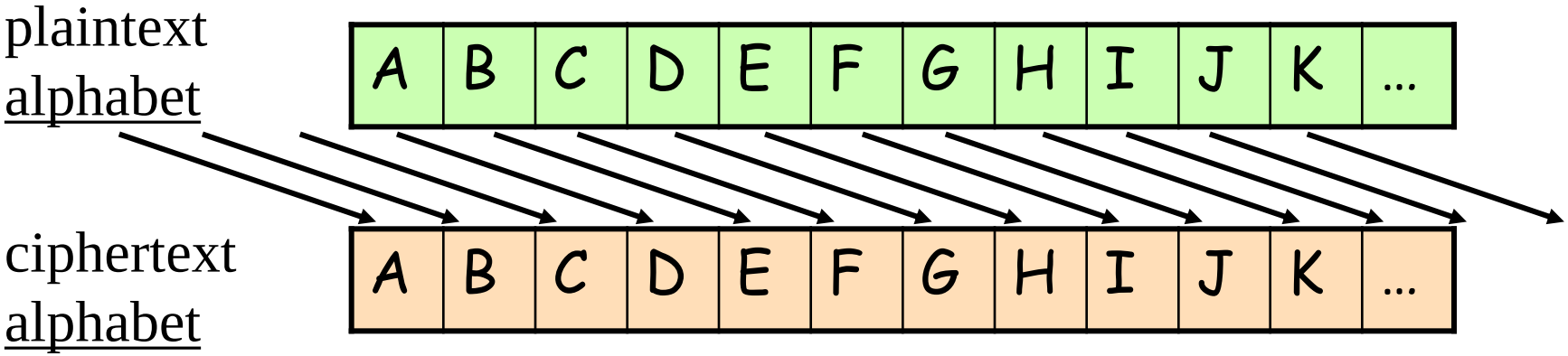
Shift Cipher

- The Key Space:
  - [0 .. 25]
- Encryption given a key K:
  - each letter in the plaintext P is replaced with the K'th letter following corresponding number (shift right)
- Decryption given K:
  - shift left

History: K = 3, Caesar's cipher

Caesar Cipher

- Replace each letter with the one 3 letters later in the alphabet
  - ex.: plaintext CAT → ciphertext FDW



Trivial to break

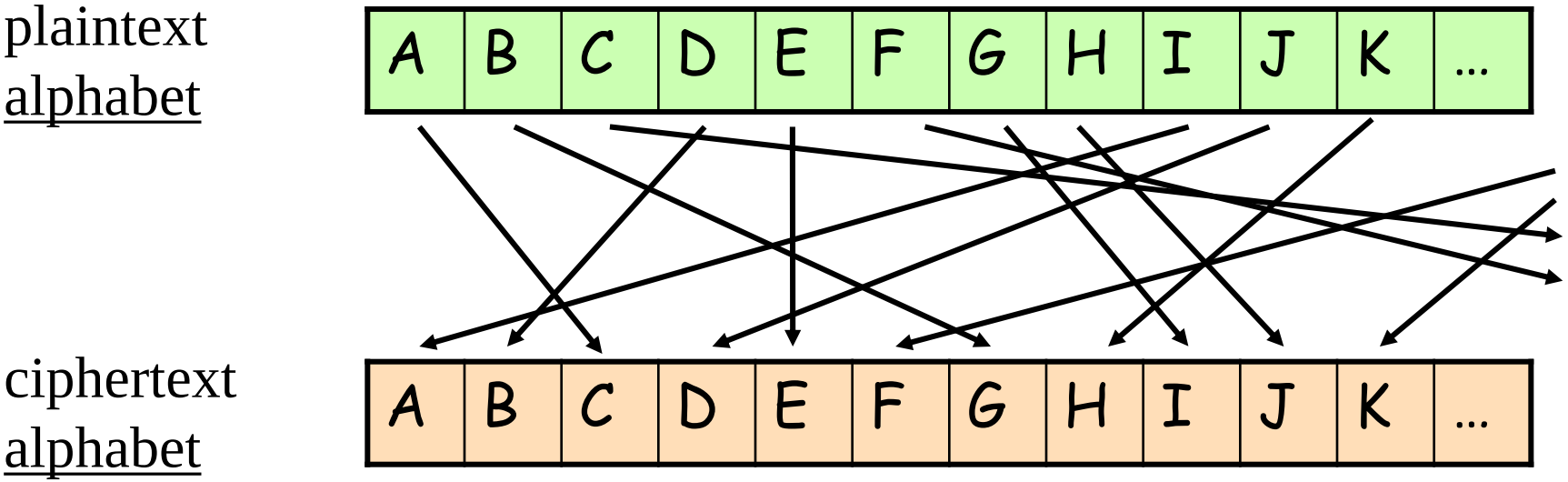
Shift Cipher: Cryptanalysis

- Can an attacker find K?
  - YES: by a bruteforce attack through exhaustive key search.
    - key space is small (<= 26 possible keys).
  - How much ciphertext is needed?

- Lessons:
  - Key space needs to be large enough.
  - Exhaustive key search can be effective.

Mono-Alphabetic Ciphers

- Generalized substitution cipher: an arbitrary (but fixed) mapping of one letter to another
  - 26! ($\approx 4.0*10^{26} \approx 2^{88}$) possibilities



Attacking Mono-Alphabetic Ciphers

- Broken by statistical analysis of letter, word, and phrase frequencies of the language
- Frequency of single letters in English language, taken from a large corpus of text:

A ≈ 8.2%	H ≈ 6.1%	O ≈ 7.5%	V ≈ 1.0%
B ≈ 1.5%	I ≈ 7.0%	P ≈ 1.9%	W ≈ 2.4%
C ≈ 2.8%	J ≈ 0.2%	Q ≈ 0.1%	X ≈ 0.2%
D ≈ 4.3%	K ≈ 0.8%	R ≈ 6.0%	Y ≈ 2.0%
E ≈ 12.7%	L ≈ 4.0%	S ≈ 6.3%	Z ≈ 0.1%
F ≈ 2.2%	M ≈ 2.4%	T ≈ 9.1%	
G ≈ 2.0%	N ≈ 6.7%	U ≈ 2.8%	

How to Defeat Frequency Analysis?

- Use larger blocks as the basis of substitution. Rather than substituting one letter at a time, substitute 64 bits at a time, or 128 bits.
  - Leads to block ciphers such as DES & AES.

- Use different substitutions to get rid of frequency features.
  - Leads to polyalphabetical substituion ciphers, and to stream ciphers such as RC4

Towards the Polyalphabetic Substitution Ciphers

- Main weaknesses of monoalphabetic substitution ciphers
  - In ciphertext, different letters have different frequency
    - each letter in the ciphertext corresponds to only one letter in the plaintext letter
- Idea for a stronger cipher (1460's by Alberti)
  - Use more than one substitutions, and switch between them when encrypting different letters
    - As result, frequencies of letters in ciphertext are similar
- Developed into an easy-to-use cipher by Vigenère (published in 1586)

The Vigenère Cipher

Treat letters as numbers: [A=0, B=1, C=2, …, Z=25]

Number Theory Notation: $Z_n = \{0, 1, \ldots, n-1\}$

Definition:

Given m, a positive integer, $P = C = (Z_{26})^n$, and $K = (k_1, k_2, \ldots, k_m)$ a key, we define:

Encryption:

$e_k(p_1, p_2 \ldots p_m) = (p_1+k_1, p_2+k_2 \ldots p_m+k_m) \pmod{26}$

Decryption:

$d_k(c_1, c_2 \ldots c_m) = (c_1-k_1, c_2-k_2 \ldots c_m-k_m) \pmod{26}$

Example:

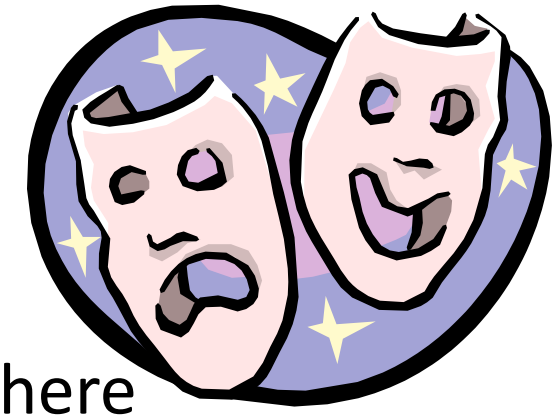
Plaintext: C R Y P T O G R A P H Y

Key: L U C K L U C K L U C K

Ciphertext: N L A Z E I I B L J J I

Security of Vigenere Cipher

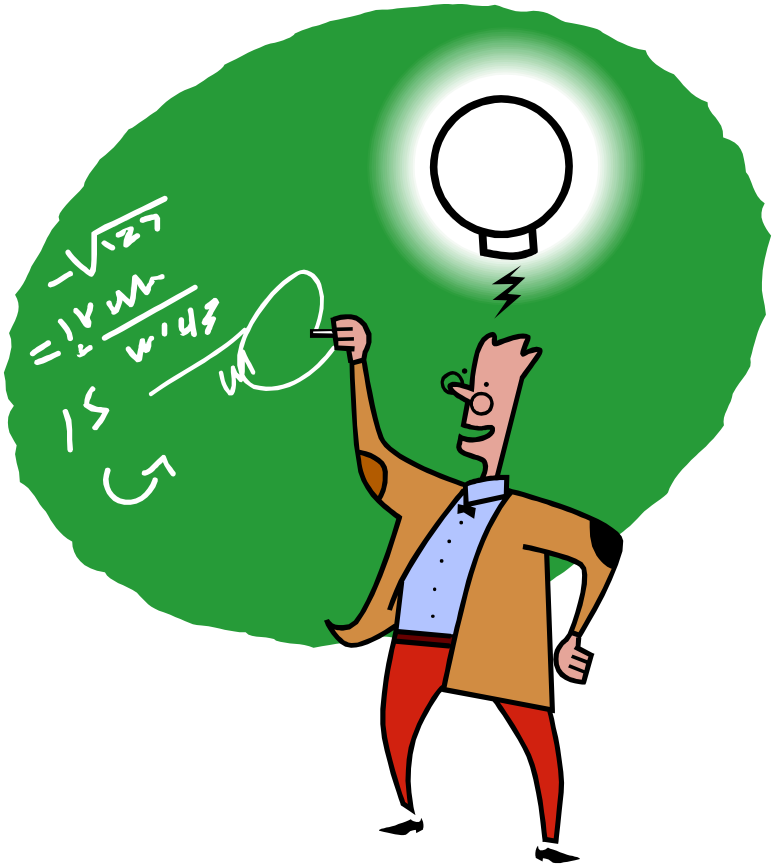
- Vigenere masks the frequency with which a character appears in a language: one letter in the ciphertext corresponds to multiple letters in the plaintext. Makes the use of frequency analysis more difficult.



- Any message encrypted by a Vigenere cipher is a collection of as many shift ciphers as there are letters in the key.

Vigenere Cipher: Cryptanalysis

- Find the length of the key.
  - Kasisky test
  - Index of coincidence (we won't cover here)
- Divide the message into that many shift cipher encryptions.
- Use frequency analysis to solve the resulting shift ciphers.
  - How?



Kasisky Test for Finding Key Length

- Observation: two identical segments of plaintext, will be encrypted to the same ciphertext, if they occur in the text at a distance Δ such that Δ is a multiple of m, the key length.
- Algorithm:
  - Search for pairs of identical segments of length at least 3
  - Record distances between the two segments: $\Delta 1$, $\Delta 2$, …
  - m divides gcd($\Delta 1$, $\Delta 2$, …)

Example of the Kasisky Test

Key	K	I	N	G	K	I	N	G	K	I	N	G	K	I	N	G	K	I	N	G	K
PT	t	h	e	s	u	n	a	n	d	t	h	e	m	a	n	i	n	t	h	e	m
CT	D	P	R	Y	E	V	N	T	N	B	U	K	W	I	A	O	X	B	U	K	W

Repeating patterns (strings of length 3 or more) in ciphertext are likely due to repeating plaintext strings encrypted under repeating key strings; thus the location difference should be multiples of key lengths.

Ciphertext Only Attacks

- Ex.: attacker can intercept encrypted communications, nothing else

- Breaking the cipher: analyze patterns in the ciphertext
  - provides clues about the encryption method/key

Known Plaintext Attacks

- Ex.: attacker intercepts encrypted text, but also has access to some of the corresponding plaintext (definite advantage)

- Makes some codes (e.g., mono-alphabetic ciphers) very easy to break

Chosen Plaintext Attacks

- Ex.: attacker can choose any plaintext desired, and intercept the corresponding ciphertext

- Allows targeted code breaking (choose exactly the messages that will reveal the most about the cipher)

The “Weakest Link” in Security

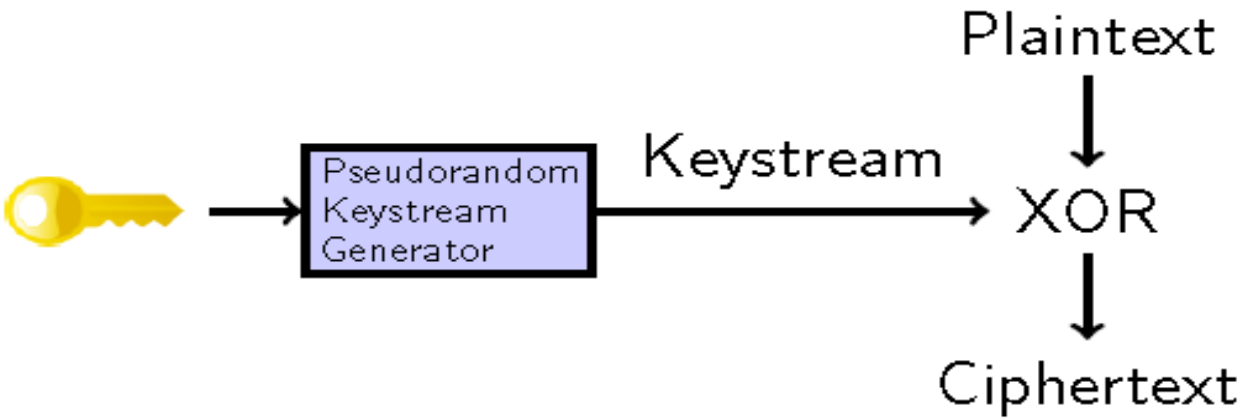
- Cryptography is rarely the weakest link
- Weaker links
  - Implementation of cipher
  - Distribution or protection of keys
  - … …

Types of secret-key cryptosystems

- Secret-key cryptosystems come in two major classes
  - Stream ciphers
  - Block ciphers

Stream ciphers

- A stream cipher is what you get if you take the One-Time Pad, but use a pseudorandom keystream instead of a truly random one



- RC4 is the most commonly used stream cipher on the Internet today

Stream ciphers

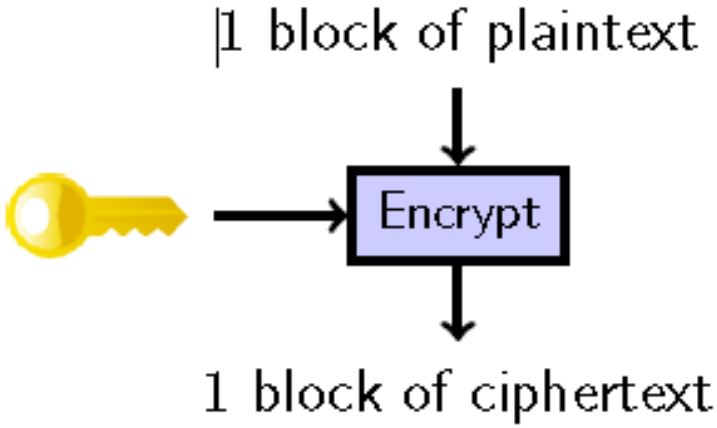
- Stream ciphers can be very fast
  - This is useful if you need to send a lot of data securely
- But they can be tricky to use correctly!
  - What happens if you use the same key to encrypt two different messages?
  - How would you solve this problem without requiring a new shared secret key for each message? Where have we seen this technique before?
- WEP, PPTP are great examples of how not to use stream ciphers

Block ciphers

- Note that stream ciphers operate on the message one bit at a time
- What happens in a stream cipher if you change just one bit of the plaintext?
- We can also use block ciphers
  - Block ciphers operate on the message one block at a time
  - Blocks are usually 64 or 128 bits long
- AES is the block cipher everyone should use today
  - Unless you have a really, really good reason

Modes of operation

•Block ciphers work like this:



•But what happens when the plaintext is larger than one block?

- The choice of what to do with multiple blocks is called the mode of operation of the block cipher

Modes of operation

•The simplest thing to do is just to encrypt each successive block separately.

This is called Electronic Code Book (ECB) mode

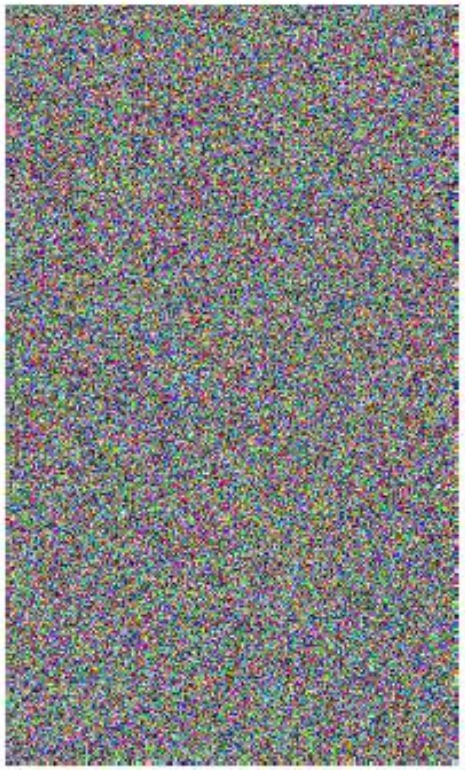
•But if there are repeated blocks in the plaintext, you'll see the same repeating patterns in the ciphertext:

Modes of operation

- There are much better modes of operation to choose from

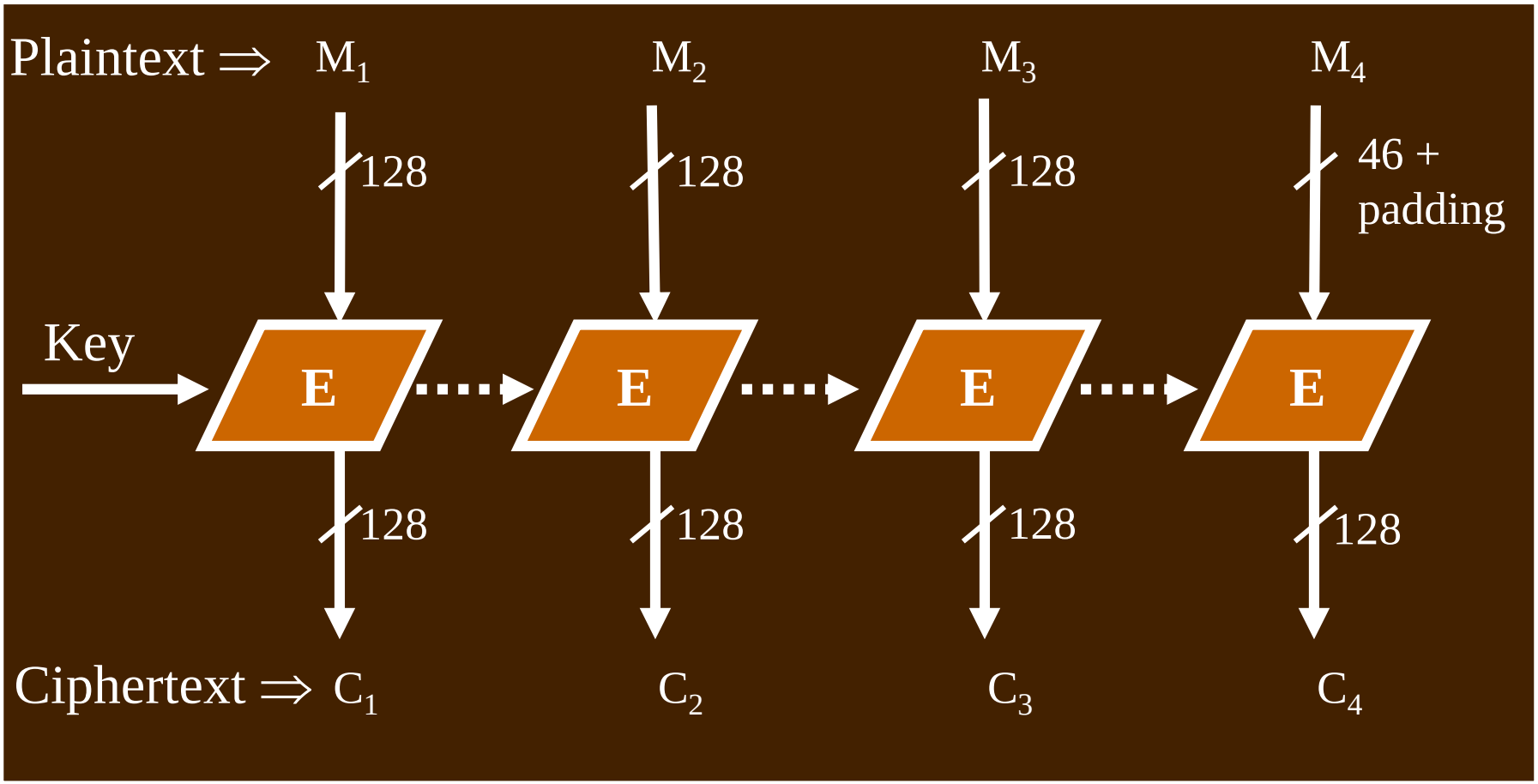
  Common ones include Cipher Block Chaining (CBC), Counter (CTR), and Galois Counter (GCM) modes

- Patterns in the plaintext are no longer exposed



- But you need an IV (Initial Value), which acts much like a salt

Electronic Code Book (ECB)



Cipher Block Chaining (CBC)