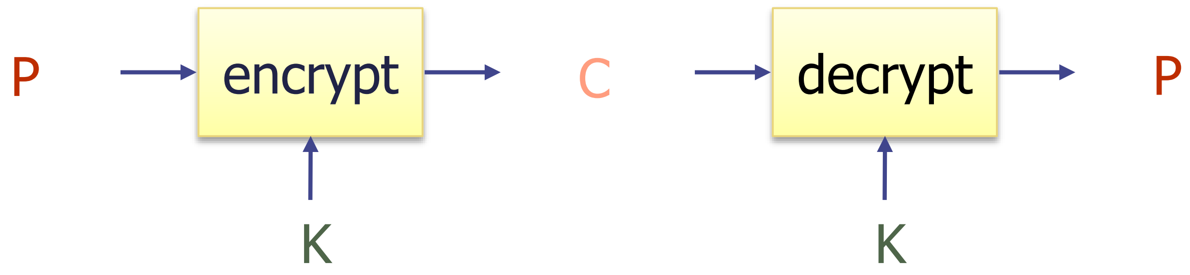
# ENEE 459-C
# Computer Security

## Symmetric key encryption

UNIVERSITY OF
MARYLAND

# Symmetric Cryptosystem

- Scenario
  - Alice wants to send a message (plaintext P) to Bob
  - The communication channel is insecure and can be eavesdropped
  - If Alice and Bob have previously agreed on a symmetric encryption scheme and a secret key K, the message can be sent encrypted (ciphertext C)
- Issues
  - What is a good symmetric encryption scheme?
  - What is the complexity of encrypting/decrypting?
  - What is the size of the ciphertext, relative to the plaintext?

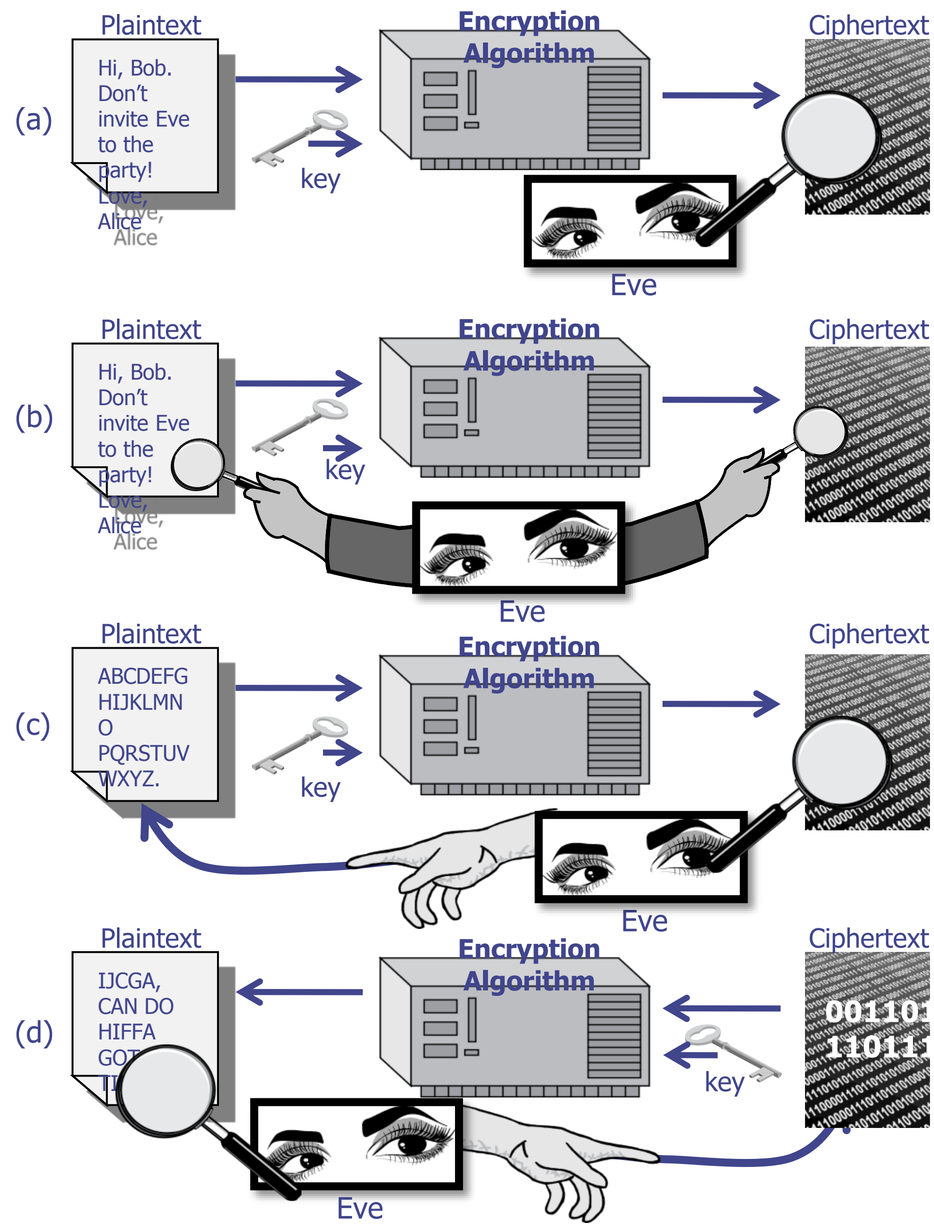


# Basics

- Notation
  - Secret key K
  - Encryption function $E_K(P)$
  - Decryption function $D_K(C)$
  - Plaintext length typically the same as ciphertext length
  - Encryption and decryption are permutation functions (bijections) on the set of all n-bit arrays
- Efficiency
  - functions $E_K$ and $D_K$ should have efficient algorithms
- Consistency
  - Decrypting the ciphertext yields the plaintext
  - $D_K(E_K(P)) = P$

# Attacks

- Attacker may have
  a) collection of ciphertexts (ciphertext only attack)
  b) collection of plaintext/ciphertext pairs (known plaintext attack)
  c) collection of plaintext/ciphertext pairs for plaintexts selected by the attacker (chosen plaintext attack)
  d) collection of plaintext/ciphertext pairs for plaintexts and ciphertexts selected by the attacker (chosen ciphertext attack or lunchtime attack)



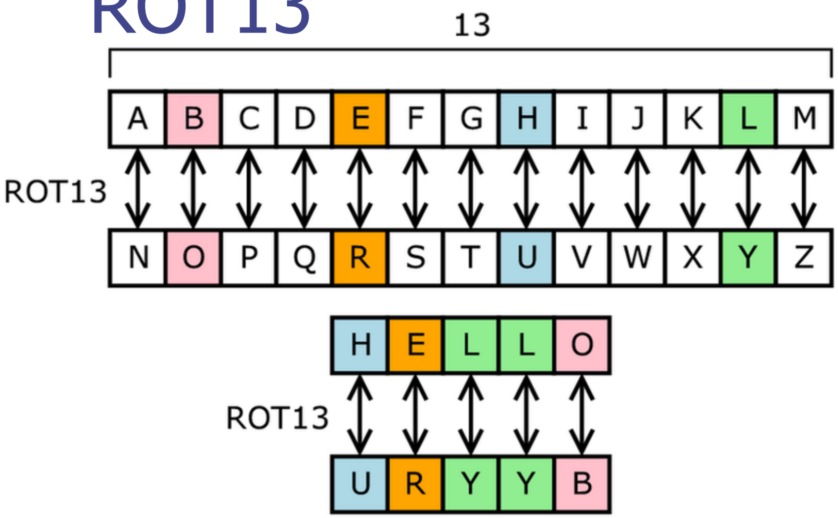
# Brute-Force Attack

- Try all possible keys K and determine if $D_K(C)$ is a likely plaintext
  - Requires some knowledge of the structure of the plaintext (e.g., PDF file or email message)
- Key should be a sufficiently long random value to make exhaustive search attacks unfeasible

Image by Michael Cote from http://commons.wikimedia.org/wiki/File:Bingo_cards.jpg

# Substitution Ciphers

- Each letter is uniquely replaced by another
- There are 26! possible substitution ciphers
- One popular substitution "cipher" for some Internet posts is ROT13




Public domain image from http://en.wikipedia.org/wiki/File:ROT13.png

# Substitution Boxes

- Substitution can also be done on binary numbers.
- Such substitutions are usually described by substitution boxes, or S-boxes.

| | 00 | 01 | 10 | 11 |
|---|---|---|---|---|
| 00 | 0011 | 0100 | 1111 | 0001 |
| 01 | 1010 | 0110 | 0101 | 1011 |
| 10 | 1110 | 1101 | 0100 | 0010 |
| 11 | 0111 | 0000 | 1001 | 1100 |

(a)

| | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| 0 | 3 | 8 | 15 | 1 |
| 1 | 10 | 6 | 5 | 11 |
| 2 | 14 | 13 | 4 | 2 |
| 3 | 7 | 0 | 9 | 12 |

(b)

**Figure 8.3:** A 4-bit S-box (a) An S-box in binary. (b) The same S-box in decimal.

# Frequency Analysis

- Letters in a natural language, like English, are not uniformly distributed
- Knowledge of letter frequencies, including pairs and triples can be used in cryptologic attacks against substitution ciphers

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| a: | 8.05% | b: | 1.67% | c: | 2.23% | d: | 5.10% |
| e: | 12.22% | f: | 2.14% | g: | 2.30% | h: | 6.62% |
| i: | 6.28% | j: | 0.19% | k: | 0.95% | l: | 4.08% |
| m: | 2.33% | n: | 6.95% | o: | 7.63% | p: | 1.66% |
| q: | 0.06% | r: | 5.29% | s: | 6.02% | t: | 9.67% |
| u: | 2.92% | v: | 0.82% | w: | 2.60% | x: | 0.11% |
| y: | 2.04% | z: | 0.06% | | | | |

Letter frequencies in the book *The Adventures of Tom Sawyer*, by Twain.

# One-Time Pads

- There is one type of substitution cipher that is absolutely unbreakable
  - The **one-time pad** was invented in 1917 by Joseph Mauborgne and Gilbert Vernam
  - We use a block of shift keys, $(k_1, k_2, \ldots, k_n)$, to encrypt a plaintext, M, of length n, with each shift key being chosen uniformly at random
- Since each shift is random, every ciphertext is equally likely for any plaintext

# Algorithms

- $K \leftarrow$ KeyGen(n): Pick a random key K of n bits
- $E_K(A)$: On input plaintext A, compute ciphertext B=A XOR K
- $D_K(B)$: On input ciphertext B, compute plaintext A=B XOR K
- Correctness: B XOR K= (A XOR K) XOR K= A XOR 0 = A
- Security?

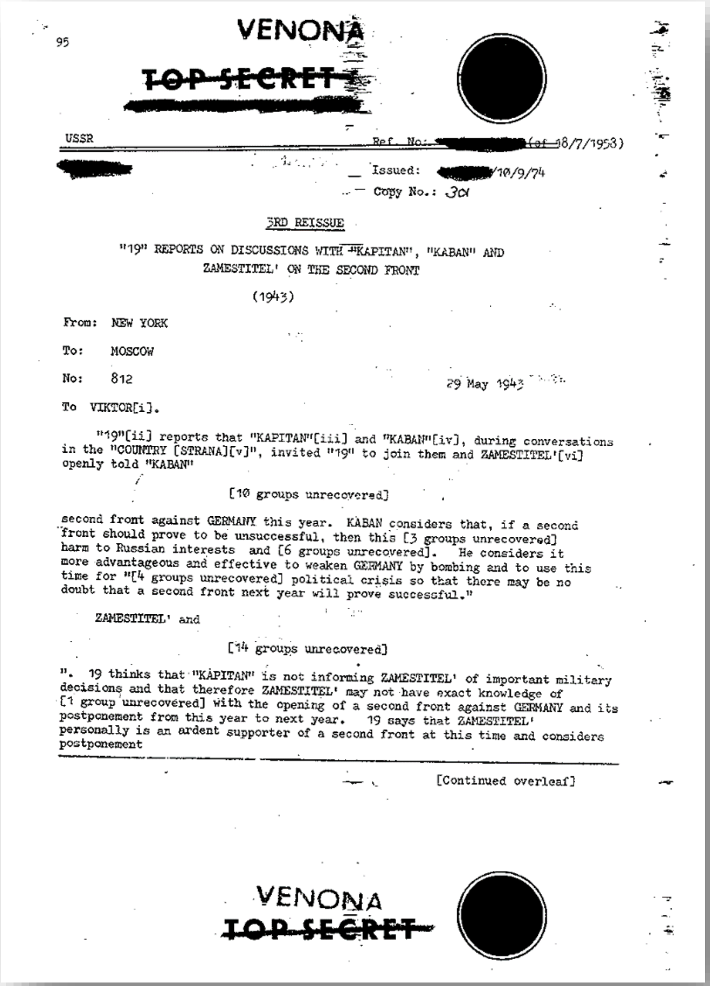
# Perfect security

- For all messages $m_1$ and $m_2$ and for all ciphertexts c
- $\Pr[K \leftarrow \text{KeyGen}(n): E_K(m_1)=c]=$

  $\Pr[K \leftarrow \text{KeyGen}(n): E_K(m_2)=c]$
- Proof
  - Note that $Enc_K(m_1)=c$ is the event $m_1$ XOR K = c which is the event K = $m_1$ XOR c
  - K is chosen at random (irrespective of $m_1$ and $m_2$, and therefore the probability is $2^{-n}$
  - Namely ciphertext does not reveal anything about the plaintext

# But…

- In spite of their perfect security, one-time pads have some weaknesses
- The key has to be as long as the plaintext
- Keys can never be reused
  - Repeated use of one-time pads compromised communications during the cold war



95

VENONA

TOP SECRET

[redacted]

[redacted]

USSR

Ref. No: [redacted] (of 18/7/1953)

[redacted]

Issued: [redacted] /10/9/74

Copy No.: 301

3RD REISSUE

"19" REPORTS ON DISCUSSIONS WITH "KAPITAN", "KABAN" AND ZAMESTITEL' ON THE SECOND FRONT

(1943)

From: NEW YORK

To: MOSCOW

No: 812

29 May 1943

To VIKTOR[i].

"19"[ii] reports that "KAPITAN"[iii] and "KABAN"[iv], during conversations in the "COUNTRY [STRANA][v]", invited "19" to join them and ZAMESTITEL'[vi] openly told "KABAN"

[10 groups unrecovered]

second front against GERMANY this year. KABAN considers that, if a second front should prove to be unsuccessful, then this [3 groups unrecovered] harm to Russian interests and [6 groups unrecovered]. He considers it more advantageous and effective to weaken GERMANY by bombing and to use this time for "[4 groups unrecovered] political crisis so that there may be no doubt that a second front next year will prove successful."

ZAMESTITEL' and

[14 groups unrecovered]

". 19 thinks that "KAPITAN" is not informing ZAMESTITEL' of important military decisions and that therefore ZAMESTITEL' may not have exact knowledge of [1 group unrecovered] with the opening of a second front against GERMANY and its postponement from this year to next year. 19 says that ZAMESTITEL' personally is an ardent supporter of a second front at this time and considers postponement

[Continued overleaf]

VENONA

TOP SECRET

[redacted]

Public domain declassified government image from
https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/books-and-monographs/venona-soviet-espionage-and-the-american-response-1939-1957/part2.htm

# Semantic security

- I give you a symmetric encryption scheme (Enc,Dec,K)
- What do you need to prove in order to say that it is secure?
- A strong notion used is "semantic security"
- We are going to define it as an interaction between the adversary **A** and a trusted party **T** that has the secret key.
- Informally:
  - **A** picks messages m_i and receives ciphertexts Enc_K(m_i) from **T**.
  - **A** picks message $m_0$ and $m_1$ and sends them to **T**.
  - **T** flips a coin b and computes $t_b$=Enc_K($m_b$).
  - **T** sends $t_b$ to the **A**.
  - The scheme is secure if **A** has no better chance of finding whether $t_b$ corresponds to $m_0$ or $m_1$ than just guessing!
- This should hold even if it is repeated many (polynomial) times

# Randomized encryption is important for semantic security

- Encryption should be randomized
  - For the same plaintext, it should output different ciphertexts
- How can we turn a deterministic encryption scheme into a randomized one?
  - Padding input with randomness
- Decryption should however always work