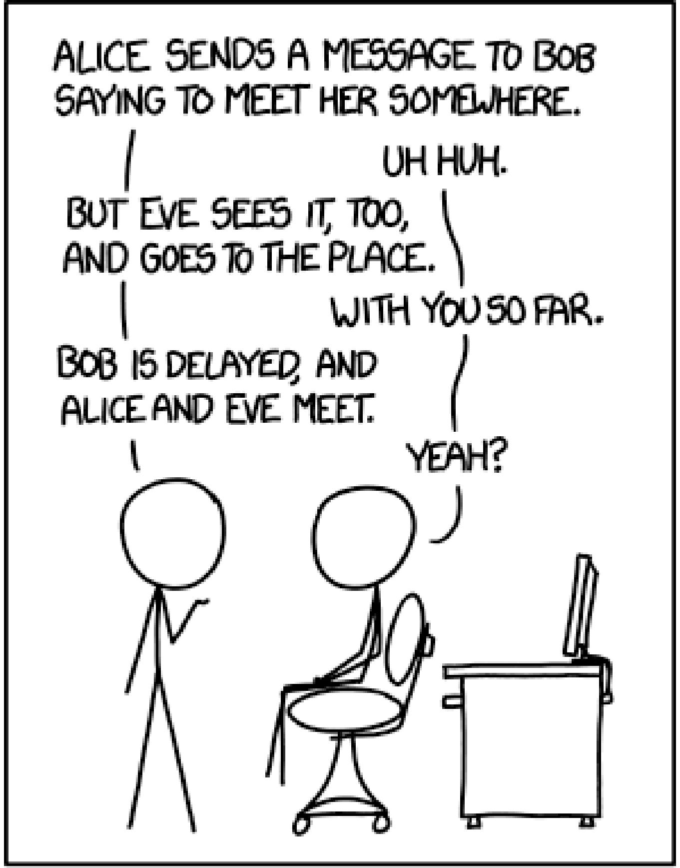
Computer and Network Security: Confidentiality Background

Kameswari Chebrolu

Outline

- **Modern Cryptography**
  - Overview
  - **Confidentiality**
    - **Background: Definition, Crypto-analysis, One Time Pads**
    - Symmetric key encryption, Block modes
    - Asymmetric key encryption
  - Integrity (includes Authentication)
    - Hashes, MAC, Digital signature



I'VE DISCOVERED A WAY TO GET COMPUTER SCIENTISTS TO LISTEN TO ANY BORING STORY.

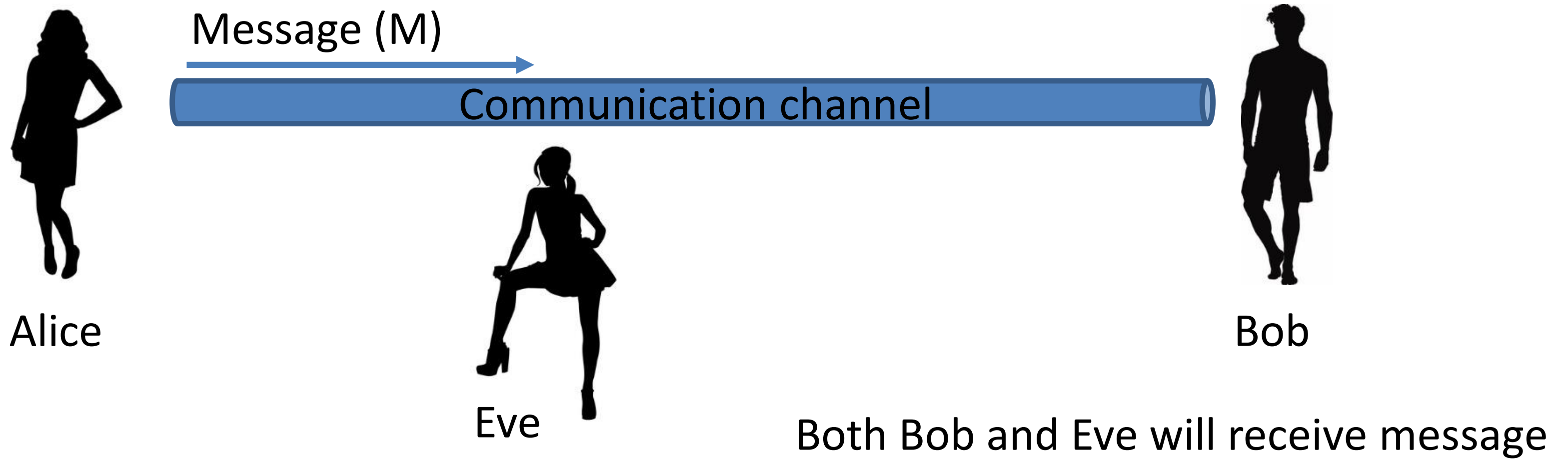
https://xkcd.com/1323/

Players

- Alice (A) and Bob (B) (lovers?)
  - In computer world: web browser/server; bank client/server; routers etc
- Eve (E, eavesdropper) (jealous ex?)
  - Passive attacker who can listen but not modify messages
- Mallory (M, malicious) or Trudy (T, intruder)
  - Active attacker who can modify, substitute, replay messages
- Goal: Alice and Bob want to communicate securely in presence of interlopers like E, M or T

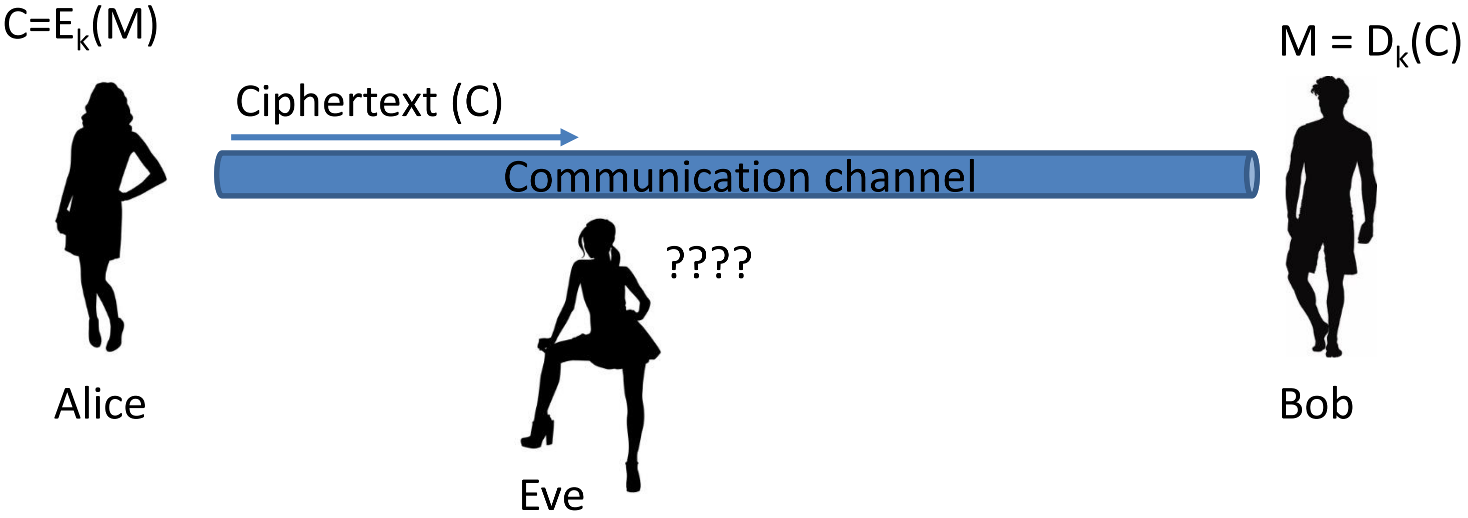
Confidentiality

- Information not available or disclosed to unauthorized entities
- Solution: Encryption and Decryption



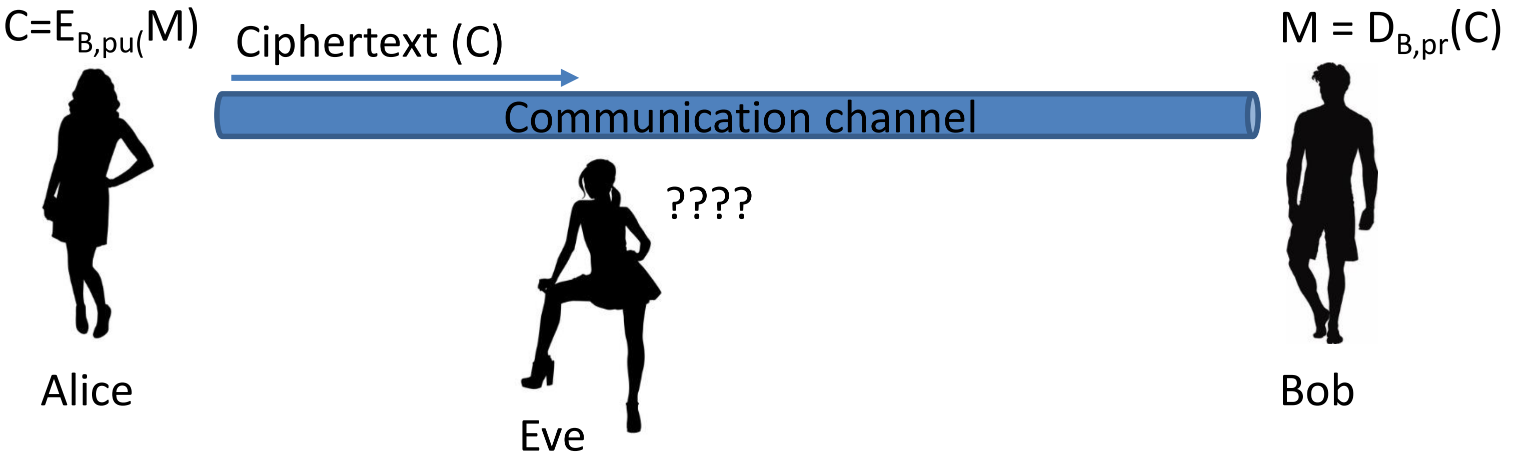
Solution

- Symmetric key:
  - Alice and Bob share a key k (how?)
  - Eve does not know the key but knows the encryption/decryption algorithm



Solution

- Asymmetric key:
  - Encryption/decryption algorithm, Bob’s public key (B,pu) open
    - Both Eve and Alice have access to this
  - Bob keeps private key (B,pr) secret
  - Alice encrypts message with Bob’s public key



Cryptoanalysis

- Science of recovering plaintext of a message without key
  - Can recover plain text or key
  - Can find weakness in implementation (side-channel attack)
- Assumption by A. Kerckhoff: Attacker knows complete details of the algorithm and implementation
  - May not be true in reality but
  - If can’t break with knowledge, cannot break without knowledge
- Also assume, eavesdroppers have complete access to communication between sender and receiver

The analyst works with:

- encrypted messages
- known encryption algorithms
- Plaintext and corresponding ciphertext
- data items known or suspected to be in a ciphertext message
- mathematical and statistical tools and techniques
- properties of languages (like English or format)
- computers
- ingenuity and luck

Breaking a Cipher

According to Lars Knudsen:

- Total Break: find key K such that $D_k(C) = M$
- Global Deduction: find alternate algorithm A equivalent to $D_k(C)$ without knowing k
- Instance Deduction: plaintext of a given ciphertext
- Information Deduction: some partial information about text or key

k should be interpreted based on context
k can be shared or public or private key

Types of Attacks

- Ciphertext-only:

  Given: $C_1 = E_k(M_1), C_2 = E_k(M_2), \ldots C_i = E_k(M_i)$
  Deduce: $M_1, M_2, \ldots M_i$ or k or
  an algorithm to infer M_{i+1} from $C_{i+1} = E_k(M_{i+1})$

- Known-plaintext:

Given: $M_1, C_1 = E_k(M_1); M_2, C_2 = E_k(M_2); \ldots M_i, C_i = E_k(M_i)$
Deduce: Either k or an algorithm to infer M_{i+1} from $C_{i+1} = E_k(M_{i+1})$

  – Not uncommon, example: letters may begin with Dear/hello, source code with #define

- Chosen-plaintext:

Given: $M_1, C_1 = E_k(M_1); M_2, C_2 = E_k(M_2); \ldots M_i, C_i = E_k(M_i)$
where the attacker can choose $M_1, M_2 \ldots M_i$
Deduce: Either k or an algorithm
to infer M_{i+1} from $C_{i+1} = E_k(M_{i+1})$

  – Example: Leak a specific message to spy

- Adaptive-chosen-plaintext: Same as above except attacker can choose subsequent plaintext based on previous encryptions

- Chosen-cipertext: Applicable to asymmetric/ public-key algorithms for digital signatures

$$\text{Given: } C_1, M_1 = D_k(C_1); C_2, M_2 = D_k(C_2), \ldots C_i, M_i = D_k(C_i)$$
$$\text{Deduce: } \mathrm{k}$$

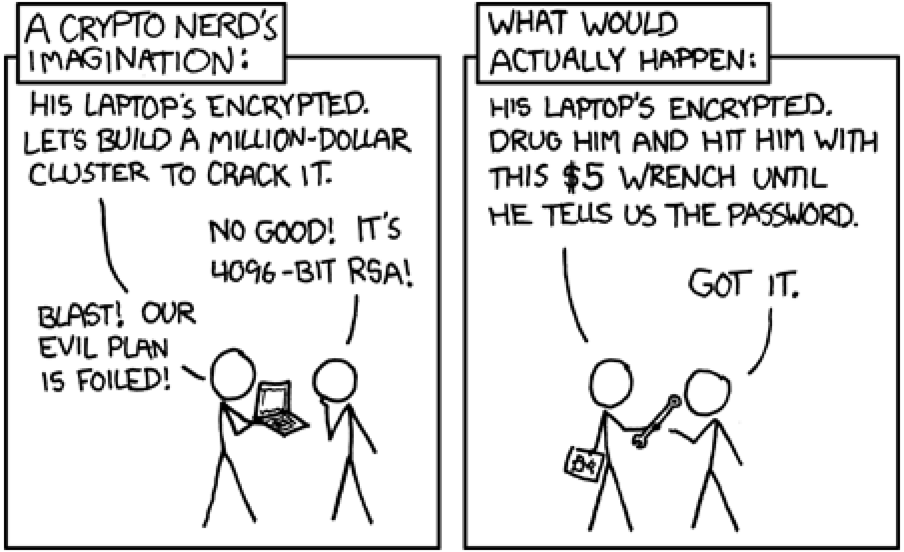
- Best and most powerful Attack: Rubber-hose cryptoanalysis
  - Torture/bribe/blackmail for key ☺

Attack Complexity

- Characterized by resources required
  - Data: Amount of input data (plain/cipher text) to attack
  - Storage: Amount of memory needed for attack
  - Time: Time (computational steps) needed for attack
- Complexity is minimum of the three factors

Attack Complexity

- Unconditionally secure: unbreakable given infinite resources (e.g. one time pad)
- Most cryptosystems breakable in cipher-text-only attack by brute-force
  - Try every possible key and look for meaningful plaintext
- **Computationally secure**: cannot be broken with available resources now or in future
  - E.g. 2^128 operations to break; 1 million computers @ 1 million operations per second → 10^{19} years (billion times the age of the universe)



https://xkcd.com/538/

What makes a good cipher?

- Encryption: $E_k(m)$ is easy to compute given message m and key k
- Decryption: $D_k(x)$ is easy to compute given encrypted content x and key k
- Attacker: Given x ($=E_k(m)$), hard to find m without k
  - Cannot be broken with available resources now or in future (computationally secure)
- Larger keyspace → stronger cipher
  - View key as an n bit string
  - Strength is non polynomial in n; e.g. extra bit doubles effort

One Time Pads (1917CE)

- **Perfect Cipher (unbreakable given infinite resources)**
  - Unconditionally secure/information-theoretically secure as opposed to computationally secure
  - Hotline between US and former Soviet Union rumored to use this
- Key: non repeating set of random letters written on a pad
  - Key used only once
  - Used pages destroyed after each use

Example

- Message: ONETIMEPAD
- Key Sequence from Pad: TBFRGFARFM
- Cipher text: IPKLPSFHGQ
- Decryption?
  - Add key sequence again

O+T mod26 = I

N+B mod26 = P

E+F mod 26 = K

……

m,k are binary strings

$$c_i = m_i \oplus k_i$$

$$m_i = c_i \oplus k_i$$

Note: Message, key and cipher-text have same length

Why Perfect?

- A given cipher-text is equally likely to correspond to any possible plain-text of equal size

Example

- Message: ONETIMEPAD
- Key Sequence from Pad: TBFRGFARFM
- Cipher text: IPKLPSFHGQ
- If key sequence was POYYAEAAZX; decrpyts to SALMONEGGS
- If key sequence was BXFGBMTMXM; decrpyts to GREENFLUID

O+T mod26 = I
N+B mod26 = P
E+F mod 26 = K
……

Why Perfect?

- A given cipher-text is equally likely to correspond to any possible plain-text of equal size

$P(M)$: probability of message M
$P(M|C)$: probability of message M after seeing C
For OTP, $P(M) = P(M|C)$
Seeing C has not helped the attacker know more about M

OTP Shortcomings

- Key sequence same length as message
  – Key distribution and storage problem for long messages
  – Two time pad insecure (key cannot be reused)
    - Check out: https://en.wikipedia.org/wiki/Venona_project
- Synchronization problem
  – Receiver off by a bit or channel drops some bits
- Malleable (can change ciphertext to alter plain text)
  – Provides confidentiality but not integrity

Summary

- Goal of Confidentiality and technique overview to achieve it
- Crypto-analysis aims at breaking ciphers to find weaknesses
  - Types of attacks and the complexity
  - Security goal: make cipher computationally secure
- Perfect Cipher: One time pads
  - But not practical