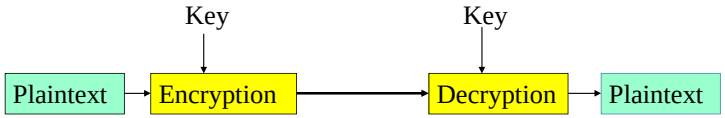
Encryption

COMP620 Information Privacy & Security

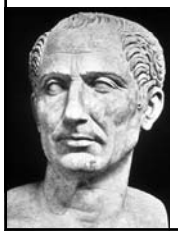
Cryptography

- Cryptography in general represents the process of encrypting a plain-text file into an unreadable cipher so that it can be stored and decrypted by the intended recipient.
- Plaintext can be any bunch of bits, text, graphics, program, etc.



Historical Encoding

- People have been writing secret messages for millennia
- The Caesar cipher *(shift cipher)* is said to have been used by Julius Caesar
- Computational efficiency was very important before computers



Encryption Media

- Encryption can be used to secure information sent over a network.
- Encryption can also be used to secure data stored on a computer.

Caesar or Shift Cipher

- The letters of the alphabet are shifted by a fixed amount
- Key is the number of letters to shift
- Can easily be defeated by trying all 26 possible shifts

Decryption by Brute Force

- `Frpsxwhuv duh ixq`
- `Eqorwvgtu ctg hwp`
- `Dpnqvufst bsf gvo`
- `Computers are fun`

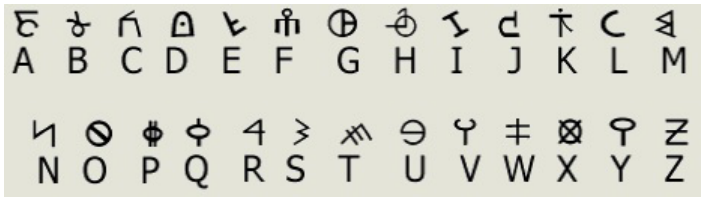
Types of Attacks

We assume that an adversary knows the encryption algorithm and has:

- **Ciphertext only** – samples of ciphertext without information about the content
- **Known plaintext** – examples of ciphertext and the corresponding plaintext
- **Chosen plaintext** – adversary can get ciphertext samples of plaintext of their choosing

Substitution Cipher

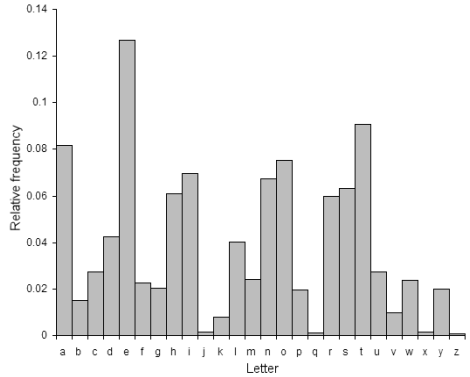
- Letters are mapped to symbols or letters
- Key – An alphabetical list of the symbols



- There are 26! = 4×10^{26} possible keys

Letter Frequency

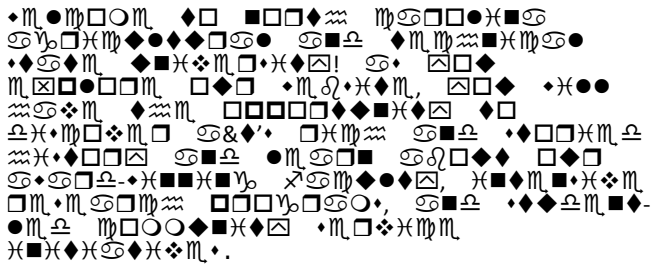
The frequency of the use of a letter in English



The frequency of letter pair, triples, and short words are also available on the web.

Cipher Text

What does this say?



Chancellor Martin

Effectiveness of Frequency Decryption

- Sample substitution cipher text was partially decrypted using only the letter frequency.

Guess what it says

welrome to nohtu raholina abhirdltdhal anc teruniral state dnipehsitg! as god ekylohe odh wevsite, god will uape tue oyyohtdnitg to cisropeh a&t's hiru anc stohiec uistohg anc leahn avodt odh awahc-winninb fardltg, intensipe heseahru yhobhams, anc stdcent-lec rommdnitg sehpire initiatipes.

Chancellor Martin

Original Text

welcome to north carolina agricultural and technical state university! as you explore our website, you will have the opportunity to discover a&t's rich and storied history and learn about our award-winning faculty, intensive research programs, and student-led community service initiatives.

Chancellor Martin

Vigenère Cipher

- Originally described by Giovan Battista Bellaso in 1553
- A text key is repeated for the length of plaintext

$C_i = (P_i + K_i) \bmod 26$ to encrypt

$P_i = (C_i - K_i) \bmod 26$ to decrypt

```
welcometonorthcarolina  plaintext
informationprivacyinfo  key
erqqfyemwbbgkpxatmtvso  ciphertext
```

Vigenère Cryptanalysis

- For long text and short keys, character frequency analysis provides a lot of information
- Sometimes the same plaintext letter will be encrypted by the same key character
- Statistical analysis on the frequency and distance between matches gives an indication of the key size

Running key cipher

- The running key cipher is similar to the Vigenère cipher, but a long, non-repeating key is used.
- Typically the key is some common publication, such as a book or periodical
- Example using well known C book

```
Plaintext:    f l e e a t o n c e w e a r e d i s c o v e r e d
Running key:  E R R O R S C A N O C C U R I N S E V E R A L P L
Ciphertext:   J C V S R L Q N P S Y G U I M Q A W X S M E C T O
```

One-Time Pad

- With one-time pad encryption, the bit stream of the message is XOR with a random key
- The key must be at least as long as the message so it is not repeated
- The key must be truly random, not just pseudo-random

Perfect Encryption

- One-time pad is a perfect encryption technique that cannot be broken
- A given cipher text can be decrypted into any possible plain text by using the appropriate key

011001010	cipher text
010101010	with key 001100000
000011111	with key 011010101

Running Key Analysis

- If the key text for the running key is perfectly random, then it is the same as one-time pad
- Usually human text is far from random providing relatively poor security
- You can subtract probably plaintext from the cipher text down the whole string and look for readable text, which is probably part of the key.

Diffusion and Confusion

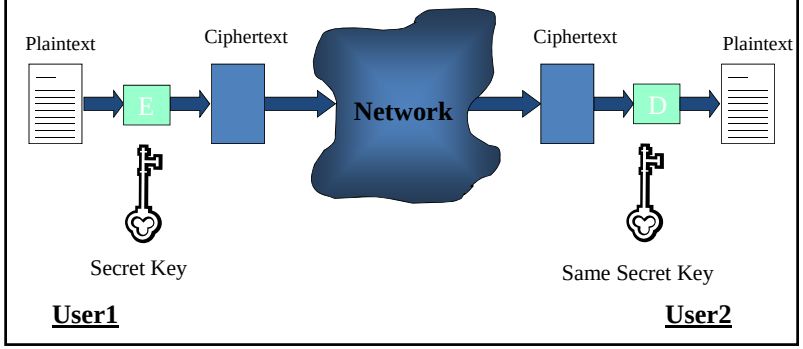
- **Diffusion** – spread the plain text data across the cipher text. A byte of plain text should impact many bytes of cipher text
- **Confusion** – change the bits of the plain text according to some rule

Types of Encryption

- Symmetric Key or Secret Key
  - The encryption key is the same as the decryption key.
  - Sender and receiver have to securely share a key.
- Asymmetric Key or Public Key
  - The key to decrypt is different, but related to, the key to encrypt.
  - The encryption key can be made public while the decryption key is kept secret.

Symmetric Key Cryptography

- Keys exchanged prior to communications. Parties verified at that time.
- Key to encrypt message is the same as key to decrypt.
- DES and AES are examples of Symmetric Key Cryptography.

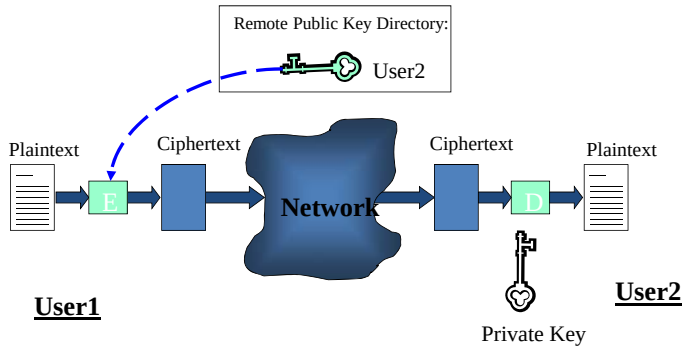


Why Publish a Standard?

- The Data Encryption Standard (DES) and Advanced Encryption Standard (AES) algorithms are published and well known
- Why not keep the algorithm secret?
- To be useful, others have to implement it.
- A good encryption algorithm will allow only those with a key to access the data. Knowing the algorithm does not give you access.

Asymmetric Key Cryptography

- Public key different from private key.
- RSA encryption is an example of Asymmetric Key Cryptography.



Encryption Performance

- RSA asymmetric key encryption is slower than DES or AES.
- DES and AES are easy to implement in hardware.
- AES can be efficiently implemented in software.
- Hybrid encryption uses both asymmetric and symmetric key systems.

Key Strength

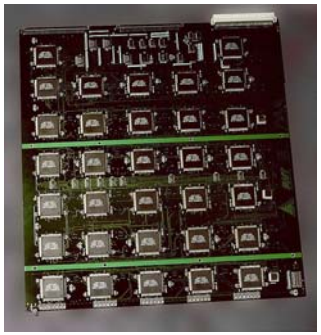
- The longer they key, the harder it is to defeat the encryption by brute force.
- If the key is n bits, it requires 2^n guesses to try all possible keys. You are likely to guess correctly in 2^{n-1} tries.
- Asymmetric key algorithms require a mathematical relation between the keys so not every bit string can be a key.

Key Lengths

- DES uses a 56 bit key
- Triple DES or DES3 uses two DES keys for a total of 112 bits
- AES uses 128, 192 or 256 bit keys.
- RSA uses variable length keys, frequently 512, 1024 or 2K bits in length.

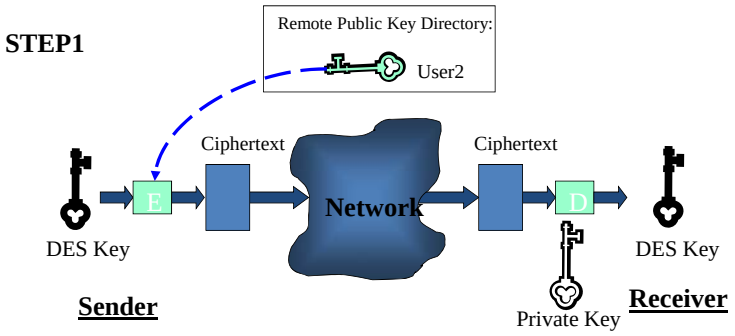
Brute Force Decryption

- Brute force tries all possible keys.
- In 1998 the Electronic Frontier Foundation built a device that could brute-force a DES key in a little more than 2 days



Hybrid Cryptography (STEP 1)

- DES key is encrypted with asymmetric key cryptography using Public Key of receiver.
- DES key sent to receiver.
- Both users end up with a shared DES key.



Hybrid Cryptography (STEP 2)

- Message is encrypted with the DES key previously sent to the receiver.
- DES key is discarded after sending the message.

STEP2