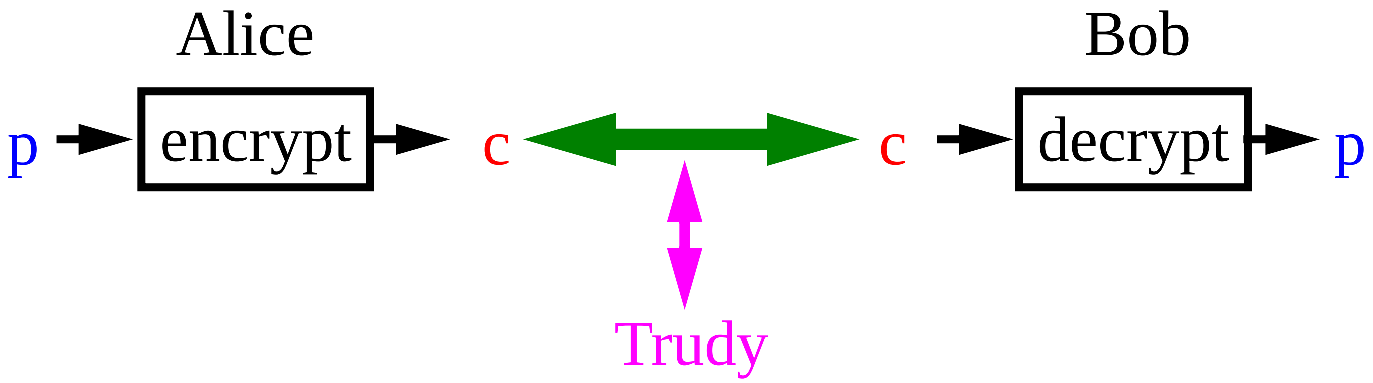
CRYPTOGRAPHY

- Secrecy
- Ciphers
- Secret Key Cryptography
- Key Exchange
- Public Key Cryptography
- Digital Signatures
- Internet applications

Secrecy

- Alice wants to send a message (plaintext p) to Bob
- The communication channel is insecure and can be eavesdropped by Trudy
- If Alice and Bob have previously agreed on an ***encryption scheme*** (***cipher***), the message can be sent encrypted (ciphertext c)



- Issues:
  - what is a good cipher?
  - what is the complexity of encrypting/decrypting?
  - what is the size of the ciphertext, relative to the plaintext?
  - if Alice and Bob have never interacted before, how can they agree on a cipher?

Traditional Cryptography

- Ciphers were already studied in ancient times
- ***Caesar's cipher***:
  - replace a with d
  - replace b with e
  - ...
  - replace z with c
- A more general ***monoalphabetic substitution cipher*** maps each letter to some other letter
- Armed with simple statistcal knowledge, Trudy can easily break a monalphabetic substitution cypher
  - most frequent letters in English: e, t, o, a, n, i, ...
  - most frequent digrams: th, in, er, re, an, ...
  - most frequent trigrams: the, ing, and, ion, ...
- The first description of the frequency analysis attack appears in a book written in the 9th century by the Arab philosopher al-Kindi

Example (S. Singh, The Code Book, 1999)

- Ciphertext

  PCQ VMJYPD LBYK LYSO KBXBJXWXV BXV ZCJPO EYPD KBXBJYUXJ LBJOO KCPK. CP LBO LBCMKXPV XPV IYJKL PYDBL, QBOP KBO BXV OPVOV LBO LXRO CI SX'XJMI, KBO JCKO XPV EYKKOV LBO DJCMPV ZOICJO BYS, KXUYPD: 'DJOXL EYPD, ICJ X LBCMKXPV XPV CPO PYDBLK Y BXNO ZOOP JOACMPLYPD LC UCM LBO IXZROK CI FXKL XDOK XPV LBO RODOPVK CI XPAYOPL EYPDK. SXU Y SXEO KC ZCRV XK LC AJXNO X IXNCMJ CI UCMJ SXGOKLU?'

  OFYRCDMO, LXROK IJCS LBO LBCMKXPV XPV CPO PYDBLK

- Frequencies

Frequency Analysis

- Identyfying comon letters, digrams and trigrams...

  PCQ VMJYPD LBYK LYSO KBXBJXWXV BXV ZCJPO EYPD KBXBJYUXJ LBJOO KCPK. CP LBO LBCMKXPV XPV IYJKL PYDBL, QBOP KBO BXV OPVOV LBO LXRO CI SX'XJMI, KBO JCKO XPV EYKKOV LBO DJCMPV ZOICJO BYS, KXUYPD: 'DJOXL EYPD, ICJ X LBCMKXPV XPV CPO PYDBLK Y BXNO ZOOP JOACMPLYPD LC UCM LBO IXZROK CI FXKL XDOK XPV LBO RODOPVK CI XPAYOPL EYPDK. SXU Y SXEO KC ZCRV XK LC AJXNO X IXNCMJ CI UCMJ SXGOKLU?'

  OFYRCDMO, LXROK IJCS LBO LBCMKXPV XPV CPO PYDBLK

- First guess: LBO is THE

  PCQ VMJYPD THYK TYSE KHXHJXWXV HXV ZCJPE EYPD KHXHJYUXJ THJEE KCPK. CP THE THCMKXPV XPV IYJKT PYDHT, QHEP KHO HXV EPVEV THE LXRE CI SX'XJMI, KHE JCKE XPV EYKKOV THE DJCMPV ZEICJE HYS, KXUYPD: 'DJEXT EYPD, ICJ X LHCMKXPV XPV CPE PYDHLK Y HXNE ZEEP JEACMPTYPD TC UCM THE IXZREK CI FXKL XDEK XPV THE REDEPVK CI XPAYEPT EYPDK. SXU Y SXEE KC ZCRV XK TC AJXNE X IXNCMJ CI UCMJ SXGEKTU?'

  EFYRCDME, TXREK IJCS THE LHCMKXPV XPV CPE PYDBTK

- More guesses ...

Solution

- Ciphertext

PCQ VMJYPD LBYK LYSO KBXBJXWXV BXV ZCJPO EYPD KBXBJYUXJ LBJOO KCPK. CP LBO LBCMKXPV XPV IYJKL PYDBL, QBOP KBO BXV OPVOV LBO LXRO CI SX'XJMI, KBO JCKO XPV EYKKOV LBO DJCMPV ZOICJO BYS, KXUYPD: 'DJOXL EYPD, ICJ X LBCMKXPV XPV CPO PYDBLK Y BXNO ZOOP JOACMPLYPD LC UCM LBO IXZROK CI FXKL XDOK XPV LBO RODOPVK CI XPAYOPL EYPDK. SXU Y SXEO KC ZCRV XK LC AJXNO X IXNCMJ CI UCMJ SXGOKLU?'

OFYRCDMO, LXROK IJCS LBO LBCMKXPV XPV CPO PYDBLK

- Code

X	**Z**	**A**	**V**	**O**	**I**	**D**	**B**	**Y**	**G**	**E**	**R**	**S**	**P**	**C**	**F**	**H**	**J**	**K**	**L**	**M**	**N**	**Q**	**T**	**U**	**W**
a	**b**	**c**	**d**	**e**	**f**	**g**	**h**	**i**	**j**	**k**	**I**	**m**	**n**	**o**	**p**	**q**	**r**	**s**	**t**	**u**	**v**	**w**	**x**	**y**	**z**

- Plaintext

Now   during this time Shahrazad had borne King Shahriyar three sons. on the thousand and first night, when she had ended the tale of Ma'aruf, she rose and kissed the ground before him, saying: 'Great King, for a thousand and one nights I have been recounting to you the fables of past ages and the legends of ancient kings. May I make so bold as to crave a favour of your majesty?'

Epilogue, Tales from the Thousand and One Nights

Secret-Key Ciphers

- A ***secret-key cipher*** uses a ***key*** to encrypt and decrypt
- Caesar's generalized cypher uses ***modular addition*** of each character (viewed as an integer) with the key:
  - $c_i = p_i + k \bmod m$
  - $p_i = c_i - k \bmod m$
- A more secure scheme is to use ***modular exponentiation*** to encrypt blocks of characters (viewed as integers):
  - $c_{[i,j]} = p_{[i,j]}^{\ k} \bmod m$
  - where m is a large prime
- Unlike modular addition, modular exponentiation is considered computationally infeasible (exponential) to invert. Thus, even if Trudy guesses a pair ($c_{[i,j]}$, $p_{[i,j]}$), e.g., she knows the plaintext starts with the words "Dear Bob", she cannot compute the key k
- Alice and Bob need to share only key k. Bob decrypts using Euler's Theorem from number theory
  - $p_{[i,j]} = c_{[i,j]}^{\ d} \bmod m$
  - where d can be easily computed from k and m using Euclid's gcd algorithm

How to Establish a Shared Key?

- What if Alice and Bob have never met and did not agree on a key?
- The ***Diffie-Hellman key exchange protocol*** (1976) allows strangers to establish a secret shared key while communicating over an insecure channel
- Briefcase with locks ...
- Alice picks her secret "***half-key***" x (a large integer) and two large primes m and g. She sends to Bob
  - (n, g, $g^x \bmod m$)
  - Even if Trudy intercepts (n, g, $g^x \bmod m$), she cannot figure out x because modular logarithms are hard to compute
- Bob picks his secret half-key y and sends to Alice
  - ($g^y \bmod m$)
  - Again, Trudy cannot figure out y.
- The shared key is
  - $g^{xy} \bmod m$
  - Bob computes it as $(g^x \bmod n)^y \bmod m$
  - Alice computes it as $(g^y \bmod n)^x \bmod m$

Algorithmic Issues

- How can we efficiently compute modular exponents for large integers?

- It is not efficient to compute $q = g^x \bmod m$ in the obvious way:
  - $p = g^x$
  - $q = a \bmod m$

- ***Repeated Squaring Algorithm***
  - represent x in binary: $x_{b-1}x_{b-2} \ldots x_1x_0$
  - **repeat** b-1 times

    $g = g^2 \bmod m$

    this yields

    $p_0 = g \bmod m$
    $p_1 = g^2 \bmod m$
    $p_2 = g^4 \bmod m$
    $\ldots$
    $p_{b-1} = g^{2^{b-1}} \bmod m$
  - **for** $i = 0$ **to** b-1

    $q = qx_ip_i \bmod m$

- The number of arithmetic operations performed is proportional to log x

The Woman-in-the-Middle Attack

- Trudy can fool Alice and Bob to share a secret key with her
- How?

Public Key Ciphers

- A pair of keys is used (e,d)
- Key e is made ***public*** and is used to encrypt
- Key d is kept ***private*** and is used to decrypt
- RSA, by Rivest, Shamir, Adleman (1978) is the most popular pubkic key cipher
  - select a pair of large primes, p and q
  - let e = pq be the public key
  - define $\Phi(e) = (p-1)(q-1)$
  - let d be the private key, where 3d mod $\Phi(e) = 1$
  - d is the inverse of 3 mod $\Phi(e)$
  - encrypt x with $c = x^3$ mod e
  - decrypt c with $x = c^d$ mod e
  - we have $x = x^{3d}$ mod e
- RSA is considered secure because the only known way to find d from e is to ***factor*** e into p and q, a problem believed to be computationally hard
- The RSA patent is to expire in September 2000

Digital Signatures

- Alice sends a message to Bob encrypting it with Bob’s public key.
- Bob decrypts the message using his private key.
- How can Bob determine that the message received was indeed sent by Alice? After all, Trudy also knows Bob’s public key.
- Alice can provide a ***digital signature*** for the message: $s = x^{d} \bmod e$
- If Bob receives both x and s, he computes
  - $y = s^{3} \bmod e = x^{d3} \bmod e = x$
- Thus, if y = x, Bob knows that Alice indeed sent x, since she is the only person who can compute s from x
- Also, Alice cannot cheat and deny to have sent message x (***nonrepudiation***)
- Using digital signatures, Alice and Bob can authenticate each other and prevent Trudy’s woman-in-the-middle attacks
- Validating a signed message requires knowledge of the other party’s public key

Internet Security

- Recall that validating a signature requires knowledge of the other party’s public key
- How do we know other people’s public keys?
- Certification Authorities (e.g., Verisign) provide ***certificates*** that bind identities to public keys
- A certificate is a pair (id, key) signed by the CA
- A user needs to know only the public key of the CA
- Some secret-key ciphers (triple DES, IDEA, BLOWFISH) are much faster than RSA
- To communicate securely, a ***two-phase protocol*** is adopted:
  - a shared secret key k is established using RSA
  - data is transfered between the parties using a secret-key cipher and the shared key k
- Examples:
  - SSH (secure shell) for secure host login
  - SSL (secure socket layer) for secure Web access (https), which uses an additional certification phase