Cryptography

Chapter 5

- Cryptography
  - The art and science of secret writing, *encrypting*, or hiding of information from all but the intended recipient.
- Encryption:
  - The conversion of data into a form, called a ciphertext, that cannot be easily understood by unauthorized people.
- Decryption
  - The process of converting encrypted data back into its original form, so it can be understood.

Background

- Why Cryptography?
  - The desire to hide information from others extends to ancient times.
    - Writing enabled individuals to share information.
    - Later, people also wanted to hide information.
  - The easiest way was not to teach others how to read and write the language.
  - As that became ineffective, methods of shifting the letters around to make the text unreadable were attempted.

3

How Cryptography Works: Codes Versus Ciphers

- Cryptography makes a message unreadable by transforming the plaintext
  - Plaintext – Original message which will be scrambled into a secret form
  - Cryptosystem – protocol employed for encrypting and decrypting a message
    - Uses a shared algorithm to process the original plaintext
  - Transformation process is called either **enciphering/ encryption / encoding**
  - Outcome is a ciphertext or a codetext
  - **Decoding/ Decryption / Deciphering** - Process of using a key to reveal the original message

4

Cryptanalysis

- Cryptanalysis
  - Cryptanalysis is the process of attempting to break a cryptographic system and return the encrypted message to its original form.
- Two methods of cryptanalysis have been developed using the computer:
  - Differential Cryptanalysis:
    - Compares the input plaintext to the output ciphertext to try and determine the key
  - Linear Cryptanalysis:
    - Also compares input plaintext to the output ciphertext.
    - Puts the plaintext to a simplified cipher to try and deduce the likely key in the full version of the cipher

5

Differential Cryptanalysis

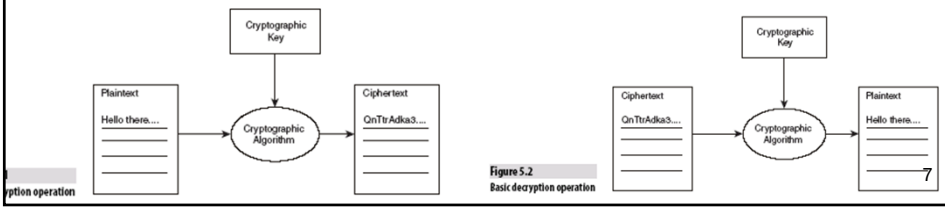
- The method searches for plaintext, ciphertext pairs whose difference is constant, and investigates the differential behavior of the cryptosystem
  - Observe the difference between the two ciphertexts as a function of the difference between the corresponding plaintexts
  - Find the highest probability differential input (called characteristic) which can be traced through several rounds
  - Assign probabilities to the keys and locate the most probable key

6

Algorithms

- Algorithm:
  - A step-by-step problem-solving procedure.
  - A recursive computational procedure for solving a problem in finite steps.

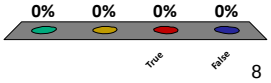
- Cryptographic algorithm
  - A set of mathematical steps for encrypting and decrypting information.
  - Commonly called as encryption algorithm or cipher
  - Used to encrypt a message- Change from plaintext to ciphertext
  - And then decrypt the message- Change from ciphertext back to plaintext



7

Original message which will be scrambled into a secret form is called as Ciphertext

1.
2.
3. True
4. False

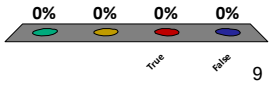


8

An algorithm is defined as a recursive computational procedure for solving a problem in infinite steps

1.
2.
3. True
4. False

Finite steps



9

Steps for Encryption

- The steps for encrypting data can be published because of the design of the systems.
  - They are designed to use a key.
- **The algorithms remain the same**.
  - Every implementation uses a **different key.**
  - This ensures that even if other know the algorithm, they cannot break the security.

While everyone knows how to use a knob to open a door, without the key to unlock the knob, that knowledge is useless

10

Keys

- Keys are special pieces of data used in both the encryption and decryption processes.
- The algorithms stay the same, but a different key is used.
  - This ensures your data is secure even if they know the algorithm.
- The more complex the key, the greater the security of the system.
- **Keyspace** is every possible key value.
  - Key complexity is achieved by giving the key a large number of possible values.
  - This is usually defined in a numeric size of bits
  - 1024 bits, meaning 2^{1024} different keys.
  - When an algorithm lists a certain number of bits as a key, it is defining the keyspace.

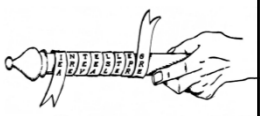
11

Types of Ciphers

- Transposition
- Shift
- Substitution
- Vigenère
- One-time pad

12

Transposition



Transposition Cipher:

Same letters are used but the order is changed

Used in Spartans Cipher

Spartans used a ribbon wrapped around a specific gauge cylinder and then wrote on the ribbon

- **The order of the letters are changed**.
  - Ex. THE UNEXAMINED LIFE IS NOT WORTH LIVING
    - Written vertically over six columns becomes:

```
TX SOV
HAL RI
EMINTN
 IFOHG
UNET
NE  L
EDIWI
```

Then, written horizontally becomes:

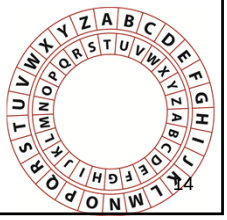
```
TX SOVHAL RIEMINTN IFOHGUNET
NE  LEDIWI
```

Transposition.xlsx

Shift Cipher

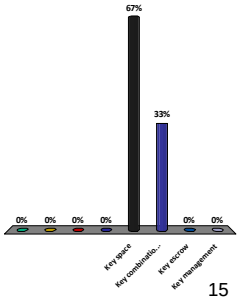
- **Shift Cipher:**
  - The algorithm specifies that you offset the alphabet either to the right (forward) or to the left (backward) for another letter
  - The key specifies how many letters the offset should
  - A classic example - Caesar's cipher.
    - Example: Every letter is rotated 19 positions in the alphabet
  - **Weakness:** The ease with which it could be broken

Class Activity 1-Ques 1



_____ refers to every possible value for a cryptographic key

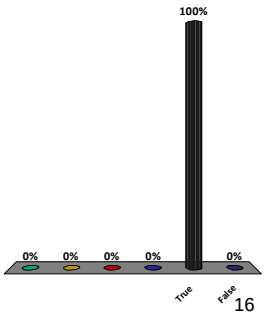
1.
2.
3.
4.
5. Key space
6. Key combinations
7. Key escrow
8. Key management



15

The shift cipher moves an alphabet a set number either to the right or left

1.
2.
3.
4.
5. True
6. False



16

Substitution Cipher

- The weakness of shift ciphers led to substitution ciphers.
  - Works on the principle of substituting a different letter for every letter.
    - For example: *a* becomes *g*, *b* becomes *d*, and so on.
    - The letters are not in order as they are in shift ciphers
    - Permits 26 possible values for every letter in a message.
  - The cipher is more complex than a standard shift cipher.
  - However,
    - Simple analysis of the cipher retrieves the key.
    - **Look for common letters and patterns that would become words.**
    - It may determine which cipher letter corresponds to which plaintext letter. This determines this system's key value.

17

Cryptanalysis of Substitution Cipher

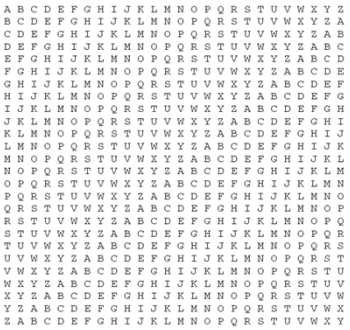
- Digrams and Trigrams
  - Two- and three-letter words.
  - There are only a limited number of one-, two-, and three-letter words.
- Look for patterns and letters that appear in multiple words.

18

Vigenère Cipher

- A polyalphabetic substitution cipher that depends on a password.
  - This is done by setting up a substitution table.

1. The cipher letter determined by matching the **plaintext character's row** with the **password character's column**.
2. Results in a single ciphertext character from where the two meet and the process is then repeated for every letter of the message.
3. Even if someone knows the table (or algorithm), without the key, the message cannot be decrypted
4. If the password is not long enough, the password is repeated until one character of the password is matched with each character of the plaintext.



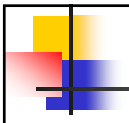
Class Activity 1-Ques 2

19

Vigenère Cipher

- The Vigenère cipher is a much more complex cipher.
- It corrects the issues with more simplistic keys.
- The Vigenère cipher system and similar systems .
  - Makes the algorithms rather simple
  - But the key rather complex, with the best keys comprising very long and very random data

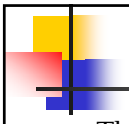
20



One-Time Pad

- Unbreakable
- Dependent on random pad generation
- Requires both parties to have the identical pad and start from the same point in the pad
- Impractical for most common applications
  - Large pads required
  - Difficult to generate truly random numbers
  - Difficult to get the pads to both parties

21



Key Complexity

- The more complex the key, the greater the security of the system.
  - Key complexity is achieved by assigning a large number of possible values to the key.
- The keyspace is the size of every possible key value.

- All encryption ciphers besides a “one-time pad” cipher are susceptible to a brute-force attack—attempting every possible key.

22

Modern Day Encryption

- Public algorithms
- Hashes
- Modern algorithms

23

Public Algorithms

- The **best algorithms are always public algorithms**.
  - They are peer reviewed by other cryptographic and mathematical experts.
  - Publication is important, as any flaws in the system can be revealed by others before actual use of the system.

24

Hashes

- Hashing functions
- Collision attacks
- Common hash algorithms
  - SHA
  - Message Digest

25

Hashing

- Hashing functions are commonly used encryption methods
- A hash is a special mathematical function that performs **one-way encryption**.
- Once the algorithm is processed, there is no way to:
  - Take the ciphertext and retrieve the plaintext that was used to generate it.
  - Generate two different plaintexts that compute to the same hash value.
- Two popular hash algorithms are the Secure Hash Algorithm (SHA) series and Message Digest (MD) hash of varying versions (MD2, MD4, MD5).
- Common Use:
  - Storing computer passwords
  - Ensuring message integrity
    - If the message is edited, hash will no longer match

26

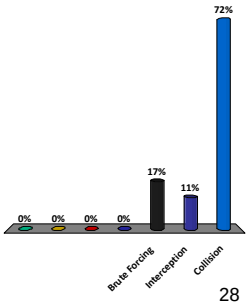
Collision Attack

- A collision attack is used to compromise a hash algorithm.
  - Occurs when an attacker finds two different messages that hash to the same value.
- This attack is very difficult and requires generating a separate algorithm that attempts to find a text that will hash to the same value of a known hash.
- This must occur faster than simply editing characters until you hash to the same value, which is a brute-force type attack.
- **Hash functions that suffers from collisions lose integrity.**
  - **A good hash function should be resistant to collision**
- An attacker that can make two different inputs hash to the same value, can trick people into running malicious code.

27

A good hash function is resistant to ____

1.
2.
3.
4.
5. Brute Forcing
6. Interception
7. Collision



28

SHA (Secure hash algorithm)

- Refers to four hash algorithms published by the National Institute of Standards and Technology (NIST) and the National Security Agency (NSA).
  - SHA-1, SHA-256, SHA-384, SHA-512
- **SHA-1** was one of the more secure hash functions.
  - But it has been found to be vulnerable to a collision attack.
- The longer versions are referred to as **SHA-2.**
  - All have longer hash results, and are more difficult to attack successfully.
  - SHA-256, SHA-384, and SHA-512
- SHA-2 does require more processing power to compute the hash.

29

Block Method

- Most algorithms use block mode to process data to create the hash.
- They break the data into sets of bits (blocks) such as 512.
- If a file were 1400 bits long, it would create three blocks with the third one being padded with zeros.
  - 2x512 is 1024, the third block would be 376 bits of the message and 136 bits of zeros.

30

SHA (Secure hash algorithm)

- Refers to a set of 4 hash algorithms developed by the National Institute of Standards and Technology for secure hashing in the U.S. Digital Signature Standard (DSS).
- Works in block mode, separating the data into words, and then grouping the words into blocks
- Four government approved hash functions:
  - SHA-1 produces a 160 bit message digest for any input < 2^{64} bits
  - SHA-256 produces a 256 bit message digest for any input < 2^{64} bits
  - SHA-384 produces a 384 bit message digest for any input < 2^{64} bits
  - SHA-512 produces a 512 bit message digest for any input < 2^{64} bits

31

Message Digest (MD)

- Message Digest (MD) is the generic version of one of the three algorithms, designed to create a message digest or hash from data input into the algorithm.
  - **MD2**
    - Generates a 128-bit message digest from a message of arbitrary length.
    - Optimized for 8-bit machines
  - **MD4**
    - Generates a 128-bit message digest from a message of arbitrary length.
    - Known vulnerabilities to collision. **Not considered secure and its use should be avoided**
  - **MD5**
    - **Structured with additional security to overcome the problems in MD4.**
    - Very similar to the MD4 algorithm, only slightly slower and more secure.
    - Generates a 128-bit message digest from a message of arbitrary length.
    - Optimized for 32-bit machines
    - **MD5 checksums are commonly used to verify the integrity of files downloaded from the internet.**
    - **Protects against both accidental and intentional tampering**

32

NTLM (NT LAN Manager)

- MD5 has been used as part of the NTLM (NT LAN Manager) challenge/response authentication protocol.
- Successful attacks on the algorithm have occurred.
- MD5 collisions can be computed in about 8 hours on a standard home PC.
  - The ability to have two entirely different Win32 executables with different functionality but the same MD5 hash.
- **This has led to people adopting a strong SHA version instead**.

33

Hashing: Summary

- Hashing functions are very common, and they play an important role in security.
  - Storing passwords
  - Signing messages
  - Maintaining message integrity
- By computing a digest of the message, less data needs to be signed by the more complex asymmetric encryption.

34

Symmetric Encryption

- Symmetric encryption
- Key management
- Trusted platform module
- Popular symmetric encryption algorithms
  - DES, 3DES, AES, CAST, RIVEST, Blowfish, IDEA

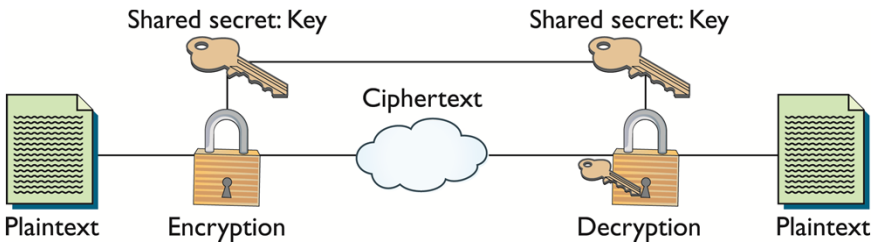
35

Symmetric Encryption

- Symmetric encryption
  - Older and simple method of encrypting information
  - Requires the sender and the receiver to have the same key.
    - All symmetric algorithms are based upon this shared secret principle.
  - Key is called ***shared secret key*** or ***secret key***
- Symmetric encryption involves a cryptographic key, requiring key management.
- The most important lesson - Store and send the key only by known secure means.

36

Symmetric Algorithm



37

Key Management

- Security of the algorithms relies on the key, as such key management is of critical concern.
  - Includes anything having to do with the exchange, storage, safeguarding, and revocation of keys.
  - A key must be current and verified.
  - If you have an old or compromised key, you need a way to check to see that the key has been revoked.

38

Key Length

- The longer the key, the greater the degree of protection
- A common attack against cryptosystems is the brute force attack
  - All possible keys are tried
  - Longer keys create an enormous number of possible combinations, frustrating brute force attacks
  - Formula used to compute the number of combinations is 2^n where n is the key length in bits

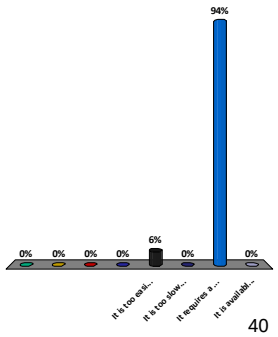
TABLE 5.1 Possible Keys of a Given Length

Key Length	Approximate Number of Possible Keys
56 bits	72,057,594,037,927,936
128 bits	3.40×10^{38}
256 bits	1.16×10^{77}
512 bits	1.34×10^{154}
1,024 bits	1.80×10^{308}
2,048 bits	3.23×10^{616}

39

What is the biggest drawback to symmetric encryption?

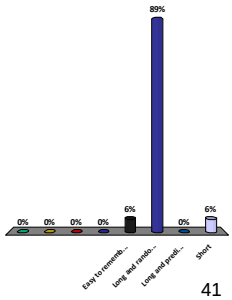
1.
2.
3.
4.
5. It is too easily broken.
6. It is too slow to be easily used on mobile devices
7. It requires a key to be securely shared
8. It is available only on UNIX.



40

What is the best kind of key to have?

1.
2.
3.
4.
5. Easy to remember
6. Long and random
7. Long and predictable
8. Short



41

Popular Symmetric Encryption Algorithms

- DES
- 3DES
- AES
- IDEA
- CAST
- Rivest
- Blowfish
- IDEA

42

DES (Data Encryption Standard)

- DES (Data Encryption Standard)
  - A block cipher, that segments the input data into blocks of 64 bits, using a 56-bit key, and outputs blocks of 64-bits.
  - The same algorithm and key are used for both encryption and decryption.
    - Performs a substitution and permutation (known as round) based on the key 16 times on every 64 bit block.
      - Produces 64 bit block of ciphertext for each plaintext block
    - After the completion of all the 16 rounds and the inverse permutation, the algorithm picks up the next 64 bits and starts all over again.
    - This is carried on until the entire message has been encrypted with DES.
  - While DES has been a common business standard for 20 years, modern computing power has made the key breakable.
  - NIST now certifies Advanced Encryption Standard (AES) to replace DES.

43

DES

- **Some concerns**:
  - **Weak keys** are keys that are less secure than the majority of keys allowed in the keyspace of the algorithm.
  - A 56-bit key is no longer considered strong enough to survive brute force attacks
  - There are also semi-weak keys, where two keys encrypt plaintext to identical ciphertext, meaning that either of the keys will decrypt the ciphertext.
- Any DES with less than 16 rounds could be analyzed more efficiently with chosen plaintext than via a brute-force attack using differential cryptanalysis.

44

Triple DES

- 3DES (Triple DES)
  - Depending on the variant, it uses either two or three keys.
  - It spins through the DES algorithm three times via what's called multiple encryption.
  - Multiple encryption can be performed in different ways.
    - **The simplest method: To stack algorithms on top of each other.**
    - **Another way: Encrypt with one key, decrypt with a second, and then encrypt with a third.**
  - 3DES is stronger than DES but has similar weakness.
  - The longer key length makes it more resistant to brute force attacks
  - A good interim step before the new encryption standard, AES.

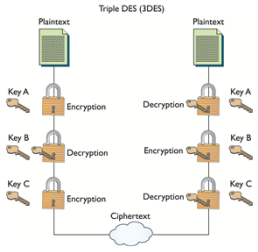
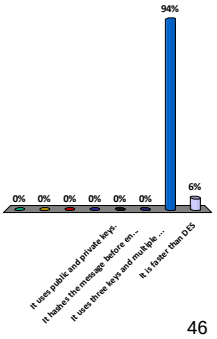


Diagram of 3DES

How is 3DES an improvement over normal DES?

1.
2.
3.
4.
5. It uses public and private keys.
6. It hashes the message before encryption
7. It uses three keys and multiple encryption and/or decryption sets
8. It is faster than DES



Advanced Encryption Standard (AES)

- AES is a block cipher that separates data input into 128-bit blocks.
  - Can also be configured to use blocks of 192 or 256 bits.
- AES can have key sizes of 128, 192, and 256 bits, with the size of the key affecting the number of rounds used in the algorithm.
  - Longer key versions are known as AES-192 and AES-256, respectively.
- No efficient attacks currently exist against AES.

47

CAST

- Designed by Carlisle Adams and Stafford Tavares
- CAST is an encryption algorithm that is similar to DES in its structure.
  - Uses 64-bit block size for 64- and 128-bit keys
  - 128-bit block size for the 256-bit key version
- CAST has undergone thorough analysis, with only minor weaknesses discovered.
- CAST should be placed with other trusted algorithms.

48

RC

- RC is the term for ciphers designed by Ron Rivest, (RC – Rivest Ciphers).
- RC2 – designed as DES replacement, 8 to 1024 bit key size, 64 bit block size, keys below 64 bit are vulnerable.
- RC5 – block cipher, multiple variable elements, RC6 is newer version.
- RC6 – 128 bit block size, keys sizes: 128, 192, 256.
  - Runs well on 32 bit computers
  - Resistant to brute force attacks
  - Should provide adequate security for some time to come
- RC4 – Stream cipher, fast, uses key lengths of 8 to 2048 bits, most vulnerable to possibility of weak keys.

49

Blowfish

- Designed in 1994 by Bruce Schneier.
- Block mode cipher, using 64-bit blocks and a variable key length from 32 to 448 bits.
- Runs well on 32-bit machines.
- Seems to be strong when implemented with the full 16 rounds.

50

International Data Encryption Algorithm (IDEA)

- Released as IDEA in 1992.
- Block mode cipher using 64-bit block size and 128-bit key.
- This algorithm is fairly new.
- Full, eight-round IDEA shows that the most efficient attack would be to brute-force the key.
- Susceptible to weak key vulnerability, but easy to mitigate.

51

Symmetric Encryption Summary

- Symmetric algorithms are important because:
  - They are comparatively fast.
  - Have few computational requirements
- Their main weaknesses:
  - Two geographically distant parties both need to have a key that matches the other key exactly.
  - Simple keys can quickly be brute-forced.
  - Secure key exchange can be an issue.

52

Asymmetric Cryptography

- Is also known as public key cryptography
- Differ from symmetric cryptography because sender and receiver use different keys
- Each user has a pair of keys
  - Public key and private key
    - A private key that is kept secret.
    - A public key that can be sent to anyone
  - Keys are mathematically related
  - Messages encrypted with public key can only be decrypted with private key
  - Public keys are freely distributed so that anyone can use them to encrypt a message
- Some of the popular asymmetric protocols are:
  - RSA, Diffie-Hellman, ECC, and ElGamal

53

Public Key Encryption

- It typically works by using hard math problems.
- A common method relies on the difficulty of factoring large numbers.
- Trapdoor functions are difficult to process without the key but easy to process when you have the key.
  - **Trapdoor functions**- A type of math problem that is difficult unless you know a specific value
- Computers can easily multiply very large primes with hundreds or thousands of digits but cannot easily factor the product.
- They also form the basis for digital signatures.

54

RSA

- Ron Rivest, Adi Shamir, and Leonard Adleman (RSA)
- One of the first public key cryptosystems invented.
  - Published in 1997
  - Used for encryption and digital signatures
  - Uses the product of two very large prime numbers (from 100 to 200 digits) to generate one key for decryption and another for encryption
  - Relies on the fact that it is extremely difficult to factor large prime numbers
- While a simple algorithm, it has withstood the test of more than 20 years of analysis.
- Does not replace symmetric encryption because RSA is 100 times slower than DES!
- Asymmetric encryption is used to exchange symmetric keys.

55

Diffie-Hellman

- Created in 1976 by Whitfield Diffie and Martin Hellman
- The protocol is one of the most common encryption protocols in use today.
- Used for:
  - Electronic key exchange method of the Secure Sockets Layer (SSL) protocol
  - TLS, SSH, and IPsec protocols
  - The protocol, like RSA, uses large prime numbers to work.
  - Enables the sharing of a secret key between two people who have not contacted each other before.
  - It is very effective because it protects a temporary, automatically generated secret key that is only good for a single communication session.
- Diffie-Hellman is still in wide use.

56

ElGamal

- Taher Elgamal designed the system in the early 1980s.
- It can be used for both encryption and digital signatures.
- This system was never patented and is free for use.
- It is used as the U.S. government standard for digital signatures.
- ElGamal has been in use for some time.
- It is used primarily for digital signatures.

57

Elliptic Curve Cryptography (ECC)

- Works on the basis of elliptic curves
- Elliptic curves are defined by this equation:

  $y^2 = x^3 + ax^2 + b$
- Elliptic curves work because they have a special property—you can add two points on the curve together and get a third point on the curve.
- Security in question due to lack of analysis.
- Unless there is a breakthrough in math, ECC will continue to be a good algorithm.

58

Symmetric Versus Asymmetric Cryptosystems

- Choice between symmetric and asymmetric cryptosystems involves the number of keys that must be generated
  - Symmetric cryptosystems don't scale well
  - Asymmetric cryptosystems are slower than symmetric ones
  - Symmetric cryptosystems are excellent for securing the ends of a communication circuit such as a Virtual Private Network
  - Asymmetric cryptosystems are more practical when there are a large number of users

59

Important

TABLE 5.2 Comparison of Symmetric and Asymmetric Cryptosystems

Symmetric Cryptosystems	Asymmetric Cryptosystems
Provide confidentiality among all participants who share the same secret key	Provide confidentiality between individual users of a cryptosystem
Provide integrity against modification by individuals who do not possess the secret key	Provide integrity against modification by anyone other than the sender of the message
Provide for authentication between two individuals when they are the only ones who possess the secret key	Provide for authentication of any individual user of the cryptosystem
Do not provide for nonrepudiation	Provide for nonrepudiation
Require shorter keys than asymmetric algorithms to achieve the same level of security	Require longer keys than symmetric algorithms to achieve the same level of security
Operate faster than asymmetric algorithms	Operate slower than symmetric algorithms
Are not easily scalable	Scale well to environments with large numbers of users
Do not facilitate the use of digital certificates	Lend themselves well to digital certificate hierarchies
Make the exchange of cryptographic keys difficult (often requiring offline exchange)	Allow for the exchange of public keys over otherwise insecure transmission media

60

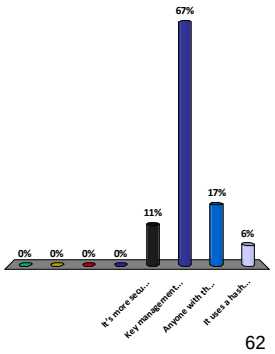
Asymmetric Encryption Summary

- Creates the possibility of digital signatures and corrects the main weakness of symmetric cryptography.
  - Key management is part of the algorithm
- Ability to send messages securely without senders and receivers having had prior contact.
- Digital signatures enable faster and more efficient exchange of all kinds of documents.
- With strong algorithms and good key lengths, security can be assured.

61

What makes asymmetric encryption better than symmetric encryption?

1.
2.
3.
4.
5. It's more secure
6. Key management is part of the algorithm.
7. Anyone with the public key can decrypt the data
8. It uses a hash.



62

Steganography

- Offshoot of cryptography technology
  - Greek word *steganos*, meaning covered
  - Invisible ink, or tattoo, on head under hair
  - Commonly hiding text message in picture file
- Difficult to detect.
- Images do not attract attention.
- Message can also be encrypted.
- Tools to detect steganography:
  - Stegdetect, StegSecret, SegSpy, and SARC tools.
- Steganography can be a nightmare for protecting an organization's sensitive information.

63

Cryptography Algorithm Use

- Confidentiality
- Integrity
- Nonrepudiation
- Aunthentication
- Key escrow
- Digital signatures
- Digital rights management
- Cryptographic applications

64

Goals of Cryptography

Important

- Four primary goals
- **Confidentiality** is most commonly addressed goal
  - The meaning of a message is concealed by encoding it
  - The sender encrypts the message using a cryptographic key
  - The recipient decrypts the message using a cryptographic key that may or may not be the same as the one used by the sender
- **Integrity**
  - When a message is sent, both the sender and recipient need to know that the message was not altered in transmission.
  - This is especially important for legal contracts.
  - The ability to independently make sure that a document has not been tampered with is very important to commerce.
  - The hash functions compute the message digests, and this guarantees the integrity of the message

65

Goals of Cryptography (continued)

- **Nonrepudiation**
  - The message sender cannot later deny that they sent the message.
  - This is important in electronic exchanges of data, especially when you are unable to meet face-to-face.
  - It is based upon public key cryptography and the principle of only you knowing your private key.
  - Nonrepudiation is tied to asymmetric cryptography and cannot be implemented with symmetric algorithms. **Why??**
- **Authentication**
- Authentication lets you prove you are who you say you are.
- Asymmetric encryption is better suited than symmetric encryption to prove one's identity
- Authentication can be accomplished in a multitude of ways: Token, digital certificates
- When you log into a secure web site, one-way authentication occurs.
  - Accomplished using digital certificates
  - Kerberos is a common cryptographic authentication system

66