Module 2: Access Control

Prof Aftab Ahmad

aftab@acm.org

Outline

- Access Control Models
  - (MAC, DAC, RBAC). (CSF.T9)
- Confidentiality, Integrity, Availability, Access, Authentication, Authorization, Non-Repudiation, Privacy. (CSF.T10)
- Security Mechanisms (e.g. Identification/ Authentication, Audit). (CSF.T13)
- Two-factor authentication
- People and security (Social engineering). (ISC.T14)

Access control purpose and concepts

NISTIR7316 Assessment of Access Control Systems

- Access control is concerned with determining the allowed activities of legitimate users
  - Mediating every attempt by a user to access a resource in the system
    - Operating systems use access control to protect files and directories
    - Database management systems DBMS apply access control to regulate access to tables
    - Most commercially available application systems implement access control
- Object – is the entity, service, information, sw, hw to be accessed
- Subject – the entity (person, process, device) that need the object
- Operation – is the act(s) of facilitating the use of object after access
  - Reading, writing, editing files in OS, getting cash in an ATM
- Privilege/permission – the allowed object-operation combinations for a subject in a given role
  - Instructor access to student records and making changes, such as, class enrollment

Access control concepts

- Access Control List (ACL) – a list of permitted and/or denied subjects associated with each object
  - Firewall is an example of an ACL
- Access Control Matrix – a table listing of entries
  - The triplet <subject, permission, object> is typical
- Separation of duties (SOD) – an access principle of require multiple subjects to complete a job
  - A person who prepares a check can't authorize its use
  - SOD can be *static* or *dynamic*
    - In dynamic SOD, role is assigned at the time of a particular access
- Access control policy – high level statement(s) for access control for all

Access Control Models

- A formal presentation of the security policy enforced by the system
  - It is a translation of organization policy in clear and technical terms
- Discretionary Access Control (DAC) leaves access in control of the owner
  - ACLs are basically DAC
  - A patient should determine who can access his medical records
  - It has several drawbacks
    - The information owner makes his own ‘policy’, no organization involved
    - The privacy of information is breached easily after it is shared – no way to stop
    - After a user shares information, no restrictions on what can be done with it
  - If Alice shares her file with Bob, and Bob corrupts it with a virus, DAC can’t prove who corrupted it; it may spread in the organization network
    - It will be Alice to be blamed

Non-discretionary Access Control (NDAC)

- This type of access rules can only be changed through administrators
  - The owner is not empowered to control access
  - It can be of many types, such as Mandatory Access Control (MAC), Role-based Access Control (RBAC), Separation of Duties in any form
- MAC is enforced by a central authority, owner can’t make changes
  - In Military, if you have clearance to see classified data, you can’t make data available to whoever you like
  - Bell LaPadula’s confidentiality model and Biba’s integrity models are MAC
- RBAC assigns access rights based on job roles (students, faculty, etc.)
  - An organization documents access rights for job positions
    - RBAC allows the implementation of *least privilege*
      - In least privilege access is provided to the minimum objects required to do the job
    - If a subject (person’s) role changes, s/he can easily be assigned new privileges

Confidentiality

- Confidentiality pertains to non-disclosure of information
- Access control is a way of keeping the object confidential
  - When object is data and it is on the Internet, there is no access control
  - Even the data in a computer can be accessed by adversary through hacking
- Confidentiality measures hide *information* even if *data* is not hidden
  - Encrypting data is the most common way of providing confidentiality
  - Sender uses a *cipher/encryption algorithm* to transform data
  - The recipient uses *decipher/decryption* to recover information
- Encryption algorithms are known to adversary
  - Secret cipher algorithm are less tested and less trust-worthy

Confidentiality - II

- A cipher algorithm uses a string of bits called *key* that is not known to adversary
  - Secrecy and security of key determines the power of encryption algorithms
- In stream cipher, the input bit stream is encrypted bit-by-bit
  - A *key stream* is generated from the key and each bit of key stream is logically combined with the input bit (using exclusive OR operation)
  - To decipher, the same key stream is generated and same logical operation used to recover information
  - Examples of stream cipher: RC4, A5/1, Trivium
- WiFi (1999) used RC4, GSM used Trivium, Bluetooth uses AES-CCM

Confidentiality - III

- Block cipher takes a block of input and generates a block of output
  - More operations can be done on blocks than individual bits
    - Two popular operations are done for *Confusion* and *Diffusion*
    - Purpose of ***confusion*** is to spread the impact of each key to many output bits
    - Purpose of ***diffusion*** is to spread the impact of each input to many output bits
  - Examples of block cipher are advanced encryption standard (AES), 3DES, IDEAS, blowfish
- Wi-Fi uses AES, IPSec uses 3DES
- Block ciphers employ *modes of operation* to provide confidentiality to a large stream of input by changing key from block to block
  - The original key acts only as a seed to generate stream of keys

Confidentiality - IV

- *Plaintext* is the data that serves as the input to a cipher algorithm
- *Ciphertext* is the output of the cipher algorithm
- *Secret key algorithms* use the same key for encrypting/decrypting
  - If N devices are using secret key cipher, they either share key or use different key for each pair of device – sharing key on the Internet is impossible
    - If N devices use different key for each device pair, they need N(N-1)/2 keys
  - They are also called *symmetric key*, *single key*, or *shared key* algorithms
  - Secret key ciphers is fast, simple, requiring short keys
- *Asymmetric key algorithms* use different keys for encrypting and decrypting
  - If N devices are using asymmetric key cipher, each of them shares one key with any one (called the *public key*) and keeps one key secret (called the *private key*)
    - If private key is used for encryption only public key can do decryption and vice versa
  - Asymmetric key ciphers are slow, complex, requiring long keys
  - They are also called two-key algorithms, or private/public key algorithms

Confidentiality - V

- Contemporary internet communications uses a combination of symmetric and asymmetric keys – called hybrid encryption
- Two-key algorithms allow two parties to share material to generate the same key on both sides on communication without transmitting it
- After generating the same keys on both sides (send/receive)
  - Symmetric cipher is used for encryption/decryption
  - The key is usually *temporal* – meaning it has an age after it is replaced with a new one
- Public keys are very large – and need to be verified
  - Companies called certificate authorities (CA)s provide certificates for public keys
- If Private key is used to encrypt then a certified public key proves that the corresponding private key was used for encryption – this is *digital signature*

Confidentiality - VI

- Two-key cipher is possible due to base-*q* algebra
- A base-*q* number system has *q* digits in it
  - For *q = 2, it's called binary, q = 8, it's called Octal, q = 10, it's called Decimal,*
  - *q = 16, it's called Hexadecimal, q = 2^n, it will be simply a base- 2^n system*
- The modulo-*q* algebra has all operations in base-*q* numbers
  - In module-2 addition: 101 $\oplus$ 101 = 000 just like exclusive-OR operation
- Example: Base-*7* system has digits: 0, 1, 2, 3, 4, 5, 6
- *Modulo-7 addition 6 $\oplus$ 1 = 0, 6 $\oplus$ 2 = 1, 6 $\oplus$ 3 = 2, 6 $\oplus$ 4 = 3, 6 $\oplus$ 6 = 5,*
  - *Modulo-7 product 6 $\otimes$ 1 = 6, 6 $\otimes$ 2 = 5, 4 $\otimes$ 3 = 5, 6 $\otimes$ 4 = 3, 2 $\otimes$ 4 = 1*
- *The Modulo answer is in fact the remainder of the decimal operation divided by the base – In programming, it's also called Modulus*
  - *In the above examples: 6 $\otimes$ 2 = Modulus (12 / 7) = 5*

Confidentiality - VI

- *Multiplicative inverse* of a number *k* in base-q system is another number, say $k^{/}$ such that $k \otimes k^{/} = 1$
  - For example in base-7 system $3 \otimes 5 = 1$ (obtained as remainder of 15/7)
- Now if we encrypt a number, say 6, by multiplying it with 3, we get 4
  - If 4 is our encrypted message, then $4 \otimes 5 = 6$ = decrypted message
  - The catch is that the recipient has to know the base and the multiplicative inverse to decrypt
- By making the base extremely large (like 2^{128} or higher), it becomes impossible calculate one number given the other
  - Of course, in practice we have more complex systems, such as in RSA algorithm
- You have learned this all in CSCI-360 – it was just a broad-brushed intro

Windows OS - Confidentiality

- If an Operating System is attacked it can still protect data stored
- Microsoft Windows uses Trusted Platform Module (TPM)
  - TPM is hardware-based encryption standard from Trusted Computing Group – (TCG) an industry alliance
    - TPM 2.0 is the latest version replacing TPM 1.2
    - It allows a lot of flexibility in choosing cipher suite and can be adopted internationally
    - TPM 1.2 depended on SHA-1 and RSA – these algorithms are losing way to new ones
- Windows uses *bitlocker* for hard drive encryption
  - Starting with Windows 10, bitlocker is on out of the box
- TPM generates and protects keys without human intervention
  - It also checks OS modules against tampering

Windows OS – Secure boot

- Windows 10 supports *secure boot* using Unified Extensible Firmware Interface (UEFI) – the firmware that has replaced earlier Basic Input Output System (BIOS)
- UEFI is an outgrowth of an Intel firmware *Extensible Firmware Interface (EFI)*
- Intel processors 6th generation and later have the ability to protect code in the main memory
  - This system is called Software Guard Extensions (SGX)
  - SGX allows applications to create *memory enclaves*
    - Code in the memory enclave is encrypted
- Using bitlocker and SGX information is protected outside the CPU
  - In memory, on hard disk, and even removable media (such as USB drive)

Integrity

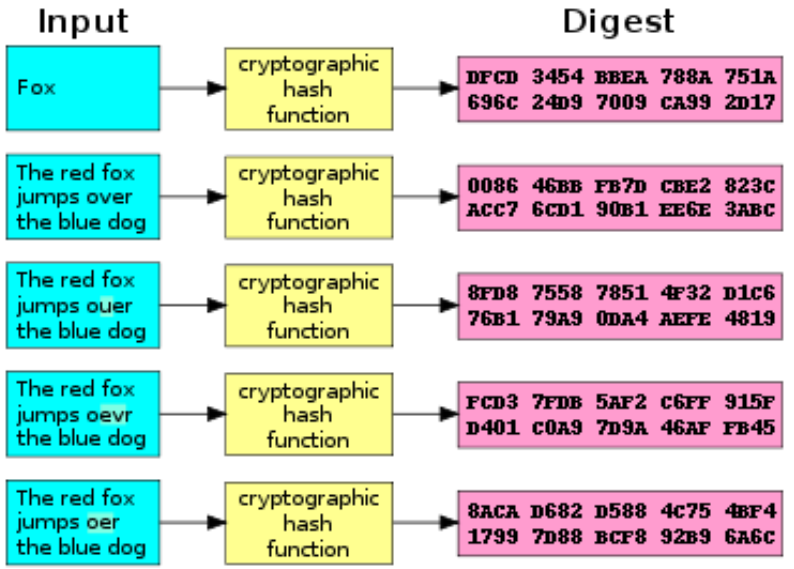
- *Data integrity* relates to embedding trust in the accuracy of data
  - Integrity breach does not require information disclosure to adversary
  - Any tampering of data, even encrypted is an attack on integrity
- In Wi-Fi (1999 ed.) linear coding called Integrity Check Vector (ICV) was used
  - Linear codes can easily be tampered without detection
- Software downloads typically use the MD5 or SHA-1 hash of the downloaded data
  - Such hashes (although, irreversible) can be regenerated for the altered data
- An encrypted hash with a strong hash algorithm is the new trend
  - Bitcoin uses SHA-512, Wi-Fi uses AES, TMP 2.0 recommends SHA-256
- Data integrity ensures that the received data is the same as was sent

Integrity - II

- In secure Hash implementation, the sender
  - Sends data (encrypted or not) plus its encrypted Hash using shared key
  - The recipient receives data (encrypted or not) plus encrypted hash
  - Recipient generates own secure Hash and compares it with the received one
    - If the two are same, recipient assumes there is no change in the data
    - If the two are different, recipient assumes there is tampering of data
- Hash algorithms map data to a fixed length value – like generating the index of an array – see diagram next slide
  - A good hash algorithm generates a different index for different input
    - Same output (index) for different inputs is called a *collision*
  - It should be impossible to regenerate input from the knowledge of hash
    - This is the *irreversibility* property of a good hash function

Integrity - III

- In this figure the hash output is called a *Digest.*
  - It is variously called as message digest, finger print, or hash (or secure hash)
- It is very easy to tamper with wireless data
  - All adversary needs is to jam the wireless channel
  - This would be an attack on *availability*



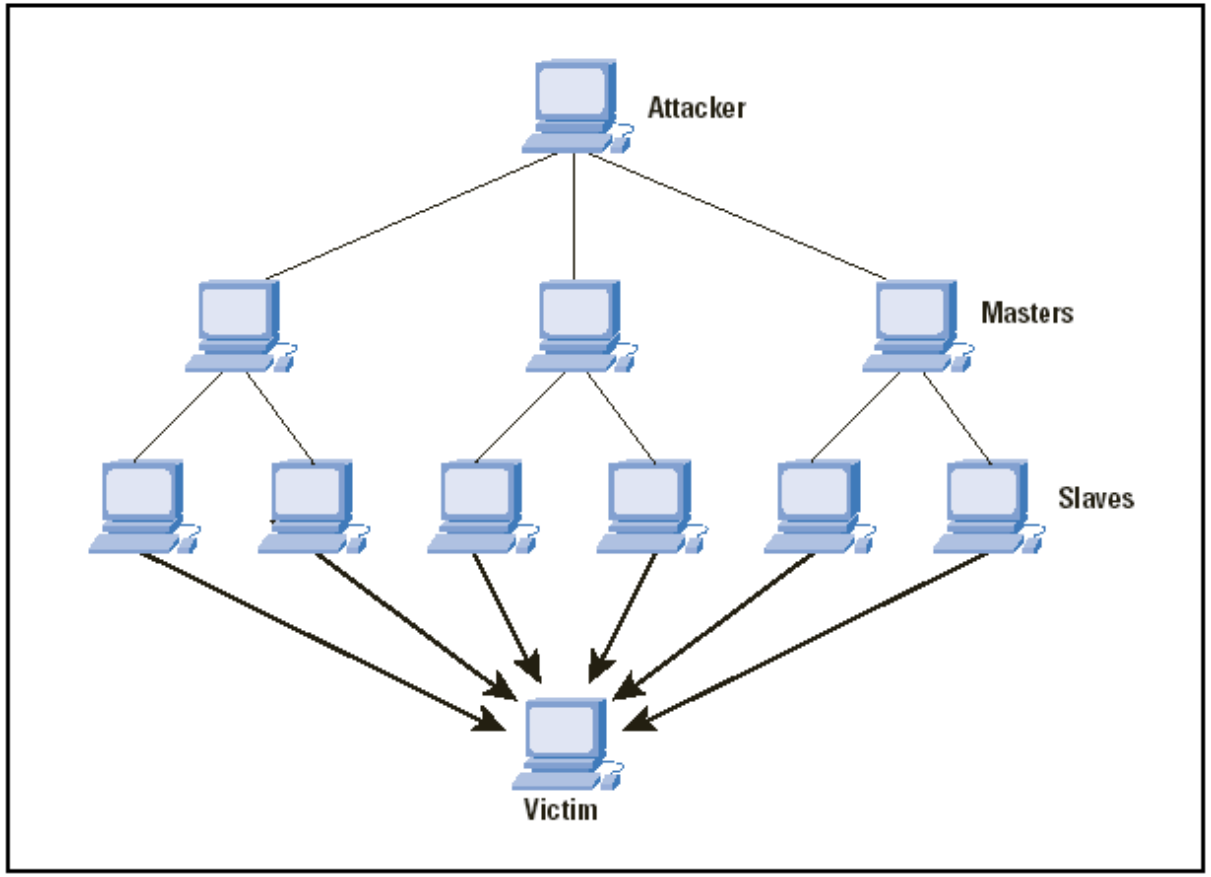
Availability

- Access control and anti-jamming are measures against attacks on availability
  - Availability service ensures that the right party can get (access to) information
- Availability relates to not just information
  - Device, software, process, protocol, personnel – all should be available
    - In a timely and efficient manner (timely = when needed and efficient = without undue delay)
- A large number of illegal connection requests sent to a server can block the server for legitimate parties
- A very large file being sent on a network can block routers
  - It can slow down the network
- Denial of Service (DoS) are attacks on availability
  - In distributed DoS (DDoS) attack a large number of requests are sent from controlled computers (zombies) by a program (bot)

Availability (DDoS) attack on Victim

- Attacks on availability are the hardest to deal with
  - Microsoft recommends ten best practices for its Windows users
    - https://msdn.microsoft.com/en-us/library/cc750213.aspx
- Having a live backup system and data can protect business from shutting down
- Education of users against phishing and ransomware is essential

Figure 4: A DDoS Attack

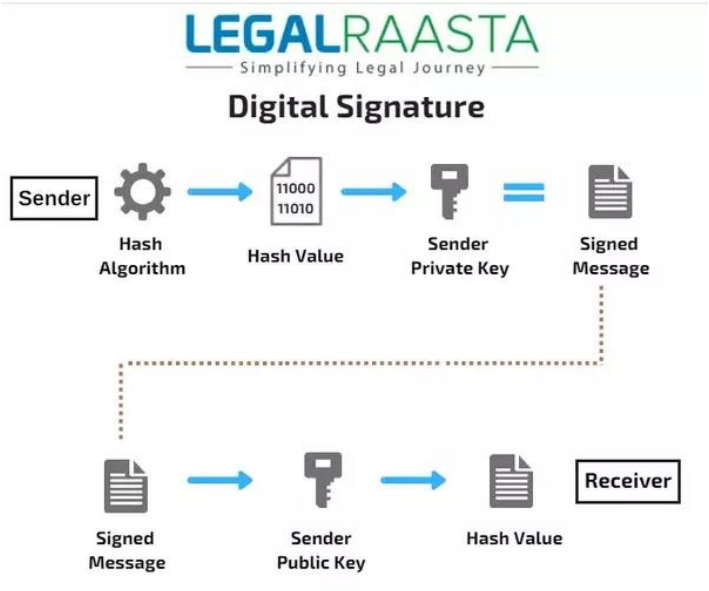


Digital Signatures (DS)

- *Digital Signatures* is an application of the asymmetric cryptography
- Encryption is done using private key and decryption using public key
  - Caveat: If data can be decrypted by any one having the public key, there is no confidentiality
  - Solution: DS is provided as an add-on and not as a type of encryption
  - The data is encrypted using symmetric key
  - Its Hash is encrypted using private key
  - Any one having public key can verify the sender (digital signature)
  - Only the recipient who has symmetric key can decrypt the data (see figure)
- Digital Signature Algorithm (DSA) is a NIST standards since 1991
  - It uses RSA algorithm for asymmetric key encryption of the Hash of data
  - The standard that uses DSA is called Digital Signatures Standard (DSS)

Digital Signatures (DS) - II

- Elliptic Curve Digital Signature Algorithm (ECDSA) is replacing RSA based DSS
  - Elliptic curve is a cipher mechanism for asymmetric ciphers
    - It is replacing the Rivest-Shamir-Adleman (RSA) for public key cryptography
- In general, a DS provides a security service called *non-repudiation*
  - The sender using DS can't decline that s/he sent the signed data



Identity and Authentication

- Identity (ID) is a claim and authentication is the verification of the claim
  - If you claim your email address is aftab@acm.org, that is identity
  - If you can send receive emails from aftab@acm.org, that email (identity) is authenticated
- Some organizations want their employees' ID not to be shown on computer screen – or partially shown
  - This adds a layer of protection, especially in social engineering scenarios
- Authentication is achieved by
  - Something (only) you know – such as a password
  - Something (only) you have – such as a school issued ID or driver's license
  - Something (only) you are – such as photo, finger prints, etc.

Identity and Authentication - II

- In Operating Systems, *identity* is automated by storing user IDs in a file
  - This file is searched to confirm (not authentications)
  - Secure passwords hashes are stored as well
    - No need to store passwords – it would be a bad practice to secure passwords
- When a user enters password, its hash is generated, encrypted and compared with the stored hash
- A second medium can be added to have *two-factor authentication*
  - Second medium can be a phone number, or an email to receive a code

Identity and Authentication - III

- Operating systems can be used to access other computer / network services
  - A separate or single authentication can be employed for each new service
  - *Single Sign On (SSO)* is authentication that can be used to access multiple services
    - Windows implements SSO using TPM
    - iOS provides SSO so that multiple Apple devices can be accessed with one authentication
- Can you think of scenarios in which 2-factor authentication is broken?
  - Breach of authentication can lead to *Identity Theft*
- MIT's Kerberos has been a popular SSO for network services
  - It uses a centralized key distribution center (KDC) to grant a ticket-granting-ticket (TGT), that can be used to get access ticket to individual services

Identity and Authentication - IV

- Windows, Linux, iOS, Android all provide two-factor authentication
  - It is implemented by adding ‘something you have’ to something you know
  - Passwords can be stolen or bypassed through phishing attacks
    - Something only you have will keep the phishing attacker away
- Android and iOS use FreeOTP
  - OTP (one time pad) is the only theoretically secure encryption
  - FreeOTP uses *authentication tokens* using open standards
    - Token can be any combination of characters
    - It can be generated with a QR code generator
    - Linux Red Hat maintains FreeOTP
- Social Engineering is the biggest threat to authentication

Privacy

- Privacy still the least understood measure against attacks
- We may want privacy from loved ones too
  - You don’t want to disclose purchase of a gift to a loved one who has your email password
    - Then you get an email from store thanking you for buying it
- It can be about information shared willfully
  - You rent books for reading and a customer prints and sells your books
- Privacy relates to confidentiality but is not the same thing
  - Confidentiality does not time factor with sharing
- Operating Systems provide file/process ownership, permissions
  - File encryption with individual password can provide data privacy

Privacy – Personally Identifiable Information

- Many organizations share personal information for public good
  - Hospitals sharing disease and cure data by removing individual names
  - Census data shared for research on demographics
  - Employees benefits, such as how many getting what type of insurance
- It is proven that data can be combined to identify the person to who it points to – called PII or *personally identifiable information*
- Many measure used for anonymizing the person
  - *k-anonymity* requires that any person's information should be represented in general groups of data – such as age group 20-30 instead of exact age
  - *L-diversity* requires that for a given value, there should be at least *L* people
    - *Such as, there should not be one person who is in 20-30 age group, at least L people*
  - *t-closeness* requires that anonymization (using *k-anonymity* and *L-diversity*) should not change the stats of the distribution of people more than a small number *t* from the general population
    - Same percentage of people in the group have diabetes as in society

Auditing

- Auditing and logging are protections against insider threats and intrusions
  - If least privilege principle is implemented, auditing provides a check for it
  - If all login activities are logged and audited forensics proof can be provided against an insider attacker
- Internal audits can be compromised if auditing party is guilty
  - External auditing a must for large organizations, sometime if not frequently
- Operating systems provide layers of auditing capabilities
  - Windows has basic and advanced auditing settings

Auditing - II

- **In Microsoft Windows 10**
- **Security Settings\Advanced Audit Policy Configuration** can help your organization audit compliance with important business-related and security-related rules by tracking precisely defined activities, such as:
  - A group administrator has modified settings or data on servers that contain finance information.
  - An employee within a defined group has accessed an important file.
  - The correct system access control list (SACL) is applied to every file and folder or registry key on a computer or file share as a verifiable safeguard against undetected access.

Social Engineering

- You get an email from your niece with the subject ‘Family photos’
  - You open the email and see a link to photos
  - You click on the link and find no photos.
  - You check your login file, it says you just logged in from another country
  - You are under phishing attack!
- You get a phone call from self-described IRS about your refund
  - S/he says you have an additional $2000 refund that can only be deposited to your bank account
    - To do that s/he needs all your account information including passwords
  - Next day you find your money has been transferred to an unknown account
    - That account has been closed
- That’s why you are asked to authenticate your device for logging in using phone or email
- User awareness and training is the best measure against social engineering
  - It should be done like a fire drill – something not done

More for reading

- Bitlocker information can be found at https://docs.microsoft.com/en-us/windows/security/information-protection/bitlocker/bitlocker-device-encryption-overview-windows-10
- TPM as used in Windows is explained at https://docs.microsoft.com/en-us/windows/security/hardware-protection/tpm/trusted-platform-module-top-node
- In the book
  - Chapter 3 is on encryption, Chapter 4 had authentication
- Also check the NIST document included with these slides