CSC12001

Data Security in Information Systems

C01 – Overview of Information Security

Dr. Phạm Thị Bạch Huệ
MSc. Lương Vĩ Minh

Information System Department – Faculty of Information Technology
Univeristy of Science, VNU-HCM

- Information Security: basic concepts
- Threats and Goals of Security
- Security Policies and Security Mechanisms
- Privacy: basic concepts and comparison with security

Information Protection - Why?

- Information are an important strategic and operational asset for any organization

- Damages and misuses of information affect not only a single user or an application; they may have disastrous consequences on the entire organization

- Additionally, the advent of the Internet as well as networking capabilities has made the access to information much easier

Confidentiality

Information Security

Integrity

Availability
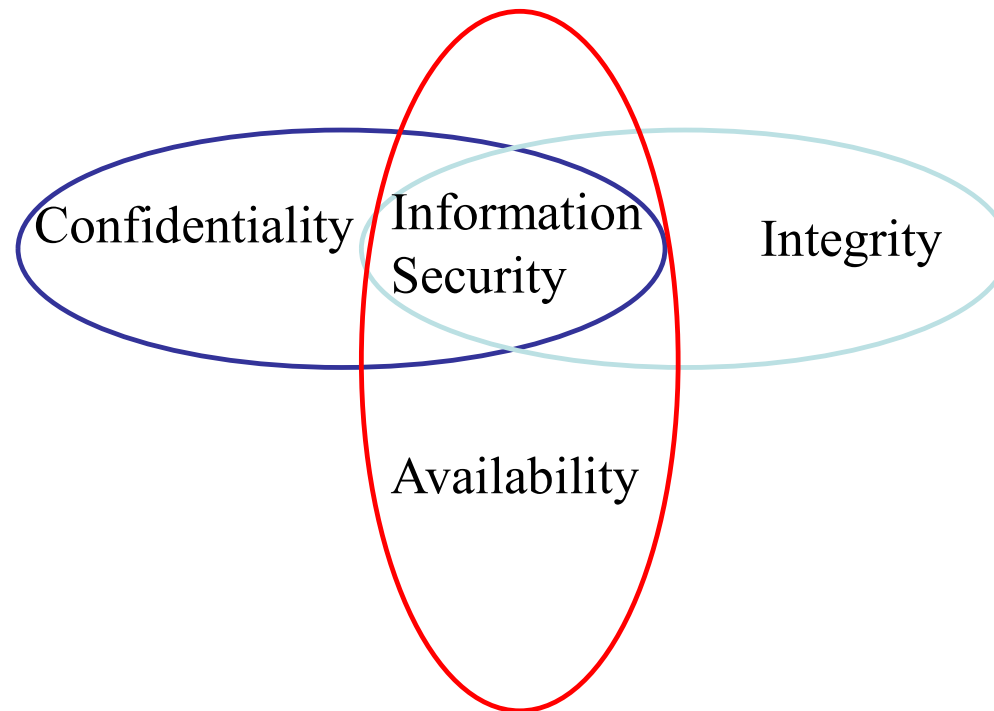
- *Confidentiality* - it refers to information protection from unauthorized read operations
  - the term *privacy* is often used when data to be protected refer to individuals
- *Integrity* - it refers to information protection from modifications; it involves several goals:
  - Assuring the integrity of information with respect to the original information (relevant especially in web environment) – often referred to as *authenticity*
  - Protecting information from unauthorized modifications
  - Protecting information from incorrect modifications – referred to as *semantic integrity*
- *Availability* - it ensures that access to information is not denied to authorized subjects

- Consider a payroll database in a corporation, it must be ensured that:
  - salaries of individual employees are not disclosed to arbitrary users of the database
  - salaries are modified by only those individuals that are properly authorized

- Our enrollment system - Moodle:
  - Only academic staff have authorizations to Insert/ Delete/ Update a course. Students can read (select) the data only.
  - Lecture of a course has authorization to update the marks of (that course) of students who enrolled the course.
  - Students can insert/ delete an enrollment.

- *Information Quality* – it is not considered traditionally as part of information security but it is very relevant

- *Completeness* – it refers to ensure that subjects receive all information they are entitled to access, according to the stated security policies

- Disclosure (Sự phơi bày)
  - Snooping (rình mò), Trojan Horses
- Deception (Lừa gạt)
  - Modification, spoofing (lừa đảo), repudiation (thoái thác) of origin, denial (phủ nhận) of receipt
- Disruption (Sự phá vỡ)
  - Modification
- Usurpation (Sự chiếm đoạt)
 - Modification, spoofing, delay, denial of service

- Prevention
  - Prevent attackers from violating security policy
- Detection
  - Detect attackers' violation of security policy
- Recovery
  - Stop attack, assess and repair damage
  - Continue to function correctly even if attack succeeds

Information Security – How?

- Information must be protected at various levels:
  - The operating system
  - The network
  - The data management system
  - Physical protection is also important

- It consists of:
    - **first** defining a *security policy*
    - **then** choosing some *mechanism* to enforce the policy
    - **finally** providing *assurance* that both the mechanism and the policy are sound
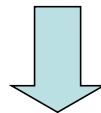
SECURITY LIFE-CYCLE

- Policy says what is, and is not, allowed
  - This defines "security" for the information
- Mechanisms enforce policies
- Composition of policies
  - If policies conflict, discrepancies may create security vulnerabilities

- *User authentication* - to verify the identity of subjects wishing to access the information
- Confidentiality is enforced by the access control mechanism

- Integrity is enforced by the access control mechanism and by the semantic integrity constraints

- Availability is enforced by the recovery mechanism and by detection techniques for DoS attacks – an example of which is query flood

- Access control
  - the matrix model and the safety problem
  - discretionary access control
  - mandatory access control
  - role-based and task-based access control
  - context-based access control

- *Information authentication* - to ensure information authenticity - it is supported by signature mechanisms

- *Encryption* - to protect information when being transmitted across systems and when being stored on secondary storage

- *Intrusion detection* – to protect against impersonation (sự mạo danh) of legitimate (hợp pháp) users and also against insider threats

Data vs Information

- Computer security is about controlling access to information and resources
- Controlling access to information can sometimes be quite elusive and it is often replaced by the more straightforward goal of controlling access to data
- The distinction between data and information is subtle (khó thấy) but it is also the root of some of the more difficult problems in computer security
- *Data* represents information. *Information* is the (subjective) interpretation of data

- Protecting information means to protect not only the data directly representing the information
- Information must be protected also against transmissions through:
  - Covert channels
  - Inference
    - It is typical of database systems
    - It refers to the derivation of sensitive information from non-sensitive data

Name	Sex	Programme	Units	Grade Ave
Alma	F	MBA	8	63
Bill	M	CS	15	58
Carol	F	CS	16	70
Don	M	MIS	22	75
Errol	M	CS	8	66
Flora	F	MIS	16	81
Gala	F	MBA	23	68
Homer	M	CS	7	50
Igor	M	MIS	21	70

- Assume that there is a policy stating that the average grade of a single student cannot be disclosed; however statistical summaries can be disclosed

- Suppose that an attacker knows that Carol is a female CS student

- By combining the results of the following legitimate queries:

  - Q1: SELECT Count (*) FROM Students WHERE Sex ='F' AND Programme = 'CS'

  - Q2: SELECT Avg (Grade Ave) FROM Students WHERE Sex ='F' AND Programme = 'CS'

  The attacker learns from Q1 that there is only one female student so the value 70 returned by Q2 is precisely her average grade

- Specification
  - Requirements analysis
  - Statement of desired functionality
- Design
  - How system will meet specification
- Implementation
  - Programs/systems that carry out design

- Policies define security, and mechanisms enforce security
  - Confidentiality
  - Integrity
  - Availability
- Importance of assurance

- Privacy is an important issue today
  - Individuals feel
    - Uncomfortable: ownership of information
    - Unsafe: information can be misused
    - (e.g., identity thefts)
  - Enterprises need to
    - Keep their customers feel safe
    - Maintain good reputations
    - Protect themselves from any legal dispute
    - Obey legal regulations

- **Privacy** is the ability of a person to control the availability of information about and exposure of him- or herself. It is related to being able to function in society anonymously (including pseudonymous or blind credential identification).

- **Types of privacy** giving raise to special concerns:
  - Political privacy
  - Consumer privacy
  - Medical privacy
  - *Information technology end-user privacy; also called data privacy*
  - Private property

Data Privacy

- Data Privacy problems exist *wherever uniquely identifiable data relating to a person or persons are collected and stored, in digital form or otherwise.* Improper or non-existent disclosure control can be the root cause for privacy issues.

- The most common sources of data that are affected by data privacy issues are:
  - Health information
  - Criminal justice
  - Financial information
  - Genetic information

Data Privacy

- The challenge in data privacy is to share data while protecting the personally identifiable information.
  - Consider the example of health data which are collected from hospitals in a district; it is standard practice to share this only in aggregate form
  - The idea of sharing the data in aggregate form is to ensure that only non-identifiable data are shared.

- The legal protection of the right to privacy in general and of data privacy in particular varies greatly around the world.

- Biometrics (DNA, fingerprints, iris) and face recognition
- Video surveillance, ubiquitous networks and sensors
- Cellular phones
- Personal Robots
- DNA sequences, Genomic Data

Approaches in Privacy-Preserving
Information Management

- Anonymization Techniques
  - Have been investigated in the areas of networks (see the Anonymity Terminology by Andreas Pfitzman) and databases (see the notion of k-anonymity by L. Sweeney)
- Privacy-Preserving Data Mining
- P3P policies
  - Are tailored to the specification of privacy practices by organizations and to the specification user privacy preferences
- Hippocratic Databases
  - Are tailored to support privacy policies
- **Fine-Grained Access Control Techniques**
- Private Information Retrieval Techniques

- Privacy is not just confidentiality and integrity of user data
- Privacy includes other requirements:
  - Support for user preferences
  - Support for obligation execution
  - Usability
  - Proof of compliance

- Chapter 1- Matt Bishop "Introduction to Computer Security"

- E. Bertino, R. Sandhu "Database Security – Concepts, Approaches, and Challenges", IEEE Transactions on Dependable and Secure Computing, 2(1), 2005.

- L. Sweeney, "k-Anonymity: a Model for Protecting Privacy" http://privacy.cs.cmu.edu/people/sweeney/cv.html#publications

- A. Pfitzman et al. "Anonymity, Unobservability, Pseudonymity and Identity Management – A Proposal for Terminology",

  http://dud.inf.tu-dresden.de/Literatur_VI.shtml

Q & A

Dr. Phạm Thị Bạch Huệ - ptbhue@fit.hcmus.edu.vn

MSc. Lương Vĩ Minh - lvminh@fit.hcmus.edu.vn