CS5322: Database Security

http://www.comp.nus.edu.sg/~cs5322

Tan Kian Lee

COM1, Level 3, 03-23

tankl@comp.nus.edu.sg

CS5322: Database Security

- Background knowledge required
  - Basic Cryptography
  - Databases
    - Database design, relational model, SQL, etc
    - “Internals” of DBMS, e.g, access methods (indexes), query processing algorithms, etc
- Read up if necesssary
  - Security in Computing (4th Edition), by Charles P. Pfleeger and Shari L. Pfleeger, Prentice Hall.
  - Database Management Systems (4th Edition), by Raghu Ramakrishnan and Johannes Gehrke, McGraw Hill.

Introduction

“Hardware is easy to protect: lock it in a room, chain it to a desk, or buy a spare. Information poses more of a problem. It can exist in more than one place; be transported halfway across the planet in seconds; and be stolen without your knowledge.”

— Bruce Schneier

Why Worry About Data Security?

- Amount of acquired data is increasing
- More sensitive data being exposed
- The advent of the Internet as well as networking capabilities has made the access to data much easier
- Damages and misuses of data affect not only a single user or an application; they may have disastrous consequences on the entire organization

Why Worry About Data Security?

"Overall, two-thirds of companies either expect a *data security* incident they will have to deal with in the next 12 months, or simply don't know what to expect."

Source:

2011 IOUG Data Security Survey

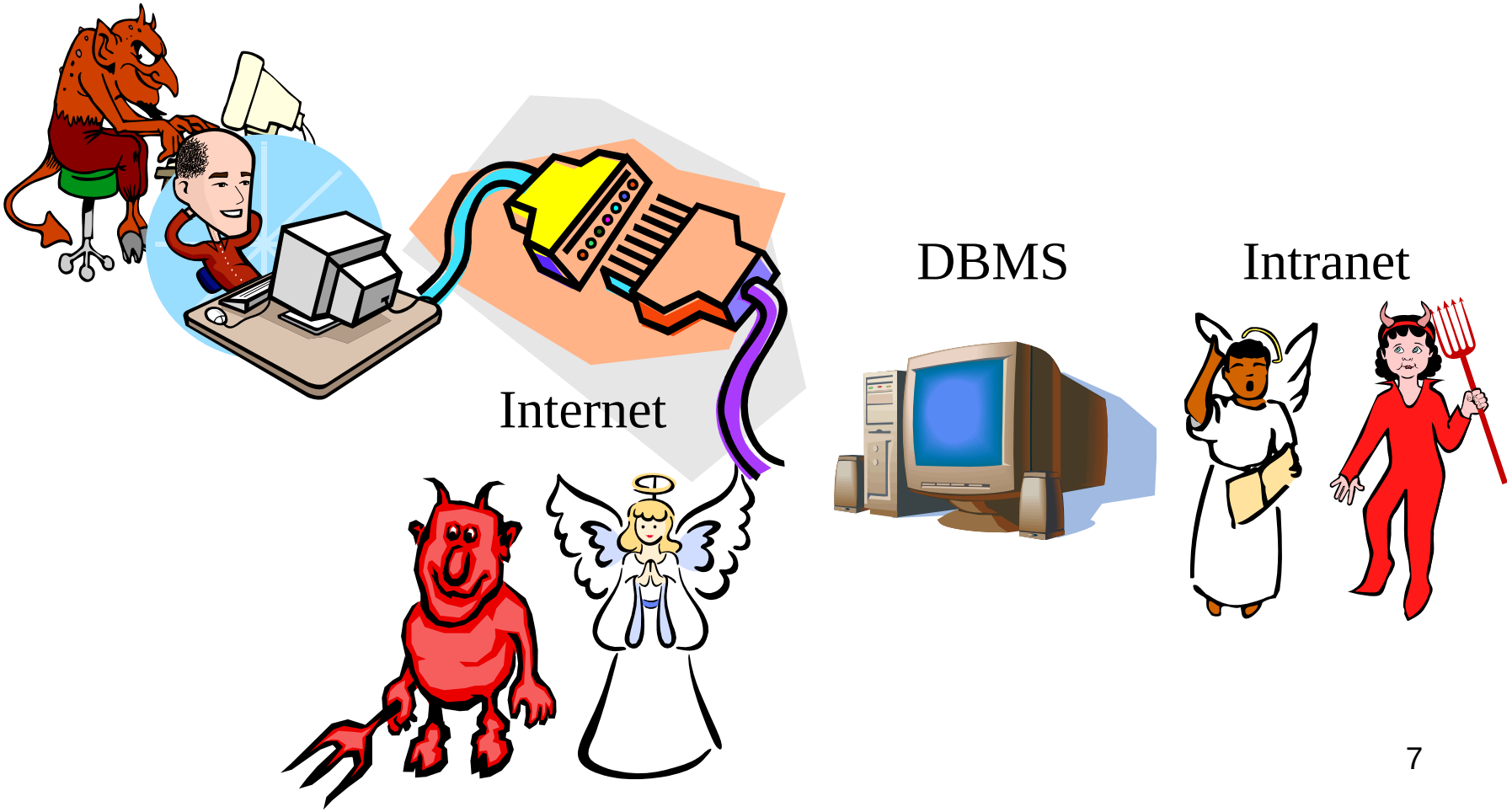
By Joseph McKendrick, Research Analyst

Produced by Unisphere Research, a division of Information Today, Inc. Oct 2011

Why Worry About Data Security?

- IOUG Survey
  - Encryption not being utilized
    - Only 22% encrypt backups and exports
  - DBA access to sensitive data
    - 76% don't have preventive controls on privileged user
    - 43% direct database access to data in database
- Google "news on database security breaches" or "SQL injection"
- Video on "SQL injection"

Why Data Security?



Well Known Security and Privacy Problems

- Computer worms (e.g., Morris worm (1988), Melissa work (1999)
- Computer virus
- Denial of service attacks
- Email spams (e.g., Nigerian scam)
- Identity theft
- Botnets
- Spyware
- Insider threat

- Excessive Privilege Abuse
- Legitimate Privilege Abuse
- Privilege Elevation
- Exploitation of vulnerable, mis-configured databases
- SQL Injection
- Weak Audit Trail

Causes of Software Security Incidents

- Buggy software and wrong configurations
  - Unsafe program languages
  - Complex programs
  - Security considered as an add-on
  - Broken access control
- Lack of awareness and education
  - Few courses in computer security
  - Programming text books do not emphasize security
- Poor usability
  - Security sometimes makes things harder to use
- Economic factors
  - Consumers do not care about security
  - Security is difficult, expensive and takes time
  - Few security audits
- Human nature

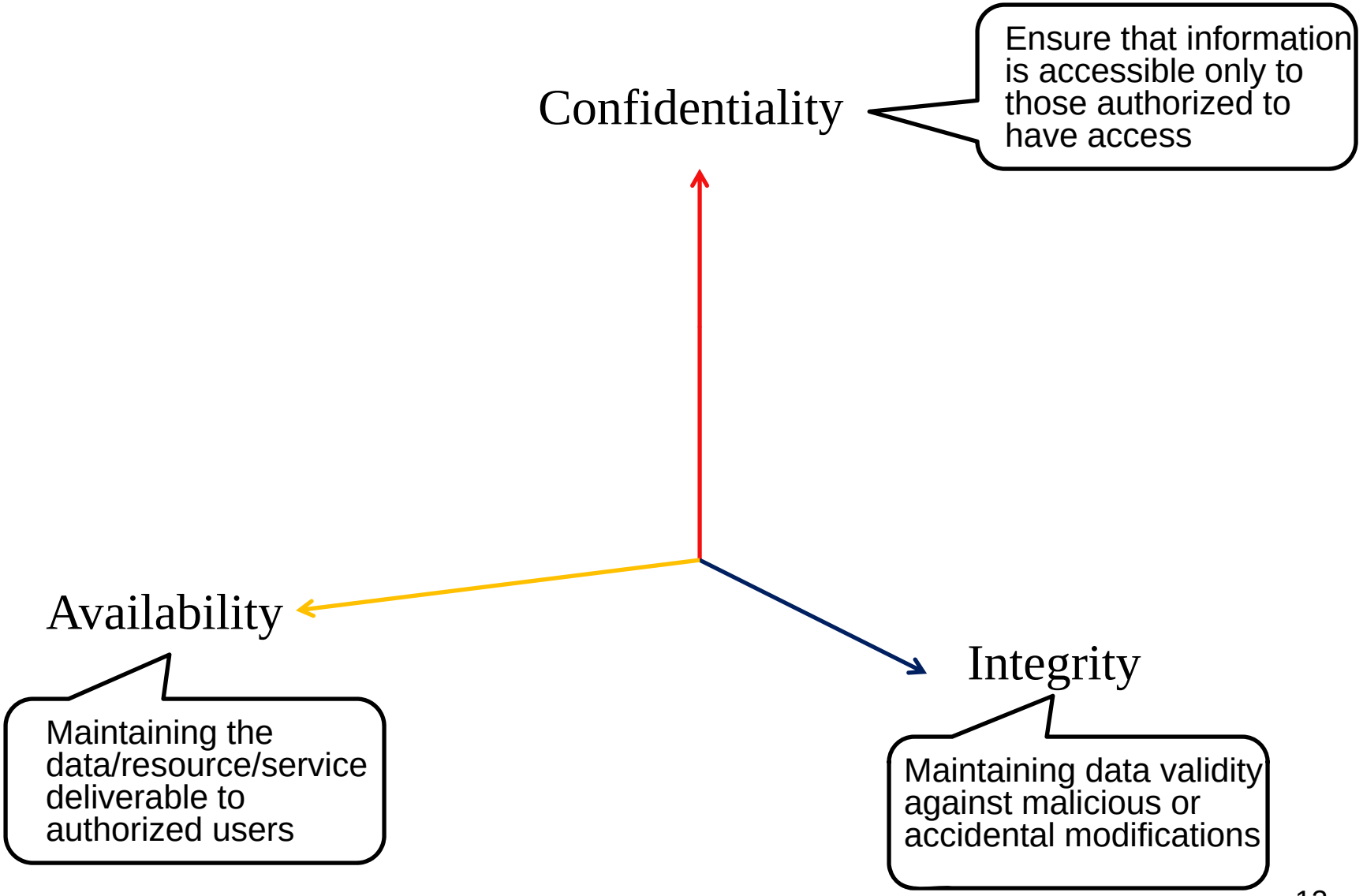
Human Factor

- Who are the attackers?
  - Bored teenagers, criminals, organized crime, organizations, rogue states, industrial, espionage, angry employees, …
- Why do they attack systems?
  - Enjoyment, curiosity, fame, profit, altruistic, …
  - Data represents an extremely valuable asset and often the main goal of attackers is to get valuable or sensitive data

CERT Vulnerabilities Reported (http://www.cert.org/stats)

Year	Total vulnerabilities cataloged
Q1-Q3, 2008	6,058
2007	7,236
2006	8,064
2005	5,990
2004	3,780
2003	3,784
2002	4,129
2001	2,437
2000	1,090
1999	417
1998	262
1997	311
1996	345
1995	171
Totals	**44,074**

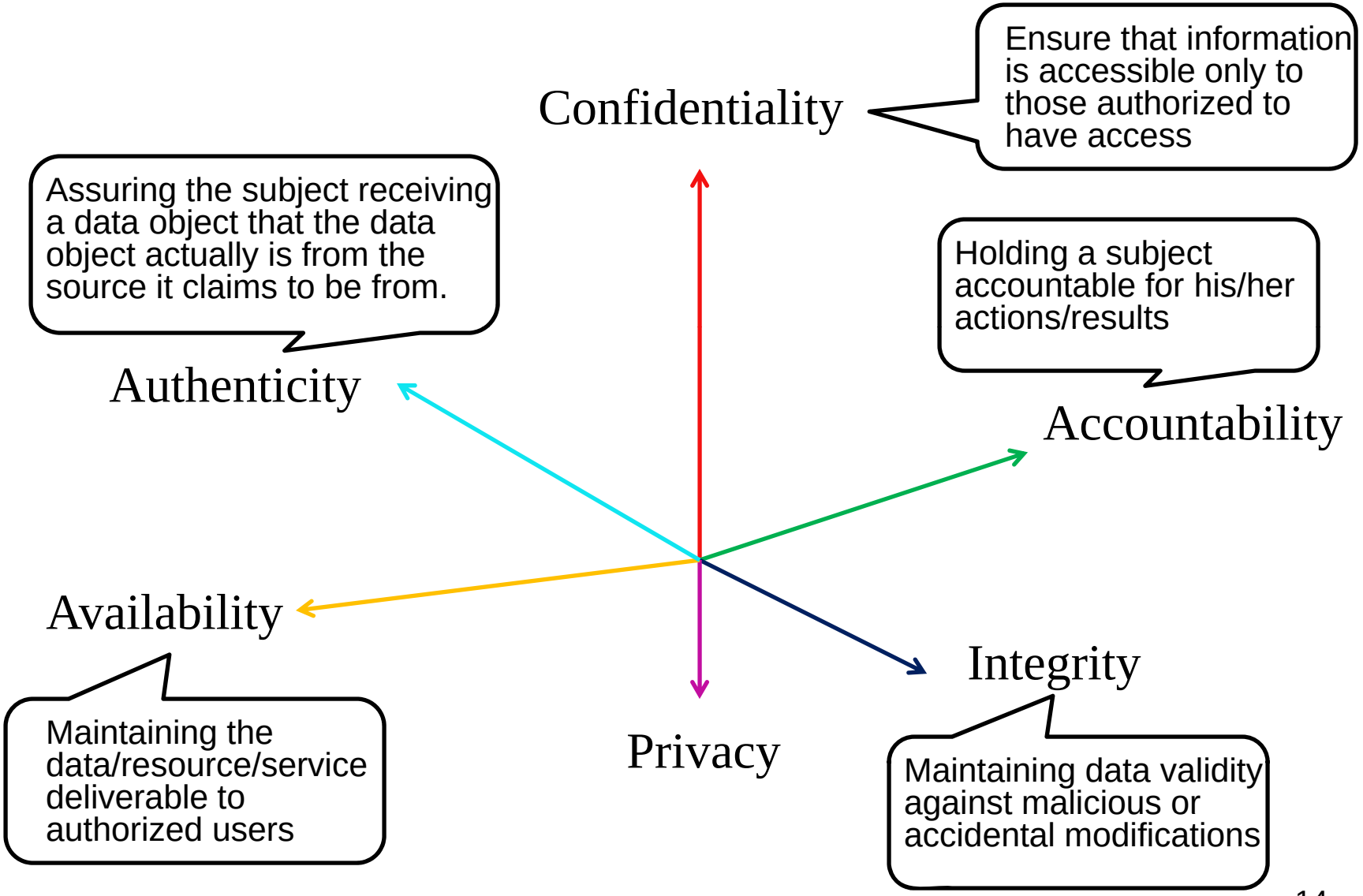
Data Security: Main Requirements



Examples

- Consider a payroll database in a corporation
  - salaries of individual employees are not disclosed to arbitrary users of the database
  - salaries are modified by only those individuals that are properly authorized
  - paychecks are printed on time at the end of each pay period
- In a health-care information system
  - patient's medical information should not be improperly disclosed
  - patient's medical information should be correct
  - patient's medical information can be accessed when needed for treatment
- In a military environment
  - the target of a missile is not given to an unauthorized user
  - the target is not arbitrarily modified
  - the missile is launched when it is fired

Data Security: Other Requirements



Data Security – additional requirements

- Non-repudiation
  - A particular case of accountability where responsibility for an action cannot be denied
  - NIST defines non-repudiation as: Assurance that the sender of information is provided with proof of delivery and the recipient is provided with proof of the sender's identity, so neither can later deny having processed the information

Privacy

- Privacy: maintaining confidentiality of *personally identifiable* information
  - Individuals feel uncomfortable (ownership of information) and 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
- The ability of an individual or organization to control the availability of information about and exposure of him/her-self or organization
  - It deals with the collection, storage, sharing and dissemination of personal and organizational information
  - It is related to being able to function in society anonymously (including pseudonymous or blind credential identification).

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.

Data Privacy

- Technologies with privacy concerns
  - Biometrics (DNA, fingerprints, iris) and face, recognition, Video surveillance, ubiquitous networks and sensors, mobile phones, Personal Robots, DNA sequences, Genomic Data
- Approaches in privacy-preserving information management
  - Anonymization Techniques
  - Privacy-Preserving Data Mining
  - P3P policies (tailored to the specification of privacy practices by organizations and to the specification user privacy preferences)
  - Hippocratic Databases (tailored to support privacy policies)
  - Fine-Grained Access Control Techniques
  - Private Information Retrieval Techniques

Privacy

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

Data Security – additional requirements

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

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

Goals of Security

- 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

Data Security – How?

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

Data Security – Mechanisms

- 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

Data Security – How? Additional mechanisms

- *User authentication* - to verify the identity of subjects wishing to access the data
- *Data authentication* - to ensure data authenticity - it is supported by signature mechanisms
- *Query (result) authentication* - to ensure query result is correct - it is supported by signature mechanisms and data structures
- *Encryption* - to protect data when being transmitted across systems and when being stored on secondary storage
- *Intrusion detection* – to protect against impersonation of legitimate users and also against insider threats

Data Security

- Data must also be protected against transmissions through:
  - Covert channels
  - Inference
    - It is typical of database systems
    - It refers to the derivation of sensitive data from non-sensitive data

Inference - Example

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

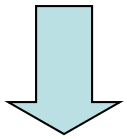
Inference - Example

- 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

Data Security: A Complete Solution

- 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

Policies and Mechanisms

- 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

Assurance

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

Other Issues

- Cost-Benefit Analysis
  - Is it more cost-effective to prevent or recover?
- Risk Analysis
  - Should we protect some information?
  - How much should we protect this information?
- Human Factor
  - Outsiders and Insiders
- Laws and Customs
  - Are desired security measures illegal?
  - Will people adopt them?

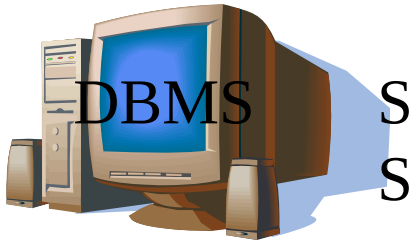
Course Overview

Querying
Encrypted Data

Privacy
Published data,
Statistical databases,
Differential privacy,
Location-based privacy

Encryption

Insider Threat/
Intrusion Detection/
SQL Injection



Steganographic
Storage

Compliance storage

Access Control
DAC, MAC, Role-based

Auditing

Query
Authentication