Database and Cloud Security

Relational Databases

Relational Databases are constructed from tables of data

- Each column holds a particular type of data
- Each row contains a specific value for each column
- Ideally has one column where all values are unique, forming an identifier/key for that row

Relational databases are designed so you can have multiple tables linked together

- They are linked by identifiers (Ex. foreign keys)

Relational databases use query language to access data (Ex. SQL)

Database Access Control

Database access control system determines:

- If the user has access to the entire database or just portions of it
- What access rights the user has (create, insert, delete, update, read, write)

Database access control can support a range of administrative policies

- Centralized administration: small number of privileged users may grant and revoke access rights
- Ownership-based administration: the create of a table may rant and revoke access rights to the table
- Decentralized administration: the owner of the table may grant and revoke authorization rights to other users, allowing them to gran and revoke access rights to the table

SQL Access Control Statements

GRANT and REVOKE statements to give/take access

Access rights: SELECT, INSERT, UPDATE, DELETE

Cascading

GRANT and REVOKE statements allow for cascading

Cascading (GRANT): when you grant access rights to one user, they can now grant that right to other users.

Cascading (REVOKE): if 1 user granted 10 other users, and they got revoked by a higher user, then the user + the 10 he granted are now all revoked.

Role-Based Access Control

RBAC is a good for database management

Categories of roles for database users:

- Application owner
- End user other than application owner
- Administrator

Two types of roles:

- Fixed roles: cannot be deleted, can only add or remove users to this role. Designed for administrative tasks
- Fixed database roles: for administrative tasks on the table level
- Users-defined roles: Two types
  - Standard: normal role, users can add others to this role
  - Application: this is for an application and not actual users, its for when an application need to access info from database

Inference

Inference attacks: the process of performing **authorized queries** and deducing **unauthorized information** from the **legitimate responses** received

- This occurs when a combination of data items is more sensitive than the individual items.

Information transfer path by which unauthorized data is obtained is called an inference channel

Two General Inference Techniques

Analyzing functional dependencies between attributes within a table or across tables

- If we have a table with three attributes about employes: NAME, RANK, and SALARY
- If we notice that all identical ranks have identical salaries then this can be abused (constraint)

Merging views with the same constraints

Inference Detection

During database design time:

- remove an inference channel by altering database when creating it

Detection at query time

- detect inference channel violation then alter/reject specific queries

Both detection method needs some sort of detection algorithm (hard problem)

Statistical Databases

Database that provides data of statistical nature (Ex. counts averages)

Two types:

1. Pure statistical database
   - Only stores statistical data
2. Ordinary database with statistical access
   - Some user have normal access other than statistical

- We are concerned with this type
- Access control objective: allow statistical queries without compromising confidentiality of individual entries
- Inference attacks are one key problem

Statistical Database Security

Uses a characteristics formula C

- A logical formula over the values of attributes

Query set X(C) of characteristic formula c, is the set of records matching C

A statistical query is a query that produces a value calculated over a query set

- There is a trade-off between security and accuracy

Tracker Attacks

Restricting query size is not enough to protect against all attacks (prevents against inference)

If an attack divide queries into parts, they can bypass query size restriction

Tracker: combination of divided parts

- Each part is an acceptable query size
- Overlap data is the desired results

Other Query Restrictions

Query set overlap control

- Limit overlap between new & previous queries
- Has problems (collusion) and overheads (accuracy)

Partitioning

- Cluster records into *exclusive groups*, only allow queries on each whole group
- Overheads (accuracy)

Query denial and information leakage

- Denials can leak information
- To counter must track queries from user

Perturbation

Perturbation: Add noise to statistics generated from data - will result in differences in statistics

Data perturbation techniques:

- Data swapping
- Generate statistics from probability distribution

Output perturbation techniques:

- Random-sample query
- Statistic adjustment

Must minimize loss of accuracy in results

Database Encryption

Databases can encrypt:

- Entire database (inflexible and inefficient)
- Individual fields (simple but inflexible, hard to query due to encryption)
- Rows (records) or columns (attributes) - better

Row encryption: map attribute values to indexes known only to client (can be randomized)

SQL Injection

SQL Injection attack:

1. Attacker finds weakness in website and send malicious code to web server
2. Web server sends it to web application server
3. Web application sends it to database server
4. Database runs and return valuable information (credit card info)
5. Web application generates page with all the info on it
6. Web server sends all credit card info to the hacker

Virtualization

Virtual Machine: give illusion of being a dedicated physical machine that is fully protect and isolated from other virtual machines

Virtual Machine Monitor: a thin layer of software that virtualizes hardware resources, exporting a virtual hardware interface that reflects the underlying machine architecture.

Cloud Security

Concern and Requirements:

- Guest OS isolation and side-channel attacks
- Security management in virtualized environment
- Advanced cryptographic systems