CSE544
Topics in Database Security

Wednesday, May 26, 2004

Outline

- Security in Relational Database Systems
- Security in Statistical Databases
- Current Trends

Discretionary Access Control in SQL

GRANT privileges ON object TO users [WITH GRANT OPTIONS]

privileges = SELECT |
INSERT(column-name) |
DELETE |
REFERENCES(column-name)
object = table | attribute

Examples

GRANT INSERT, DELETE ON Reserves TO Yuppy WITH GRANT OPTIONS

GRANT SELECT ON Reserves TO Michael

GRANT SELECT ON Sailors TO Michael WITH GRANT OPTIONS

GRANT UPDATE (rating) ON Sailors TO Leah

GRANT REFERENCES (bid) ON Boats TO Bill

Views and Security

- David has SELECT rights on table Students
- Creates a VIEW BrightStudents
- Grants SELECT rights on BrightStudents to Dan

Revocation

REVOKE [GRANT OPTION FOR] privileges
ON object FROM users { RESTRICT | CASCADE }

Administrator says:

REVOKE SELECT ON Students FROM David CASCADE

Dan loses SELECT privileges on BrightStudents

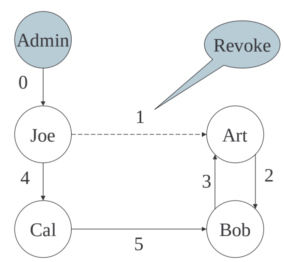
Revocation

Joe: GRANT [….] TO Art …
Art: GRANT [….] TO Bob …
Bob: GRANT [….] TO Art …
Joe: GRANT [….] TO Cal …
Cal: GRANT [….] TO Bob …
Joe: REVOKE [….] FROM Art CASCADE

Same privilege,
same object,
GRANT OPTION

What happens ??

Revocation



According to SQL everyone keeps the privilege

Attacks

- SQL injection (in class)

Security in Statistical Databases

Goal:

- Allow aggregate queries
- Hide confidential data

Why it's hard:

- Allow arbitrary aggregate queries, as long as no compromize

Security in Statistical Databases

Table

Age	Sex	Employer	Diagnosis
42	M	ABC	Schizophrenia
25	F	XYZ	Depression
42	F	XYZ	Depression
. . .			

Queries

count, avg
sum, max, min

```
SELECT count(*)
FROM    Table
WHERE Age=42 and Sex='M' and Employer='ABC'
```

Allow arbitrary
conditions

Attacks

- Mallory knows about John Smith:

  Age=42 & Sex='M' & Employer='ABC'

- Query 1:

  count(Age=42 & Sex='M' & Employer='ABC')

  Answer= 1 we're lucky !
- Query 2:

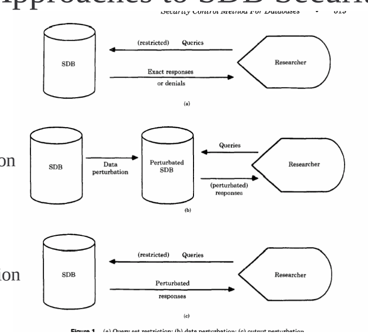
  count(Age=42 & Sex='M' & Employer='ABC' & Diagnosis = 'Schizophrenia')

Approaches to SDB Security

Query restriction

Data perturbation

Query perturbation



Current Research

- Data privacy
- Security in global information sharing
  - Secrecy
  - Integrity

Data Privacy

- The right of individuals to determine for themselves when, how, and to what extent information about them is communicated to others

- US Privacy Act 1974
- US Health Insurance Portability and Accountability Act (HIPPA) 1996
- OECD

Data Privacy

Privacy policies = complex access control:
- Data: e.g. name, SSN, email, disease
- Purpose: e.g. solicitation, treatment, statistics, research
- Recipient: e.g. owner, commercial organization, charity organization
- Condition: e.g. 'opt in', 'opt out'

- Standards: P3P, EPAL

Data Privacy

Attitudes to your own data privacy:

- Paranoid
- Pragmatist
- Indifferent

Which one describes you best ?

Hippocratic Databases

- For the pragmatists
- IBM Almaden [Agrawal et al.]
- Hippocratic Oath: "…I will remain silent…"
- Hippocratic Databases: ten principles:
  - Purpose specification
  - Consent
  - Limited collection
  - Limited use
  - etc

Security in Data Exchange

- Secrecy: make sure you don't give away data when you don't mean to
- Integrity: how can you verify that the data you download is unchanged from its original form ?

Latanya Sweeney's Finding

- In Massachusetts, the Group Insurance Commission (GIC) is responsible for purchasing health insurance for state employees
- GIC collects data, and since it's "private", it publishes it:

GIC(**zip, dob, sex**, diagnosis, procedure, ...)

Latanya Sweeney's Finding

- Sweeney paid $20 and bought the voter registration list for Cambridge Massachusetts:

GIC(**zip, dob, sex**, diagnosis, procedure, ...)
VOTER(name, party, ..., **zip, dob, sex**)

Latanya Sweeney's Finding

- William Weld (former governor) lives in Cambridge, hence is in VOTER
- 6 people in VOTER share his dob
- only 3 of them were man (same sex)
- Weld was the only one in that zip
- Sweeney learned Weld's medical records !

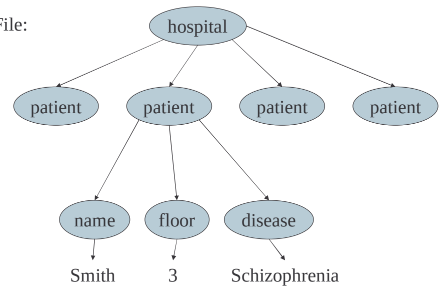
Current proposed solution: k-anonymity

Secrecy in Data Exchange

- Enforce access control policies with encryption
- Start with the plain XML document, then encrypt all fragments that need to stay secret
- Only users having the right key have access
- Problem: multiple policies

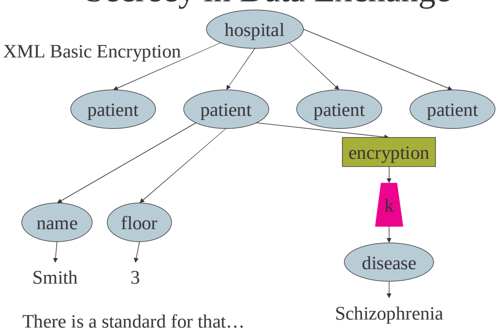
Secrecy in Data Exchange

XML File:



Secrecy in Data Exchange

XML Basic Encryption



There is a standard for that…

Secrecy in Data Exchange

Mr. Smith's disease is accessible to:

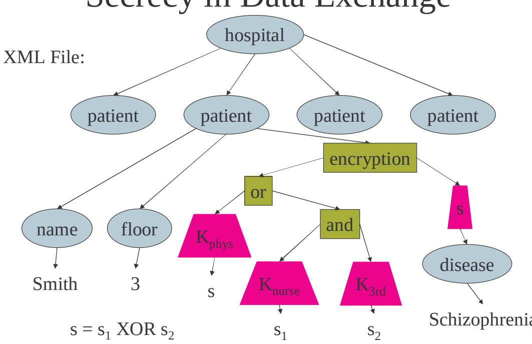
- Physicians
- Nurses working on the 3rd floor

Keys: $K_{physician}$ K_{nurse} K_{3rd}

How do we encrypt ? Need $K_{physician} \vee K_{nurse} \wedge K_{3rd}$

Secrecy in Data Exchange

XML File:



Secure Information Sharing

- Agrawal, Evfimievski, Srikant [SIGMOD'2003]
- Example: two competing companies agree to share their list of their customers with a poor payment record but nothing else

Secure Information Sharing

Formally:

- Alice has A = {x1, …, xn}
- Bob has B = {y1, …, ym}
- They want to find out A ∩ B, and not reveal anything else

Secure Information Sharing

Attempt 1:

- Alice computes HA = h(A) sends to Bob
- Bob computes HB = h(B) sends to Alice
- Now each computes $A \cap B$

- What's wrong ?

Secure Information Sharing

- Solution: use commutative encryption

$$\boxed{E_k (E_{k'}(x)) = E_{k'}(E_k(x))}$$

- Example: $E_k(x) = x^k \bmod p$

Secure Information Sharing

Solution:

1. Alice computes $YA = \{E_a(x) \mid x \in A\}$
2. Bob computes $YB = \{E_b(y) \mid y \in B\}$
3. Exchange YA, YB. ORDERED !
4. Bob computes $\{(E_a(x), E_b(E_a(x))) \mid x \in A\}$
5. Alice computes $\{E_a(E_b(y)) \mid y \in B\}$
6. Bob sends that to Alice
7. Alice can now compute $A \cap B$

HA,HB instead

Integrity in Data Sharing

- Merkle Trees (in class)