

Ασφάλεια συστημάτων βάσεων δεδομένων

- Η αξία της πληροφορίας είναι το κύριο «περιουσιακό στοιχείο» των πληροφοριακών συστημάτων
- Πάνω από το 90% των σύγχρονων πληροφοριακών συστημάτων περιλαμβάνει κάποιο είδος βάσης δεδομένων
- Η ασφάλεια των συστημάτων βάσεων δεδομένων αποκτά ιδιαίτερη σημασία

Συστήματα βάσεων δεδομένων

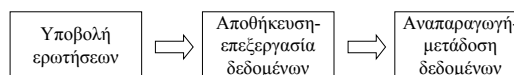
- Συστήματα βάσεων δεδομένων
 - οι ίδιες οι βάσεις δεδομένων
 - το λογισμικό που διαχειρίζεται τις βάσεις δεδομένων
 - » εκτελείται πάνω από το λειτουργικό σύστημα
 - » υλοποιεί τις ειδικές λειτουργίες που χρειάζονται
- Μοντέλα δεδομένων
 - σχεσιακό
 - αντικειμενοστραφές
 - ιεραρχικό
 - δικτυακό

Παράμετροι ασφάλειας

- Βασικές διαστάσεις
 - ακεραιότητα
 - έλεγχος προσπέλασης
 - εμπιστευτικότητα
 - διαθεσιμότητα
 - έλεγχος (audit) κ.λπ.
- Νέες διαστάσεις
 - διακριτότητα (granularity)
 - συμπερασμός ή έμμεση προσπέλαση (inference)
 - συνάθροιση (aggregation)
 - φιλτράρισμα (filtering)
 - καταγραφή (journaling)

Γενικές αρχές (1/2)

- Η βάση δεδομένων χρησιμοποιείται ως εργαλείο για αποθήκευση, επεξεργασία και μετάδοση των πληροφοριών



- Η αξιοπιστία της μετάδοσης ελέγχεται από ειδικά πρωτόκολλα που εγγυώνται
 - την ολοκλήρωση των δοσοληψιών (transactions)
 - την εφαρμογή κανόνων ακεραιότητας (integrity constraints)

Γενικές αρχές (2/2)

- Η ασφάλεια μιας βάσης δεδομένων πρέπει να λαμβάνει υπόψη το συνολικό περιβάλλον υλικού και λογισμικού που σχετίζεται μ' αυτή
- Η ακεραιότητα είναι βασική απαίτηση
 - αντοχή σε βλάβες υλικού και δυσλειτουργίες λογισμικού
 - τροποποιήσεις μόνο από εξουσιοδοτημένους χρήστες
 - αν υπάρχει παραβίαση ακεραιότητας, άμεση ενημέρωση του χρήστη
- Διαθεσιμότητα για τους εξουσιοδοτημένους χρήστες
- Οι έλεγχοι ορθότητας (audit) να είναι αναλυτικοί και διεξοδικοί, αλλά να μην επηρεάζουν την απόδοση
- Εμπιστευτικότητα – τα δεδομένα είναι διαθέσιμα μόνο σε εξουσιοδοτημένους χρήστες

Απαιτήσεις ασφάλειας Συστημάτων ΒΔ

- Φυσική ακεραιότητα της βάσης
 - από πτώσεις τάσης, σφάλματα δίσκων κ.λπ.
- Λογική ακεραιότητα της βάσης
 - σε επίπεδο πεδίων (οι επί μέρους τιμές είναι σωστές)
 - σε συνδυαστικό επίπεδο (αναφορική ακεραιότητα, συνθήκες μεταξύ πεδίων, η μεταβολή ενός πεδίου δεν επηρεάζει τα άλλα εκτός αν αυτό έχει προβλεφθεί κ.ά.)
- Διακρίβωση ταυτότητας (αυθεντικοποίηση – user authentication)
 - η βάση δεδομένων αναγνωρίζει την ταυτότητα του χρήστη πριν του επιτρέψει την προσπέλαση
- Έλεγχος προσπέλασης
 - ανάλογα με την ταυτότητα του χρήστη δίνονται κατάλληλα δικαιώματα πρόσβασης σε υποσύνολα δεδομένων
- Διαθεσιμότητα
 - οι εξουσιοδοτημένοι χρήστες μπορούν άμεσα να προσπελάσουν τα δεδομένα για τα οποία έχουν το σχετικό δικαίωμα

Φυσική ακεραιότητα της βάσης

- Διαδικασίες τήρησης εφεδρικών αντιγράφων
 - έμφαση στην τήρηση εφεδρικών αντιγράφων εν λειτουργία
 - Ιδιαίτερα επιθυμητή η ύπαρξη γρήγορων διαδικασιών ανάκαμψης
- Υποστήριξη διατάξεων πολλαπλών αντιγράφων των δεδομένων (RAID)
- Διαδικασίες για τον ασφαλή τερματισμό της λειτουργίας της βάσης δεδομένων

Λογική ακεραιότητα της βάσης (1/4)

- Έλεγχος τιμών πεδίων
 - σε επίπεδο πεδίου ορισμού (ακέραιος, πραγματικός κ.ά.)
 - » price number(6, 2)
 - » quantity number(5, 0)
 - αποδοχή-μη αποδοχή τιμών null
 - » customerId varchar(20) not null
 - » faxNo varchar(30)
 - σε επίπεδο εύρους τιμών
 - » grade number(2) check(grade >= 0 and grade <= 10)
 - » isSecure char(1) check(isSecure = 'y' or isSecure = 'n')

Λογική ακεραιότητα της βάσης (2/4)

- Έλεγχος σχέσης πεδίων
 - check (endDate is NULL or endDate > startDate)
- Έλεγχος μοναδικότητας τιμών
 - customerId number(10, 0) unique
 - countryId varchar(3), passportNo varchar(10)
primary key(countryId, passportNo)
- Έλεγχος αναφοράς τιμών
 - countryId varchar(3) references country(countryId)
 - foreign key(countryId, passportNo)
references immigrant(countryId, passportNo)
on update cascade
- Έλεγχος μεταβάσεων
 - ελέγχεται τόσο η τρέχουσα όσο και η νέα κατάσταση των δεδομένων
 - » if (new.salary / old.salary > 1.2) then
raise_application_error(100,
'only raises up to 20% are allowed')

Λογική ακεραιότητα της βάσης (3/4)

- Εξασφάλιση της ακεραιότητας μέσω δοσοληψιών

```
begin transaction;
update account set balance = balance - :amount
where accountId = :account1;
update account set balance = balance + :amount
where accountId = :account2;
end transaction;
```

 - Ημερολόγια αναίρεσης ή επανάληψης ενεργειών με πρότερη ή ύστερη εγγραφή δεδομένων
- Εξασφάλιση της ακεραιότητας μέσω ελέγχου ταυτοχρονισμού

```
select sum(balance) from account;
update account
set balance = balance + interest, interest = 0
```

 - Κλειδώματα δεδομένων ή χρήση χρονσήμων

Λογική ακεραιότητα της βάσης (4/4)

- Οι περισσότεροι έλεγχοι συνέπειας μπορούν να διενεργηθούν άμεσα (άμεσοι έλεγχοι – immediate checks)
 - π.χ. περιοχή τιμών πεδίου
- Ορισμένοι έλεγχοι μπορούν να διενεργηθούν μόνο στο τέλος των δοσοληψιών (μεταχρονολογημένοι έλεγχοι – deferred checks)
 - π.χ. μία εταιρεία πρέπει να έχει πάντα ακριβώς έναν εργαζόμενο με θέση πρόεδρος. Τι συμβαίνει στην αλλαγή προέδρου;
» select count(*) into :numPresidents
from employee where position = 'President';
if (:numPresidents <> 1) then /* error */ ...
» update employee set position = 'Retired'
where position = 'President';
» update employee set position = 'President'
where empId = :newPresident;
 - Οι μεταχρονολογημένοι έλεγχοι διενεργούνται κατά την επικύρωση της οικείας δοσοληψίας

Διακρίβωση ταυτότητας χρηστών (1/3)

- Διακρίβωση με όνομα χρήστη-συνθηματικό

```
create user auser identified by apassword;
```

 - Το ΣΔΒΔ φυλάσσει την αντιστοιχία μεταξύ ονόματος χρήστη και συνθηματικού
 - Ο χρήστης κατά τη σύνδεσή του οφείλει να παρουσιάσει το συνδυασμό των διαπιστευτηρίων σύνδεσης
 - Δεν είναι απαραίτητο να υπάρχει οποιαδήποτε συσχέτιση μεταξύ των διαπιστευτηρίων για το λειτουργικό σύστημα και των διαπιστευτηρίων της βάσης

Διακρίβωση ταυτότητας χρηστών (2/3)

■ Διακρίβωση ταυτότητας χρήστη από το λειτουργικό σύστημα

`create user auser identified externally;`

- Το ΣΒΔ επαφίεται στο λειτουργικό σύστημα για τη διακρίβωση ταυτότητας
- Οι χρήστες δεν είναι απαραίτητο να γνωρίζουν πρόσθετα ονόματα χρηστών ή συνθηματικά πέρα από αυτά που χρησιμοποιούν για τη σύνδεση με το σύστημα
- Ο χρήστης της βάσης πρέπει να έχει λογαριασμό στο λειτουργικό σύστημα
- Πρέπει να χρησιμοποιείται ΜΟΝΟ όταν το λειτουργικό σύστημα έχει επαρκώς αξιόπιστους μηχανισμούς διακρίβωσης ταυτότητας

Διακρίβωση ταυτότητας χρηστών (3/3)

■ Διακρίβωση ταυτότητας χρήστη μέσω καθολικής υπηρεσίας διακρίβωσης ταυτότητας (X509, DCE, Kerberos, LDAP κ.λπ.)

- Το ΣΒΔ συνεργάζεται την υπηρεσία για τη διακρίβωση της ταυτότητας του χρήστη
- Προωθεί τη χρήση κεντρικού σημείου φύλαξης διαπιστευτηρίων σύνδεσης

Έλεγχος προσπέλασης

■ Βάσει της διακριβωμένης ταυτότητάς τους οι χρήστες έχουν συγκεκριμένα προνόμια όπως καθορίζονται από την πολιτική ασφάλειας

- προνόμια συνολικά επί του συστήματος
- προνόμια επί συγκεκριμένων αντικειμένων
 - » κατ' επιλογήν έλεγχος προσπέλασης (discretionary access control)
 - » υποχρεωτικός έλεγχος προσπέλασης (mandatory access control)

Έλεγχος προσπέλασης – προνόμια συστήματος

- Προνόμια που καθορίζουν τις γενικές δυνατότητες που έχει ο χρήστης σε σχέση με το σύστημα βάσεων δεδομένων
 - Δυνατότητα δημιουργίας συνόδου (create session), δυνατότητα χρήσης πόρων (resource), δυνατότητα δημιουργίας πινάκων (create table), δυνατότητα δημιουργίας δεικτών (create index), δυνατότητα δημιουργίας διαδικασιών (create procedure)
 - Όρια χρήσης χώρου σε ενότητες αποθήκευσης (quota)
 - Όρια χρόνου εκτέλεσης ερωτήσεων, όρια εισόδου-εξόδου)
- Τα προνόμια παραχωρούνται με την εντολή grant, ανακαλούνται με την εντολή revoke – τα όρια καθορίζονται με την παράμετρο quota της εντολής alter
 - grant create session to user1;
 - grant create table to user1 with admin option;
 - revoke create table from user1;
 - alter user user1 quota 20M on users;

Έλεγχος προσπέλασης – προνόμια επί συγκεκριμένων αντικειμένων (1/3)

■ Η γλώσσα SQL υποστηρίζει άμεσα τον κατ' επιλογήν έλεγχο προσπέλασης

- ο δημιουργός ενός αντικειμένου είναι ο ιδιοκτήτης του αντικειμένου
- ο ιδιοκτήτης έχει όλα τα δικαιώματα επί του αντικειμένου και μπορεί επίσης να παραχωρεί προνόμια σε άλλους χρήστες ή να τους αφαιρεί τα παραχωρηθέντα
- τα συγκεκριμένα δικαιώματα εξαρτώνται από τη φύση του αντικειμένου
 - » πίνακες: επιλογή, εισαγωγή, διαγραφή, ενημέρωση, τροποποίηση, δημιουργία δεικτών, δημιουργία αναφορών
 - » όψεις: επιλογή, εισαγωγή, διαγραφή, ενημέρωση, τροποποίηση
 - » διαδικασίες: εκτέλεση
 - » δείκτες: αλλαγή δομής αποθήκευσης

Έλεγχος προσπέλασης – προνόμια επί συγκεκριμένων αντικειμένων (2/3)

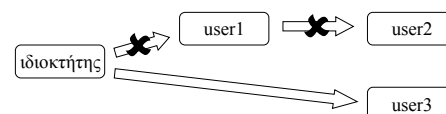
- Εντολές grant και revoke
 - grant select, delete on table1 to user1;
 - grant insert on table1(col1, col2, col3) to user1;
 - grant update on table1(col3) to public;
 - grant select on table1 to user2 with grant option;
 - revoke select on table1 from user2;
 - grant references on table1(col1) to user1, user2;
 - Η παροχή προνομίου insert που δεν περιλαμβάνει όλες τις υποχρεωτικές στήλες είναι άσκοπη
- Οι προσδιορισμοί select, delete δεν επιδέχονται καθορισμό στηλών
 - Για παραχώρηση προνομίου επιλογής σε συγκεκριμένες στήλες πρέπει να δημιουργηθεί όψη
 - create view view1 as select col1, col2, col3 from table1;
 - grant select, insert on view1 to user1, user2;
 - Ο ορισμός όψης μπορεί να περιέχει αναφορές στον χρήστη
 - create view sal as select * from salary where user = uid();
 - grant select on sal to public;

Έλεγχος προσπέλασης – προνόμια επί συγκεκριμένων αντικειμένων (3/3)

- Δικαίωμα εκτέλεσης σε διαδικασίες
 - `grant execute on procedure1 to user1, user2;`
 - Οι διαδικασίες μπορούν να ενσωματώνουν οποιουδήποτε ελέγχους μπορεί να εκφράσει η σχετική γλώσσα
 - Μία διαδικασία κατά την εκτέλεσή της μπορεί να προσπελάσει τα αντικείμενα που έχει δικαίωμα να προσπελάσει ο ιδιοκτήτης της διαδικασίας
 - Είναι συνήθης πρακτική να μην παρέχονται προνόμια απ' ευθείας σε πίνακες αλλά να δημιουργούνται οι σχετικές διαδικασίες και να παραχωρούνται προνόμια εκτέλεσης σ' αυτές

Ανάκληση των προνομίων

- Ο ιδιοκτήτης ενός αντικειμένου παραχωρεί στον χρήστη user1 συγκεκριμένα προνόμια με δικαίωμα περαιτέρω παραχώρησης τους
- Ο χρήστης user1 παραχωρεί τα προνόμια σε άλλους χρήστες user2, user3 κ.λπ.
- Κατόπιν ο ιδιοκτήτης του αντικειμένου ανακαλεί τα προνόμια από τον χρήστη user1
 - το σύστημα πρέπει να ανακαλέσει τα προνόμια από τους χρήστες user2, user3 κ.λπ.



Διαχείριση προνομίων με ρόλους

- Ρόλος: διαχειριστική οντότητα στην οποία παραχωρούνται προνόμια
- Οι χρήστες είναι δυνατόν να συσχετίζονται με ρόλους
- χρήστης που έχει συσχετισθεί με ρόλο αποκτά αυτόματα όλα τα προνόμια που έχουν παραχωρηθεί στον ρόλο
 - Η συσχέτιση μπορεί να είναι άνευ συνθήκης, όταν γίνεται από τον διαχειριστή ή να προαπαιτεί τη γνώση συνθηματικού από τον χρήστη
- Η τροποποίηση των προνομίων που έχουν παραχωρηθεί σε κάποιον ρόλο οδηγεί στην αυτόματη τροποποίηση των δικαιωμάτων όλων των χρηστών που έχουν συσχετισθεί με τον ρόλο αυτό
 - `create role personnel identified by secret;`
 - `create role accountant identified by anotherSecret;`
 - `grant all on tbl1, tbl2, tbl3 to personnel;`
 - `grant all on tbl1, tbl5, tbl6 to accountant;`
 - `grant personnel to user1, user2;`
 - `grant accountant to user3, user5;`
 - `revoke update, delete on tbl1 from accountant;`

Ευαίσθητα δεδομένα

- Δεδομένα τα οποία δεν πρέπει να αποκαλυφθούν
 - εξαρτάται από τη βάση δεδομένων και τη σημασιολογία των δεδομένων
 - » π.χ. ο Χρυσός οδηγός δεν έχει ευαίσθητες πληροφορίες
 - » το αρχείο καινοτόμων προϊόντων μιας εταιρίας είναι εξ ολοκλήρου απόρρητο
 - » Πιο δύσκολες είναι οι ενδιάμεσες καταστάσεις
 - Παράδειγμα: έστω βάση δεδομένων φοιτητών
 - » το όνομα και η διεύθυνση δεν είναι ευαίσθητα
 - » η οικονομική βοήθεια και το ιατρικό ιστορικό είναι απόρρητα
 - » το φύλλο, η ηλικία και η φυλή είναι στο ενδιάμεσο
 - » Όλοι θα έχουν πρόσβαση σε όνομα και διεύθυνση, κάποιος στο φύλλο, ηλικία & φυλή, ελάχιστοι στην οικονομική βοήθεια & το ιατρικό ιστορικό
 - » Η πολιτική ασφάλειας μπορεί να ορίζει ότι κανείς δεν πρέπει να έχει πρόσβαση σε όλα τα στοιχεία

Καθορισμός ευαισθησίας δεδομένων

- Τα δεδομένα μπορεί να είναι ευαίσθητα:
 - *εκ της φύσεως τους*: η τοποθεσία των οπλικών συστημάτων ή το μέσο εισόδημα των πεταλωτών αλόγων σε μία πόλη με μόνο έναν πεταλωτή
 - *λόγω της πηγής τους*: η αποκάλυψη συγκεκριμένων δεδομένων μπορεί να καταδείξει την πηγή της πληροφορίας, η οποία πρέπει να παραμείνει μυστική
 - *ρητώς χαρακτηρισμένα*: π.χ. διαβαθμισμένα στρατιωτικά μυστικά
 - *τμήμα ευαίσθητου δεδομένου ή ευαίσθητης εγγραφής*: ο βασικός μισθός ενός εργαζόμενου (τμήμα της μισθοδοσίας) ή μια εγγραφή που περιγράφει μία απόρρητη αποστολή
 - *απόρρητο λόγω πληροφοριών που αποκαλύφθηκαν προηγουμένως*: αν αποκαλυφθεί το γεωγραφικό πλάτος ενός ορυχείου δεν πρέπει να αποκαλυφθεί το γεωγραφικό μήκος

Είδη αποκαλύψεων (1/2)

- *Αποκάλυψη επακριβών τιμών*: η πιο σοβαρή περίπτωση όπου τα ευαίσθητα δεδομένα αποκαλύπτονται
- *Αποκάλυψη ορίων*: αποκαλύπτεται ότι για την τιμή ενός ευαίσθητου δεδομένου x ισχύει $min \leq x \leq max$
 - ιδιαίτερα επικίνδυνο αν μπορεί να αξιοποιηθεί αναδρομικά αποκαλύπτοντας ότι $min \leq x \leq (min + max) / 2$ κ.ο.κ.
 - η αποκάλυψη και μόνο ότι το πλήθος γιατρών με ειδικευση στον βιολογικό πόλεμο είναι πάνω από ένα όριο είναι εν δυνάμει επικίνδυνη
- *Αρνητικός συμπερασμός*: η αντίληψη της πληροφορίας ότι κάποιον ευαίσθητο δεδομένο δεν έχει κάποια τιμή
 - π.χ. πλήθος καταδικών για κακουργήματα διαφορετικό από μηδέν – η διαφορά μεταξύ 99 και 100 είναι αδιάφορη, από 0 σε 1 σημαντική
 - η αποκάλυψη ότι κάποιος φοιτητής δεν αποφοίτησε με άριστα είναι λιγότερο σημαντική – το δυνατό εύρος τιμών 5.00 – 8.49 είναι ιδιαίτερα μεγάλο

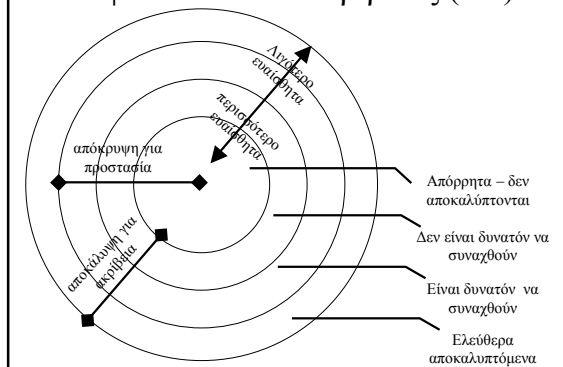
Είδη αποκάλυψης (2/2)

- **Υπαρξη:** η αποκάλυψη ότι συνολικά *υπάρχει* ένα είδος πληροφορίας μπορεί να είναι σημαντική
 - πεδίο πολιτικά φρονήματα στη βάση δεδομένων προσωπικού
- **Πιθανή τιμή:** συμπερασμός ότι κάποιο δεδομένο έχει συγκεκριμένη τιμή
 - *ερώτηση1:* πόσοι κατοικούν στη διεύθυνση «Αρίων 03»; → 8
 - *ερώτηση2:* πόσοι κατοικούν στη διεύθυνση «Αρίων 03» και έχουν καταδικασθεί για φοροδιαφυγή; → 1

Ασφάλεια έναντι ακρίβειας (1/2)

- Η απόλυτη ασφάλεια επιτυγχάνεται αν δεν αποκαλύπτονται καθόλου ευαίσθητα δεδομένα
- Οι χρήστες πιθανώς θα ήθελαν τα δεδομένα για παραδεκτούς λόγους που δεν διακυβεύουν την ασφάλεια
 - » π.χ. βαθμοί και ιατρικό ιστορικό των φοιτητών για ιατρικές μελέτες
 - » λίστα μισθών και φύλλου για στατιστική
- οι ερωτήσεις δεν αποκαλύπτουν ταυτότητες υποκειμένων

Ασφάλεια έναντι ακρίβειας (2/2)



Το πρόβλημα του συμπερασμού

- Αποκάλυψη ευαίσθητων δεδομένων διαμέσου μη ευαίσθητων δεδομένων

Name	Address	Sex	Race	Age	FinAid	Drugs
Adams	Holmes	M	C	32	5000	1
Bailey	Grey	M	B	28	0	0
Chin	West	F	A	27	3000	0
Dewitt	Grey	M	B	28	1000	3
Earhart	Holmes	F	C	31	2000	1
Fein	West	F	C	26	1000	0
Groff	West	M	C	34	4000	3
Hill	Holmes	F	B	23	5000	2
Koch	West	F	C	21	0	1
Liu	Grey	F	A	28	0	2
Majors	Grey	M	C	22	2000	2

Ευθεία επίθεση για συμπερασμό

- Άμεση αναφορά σε ευαίσθητα δεδομένα στη συνθήκη επιλογής
 - `select name from students where sex = 'M' and drugs = 1`
 - Η απαίτηση χρήσης μη ευαίσθητων δεδομένων στη συνθήκη επιλογής δεν αντιμετωπίζει το πρόβλημα
 - `select name from students where (sex = 'M' and drugs = 1) OR (sex <> 'M' and sex <> 'F') OR (address = 'ATX0NN@45Q!46')`
 - Κανόνας *ν αντικείμενα άνω του κ%*: δεν δίνεται απάντηση αν μικρό πλήθος αντικειμένων αποτελούν μεγάλο ποσοστό της απάντησης
 - » πιο ισχυρό από το περισσότερες από *ν* γραμμές απάντησης

Έμμεση επίθεση για συμπερασμό

- Διάφοροι οργανισμοί επιτρέπουν μόνο εξαγωγή «ουδέτερων» στατιστικών μεγεθών (πλήθος, μέσος όρος, άθροισμα) – όχι συγκεκριμένες τιμές
- Συμπερασμός συγκεκριμένων τιμών διαμέσου ενδιάμεσων αποτελεσμάτων

Έμμεση επίθεση με αθροίσματα

- Συμπερασμός στοιχείων μέσω κατάλληλα διαμορφωμένων αθροισμάτων
 - Σύνολο οικονομικής ενίσχυσης κατά φύλο και κατοικία

	Holmes	Grey	West	Σύνολο
M	5000	3000	4000	12000
F	7000	0	4000	11000
Σύνολο	12000	3000	8000	23000

Καμία γυναίκα που μένει στο Grey δεν λαμβάνει οικονομική ενίσχυση

Έμμεση επίθεση με πληθάριθμους

- Συνήθως σε συνδυασμό με αθροίσματα
 - Πλήθος ατόμων ανά διεύθυνση διαμονής και φύλο

	Holmes	Grey	West	Σύνολο
M	1	3	1	5
F	2	1	3	6
Σύνολο	3	4	4	11

Σε συνδυασμό με τον προηγούμενο πίνακα δίνει επακριβώς την οικονομική ενίσχυση

Έμμεση επίθεση με μέσους όρους

- Εύρεση συνόλων που διαφέρουν κατά ένα μέλος και εξαγωγή μέσου όρου για τα σύνολα αυτά
 - Η διαφορά είναι η τιμή του ευαίσθητου δεδομένου για το συγκεκριμένο μέλος
 - » `select avg(finAid) from students where sex = 'M'`
 - » `select avg(finAid) from students where drugs = 2`
 - Η διαφορά είναι η οικονομική ενίσχυση για τον Majors

Επιθέσεις μέσω ερωτήσεων εντοπισμού

- Συστήματα πιθανώς να αρνηθούν να δώσουν απαντήσεις αν το πλήθος των εγγράφων που χρησιμοποιούνται για τον υπολογισμό είναι μικρότερο από ένα όριο
 - Αντί να υποβάλλουμε ερώτηση που δίνει στοιχεία για ένα αντικείμενο, υποβάλλουμε ερώτηση που επιστρέφει ν-/αντικείμενα και ερώτηση που επιστρέφει ν αντικείμενα
 - Η διαφορά είναι η τιμή του ευαίσθητου δεδομένου που αναζητούμε

Γραμμική ευπάθεια συστημάτων

- Γενίκευση των ερωτήσεων εντοπισμού

$$\begin{aligned}
 q_1 &= c_1 + c_2 + c_3 + c_4 + c_5 \\
 q_2 &= c_1 + c_2 + c_4 \\
 q_3 &= c_3 + c_4 \\
 q_4 &= c_4 + c_5 \\
 q_5 &= c_2 + c_5
 \end{aligned}$$

- Καμία ερώτηση δεν αποκαλύπτει μεμονωμένη τιμή
- Λύνοντας το γραμμικό σύστημα βρίσκουμε όλες τις τιμές

Αντιμετώπιση συμπερασμού

- Δύο μέθοδοι:
 - Έλεγχος των ερωτήσεων – η ερώτηση δεν πρέπει να αποκαλύπτει ευαίσθητα δεδομένα
 - » Δύσκολο να καθοριστεί αν αποκαλύπτονται ευαίσθητα δεδομένα ή όχι
 - » Κυρίως χρήσιμη έναντι της ευθείας επίθεσης
 - Έλεγχος σε επίπεδο μεμονωμένου στοιχείου βάσης δεδομένων
 - » Απαγόρευση (suppression) – ερωτήσεις που βασίζονται σε ευαίσθητα δεδομένα απορρίπτονται και δεν δίνεται απάντηση
 - » Απόκρυψη (concealing) – όταν χρειάζεται τιμή ευαίσθητου δεδομένου το σύστημα επιστρέφει μία τιμή κοντά αλλά όχι ακριβώς ίση με την πραγματική τιμή
 - Η απαγόρευση δίνει ακριβή αποτελέσματα όταν απαντά, αλλά απορρίπτει πολλές ερωτήσεις
 - Η απόκρυψη δίνει λιγότερο ακριβή αποτελέσματα, αλλά απαντά σε περισσότερες ερωτήσεις
 - Η επιλογή εξαρτάται από τη σημασιολογία της Β.Δ.

Παραδείγματα

- Απαγόρευση αποκαλυπτικών απαντήσεων
 - Πλήθος ατόμων ανά διεύθυνση διαμονής και φύλο

	Holmes	Grey	West	Σύνολο
M	1	3	1	5
F	2	1	3	6
Σύνολο	3	4	4	11

- Η απαγόρευση μόνο των κελιών με «1» δεν ωφελεί – μπορούν να συναχθούν αφαιρώντας από το σύνολο το άλλο κελί της στήλης
- Αν παρουσιάζονται αθροίσματα πρέπει να αποκρυφθούν τουλάχιστον δύο κελιά ανά στήλη/γραμμή – αν όχι, ένα κελί αρκεί

Παραδείγματα

- Συνδυασμός αποτελεσμάτων - στήλες ή γραμμές συγχωνεύονται για να διασφαλιστεί η ασφάλεια των δεδομένων

– Π.χ. πλήθος ατόμων ανά φύλο και κατηγορία ιατρικού ιστορικού

	0	1	2	3		0 - 1	2-3
M	1	1	1	2	M	2	3
F	2	2	2	0	F	4	2

- Παρουσίαση τιμών σε περιοχές
 - π.χ. αντί να δίνονται ακριβή στοιχεία για την οικονομική βοήθεια, δίνονται στοιχεία για τις περιοχές 0-1999, 2000-3999 και ≥ 4000
- Στρογγύλευση τιμών
 - αντί να χρησιμοποιούνται οι ακριβείς τιμές, στρογγυλεύονται πρώτα π.χ. στην πλησιέστερη δεκάδα

Παραδείγματα

- Υπολογισμός βάσει τυχαίου δείγματος
 - Η απάντηση δεν υπολογίζεται έναντι όλης της βάσης δεδομένων αλλά έναντι ενός τυχαίου δείγματος
 - Το δείγμα πρέπει να είναι αρκετά μεγάλο και αντιπροσωπευτικό για να δώσει έγκυρα αποτελέσματα
 - Αχρηστεύει τους συμπερασμούς με μέσο όρο διότι δεν χρησιμοποιείται πάντα το ίδιο δείγμα
- Εισαγωγή τυχαίου «θορύβου»
 - Σε κάθε στοιχείο της βάσης προστίθεται ένα τυχαίο σφάλμα ϵ με $-\epsilon \leq \epsilon \leq \epsilon$
 - Στατιστικώς το αποτέλεσμα δεν επηρεάζεται
- Μνήμη αποτελεσμάτων ερωτήσεων
 - Απομνημόνευση των συνόλων που συνθέτουν την απάντηση σε ερωτήσεις και απαγόρευση εμφάνισης απαντήσεων που βασίζονται σε παραπλήσια σύνολα

Παραδείγματα

- Συνάθροιση
 - Συλλογή στοιχείων από διαφορετικές βάσεις δεδομένων ή διαφορετικούς χρήστες και κατόπιν συγκερασμός των αποτελεσμάτων
 - Ιδιαίτερο ενδιαφέρον μετά την εμφάνιση τεχνολογιών εξόρυξης δεδομένων

Υποχρεωτικός έλεγχος προσπέλασης με πολυεπίπεδη ασφάλεια

- Υποχρεωτικός έλεγχος προσπέλασης: κάθε δεδομένο έχει μία διαβάθμιση και κάθε χρήστης ένα επίπεδο εξουσιοδότησης
 - Η ανάγνωση επιτρέπεται μόνο αν η εξουσιοδότηση του χρήστη είναι μεγαλύτερη ή ίση από τη διαβάθμιση του δεδομένου
 - Η εγγραφή επιτρέπεται μόνο αν η εξουσιοδότηση του χρήστη είναι μικρότερη ή ίση από τη διαβάθμιση του δεδομένου
 - παραδείγματα διαβάθμισης-εξουσιοδότησης (άκρως απόρρητο, απόρρητο, εμπιστευτικό, αδιαβάθμιτο)

Διακριτότητα χαρακτηρισμών ασφάλειας

- Ο χαρακτηρισμός ασφάλειας για τα δεδομένα εφαρμόζεται:
 - σε επίπεδο πλειάδας
 - » η πλειάδα που περιγράφει την *Επιχείρηση Moonraker* είναι άκρως απόρρητη, αυτή που περιγράφει την *Καθαροί δρόμοι 2003* είναι αδιαβάθμιτη
 - Σε επίπεδο ιδιοχαρακτηριστικού (attribute)
 - » Το ιδιοχαρακτηριστικό *προσπολισμός* για την πλειάδα *ειδικά κονδύλια υπουργείου Άμυνας* είναι άκρως απόρρητο – το ίδιο ιδιοχαρακτηριστικό για την πλειάδα *συνδεδεμένες και συρραπτικά υπουργείου Άμυνας* είναι αδιαβάθμιτο
 - Σε επίπεδο συνδυασμού ιδιοχαρακτηριστικών
 - » Έστω ιδιοχαρακτηριστικά *εργοδότης* και *θέση* με δυνατές τιμές (Ocean Travel, CIA) και (γραμματέας, πράκτορας, διευθυντής). Το κάθε ιδιοχαρακτηριστικό είναι εμπιστευτικό, ο συνδυασμός τους είναι άκρως απόρρητο
 - Σε επίπεδο στατιστικών μεγεθών
 - » μέσοι όροι, μέγιστα, ελάχιστα, αθροίσματα μπορεί να έχουν διαφορετικό επίπεδο διαβάθμισης από τα επί μέρους στοιχεία βάσει των οποίων υπολογίζονται – π.χ. ο συνολικός προϋπολογισμός του υπουργείου Εθνικής Άμυνας είναι αδιαβάθμιτος, το κονδύλι που κατανέμεται στα πυρηνικά όπλα είναι άκρως απόρρητο. Οι επί μέρους προμήθειες καυσίμων για τους πυραύλους είναι εμπιστευτικές

Γενικό σχήμα για πολυεπίπεδη ασφάλεια

- Συμπλήρωση σχήματος σχέσης με χαρακτηρισμό ασφαλείας ανά ιδιοχαρακτηριστικό και ανά πλειάδα
 - $R(A_1, C_1, A_2, C_2, \dots, R_n, C_n, C_T)$
 - $C_T \geq C_i, 1 \leq i \leq n$
- **Φαινόμενο κλειδί:** τα ιδιοχαρακτηριστικά που θα αποτελούσαν το πρωτεύον κλειδί της σχέσης σε περιβάλλον χωρίς πολυεπίπεδη ασφάλεια
- Τα περιεχόμενα της βάσης παρουσιάζονται διαφορετικά σε χρήστες με διαφορετικό επίπεδο εξουσιοδότησης – διαδικασία *φιλτραρίσματος*

Παράδειγμα

■ Σχέση Εργαζόμενος

Όνομα	Μισθός	Θέση	C_T
Moneypenny	U 5000	C Secretary	U C
Bond, James	C 7000	S Secret Agent	TS TS

■ Επιλογή από χρήστη με επίπεδο εξουσιοδότησης C

Όνομα	Μισθός	Θέση	C_T
Moneypenny	U 5000	C Secretary	U C
Bond, James	C null	C null	C C

■ Επιλογή από χρήστη με επίπεδο εξουσιοδότησης U

Όνομα	Μισθός	Θέση	C_T
Moneypenny	U null	U Secretary	U U

Πολυστιγμιουτυπία

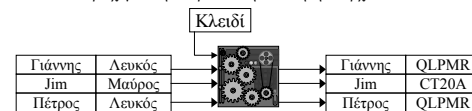
- Χρήστης με επίπεδο εξουσιοδότησης U εισάγει την πλειάδα ('Bond, James', U, 6000, U, 'Commander', U, U)
 - Αν το σύστημα απορρίψει την εισαγωγή «προδίδει» την ύπαρξη μιας πλειάδας με υψηλότερο επίπεδο ασφαλείας
 - Τελικά καταλήγουμε σε ύπαρξη περισσότερων πλειάδων με διαφορετικά επίπεδα ασφαλείας
 - Αντίστοιχα αν ο ίδιος χρήστης επιχειρήσει να ενημερώσει τον μισθό της Moneypenny

Όνομα	Μισθός	Θέση	C_T
Moneypenny	U 5000	C Secretary	U C
Bond, James	C 7000	S Secret Agent	TS TS
Bond, James	U 6000	U Commander	U U
Moneypenny	U 4000	U Secretary	U U

- Η πολυστιγμιουτυπία εισάγει απρόβλεπτο βαθμό πλεονασμού

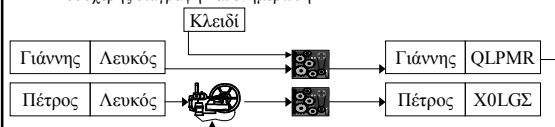
Τεχνικές σε ασφαλείς βάσεις δεδομένων – Κρυπτογραφία (1/2)

- **Κρυπτογραφία** – κρυπτογράφηση των ευαίσθητων δεδομένων ώστε να μην είναι χρήσιμα σε μη εξουσιοδοτημένα άτομα, ακόμη και αν αποκαλυφθούν
 - Χρήση του ίδιου κλειδιού για κρυπτογράφηση όλων των εγγραφών
 - » για πεδία με μικρό πλήθος διακριτών τιμών, είναι εύκολο να βρεθεί η αντιστοίχια μεταξύ κρυπτογραφημένου και μη κρυπτογραφημένου κειμένου (επίθεση με επιλεγμένο μη κρυπτογραφημένο κείμενο)
 - » Με γνωστή την αντιστοίχια ο «ακακός» μπορεί να αποκωδικοποιήσει τα περιεχόμενα ή να τροποποιήσει τα ήδη υπάρχοντα



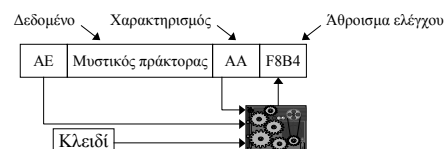
Τεχνικές σε ασφαλείς βάσεις δεδομένων – Κρυπτογραφία (2/2)

- Χρήση διαφορετικού κλειδιού για κάθε εγγραφή
 - απαιτείται σχήμα δημιουργίας-αποθήκευσης-ανάκτησης κλειδιών, επιπλέον αποθηκευτικός χώρος και πρόσθετη επιβάρυνση εισόδου-εξόδου
- **Αλυσιδωτή κρυπτογράφηση**
 - Το δεν κρυπτογραφείται η ίδια η τιμή του ιδιοχαρακτηριστικού A_3 της εγγραφής $v+1$, αλλά το $f(A_3[v], \text{cipher}(A_3[v]))$
 - απαιτείται να διαβαστούν και να αποκρυπτογραφηθούν όλες οι προηγούμενες εγγραφές
 - δυσχερής διαγραφή και ενημέρωση



Τεχνικές σε ασφαλείς βάσεις δεδομένων – Κλειδωμα χαρακτηρισμού ευαισθησίας

- Προστασία του χαρακτηρισμού ευαισθησίας από τροποποίηση
- Συνάρτηση μοναδικού προσδιοριστή (π.χ. αριθμός εγγραφής), χαρακτηρισμού ευαισθησίας και κρυπτογραφικού κλειδιού



Τεχνικές σε ασφαλείς βάσεις δεδομένων – Κλείδωμα ακεραιότητας

- Κάθε ευαίσθητο δεδομένο έχει το χαρακτηρισμό επιπέδου ευαισθησίας που φυλάσσεται μαζί με αυτό
- Η τιμή του δεδομένου φυλάσσεται σε μη κρυπτογραφημένη μορφή για λόγους απόδοσης
- Υπολογίζεται ένα άθροισμα ελέγχου πάνω στο ζεύγος (τιμή, χαρακτηρισμός ευαισθησίας) που διαφυλάσσει την ότι δεν θα τροποποιηθούν αυθαίρετα τα στοιχεία

