### 

### **3413ICT Network Security**

### **Workshop – 12A**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Review Questions:**  The review questions are based on the lecture of database security.   1. By an example, explain why inference attacks rely on certain “metadata”.   CREATE view V1 AS  SELECT Availability, Cost FROM Inventory WHERE Department = ”hardware”  CREATE view V2 AS  SELECT Item, Department  FROM Inventory  WHERE Department = ”hardware”   |  |  |  |  | | --- | --- | --- | --- | | **Item** | **Availability** | **Cost ($)** | **Department** | | Shelf support | in-store/online | 7.99 | hardware | | Lid support | online only | 5.49 | hardware | | Decorative chain | in-store/online | 104.99 | hardware | | Cake pan | online only | 12.99 | housewares | | Shower/tub cleaner | in-store/online | 11.99 | housewares | | Rolling pin | in-store/online | 10.99 | housewares |  1. Inventory table  |  |  | | --- | --- | | **Availability** | **Cost ($)** | | in-store/online | 7.99 | | online only | 5.49 | | in-store/online | 104.99 | | **Item** | **Department** | | Shelf support | hardware | | Lid support | hardware | | Decorative chain | hardware |  1. Two views  |  |  |  |  | | --- | --- | --- | --- | | **Item** | **Availability** | **Cost ($)** | **Department** | | Shelf support | in-store/online | 7.99 | hardware | | Lid support | online only | 5.49 | hardware | | Decorative chain | in-store/online | 104.99 | hardware |   (c) Table derived from combining query answers   1. Consider the following table.   ***RowID***  **Name**  **Position**  **Section**  **Age**  **Gender**  **Salary**  ***R1***  Adam  Officer  Marketing  29  M  48500  ***R2***  Natasha  Manager  Operations  33  F  62000  ***R3***  Sarah  Officer  Operations  27  F  51000  ***R4***  Smith  Clark  Marketing  24  M  42000  ***R5***  Liz  Officer  Operations  27  F  38500  ***R6***  Martha  Officer  Marketing  32  F  44500  ***R7***  Brown  Clark  Marketing  27  M  33000  ***R8***  Carlyle  Manager  Marketing  35  M  58000  ***R9***  Diane  Clark  Operations  30  F  34500  ***R10***  Joe  Officer  Operations  34  M  37500    The salary of a staff is regarded as sensetive information. Any query asking the system to return the salary of an inidividual staff member will be denied. Under this restriction, can you still get the salary of Joe by inference? If so, explain how you can do so.  select SUM(Salary)  from Employee  where Section = ‘Operations’ AND Gender = ‘M’;   1. For the same table in Question 2, describe a second method to access Joe’s salary. What extra knowledge would you need for this method to be successful?   We would need to know if Joe is the only male on his team. Inference can be used to do 2 queries to find his salary:  select SUM(Salary)  from Employee  where Section=‘Operations’ AND Position=‘Officer’ AND  Gender=‘F’;  select SUM(Salary)  from Employee  where Section = ‘Operations’ AND Position = ‘Officer’;  And then subtract the second query and the first to find Joe’s salary.   1. Name and explain two inference detection approaches. What is the main difference between these two approaches?   **Inference detection during database design:** This approach removes an inference channel by altering the database structure or by changing the access control regime to prevent inference. Examples include removing data dependencies by splitting a table into multiple tables or using more fine-grained access control roles in an RBAC scheme. Techniques in this category often result in unnecessarily stricter access controls that reduce availability.  **Inference detection at query time:** This approach seeks to eliminate an inference channel violation during a query or series of queries. If an inference channel is detected, the query is denied or altered.   1. Differentiate between the two types of statistical databases. Would you think the table given in Question 2 is from a pure statistical database? Why or why not?   **Pure statistical database:** This type of database only stores statistical data. An example is a census database. Typically, access control for a pure SDB is straightforward: certain users are authorized to access the entire database.  **Ordinary database with statistical access:** This type of database contains individual entries; this is the type of database discussed so far in this chapter. The database supports a population of nonstatistical users who are allowed access to selected portions of the database using discretionary access control (DAC), role-based access control (RBAC), or mandatory access control (MAC). In addition, the database supports a set of statistical users who are only permitted statistical queries. For these latter users, aggregate statistics based on the underlying raw data are generated in response to a user query, or may be precalculated and stored as part of the database.  The table above is not a purse statistical databse because it contains other data such as   1. Explain the following:  * Query restriction   Rejects a query that can lead to a compromise. The answers provided are accurate. Query restriction techniques defend against inference by restricting statistical queries so that they do not reveal user confidential information. Restriction in this context simply means that some queries are denied. The simplest form of query restriction is query size restriction. For a database of size *N* (number of rows, or records), a query *q*(*C*) is permitted only if the number of records that match *C* satisfies: k <= |X(C)| <= N-k, where *k* is a fixed integer greater than 1. Thus, the user may not access any query set of less than *k* records. Note that the upper bound is also needed. The upper bound of *N* – *k* guarantees that the user does not have access to statistics on query sets of less than *k* records. In practice, queries of the form *q*(*All*) are allowed, enabling users to easily access statistics calculated on the entire database. Query size restriction counters attacks based on very small query sets.   * Perturbation: Provides answers to all queries, but the answers are approximate, due to the addition of noise to the statistics generated from the original data. This can be done in one of two ways: the data in the SDB can be modified (perturbed) so as to produce statistics that cannot be used to infer values for individual records( **data perturbation)**. Alternatively, when a statistical query is made, the system can generate statistics that are modified from those that the original database would provide, again thwarting attempts to gain knowledge of individual records (**output perturbation)**. Regardless of the specific perturbation technique, the designer must attempt to produce statistics that accurately reflect the underlying database.        1. Explain why query restriction itself can cause some security issue.   The denial of a query may provide sufficient clues that an attacker can deduce underlying information. This is generally described by saying that query denial can leak information.   1. Is the query set overlap control a good solution to database security in an e-commerce system for which efficiency is the first priority? Why or why not?   Nope, additional overhead is required to process the queries which is not ideal for an eCommerce system.   1. Name and explain an attack which can break query size restriction.   This control mechanism ineffective for preventing the cooperation of several users to compromise the database. |

**Challenging task:**

Discuss how eBay database was recently compromised? What were the risks and what controls were in place? Clearly explain how this could have been mitigated? List at least three good references (300 words).