



Sicherheit bei Informationsflüssen

Xuejun Li

Universität Karlsruhe

Sommersemester 2006

Gliederung

1. Einstieg
2. Sicherheitsmodelle
 - Gittermodell
 - Bell-LaPadula Modell
3. Einschränkungen von Statistiken
 - Query-set-size-control
 - Query-set-overlap-control
 - Zellenunterdrückung
 - Partitionierung

Gliederung

4. Modifikationen von Statistiken

- Allgemeines
- Datenstörung
- Austausch von Daten

5. Zusammenfassung

6. Diskussion

Einstieg

Motivation

- Gewährleistung für Datenschutz in Informationssystem
- Analyse von explizite und implizite Informationsfluss
- Eine sicherere Anwendung mit Sicherheitsmodell aufzubauen
- Sichere Anwendung in statischen Datenbank

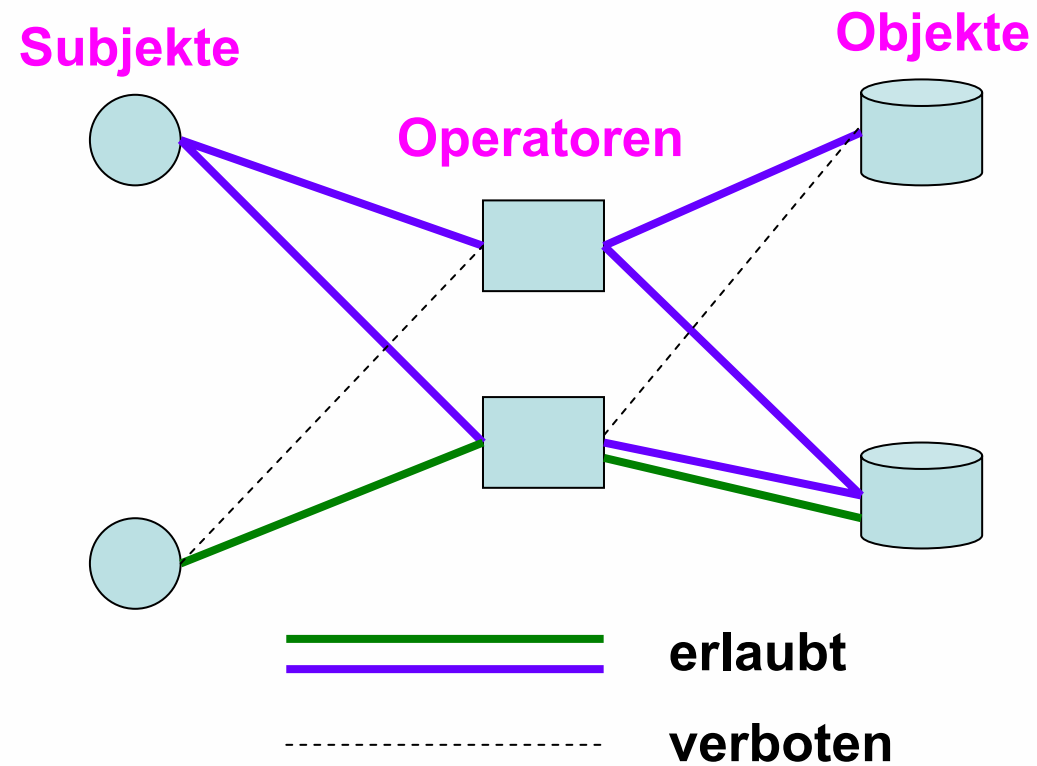
Einstieg

1. Zugriffskontrollstrategien

- Benutzerbestimmte Zugriffskontrolle
 - Discretionary Access Control (DAC)
- Systembestimmte Zugriffskontrolle
 - Mandatory Access Control (MAC)
- Rollenbasierte Zugriffskontrolle
 - Role Based Access Control (RBAC)

Einstieg

2. Grundmodell Zugriffskontrolle



Sicherheitsmodelle

1. Gittermodell(Lattice Model) oder Verbandsmodell

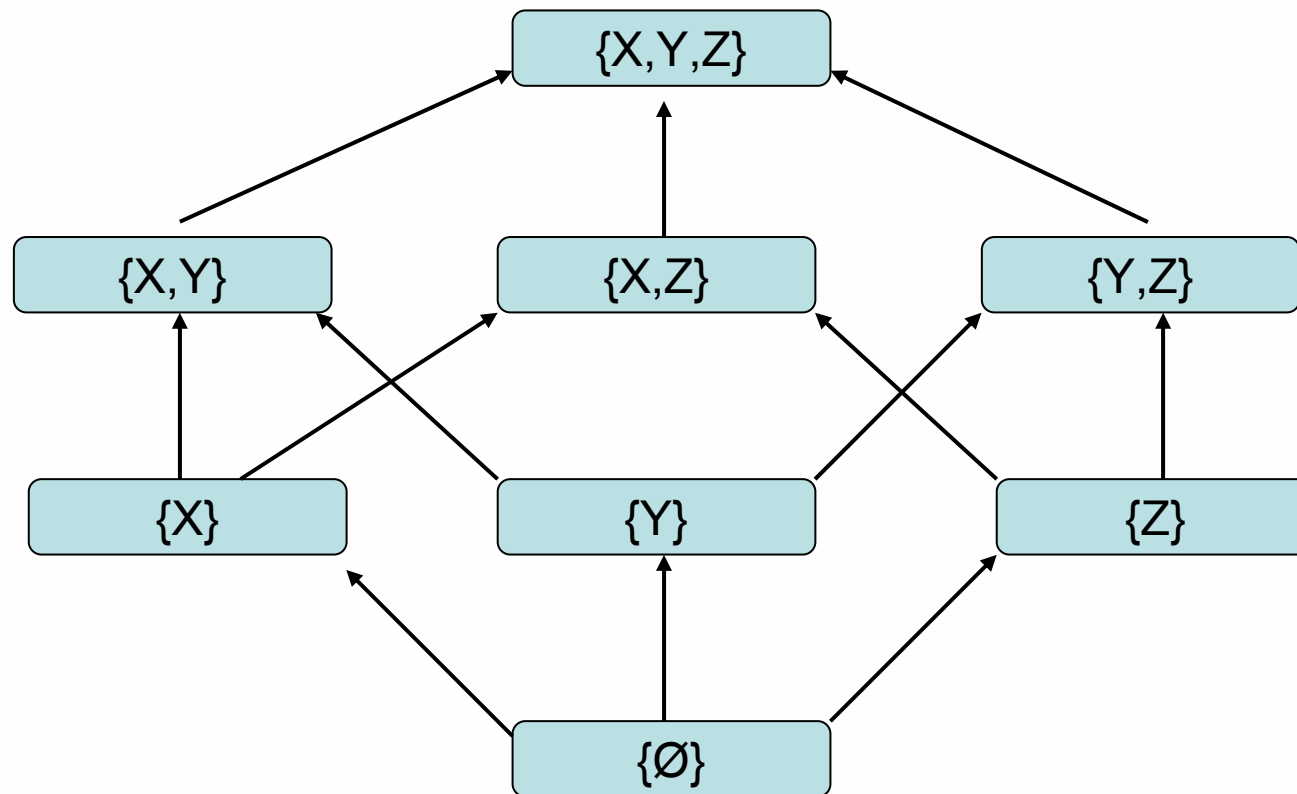
Voraussetzung:

- (M, \leq) eine geordnete Menge und $A \subseteq M$
- s obere Grenze oder Supremum von A , $s = \sup A$
- t untere Grenze oder Infimum von A , $t = \inf A$

Definition: Eine geordnete Menge (M, \leq) heißt ein Verband, wenn jede aus zwei Elementen bestehende Teilmenge von M ein Supremum und ein Infimum besitzt.

Sicherheitsmodelle

Gitterbeispiel



Sicherheit bei Informationsflüssen

Sicherheitsmodelle

Zugriffsmatrix-Modell(ZM)(engl.Access Matrix Model)

Charakteristika

(Dynamische)Menge von **Objekten** O_t

(Dynamische)Menge von **Subjekten** S_t mit: $S_t \subseteq O_t$

Menge von Rechten R

Zugriffsmatrix $M_t : S_t \times O_t \rightarrow 2^R$

$$M_t(S_2, O_2) = \{R_1, R_2\}$$

d.h. zum Zeitpunkt t
besitzt **Subjekt** S_2 die
Rechte R_1, R_2 an
Objekt O_2

	Objekte		
Subjekte	O_1	O_2	O_3
S_1			
S_2		$\{R_1, R_2\}$	
S_3			

Sicherheitsmodelle

Bell-LaPadula-Modell(1973)

Erstes vollständig formalisiertes Sicherheitsmodell

Zugriffsrechte

$R = \{\text{read-only, append, execute, read-write, control}\}$

Menge von Sicherheitsklassen SC , $X \in SC$ mit $X=(A,B)$

A ist Sicherheitsmarke, z.B. vertraulich, geheim

B ist eine Menge von Kategorien, z.B. Arzt, Schwester

$\forall s \in S : Clearance : SC(s) \in SC$; maximale $SC(s)_{MAX}$ und
aktuelle $SC(s)_{AKT}$

$\forall o \in O : Classification : SC(o) \in SC$

Bell-LaPadula Modell

Partielle Ordnung auf $SC : (SC, \leq)$, für $X, Y \in SC$ gilt :

$$X = (A, B), Y = (A', B') \quad X \leq Y \Leftrightarrow A \leq A' \wedge B \subseteq B'$$

Beispiel: Krankenhaus, Umgang mit Patientenakten

Sicherheitsmarken:

{unklassifiziert, vertraulich, geheim, streng geheim}

Mit Ordnung:

unklassifiziert \leq vertraulich \leq geheim \leq streng geheim

Sicherheitskategorien:

{Arzt, Schwester, Patient, Verwaltung}

Menge der Sicherheitsklassen SC:

SC: {(geheim, \emptyset), (vertraulich, \emptyset),
(vertraulich, {Arzt, Schwester}), ..., }

Bell-LaPadula-Modell

Wegen Ordnungsrelation gilt u.a.

$(\text{geheim}, \emptyset) \succeq (\text{vertraulich}, \emptyset)$

$(\text{vertraulich}, \{\text{Arzt}, \text{Schwester}\}) \succeq (\text{vertraulich}, \{\text{Schwester}\})$

Bell-LaPadula-Modell

Systembedingte Zugriffsbeschränkungen

No-read-up (Simple-Security)-Regel

Lese- oder Execute-Zugriff auf ein Objekt o durch Subjekt s nur zulässig, wenn s das entsprechende Zugriffsrecht r besitzt und die Objektklassifikation kleiner oder gleich der Subjekt-Clearance ist.

$$r \in M_t(s, o) \wedge SC(s) \geq SC(o)$$

No-write-down-Regel(*-Eigenschaft)

Append-Zugriff auf ein Objekt o durch Subjekt s nur zulässig wenn die Objektklassifikation mindestens so hoch wie die Subjekt-Clearance ist.

$$\text{append} \in M_t(s, o) \wedge SC(s) \leq SC(o)$$

Bell-LaPadula-Modell

No-write-down-Regel(*-Eigenschaft)

Lese-Schreib-Zugriff auf ein Objekt o durch Subjekt s nur zulässig, wenn die Objektklassifikation gleich der Subjekt-Clearance ist.

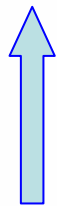
$$\text{read - write} \in M_t(s, o) \wedge SC(s) = SC(o)$$

Bell-LaPadula-Modell

Die zwei systembedingten Regeln sind leicht zu überprüfende Bedingungen

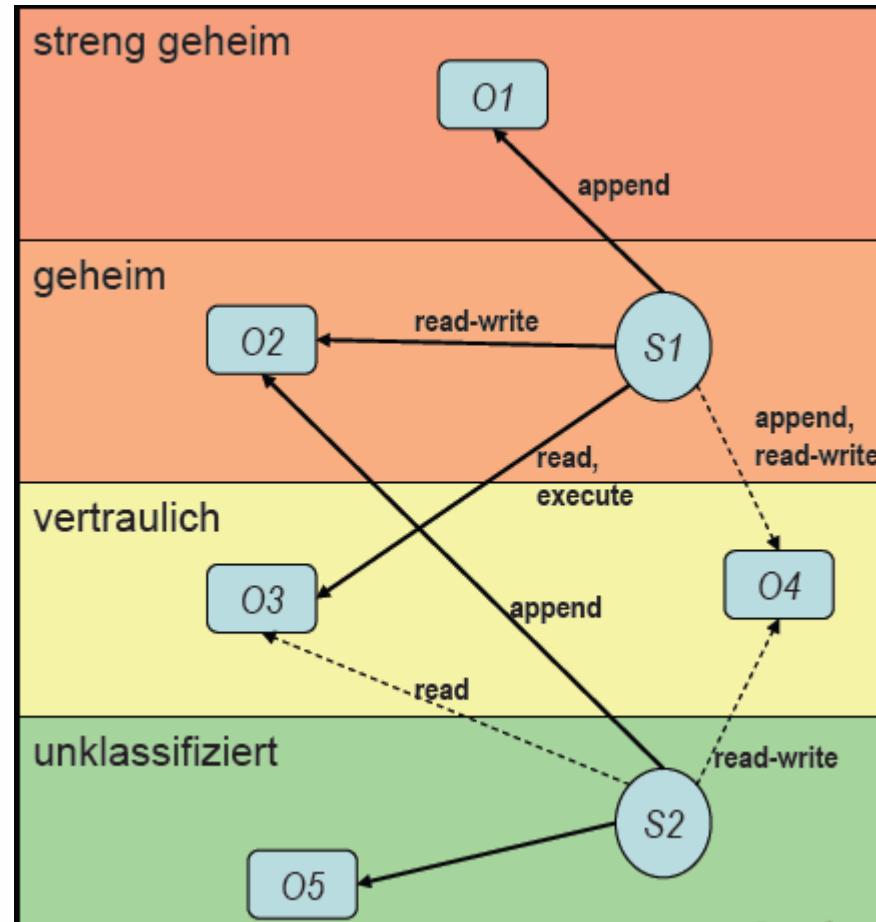
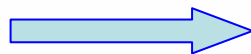
Informationsflüsse

höchstens von



- unten nach oben gemäß der Partiellen Ordnung oder

- Innerhalb einer Sicherheitsklasse



Bell-LaPadula-Modell-Grenze

Bell-LaPadula-Modell häufig in der Praxis eingesetzt, hat aber gravierende Mängel:

Problem des blinden Schreibens

- Systembestimmte Regeln erlauben Schreiben in ein höher eingestuftes Objekt, aber kein (Kontroll-)Lesen der Veränderungen
- Integritätsproblem

Sukzessive Höherstufung von Informationen

- Subjekte können Information von Objekten niedriger Klasse aufnehmen, diese jedoch nicht an sie zurückfließen lassen
- Zwei Möglichkeiten, wenn solcher Nachrichtenfluss explizit gewünscht
 - Temporäres Herabstufen für solche Subjekte (downgrade)
 - Einführung von "Trusted Subjects" wie Systemprozessen, die die No-Write-Down-Regel umgehen können.

Bell-LaPadula-Modell

Anwendung

- Viele Betriebssysteme bieten BLP-Erweiterungen: u.a.
 - Sun Trusted Solaris 7, Trusted HP-Unix, Linux-Derivate
 - Erweiterter Referenz-Monitor zur Überprüfung der BLP-Regeln
- System Z (John McLean):
 - Erweitertes BLP - System
 - Klassifikation einer Datei kann temporär herabgesetzt werden

Einschränkungen von Statistiken

Abfragensatz-Größe-Kontrolle (Query-set-size-control)

Annahme: Nutzer kennt ein Individuum I , welches in der Datenbank repräsentiert wird und eine charakteristische Formel C erfüllt

$\text{count}(C) = 1 \Rightarrow I$ ist eindeutig identifizierbar

$\text{count}(C \bullet D) = 1 \Rightarrow I$ hat D

$\text{count}(C \bullet D) = 0 \Rightarrow I$ hat D nicht

$\text{sum}(C, A) =$ der Wert A für I

Name	Sex	Major	Class	SAT	GP
Allen	<i>Female</i>	<i>CS</i>	1980	600	3.4
Baker	<i>Female</i>	<i>EE</i>	1980	520	2.5
Cook	<i>Male</i>	<i>EE</i>	1978	630	3.5
Davis	<i>Female</i>	<i>CS</i>	1978	800	4.0
Evans	<i>Male</i>	<i>Bio</i>	1979	500	2.2
Frank	<i>Male</i>	<i>EE</i>	1981	580	3.0
Good	<i>Male</i>	<i>CS</i>	1978	700	3.8
Hall	<i>Female</i>	<i>Psy</i>	1979	580	2.8
Iles	<i>Male</i>	<i>CS</i>	1981	600	3.2
Jones	<i>Female</i>	<i>Bio</i>	1979	750	3.8
Kline	<i>Female</i>	<i>Psy</i>	1981	500	2.5
Lane	<i>Male</i>	<i>EE</i>	1978	600	3.0
Moore	<i>Male</i>	<i>CS</i>	1979	650	3.5

Einschränkungen von Statistiken

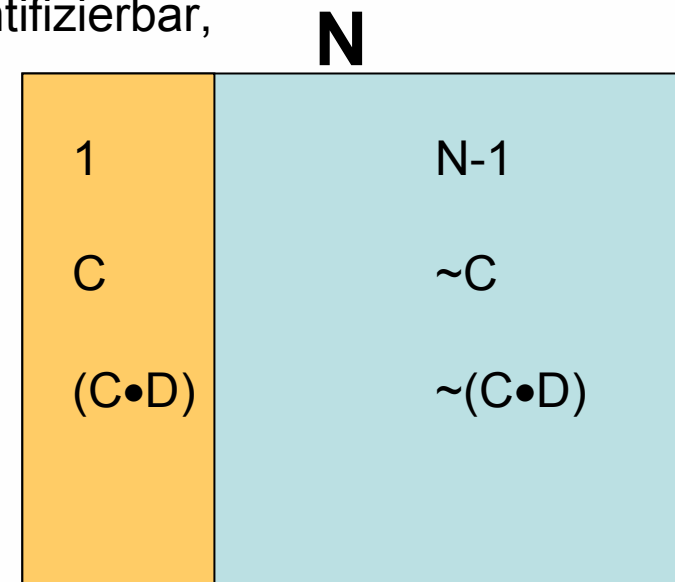
Abfragensatz-Größe-Kontrolle (Query-set-size-control)

- kleine query sets müssen gesperrt werden
aber: auch große query sets müssen gesperrt werden, da:

$\text{count}(\sim C) = N - 1 \Rightarrow I$ ist eindeutig identifizierbar,
wobei $N = \text{count}(All)$

$\text{count}(\sim(C \bullet D)) = N \Rightarrow I$ hat D nicht
 $\text{count}(\sim(C \bullet D)) = N - 1 \Rightarrow I$ hat D

$\text{sum}(C, A) = \text{sum}(All, A) - \text{sum}(\sim C, A)$



Einschränkungen von Statistiken

Abfragensatz-Größe-Kontrolle (Query-set-size-control)

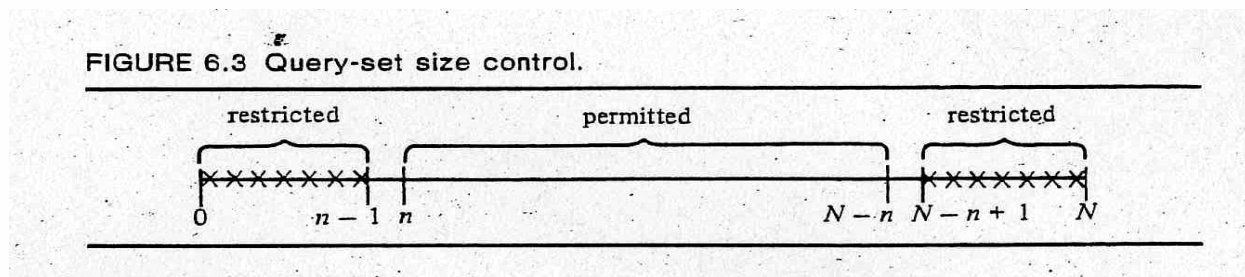
- Statistik ist erlaubt, wenn

$$n \leq |C| \leq N - n \quad \text{gilt}$$

N: Anzahl der Entities von SDB

|C| : Größe der charakteristischen Formel

$1 \leq n \leq N/2$: Parameter der Datenbank



- sehr geschätzt, einfach Implementierbar, aber ungenügend

Einschränkungen von Statistiken

Abfragensatz-Überlagerungs-Kontrolle (Query-set-overlap-control)

- Eine Statistik $q(C)$ ist nicht erlaubt, wenn

$$q(C) \geq \{\text{untere Grenze von } (1+(K-1)/r)\},$$

wobei $K = \min\{q(C), q(D), \dots\}$ und $r = \{q(C) \cap q(D) \cap \dots\}$

- Verschlechterung des Gebrauchs von Datenbanken
- Methode verhindert nur wenige Angriffe

Einschränkungen von Statistiken

Partitionierung

- Unterteilung der dynamischen Datenbank in zusammenhanglose Gruppen
 - Jede Gruppe G besitzt $g = |G|$ Einträge, wobei $g=0$ oder $g \geq n$ und g ist gerade
 n ist die Anzahl von Teilmengen in G
 - Paarweises Hinzufügen oder Löschen von Einträgen in G
 - Abfragesätze müssen die gesamte Gruppe umschließen

Einschränkungen von Statistiken

Partitionierung

Beispiel

TABLE ASE		AGE			
		0-20	21-45	46-65	>65
UNEMPLOYED	M	24	2	9	49
	F	26	0	1	51
ABC-COMPANY	M	0	1	9	0
	F	0	16	0	0
XYZ-INC	M	1	20	48	0
	F	1	0	52	0

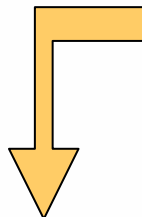


TABLE ASE		AGE			
		0-20	21-45	46-65	>65
UNEMPLOYED	M	24	2	10	49
	F	26	0		51
ABC-COMPANY	M	0	17	9	0
	F	0		0	0
XYZ-INC	M	2	20	48	0
	F		0	52	0

Einschränkungen von Statistiken

Zellenunterdrückung

- findet Anwendung bei der Volkszählung um veröffentlichte Daten zu schützen
- Limitierung der Zellunterdrückung durch die Computerkomplexität der Ausführungsprozedur erfolgreich bei 2- und 3-dimensionalen Tabellen hinzugefügt

Einschränkungen von Statistiken

Zellenunterdrückung

Zum Beispiel:

TABLE 6.3 Student counts by *Sex* and *Class*.

Sex	Class				Sum	Total
	1978	1979	1980	1981		
<i>Female</i>	1	2	2	1	6	
<i>Male</i>	3	2	0	2	7	
Sum	4	4	2	3	13	

TABLE 6.4 Total SAT scores by *Sex* and *Class*.

Sex	Class				Sum	Total
	1978	1979	1980	1981		
<i>Female</i>	800	1330	1120	500	3750	
<i>Male</i>	1930	1150	0	1180	4260	
Sum	2730	2480	1120	1680	8010	

Einschränkungen von Statistiken

Zellenunterdrückung

TABLE 6.5 Total SAT scores by Sex and Class.

Sex	Class				Sum
	1978	1979	1980	1981	
<i>Female</i>	—	1330	1120	—	3750
<i>Male</i>	1930	1150	0	1180	4260
Sum	2730	2480	1120	1680	8010
					Total

TABLE 6.6 Total SAT scores by Sex and Class.

Sex	Class				Sum
	1978	1979	1980	1981	
<i>Female</i>	—	1330	1120	—	3750
<i>Male</i>	—	1150	0	—	4260
Sum	2730	2480	1120	1680	8010
					Total

Modifikationen von Statistiken

Allgemeines

- Einschränkungen von Statistiken können teilweise sehr kostspielig und ungenau sein
- Interesse an der Entwicklung von Methoden, welche die Enthüllung durch hinzufügen von Lärm kontrollieren
 - generell effizienter
 - erlauben Freigrabe von mehr nicht empfindlichen Statistiken

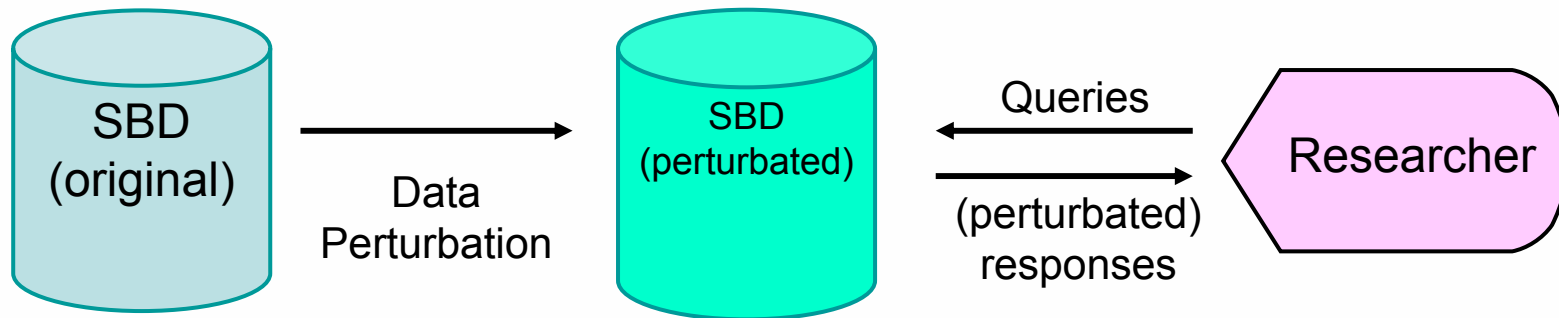
Modifikationen von Statistiken

Datenstörung

- direkt hinzufügen von Geräuschen zu den Datenwerten durch permanente Modifizierung der gespeicherten Daten in der Datenbank oder Störung der Daten wenn sie gerade für die Kalkulation verwendet werden
- beinhaltet die Störung aller Datenwerte x_i , die benötigt werden um eine Statistik $q(C)$ zu einer Funktion $f(x_i)$ zu generieren
 - zur Berechnung wird $x_i' = f(x_i)$ anstelle von x_i benutzt

Modifikationen von Statistiken

Datenstörung



Modifikationen von Statistiken

Austausch von Daten

- Entwicklung eines Datentransformationsschemas, das auf dem Vertauschen der Werte in den Einträgen beruht
- Schlörer definierte eine Datenbank als t -Ordnung Statistik, wenn mindestens noch eine andere Datenbank D' existiert mit
 - D und D' besitzen gleiche Elemente in K
 K ist der Wertbereich von originaler Datenbank
 - Element in D' wird wahllos in K ausgesucht
am Besten besitzt die gleiche Wahrscheinlichkeitsverteilung wie zur originale Datenbank

Modifikationen von Statistiken

Austausch von Daten

TABLE 6.13 A 2-transformable database.

Record	<i>D</i>			<i>D'</i>		
	Sex	Major	GP	Sex	Major	GP
1	<i>Female</i>	<i>Bio</i>	4.0	<i>Male</i>	<i>Bio</i>	4.0
2	<i>Female</i>	<i>CS</i>	3.0	<i>Male</i>	<i>CS</i>	3.0
3	<i>Female</i>	<i>EE</i>	3.0	<i>Male</i>	<i>EE</i>	3.0
4	<i>Female</i>	<i>Psy</i>	4.0	<i>Male</i>	<i>Psy</i>	4.0
5	<i>Male</i>	<i>Bio</i>	3.0	<i>Female</i>	<i>Bio</i>	3.0
6	<i>Male</i>	<i>CS</i>	4.0	<i>Female</i>	<i>CS</i>	4.0
7	<i>Male</i>	<i>EE</i>	4.0	<i>Female</i>	<i>EE</i>	4.0
8	<i>Male</i>	<i>Psy</i>	3.0	<i>Female</i>	<i>Psy</i>	3.0

Modifikationen von Statistiken

Willkürliches Antworten

- bei Befragungen antworten viele Leute nicht wahrheitsgemäß, weil sie Angst vor einem Eingriff in ihre Privatsphäre haben
- Warner entwickelte Technik dies zu umgehen
- Technik wird zum Zeitpunkt der Befragung angewandt



Zusammenfassung

Diskussion



Danke für die Aufmerksamkeit



Zusammenfassung

Diskussion