



Tax Fraud and Investigations

– everybody, everywhere, everytime

Hans - J. Lenz
Inst. f. Statistik und Ökonometrie
Freie Universität Berlin

hans-j.lenz@fu-berlin.de

First comes the Food
and later the Profit.

B. Brecht



Belastingdienst





Agenda

- **Actors, sparse Data, Roles**
- **Sparsity and Plausibility of data**
- **Framework of a model driven Tax Investigation**
– step by step
- **Investigation Procedures**
- **Case Studies of Tax Fraud**
- **Perspectives**



I Actors



an das Finanzamt
Selbstanzeige

CAYMAN-CONTAIN
von Kingston nach Zürich



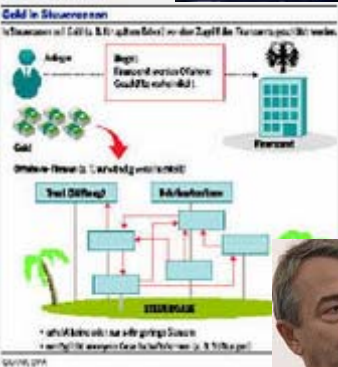
Griechische Reeder



Sehr geehrte Dame



Passat TDI Clean Diesel



Steuerdaten



Hans-J. Lenz ICISSP



Actors, Roles (cont.)





II Sparsity and Plausibility of data

- **Sparse data:** Hunting for hints and facts
(Belief → hints / facts → belief update)
- **Plausibility checks** (cf. Fermi assessment in science)
 - of normative / actual business figures of corporate entities
 - of fiscal income and declared estate with reality and life style of natural persons
- **Rule of Thumb:** Examine appearance and reality
(Schein and Sein) Weinheim (2011)

Problem: Classify a susceptible as „guilty“ or „not guilty/negligently“
given all discovered facts and background information

Tax fraud* = $\max\{0, \text{excisable income} - \text{declared income}\}$

Plausibility of data (cont.)

- **Checks of stocks and flow**

Ex.: $\text{Inventory}_t = \text{Inventory}_{t-1} + \text{Inflow}_t - \text{Outflow}_t$

Net Income = 2000 €m > Penthouse Rent Rate = 1800 €m

- **Checks based on I/O-Relations**

(OFD Köln; Erhard & Wenzig 1985)

Ex.: recorded annual mileage (km/a) of a taxi by speedometer in period t (F);

mean consumption rate $l/100\text{km}$ of taxi type (v);

recorded annual fuel consumption in period t , l/a (V)

Estimate $\hat{V} = F * v$

- **Checks of (pooled) longitudinal and cross-sectional Data**

Ex.: Time-Series of annual fuel consumption rates, $(V_t)_{t=1,2,\dots,n}$

- **Checks** of income difference = excisable income $_t$ - declared income $_t$

- **Alternative (qual.) Checks** Observation of business processes, house to house /room to room or face to face inspection

Fermi assessment

„How many piano tuner in Chicago?”



Enrico Fermi
29.9.1901 –
28.11.1954

No of piano tuners in Chicago (AS) = ?

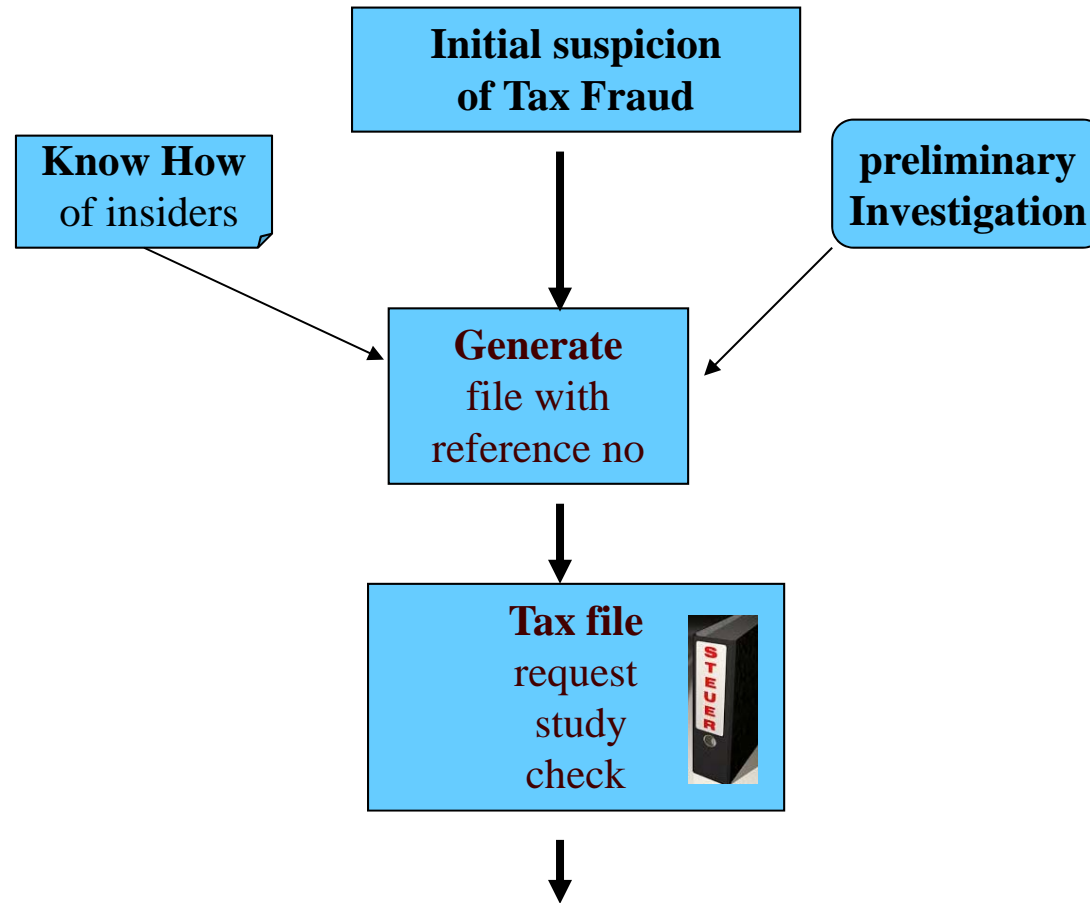
- *Assumption 1*: AE = 4.000.000 population size citizen
- *Assumption 2*: h = 4 mean size of households citizen/h
- *Assumption 3*: k = 1/5 piano density piano/h
AK = AE k/h = 200.000 no of pianos piano
- *Assumption 4*: f = 1/4 annual tuning frequency a⁻¹
JLK = f AK = 50.000 annual tuning load piano/a
- *Assumption 5*: p=4 daily tuning productivity piano/d*tuner
- *Assumption 6*: AA = 250 annual no of working days d/a

$$AS = JLK / (p * AA) = \mathbf{50} \text{ tuners in Chicago}$$

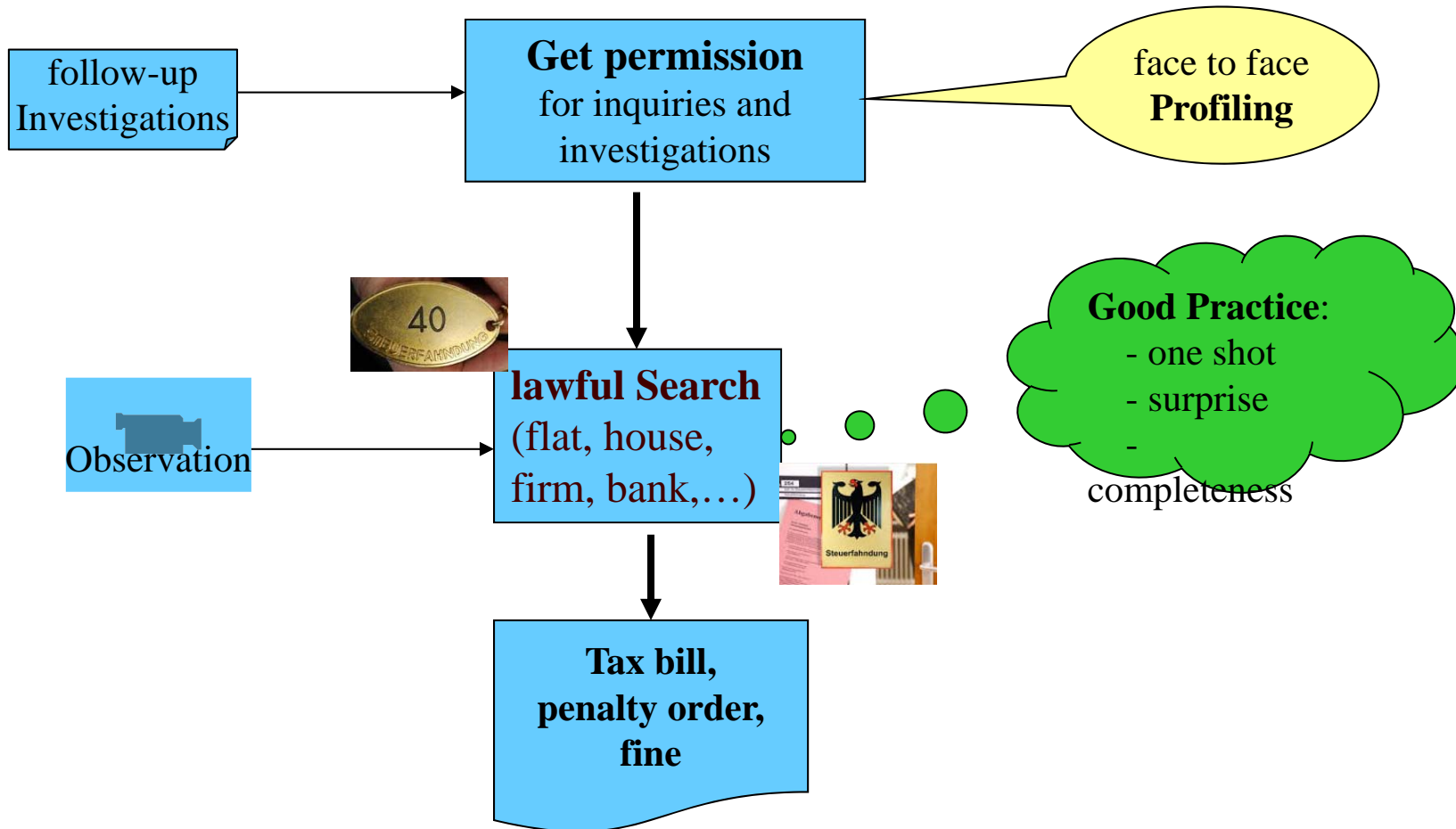
Hans-J. Lenz ICISSP16 Roma 21 Feb 2016

8

III Tax Investigation – step by step (workflow)



Tax Investigation (cont.)



Tax fraud state space

$$\Theta = \mathbf{R}^4_{+} \times N_0 \times \{0,1\}$$

θ_1 : Delinquent tax (€) ! main variable

θ_2 : Penalty (€)

θ_3 : Debit interest charge (€)

θ_4 : Legal Charges (€)

θ_5 : Imprisonment (m)

θ_6 : on probation (yes/no: (1/0))

Ex.: Ulli Hoeness, Munich: (28,4 Mio €, ?, ?, ?, 42 m, 0)



Reference Table

of measures of uncertainty

Probability 0-100%	Linguistic Term
0 %	impossible (excluded)
5 %	extremely impossible
10 %	fairly impossible
30 %	doubtful
50 %	odds are 1:1 (indifferent)
75 %	likely
90 %	almost certain
95 %	very certain (confident)
99 %	extremely certain
100 %	with certainty (totally sure)



IV Investigation Procedures

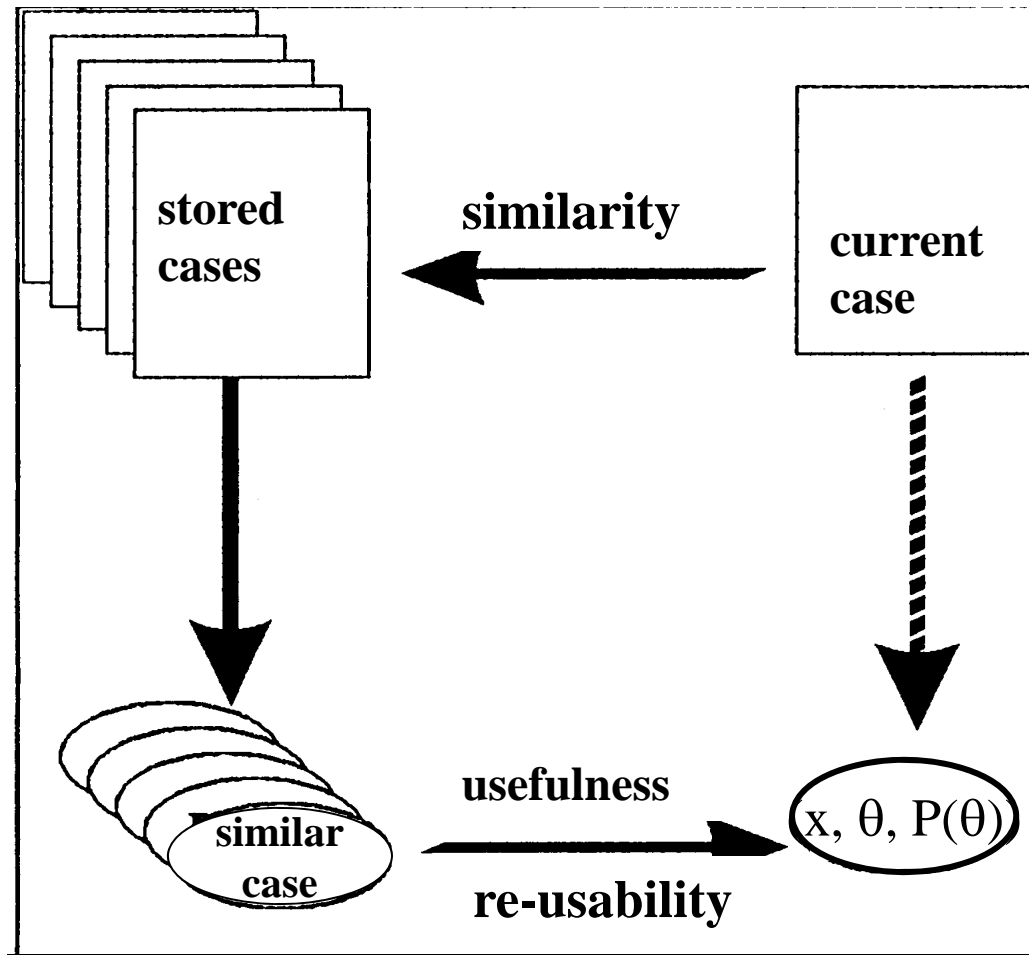
- Rule based Systems (RBS)
- Case based Reasoning (CBR)
- (social) Network Analysis (sNA)
- **Bayesian Learning (BL)**

IVa Rule based systems

(KB,FB,H,WHY/HOW)

- *if* <condition> *then* <conclusion>
- Predicate logic \approx Boolean logic (\wedge, \vee, \neg)
& variables & constants & quantifiers
- KB Knowledge Base (rule set)
- FB Fact Base (data)
- H Hypothesis (suspicion)
- WHY/HOW self explanation facility
- PROLOG Inference component

IVb Case based reasoning (CBR)

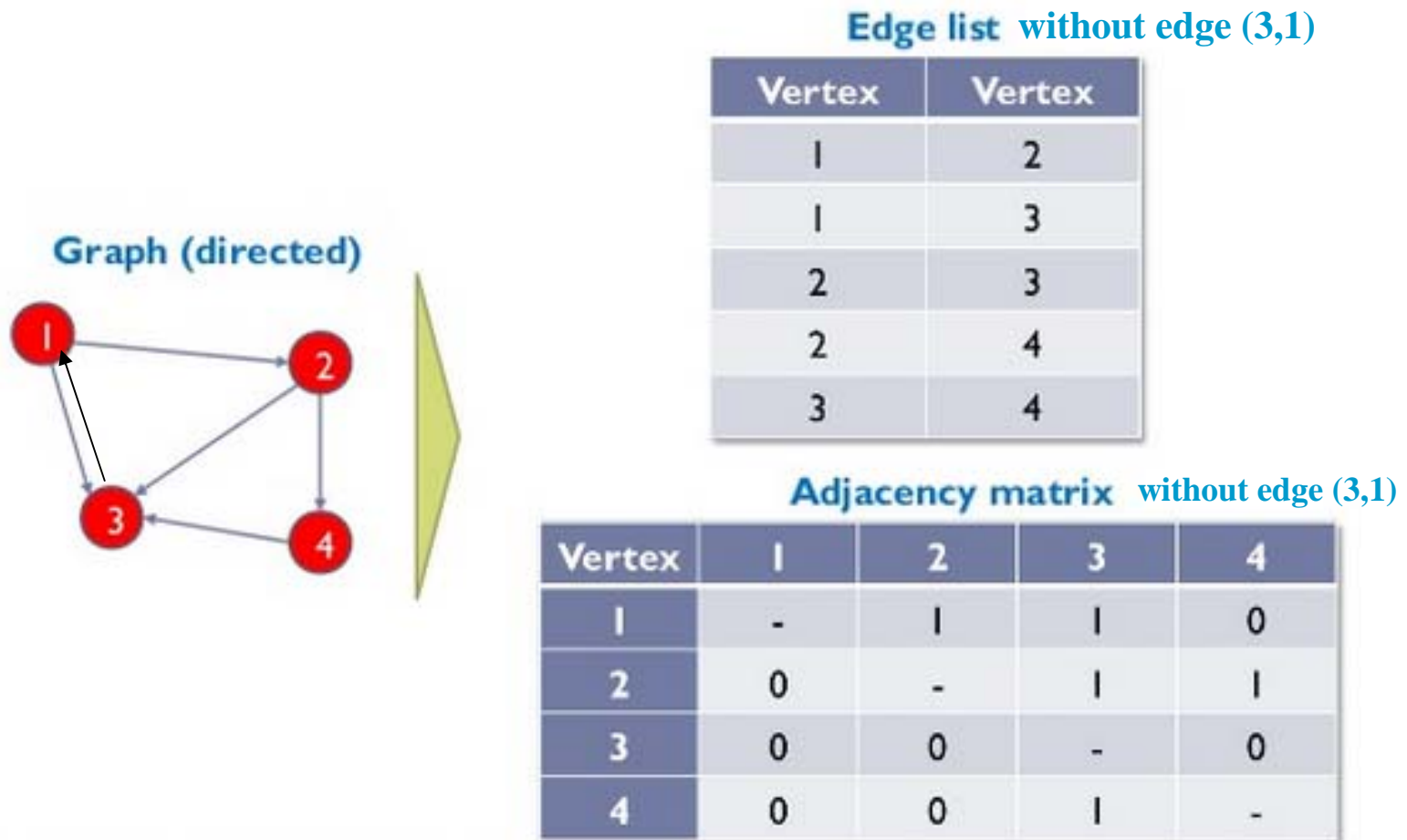


CBR in NL: ~ 150 Business Rules to score Dutch Network of businesses, people, addresses, phone-numbers, bank-accounts etc.

Source: Driessen (2015), Business Intelligence & Analytics, Belastingdienst Netherlands

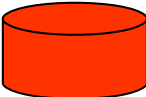
IVc Network Analysis

by weighted *DiGraphs*



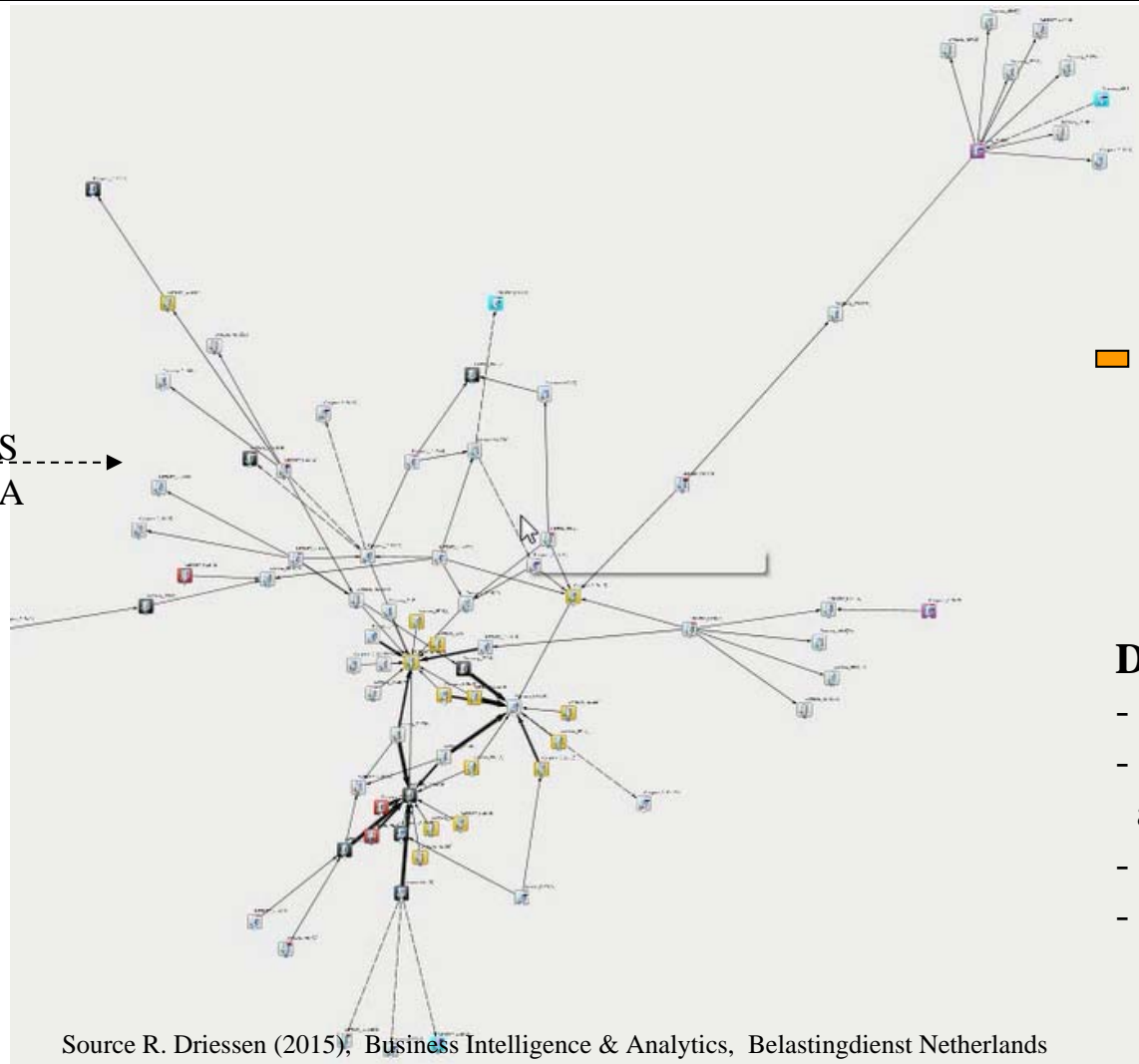
Network Analysis

„Carousel Fraud“ in Netherlands



DB_{NL}
~450 TB

SAS
SNA



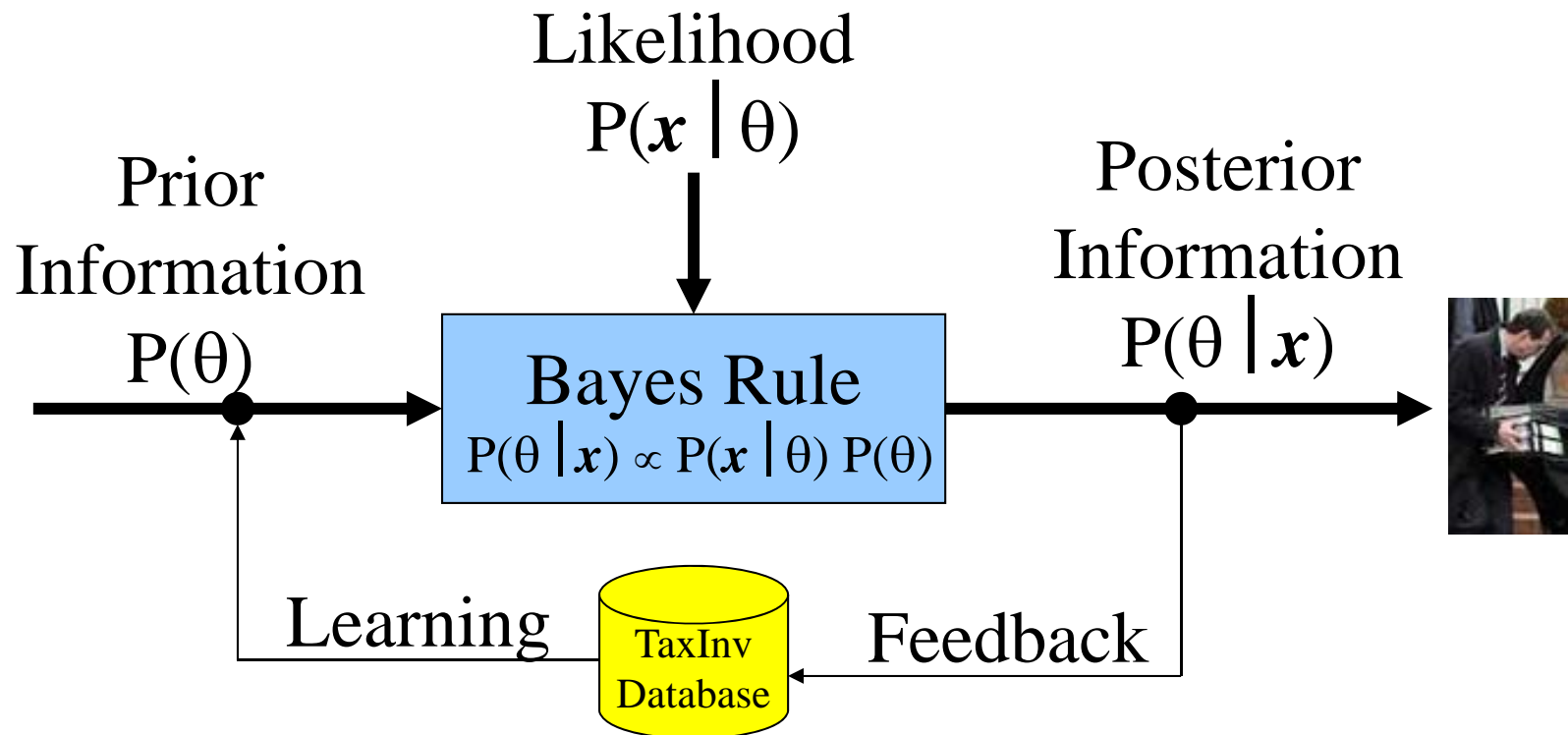
known fraudsters

Data sources:

- VAT-declarations
- client data
addresses/names etc.)
- fraud lists
- communication data

IVd Bayesian Learning

Belief $\pi^* = P(\theta^* | \mathbf{x})$ where $\theta^* = \arg \max_{\theta \in \Theta} P(\theta | \mathbf{x})$



V Case Studies of Tax Fraud

Case 1: Al Capone 1930, 1931



- **Initial Suspicion:** Covering-up tax fraud by investing illegal revenues in laundries (“Money Laundering”),
 $\theta_1 = 200.000 \text{ US } \$$ prior Belief= “extremely certain” = $P(\theta_1) \approx 1$
- **Facts x:** Business of illegal gambling (x_{i1}), prostitution (x_{i2}), protection racket (x_{i3}), illegal dealing with alcohol (x_{i4}) observed but NOT provable: $P(\theta_1 \mid x_{i1}, \dots, x_{i4}) = 0 !$
- **Assessment:** declared income $\Sigma E_{T-\tau} < V_T - V_0 \approx \Sigma_j x_j$ value added wealth in $[0, T]$: $P(\theta_1 \mid (x_i)_{i=T, T-1, \dots}) > 99\%$
- **Punishment:** Tax fine \$ 50.000, 11 years of jail, legal charges \$8.000 $\Rightarrow \theta_{AC} = (?, 50 \text{ T}\$, 0, 8 \text{ T}\$, 132\text{m}, 0)$

Case 2: Burglary



- **Initial Suspicion** (by self report):

Breaking news of media: Plundering objects d'art and assets in flats and houses and acting as „Beagle Boys“,
 $\theta_1 \gg 0$ prior Belief = “likely”

- **Facts:** Declared total income / wealth deviate strongly from size from declared value of stolen goods, i.e.

$X_{\text{Plunder}} \gg X_{\text{income}}$ posterior Belief = “confident”

- **Punishment:** Fine \sim (estimated) delinquent tax

$\theta = (\tilde{\theta}_1, \tilde{\theta}_1, 0, 0, 0, 0)$



Case 3: Insurance fraud of a craftsman by feigned burglary



- **Initial Suspicion** (by tax authority):
Divergence of after-tax profit of income and value-added taxes with lifestyle of a craftsman
(real estate, house, swimming pool, premium class car, ...)
 $\theta_1 \approx 100.000$ DM prior Belief = “likely”
- **Facts:** File inspection gives hints to black money in hand up to ~ 100.000 DM used for investments, but where? $x_{(1)} = \text{null}$ updated Belief = “very certain”

Case 3: Insurance fraud (cont.)

- **Inspection:** Executing a search warrant gives evidence of an insurance fraud by a fictitious break-in,
 $x_{(2)} = \text{null}$ updated Belief = “extremely certain”
- **Facts:** Follow-up audit of the single entry bookkeeping system initiated by pretending money sources like Arabic creditor, lottery, casino;
 $x_{(3)} = \text{null}$ posterior Belief = “extremely certain”
- **Punishment:** Enforceable legal claim on real estate (“dinglicher Arrest”) of 2 Mio DM due to craftsman’s disappearing, suspended sentence and payments of tax arrears summing up to 100.000 DM
 $\theta = (100\text{TDM}, 0, 0, 0, 0, 0)$

Case 4: Business data manipulated by a plumber and heating installer



- **Initial Suspicion** (self-denunciation by the deceived spouse): Missing invoices, (Spouse responsible for bookkeeping, but not subject to prosecution due to self-denunciation (§371 German Law (AO)) $\theta_1 \gg 0$ $P(\theta_1) = 95\%$
- **Facts** (illicit actions):
 - Private acquisitions (antiques, entertainment electronics) and repairs booked as operating costs
 $x_1 = 300$ TDM $P(\theta_1 | x_1) > P(\theta_1)$
 - Depreciation of antiques, fictitious car pool
 - Safe with black money and jewelry, foreign bank account, manipulated accounts, receipts file, notebook with black money entries, $x_2 \approx 90$ TDM
 $P(\theta_1 | x_1, x_2) > P(\theta_1 | x_1)$
- **Punishment:** Payment of arrears 500.000 €, fine 280.000 €
 $\theta = (500 \text{ TDM}, 280 \text{ TDM}, 0, 0, 0, 0)$ Belief = “with certainty”



Case 5: Interception of dental gold



- **First suspicion** (Report by the spouse):
The dentist collected dental gold of his patients over years
 $\theta_1 \approx 100.000 \text{ DM} / 10$ a prior Belief = “doubtful”
- **Facts and Assessment** of θ_1 :
Gold sold at a gold parting factory and paid by crossed checks.
No bookkeeping, but money transfer to a bank in CH for later use
after retirement
Evidence of a well operated dental surgery with dental lab →
Assessment of annual revenues $x = N * g = 285 * 35 \approx 10.000 \text{ DM/a}$
 $P(\theta_1 | x) = 0.99$ or posterior Belief = “extremely certain”
- **Punishment:** Assumed 10-years-tax fraud =
 $10 * \text{annual revenues} = 100.000 \text{ DM}$
 $\theta = (100 \text{ TDM}, 100 \text{ TDM}, 0, 0, 0, 0, 0)$

Case 6: Tax saving at a doctor's surgery



- **Initial Suspicion** (by a excise officer)
Profitable doctor's office + rents from inherited houses.
His spouse manages bookkeeping.
 $\theta_1 = 0$ DM prior Belief = "1:1 odds"
- **Facts:** Tax inspector confirms annual tax balance sheet and profit and loss account. Examining single bookkeeping transactions of the last year he detects transposed digits. Addition gives an annual difference $x_t = 45.000$ DM (149.000 instead of 194.000)
updated Belief = "likely"
- Audit of the last eight years
Update of suspicion: $\theta_1 \gg 45.000$ DM Belief = "extremely certain"
- **Punishment:** Payment of arrears 450 TDM; no fine although digits transposed with intent. But *intention / negligence* difficult to discriminate

$$\theta = (450 \text{ TDM}, 0, 0, 0, 0, 0)$$

Source: Wehrheim 95 (2011)

Case 7: Subsidy fraud construction machine manufacturer



- **First suspicion** (by excise officer)
Subsidy fraud when selling construction machines outside the limited subsidy period.
 $\theta_1 \approx c * 15 \text{ TDM}$ ($c > 0$) Belief = “almost certain”
- **Facts** (confessed by bookkeeper): During audits n excise officer detects two kind of identifiers of invoices:
 1. The values of the sales invoice identifier are sequenced
 2. Invoices backdated to the period of subsidy have the extra label A, B, C,... $x_1 = 20 \times 15000 = 300 \text{ TDM}$ Belief = “totally sure”
- **Punishment:** Payment of arrears 300 TDM and fine 30 TDM
 $\theta = (300 \text{ TDM}, 300 \text{ TDM}, 0, 0, 0, 0)$



Moral of Tax Fraud

- No tax fraudster is able to create a world of facts and figures without any contradiction.
- Truth will come out.
- „In the long run we get all of them.“

F. Wehrheim (2011)

Data Fraud is no trivial offense

