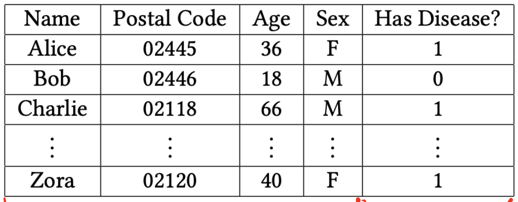
Lecture 3. Reconstruction Attacks. (Part 2)

- Recap on linear reconstruction attacks.
- Reconstruction Attacks w/ less queries
- More efficient attacks
- Reconstruction Attack in practice. (Reading)

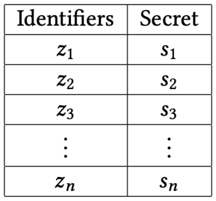
→ Announcement.

Linear Reconstruction Attack

- Introduced by Dinur&Nissim in 2003



Name	Postal Code	Age	Sex	Has Disease?
Alice	02445	36	F	1
Bob	02446	18	M	0
Charlie	02118	66	M	1
⋮	⋮	⋮	⋮	⋮
Zora	02120	40	F	1



Identifiers	Secret
z_1	s_1
z_2	s_2
z_3	s_3
⋮	⋮
z_n	s_n

Z : identifiers Secret bit

Release count statistics : # people satisfy some property

- How many people $\underbrace{\text{are older than 40}}_{\varphi(z_j)}$ & have secret bit =1 ?

inner/dot product →

$$f(X) = \sum_{j=1}^{n} \varphi(z_j)\, s_j \qquad \text{for some } \varphi : Z \longmapsto \{0,1\}$$

Boolean function

$$f(X) = (\varphi(z_1), \varphi(z_2), \dots, \varphi(z_n)) \cdot \underbrace{(s_1, \ . \ . \ . \ , s_n)}_{\text{Secret bits}}$$

bit vector $\in \{0,1\}^n$

Releasing k linear Statistics

Released Statistics →

$$\begin{bmatrix} f_1(X) \\ \vdots \\ f_k(X) \end{bmatrix} = \underbrace{\begin{bmatrix} \varphi_1(z_1) & \cdots & \varphi_1(z_n) \\ \vdots & F_i & \vdots \\ \varphi_k(z_1) & \cdots & \varphi_k(z_n) \end{bmatrix}}_{F} \begin{bmatrix} s_1 \\ \vdots \\ s_n \end{bmatrix}$$

← Secret bits

F : query matrix

$$f_i(X) = F_i \cdot s$$

Examples :

$\varphi_1(z_j) = 1$: z_j is older than 40

$\varphi_2(z_j) = 1$: z_j is older than 40 and male

$\varphi_3(z_j) = 1$: z_j is older than 20 and male

First Reconstruction Attack

"You can't release all count statistics with non-trivial accuracy."

Queries : $k=2^n$

For every $v\in\{0,1\}^n$, $F_v = v$

"subset in the dataset"

$$\overset{2^n \times n}{k=2^n \underset{F}{\begin{bmatrix} \text{—} F_v \text{—} \\ \\ \\ \end{bmatrix}}} \cdot \underset{\text{Secret bits}}{\begin{bmatrix} s \\ \\ \end{bmatrix}}$$

Reconstruction :

Suppose the answers $(a_v)_{v\in\{0,1\}^n}$, $\forall v\in\{0,1\}^n$, $|F_v \underset{\text{true answer}}{\cdot} s - \underbrace{a_v}_{\text{Released answer.}}| \le \underline{\alpha n}$

Choose $\tilde{s}\in\{0,1\}^n$, $\forall v$, $\boxed{|F_v\cdot\tilde{s} - a_v| \le \alpha\cdot n}$

Theorem. $\|s-\tilde{s}\|_1 \le 4\alpha n$

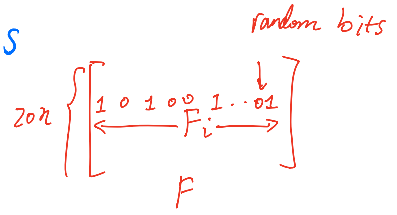
Theorem. If all $\boxed{2^n}$ counts are within $\overset{\alpha = 5\% \, \downarrow}{\alpha n}$ error, then $s, \tilde{s}$ disagree on $\le \underline{4\alpha n \text{ bits.}}$

$\le 20\%$

Not practical : 2^n.

Reconstruction Using Fewer Queries

Released Statistics $<< 2^n$.



Attack : Choose $\boxed{k = 20n}$ random $\varphi_i : Z \mapsto \{0,1\}$, $\forall i \in [k]$.

$\Longrightarrow$ k random vectors/queries $F_i \in \{0,1\}^n$

Suppose that answers : $\forall i \in [k]$, $|F_i \cdot s - a_i| \leq \alpha n$

Find $\tilde{s} \in \{0,1\}^n$ such that: $\forall i \in [k]$, $|F_i \cdot \tilde{s} - a_i| \leq \alpha n$

"just a constant"

Theorem. $\|s - \tilde{s}\|_1 \leq \overset{\downarrow}{256}\, \alpha^2 n^2$.

with high probability (> 99% of the time)

previously $4\alpha n$

$(a = s - \tilde{s})$ $\|a\|_1 = \sum_{j=1}^{n} |a_j|$. ell one norm.

$\|a\|_2 = \sqrt{\sum_{j=1}^{n} a_j^2}$ ell two norm

Theorem. If we ask $\overset{\text{think } c\cdot n}{O(n)}$ random queries $F \in \{0,1\}^n$ and all answers have error $\le \alpha n$, then reconstruct $\tilde{s}$ such that $\|s-\tilde{s}\|_1 \le O(\alpha^2 n^2)$. v.s. $O(n)$

$c \cdot \alpha^2 n^2$

How to parse this?

- Improvement $O(n) \ll 2^n$
- when $\alpha n \ll \sqrt{n}$, then $\alpha^2 n^2 \ll n$. $\|s-\tilde{s}\| \ll n$.

  For example, $\alpha = 10\%$, $\boxed{\alpha n \le \frac{\sqrt{n}}{10}}$, $\alpha^2 n^2 \le \frac{n}{100}$

  $\|s-\tilde{s}\|_1 \le O\left(\frac{n}{100}\right)$

Why is this an interesting case?

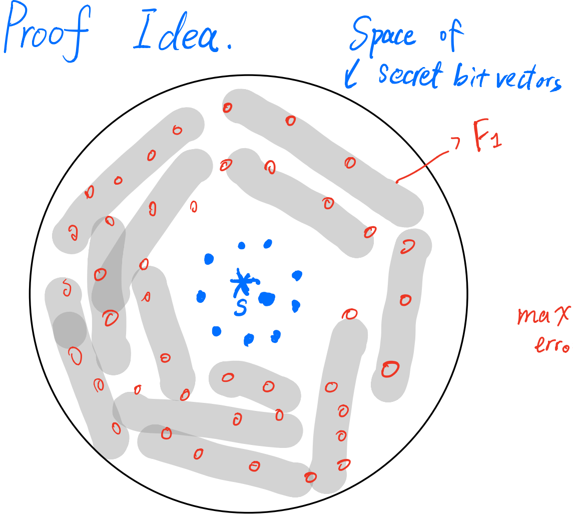
Population
draw data
Data Set

$$\mathbb{E}[\varphi(z)\cdot s] = p \xleftarrow{\text{How close ?}} \frac{1}{n}\sum_{j=1}^{n} \varphi(z_j)\cdot s_j$$

$$\boxed{\frac{1}{\sqrt{n}} := \text{Sampling error.}}$$

Sampling error in $\sum_{j=1}^{n} \varphi(z_j)\cdot s_j$ is roughly $\sqrt{n}$.

Proof Idea.



- ● good $\tilde{s}$: $\|s-\tilde{s}\|_1 \le \alpha^2 n^2$
- ○ bad $\tilde{s}$: $\|s-\tilde{s}\|_1 > \alpha^2 n^2$

▬

↑ given by a random query F_i.

Reconstruction Method

Given queries $F_1, \dots, F_k$, ← random

answers $a_1, \dots, a_k$

Find $\tilde{s} \in \{0,1\}^n$ that $\boxed{\text{minimizes}}$

max error. → $\max_{i\in\{1,\dots,k\}} \left| F_i \cdot \tilde{s} - a_i \right|$ ← want small.

↑ answer evaluated on $\tilde{s}$

↑ Released answer.

Output $\tilde{s}$.

Recall :

$$\max_i \left| F_i \cdot s - a_i \right| \le \alpha n$$

Find $\tilde{s}$ such that

$$\forall i \in \{1, .., k\},\ \left| F_i \cdot \tilde{s} - a_i \right| \le \alpha n.$$

Feasible because s^* satisfies all of them

Proof Idea.

① $\tilde{s}$ satisfies

$$\max_i \left| F_i \cdot \tilde{s} - a_i \right| \le \alpha n$$

② $\tilde{s}$ is eliminated if

$\exists F_i$ s.t. $|F_i \cdot \tilde{s} - a_i| > \alpha n$

($\tilde{s}$ is eliminated by F_i)

Reconstruction Method

Given queries $F_1, \dots, F_k$,

answers $a_1, \dots, a_k$

Find $\tilde{s} \in \{0,1\}^n$ that minimizes

$$\max_{i \in \{1,\dots,k\}} \left| F_i \cdot \tilde{s} - a_i \right|$$

Output $\tilde{s}$.

③ For every bad $\tilde{s}$,

Some random query eliminates $\tilde{s}$ with high probability.

Proof.

"there exists"

$$\mathbb{P}\left(\exists \text{ some bad } \tilde{s} \text{ not eliminated}\right)$$

$$\leq \sum_{\text{bad } \tilde{s}} \underline{\mathbb{P}[\tilde{s} \text{ not eliminated})}$$

$$\mathbb{P}[\tilde{s} \text{ not eliminated}]$$

$$= \mathbb{P}[\forall i, \ \tilde{s} \text{ is not eliminated}]$$

"for all"

$$= \mathbb{P}[\tilde{s} \text{ not eliminated by } F_i]^k$$

$$\leq \underbrace{\mathbb{P}[\ |F_i \cdot \tilde{s} - F_i \cdot \underline{s}| \leq 4\alpha n]^k}_{\leq \frac{9}{10}} \ \boxed{\leq} \left(\frac{9}{10}\right)^k \leq 2^{-2n}$$

Key Step to be shown

Reconstruction Method

Given queries $F_1, \ldots, F_k$,

answers $a_1, \ldots, a_k$

Find $\tilde{s} \in \{0,1\}^n$ that minimizes

$$\max_{i \in \{1,\ldots,k\}} |F_i \cdot \tilde{s} - a_i|$$

Output $\tilde{s}$.

$k = \underline{20n}$.

Proof.

Key Lemma.

bad candidate

If $s, \tilde{s} \in \{0,1\}$ s.t.

think $\gg 2^2 n^2$

$\|s - \tilde{s}\|_1 = m$ (differ on m coordinates)

Let $F \in \{0,1\}^n$ be random,

then

$$\mathbb{P}\left[\, |F \cdot (s-\tilde{s})| \leq \frac{\sqrt{m}}{10} \right] \leq \frac{9}{10}$$

$$\mathbb{P}\left[\, \underbrace{|F \cdot (s-\tilde{s})| > \frac{\sqrt{m}}{10}}_{\text{sufficient prob. mass}} \right] > \frac{1}{10}.$$

Intuition:

$t = s - \tilde{s} \in \{-1, 0, 1\}^n$

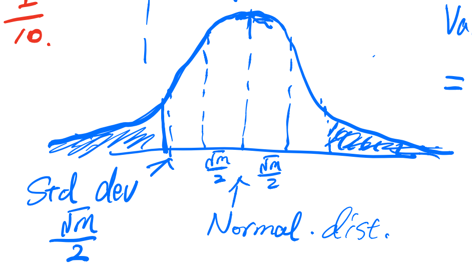
If $t_j = 1$,

$$F_j t_j = \begin{cases} 1 & \text{w.p. } \frac{1}{2} \\ 0 & \text{w.p. } \frac{1}{2} \end{cases}$$

If $t_j = -1$

$$F_j t_j = \begin{cases} -1 & \text{w.p. } \frac{1}{2} \\ 0 & \end{cases}$$

$$F \cdot t = \underbrace{\sum_{j : s_j \neq \tilde{s}_j}}_{m \text{ terms}} F_j t_j$$



$$\begin{aligned} &\mathrm{Var}(F \cdot t) \\ &= \mathrm{Var}\Big(\sum_j F_j \cdot t_j\Big) \\ &= \sum_j \underbrace{\mathrm{Var}(F_j \cdot t_j)}_{\frac{1}{4}} \\ &= \frac{m}{4} \end{aligned}$$

Efficient Reconstruction.

Reconstruction Method

Given queries $F_1, \dots, F_k$,

answers $a_1, \dots, a_k$

Find $\boxed{\tilde{s} \in \{0,1\}^n}$ that minimizes

$$\max_{i\in\{1,\dots,k\}} \left|F_i \cdot \tilde{s} - a_i\right|$$

Output $\tilde{s}$.

NP-hard.

Constraint Satisfaction Problem.

Linear Programming

$$\max_{x\in\mathbb{R}^d} \; c \cdot x$$

s.t.

$$\forall i \in [k], \; v_i \cdot x \le b_i$$

Can solve in polynomial time.

Relax

$\hat{s} \in [0,1]^n$

$\hat{s} \xrightarrow{\text{rounding}} \tilde{s} \in \{0,1\}^n$

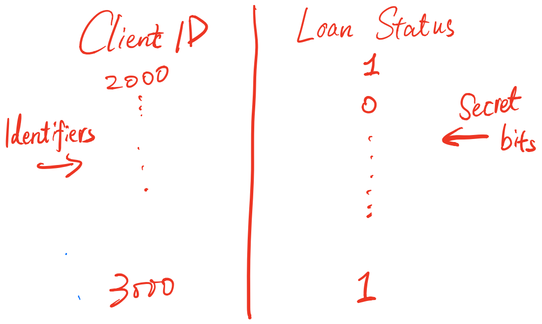
$\hat{s}_j = 0.6 \longrightarrow \tilde{s}_j = 1$ with prob. 0.6

Attacking Diffix

private analytics product by Aircloak

Check out the Diffix Challenge!

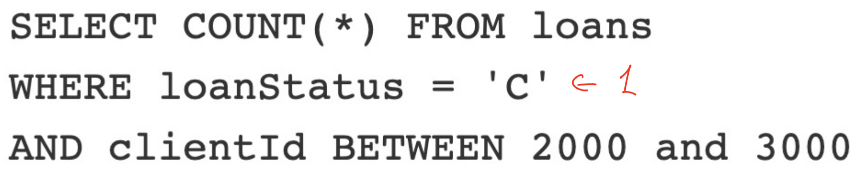
```
SELECT COUNT(*) FROM loans
WHERE loanStatus = 'C'
AND clientId BETWEEN 2000 and 3000
```



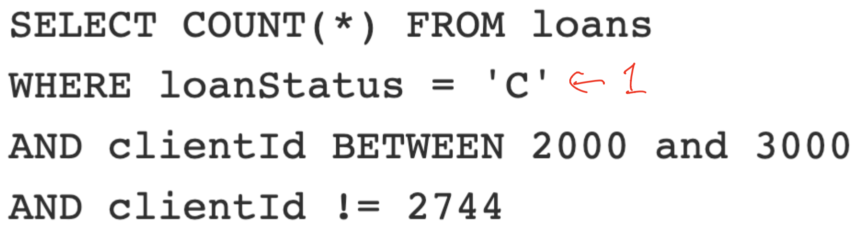
Count query

$$\sum_{iD=2000}^{3000} \text{Loan Status}(iD)$$

Difference Attack.



```
SELECT COUNT(*) FROM loans
WHERE loanStatus = 'C' ← 1
AND clientId BETWEEN 2000 and 3000
```

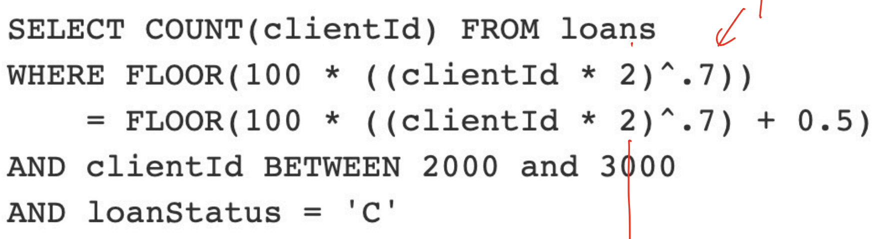


```
SELECT COUNT(*) FROM loans
WHERE loanStatus = 'C' ← 1
AND clientId BETWEEN 2000 and 3000
AND clientId != 2744
```

```
SELECT COUNT(*) FROM loans
WHERE loanStatus = 'C'
AND clientId BETWEEN 2000 and 3000
```

Attack by Kobbi Nissim & Aloni Cohen 2018.

prime



```
SELECT COUNT(clientId) FROM loans
WHERE FLOOR(100 * ((clientId * 2)^.7))
    = FLOOR(100 * ((clientId * 2)^.7) + 0.5)
AND clientId BETWEEN 2000 and 3000
AND loanStatus = 'C'
```

prime

Dick- Joseph- Schutzman.

```
SELECT COUNT(*) FROM rides
WHERE FLOOR(pickup_latitude ^  8.789 + 0.5)
    = FLOOR(pickup_latitude ^  8.789)
AND trip_distance IN (0.87, 1.97, 2.75)
AND payment_type = 'CSH'
```

Announcement :

- HW 0
- Recitation on Friday
- Office Hours