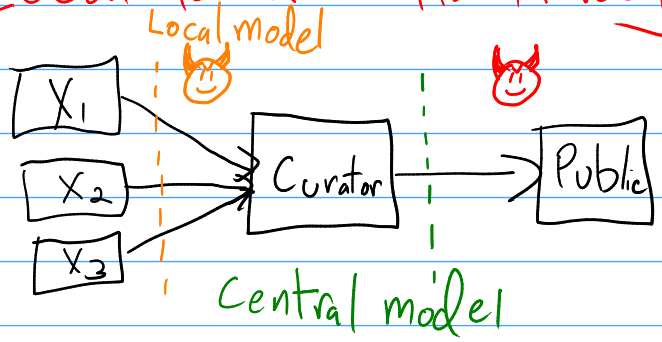
Lecture 17
Deployments of DP: Local Differential Privacy

Local Differential Privacy



$X_1, \dots, X_n \in \{0,1\}, \mu = \frac{1}{n}\sum X_i$.

$\hat{\mu} = \frac{1}{n}\sum X_i + Lap\left(\frac{1}{\varepsilon n}\right)$.

$|\hat{\mu} - \mu| \le \frac{1}{\varepsilon n}$ $\Leftrightarrow$ If want $|\hat{\mu} - \mu| \le \alpha$, $n \ge \frac{1}{\alpha\varepsilon}$.

Randomized Response $\frac{1}{2}+\gamma, \frac{1}{2}-\gamma$

$$Y_i = \begin{cases} X_i & \text{w.p. } \frac{e^\varepsilon}{1+e^\varepsilon} \\ 1-X_i & \text{w.p. } \frac{1}{1+e^\varepsilon} \end{cases} \qquad \left(\frac{e^\varepsilon}{1+e^\varepsilon}\right) \Big/ \left(\frac{1}{1+e^\varepsilon}\right) = e^\varepsilon$$

$$\hat{\mu} = \frac{1}{n}\sum\left(\frac{e^\varepsilon+1}{e^\varepsilon-1}\cdot Y_i - \frac{1}{e^\varepsilon-1}\right)$$

$$E[\hat{\mu}] = \frac{1}{n}\sum E\left[\frac{e^\varepsilon+1}{e^\varepsilon-1}\cdot Y_i - \frac{1}{e^\varepsilon-1}\right]$$

$$= \frac{1}{n}\sum\left(\frac{e^\varepsilon+1}{e^\varepsilon-1}\cdot\left(X_i\cdot\frac{e^\varepsilon}{1+e^\varepsilon} + (1-X_i)\cdot\frac{1}{1+e^\varepsilon}\right) - \frac{1}{e^\varepsilon-1}\right)$$

$$= \frac{1}{n}\sum\left(\frac{1}{e^\varepsilon-1} - \frac{1}{e^\varepsilon-1}\,\frac{X_i\cdot e^\varepsilon - X_i}{e^\varepsilon-1}\right)$$

$$= \frac{1}{n}\sum X_i$$

$$= \mu$$

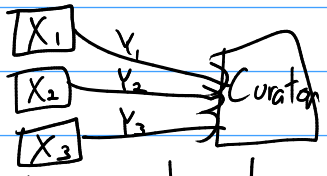
$e^\varepsilon \approx 1+\varepsilon$

$$Var[\hat{\mu}] \doteq Var\left[\frac{1}{n}\sum\frac{2+\varepsilon}{\varepsilon}\cdot Y_i\right] = \frac{1}{n^2}\cdot n\cdot\left(\frac{2+\varepsilon}{\varepsilon}\right)^2\cdot Var(Y_i)$$

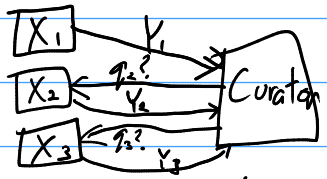
$$= O\left(\frac{1}{n\varepsilon^2}\right)$$

Central DP: $|\hat{\mu} - \mu| \le \frac{1}{\varepsilon n}$ $\Leftrightarrow$ If want $|\hat{\mu} - \mu| \le \alpha$, $n \ge \frac{1}{\alpha \varepsilon}$.

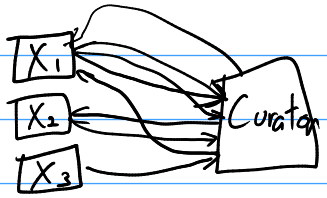
LDP: $|\hat{\mu} - \mu| \le \frac{1}{\varepsilon \sqrt{n}}$. $\Leftrightarrow$ $|\hat{\mu} - \mu| \le \alpha$ reqs $n \ge \frac{1}{\alpha^2 \varepsilon^2}$



Non interactive



Seq. Interactive



Interactive

ε-LDP $\approx (\varepsilon, \delta)$-LDP?

Warner '65
Evfimievski, Gehrke, Srikant '03
Kasiviswanathan, Lee, Nissim, Raskhodnikova, Smith '08
KLeeNRS
Duchi, Jordan, Wainwright '13

Local DP at Google

RAPPOR - Erlingsson, Pihur, Korolova '14

$\varepsilon = \ln 3 \Rightarrow RR = \begin{cases} \text{truth} & \text{w.p } 3/4 \\ \text{lie} & \text{w.p } 1/4. \end{cases}$

RR = yes w.p. $\frac{1}{4}$ if user's value is no

100 days in a row

RR = 75 yes's, 25 no's. w.p. 1.39×10^{-24} if val = no

$\varepsilon = \ln 3 \approx 1.1 \Rightarrow 100\varepsilon \approx 110, \quad e^{110}$

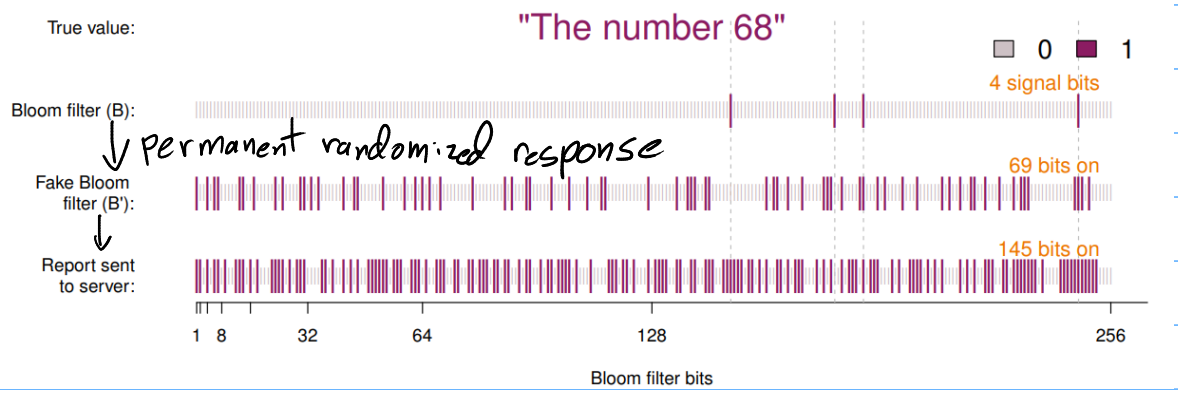
Memoization: Don't re-randomize ans to same q.

RR 0,1

RAPPOR $v \in \mathcal{X} \to \{0,1\}^k$ $\quad B_j = 1$ if $\exists i$ s.t. $f_i(v) = j$, 0

B

h hash fns $f_1, \ldots, f_h$ $\mathcal{X} \to [k]$ $\quad h = 1 \to$ one-hot encoding



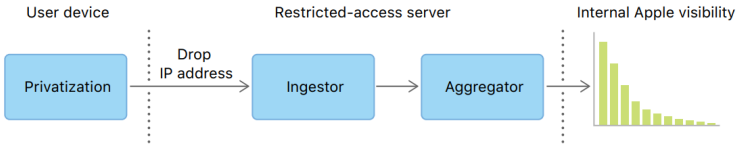
10000, 10001, 10002

ε $\quad \varepsilon$ $\quad \varepsilon$

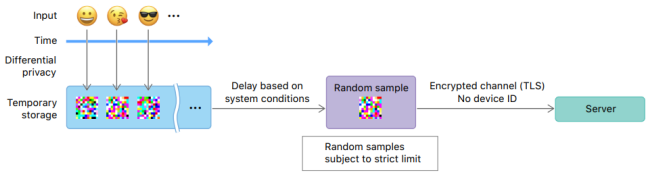
Chrome
- malware
- process

Local DP at Apple '17

Tang et al. '17



System Overview



Privatization Stage

Charikar, Chen, Farach-Colton '02

Communication.

aaaaa aaaaa → [00 1000000 1 0 ...]

aaaaaaaaab →

ĉ →

Sequence Fragment Puzzle

1. Discover new words
2. Compute freqs of words

waterloo. f: word → $\{0, \ldots, 255\}$

f(waterloo) = 42.

(1, wa, 42) (3, te, 42) (5, rl, 42) (7, oo, 42)

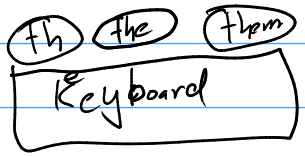
waterloo.

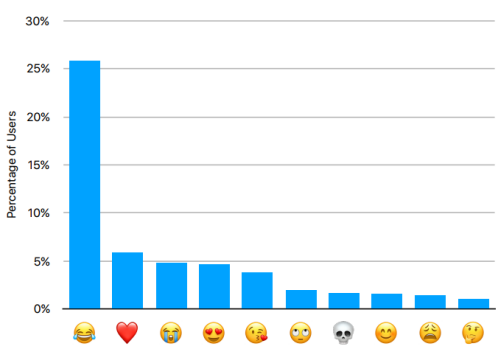
iPhone Messages

wyd wbu idc

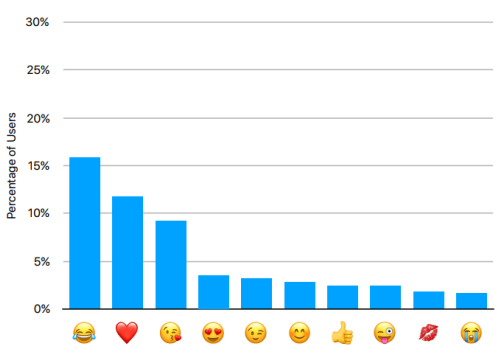
bruh bae

lov , th





(a) English



(b) French

Differential Privacy at Microsoft

Ding, Kulkarni, Yekhanin '17

~~Communication~~

Average ~~$[0,1]$~~ $[0,m]$ ε-LDP

$X_i \in [0,m] \to Y_i = X_i + \text{Lap}\left(\frac{m}{\varepsilon}\right)$

$\hat{\mu} = \frac{1}{n}\sum\left(X_i + \text{Lap}\left(\frac{m}{\varepsilon}\right)\right)$

$\hat{\mu} - \mu = \frac{1}{n}\sum \text{Lap}\left(\frac{m}{\varepsilon}\right)$

$\text{mag} \approx \frac{m}{\varepsilon\sqrt{n}}$. $n \geq \frac{m^2}{\alpha^2\varepsilon^2} \to \text{err} \leq \alpha$.

optimal

Comm: $\Omega(\log m)$ bits

1 bit protocol.

$X_i \in [0,m]$. $Y_i = \begin{cases} 1 & \text{w.p. } \frac{1}{e^\varepsilon+1} + \frac{X_i}{m}\cdot\frac{e^\varepsilon-1}{e^\varepsilon+1} \\ 0 & \text{o.w.} \end{cases}$

$\hat{\mu} = \frac{m}{n}\sum \frac{Y_i(e^\varepsilon+1)-1}{e^\varepsilon-1}$

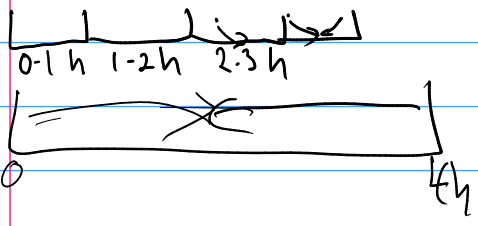
$E[\hat{\mu}] = \frac{m}{n}\sum \frac{\left(\frac{1}{e^\varepsilon+1} + \frac{X_i}{m}\cdot\frac{e^\varepsilon-1}{e^\varepsilon+1}\right)(e^\varepsilon+1) - 1}{e^\varepsilon - 1} = \frac{m}{n}\sum \frac{\left(1 + \frac{X_i(e^\varepsilon-1)}{m}\right) - 1}{e^\varepsilon - 1}$

$= \frac{1}{n}\sum X_i = \mu$

Chernoff-Hoeffding $\Rightarrow \text{err} \leq O\left(\frac{m}{\varepsilon\sqrt{n}}\right) \iff n \geq \frac{m^2}{\alpha^2\varepsilon^2} \to \text{err} \leq \alpha$.

but only 1 bits

Memoization
- Doesn't handle small val changes



$X_i \in [0, m]$

Pretend val $X_i = 0$ $X_i = m$

1. Memoize: Precompute $A_i(0)$ and $A_i(m)$
2. Select offset: $\alpha_i \in \{0, \ldots, m-1\}$ u.a.r.
3. If $X_i + \alpha_i \leq m$, report $A_i(0)$
   else, $A_i(m)$

$A_i(0), A_i(0), A_i(0), \ldots, A_i(m), A_i(m), \ldots$