Differentially Private Password Frequency Lists

Jeremiah Blocki
MSR/Purdue

Anupam Datta
CMU

Joseph Bonneau
Stanford/EFF

Differentially Private Password Frequency Lists

Or, How to release statistics from 70 million passwords (on purpose)

Jeremiah Blocki
MSR/Purdue

Anupam Datta
CMU

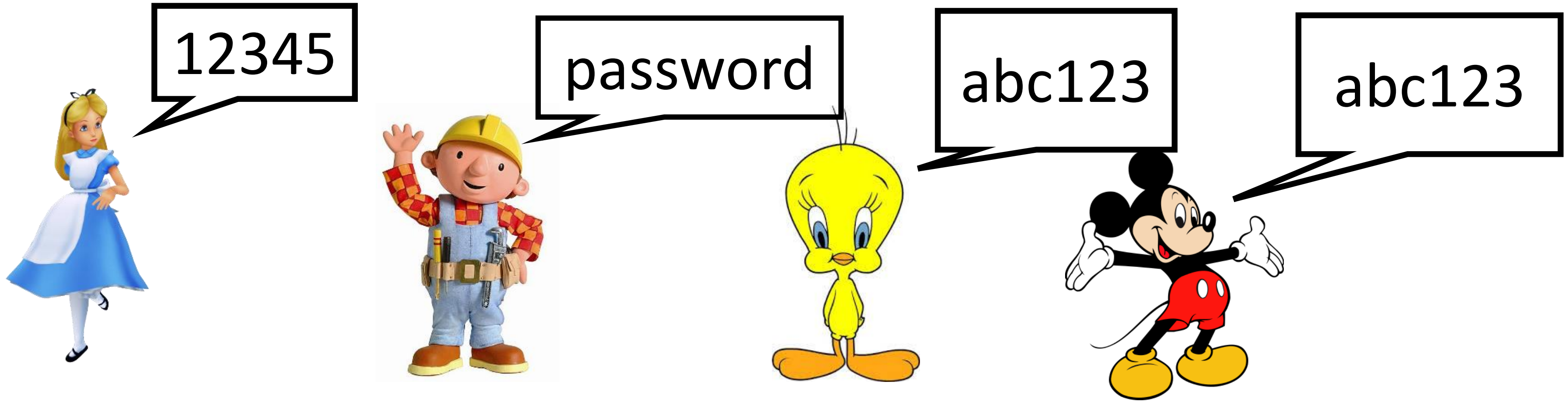
Joseph Bonneau
Stanford/EFF

Outline

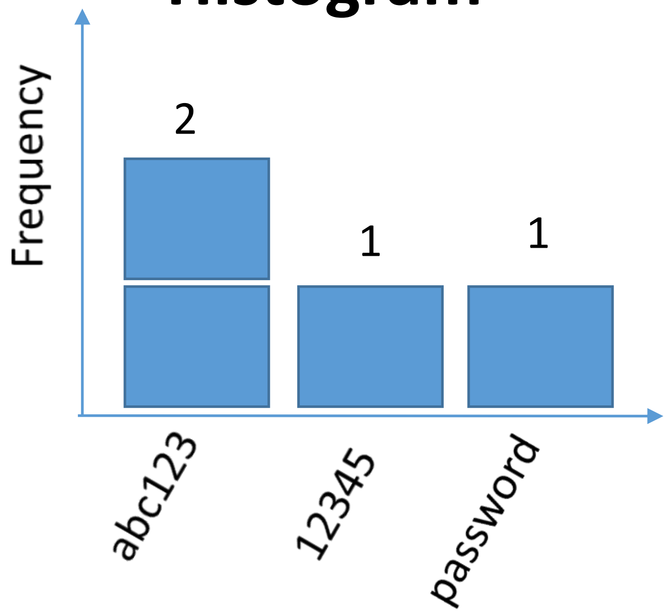
- Password Frequency List
- Potential Security Concerns
- Differential Privacy
- A DP Algorithm with Minimal Distortion
- Released Yahoo! Frequency List

What is a Password Frequency List?

Password Dataset:
(N users)

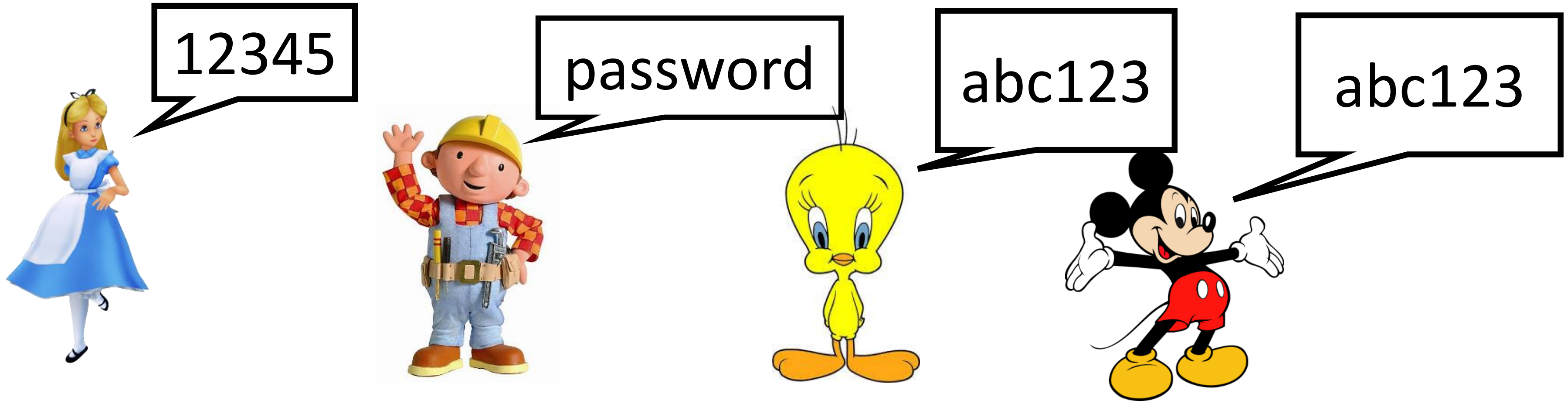


Histogram

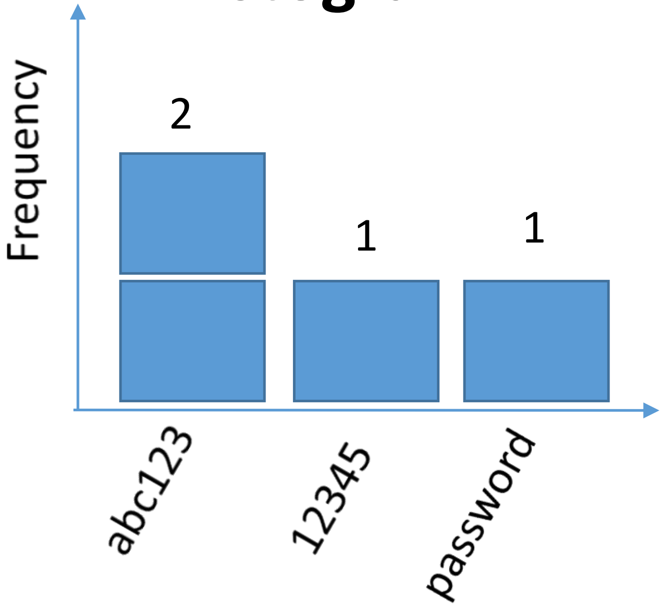


What is a Password Frequency List?

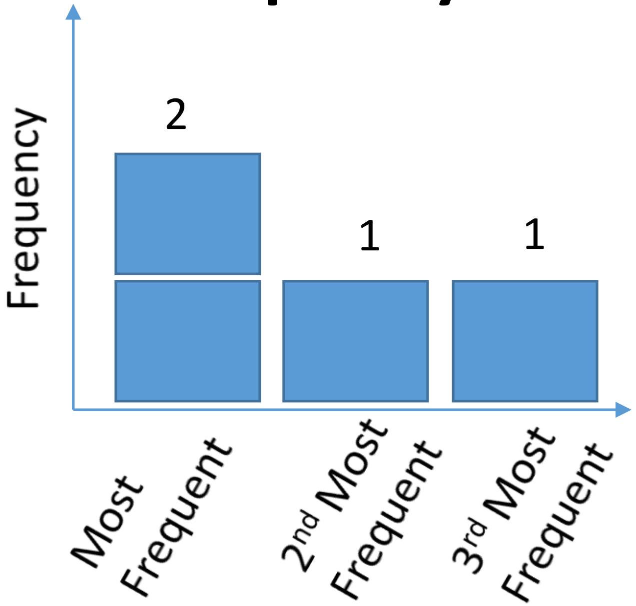
Password Dataset:
(N users)



Histogram

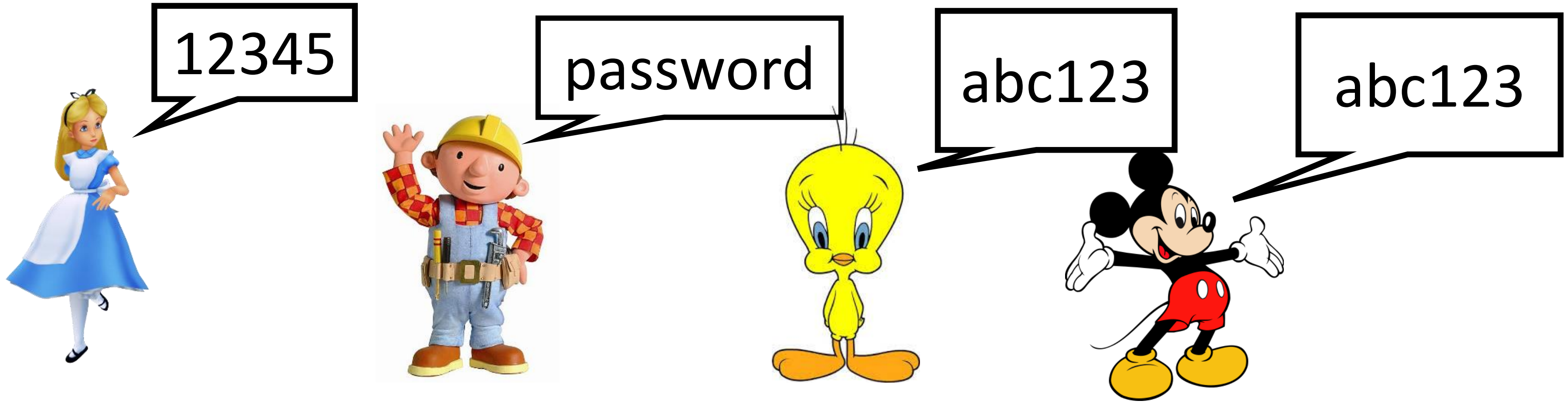


Frequency List

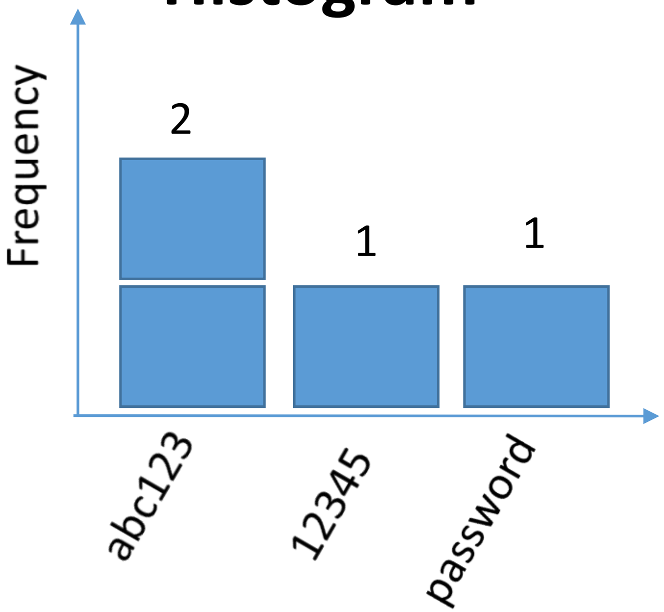


What is a Password Frequency List?

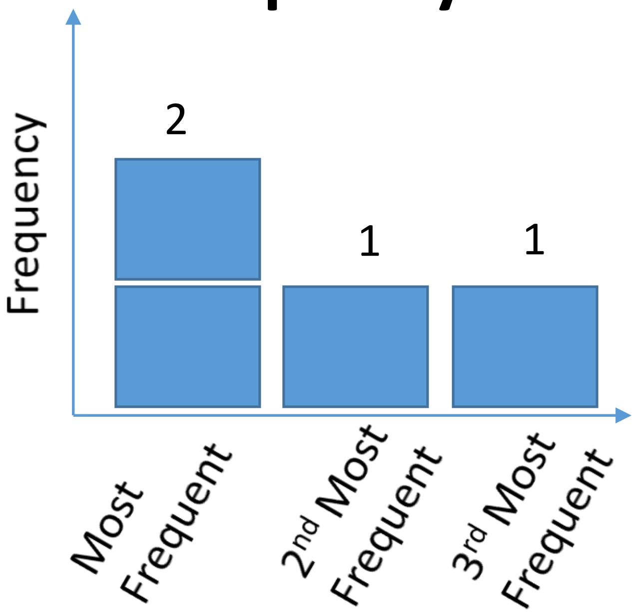
Password Dataset:
(N users)



Histogram



Frequency List



Formal Notation:

$\mathbf{f} = (f_1, \ldots, f_N)$ such that

- $f_1 \geq f_2 \geq \cdots \geq f_N \geq 0$
- $N = \sum_{i=1}^{N} f_i$

Password Frequency List (Application 1)

Estimate #accounts compromised by attacker with β guesses per user

- Online Attacker (β small)
- Offline Attacker (β large)

$$\lambda_\beta = \sum_{i=1}^{\beta} f_i$$

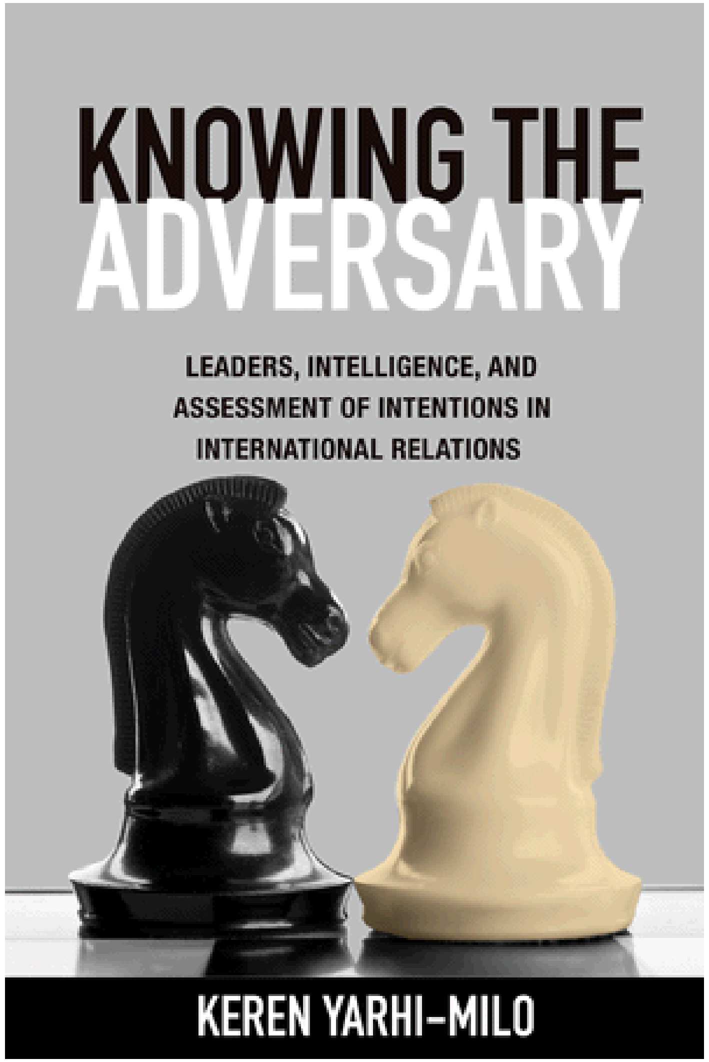
Password Frequency List (Application 2)

Quantify Benefits from Key-Stretching

Halting Condition (Rational Offline Adversary):
- Marginal Guessing Cost $\geq$ Marginal Benefit

Password Frequency Lists allow us to estimate
- Marginal Guessing Cost (MGC)
- Marginal Benefit (MB)
- Rational Adversary: MGC = MB

Can estimate when the offline adversary will give up.



Available Password Frequency Lists

Site	#User Accounts (N)	How Released
RockYou	32.6 Million	Data Breach*
LinkedIn	6	Data Breach*
….	…	…

* entire frequency list available due to improper password storage

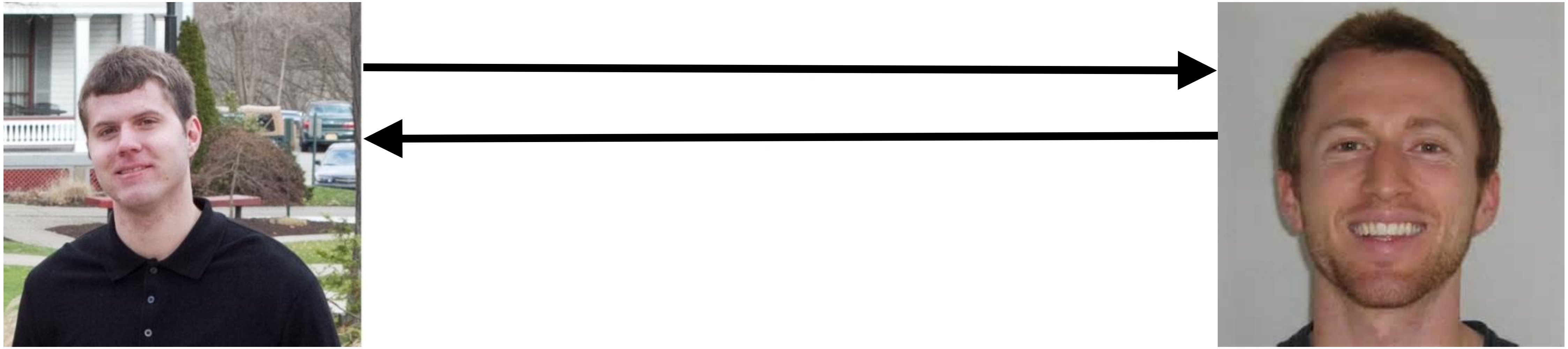
How the project started

Would it be possible to access the Yahoo! data? I am working on a cool new research project and the password frequency data would be very useful.

How the project started

I would love to make the data public, but Yahoo! Legal has concerns about security and privacy. They won't let me release it.

How the project started



I would love to make the data public, but Yahoo! Legal has concerns about security and privacy. They won't let me release it.

Available Password Frequency Lists

Site	#User Accounts (N)	How Released
RockYou	32.6 Million	Data Breach*
LinkedIn	6	Data Breach*
….	…	…
Yahoo! [B12]	70 Million	With Permission**

* entire frequency list available due to improper password storage

** frequency list perturbed slightly to preserve differential privacy.

Yahoo! Frequency data is now available online at:
https://figshare.com/articles/Yahoo_Password_Frequency_Corpus/2057937

Why not just publish the original frequency lists?

- Heuristic Approaches to Data Privacy often break down when the adversary has background knowledge
  - Massachusetts Group Insurance Medical Encounter Database [SS98]
    - Background Knowledge: Voter Registration Record

Commonwealth of Massachusetts
Group Insurance Commission

Why not just publish the original frequency lists?

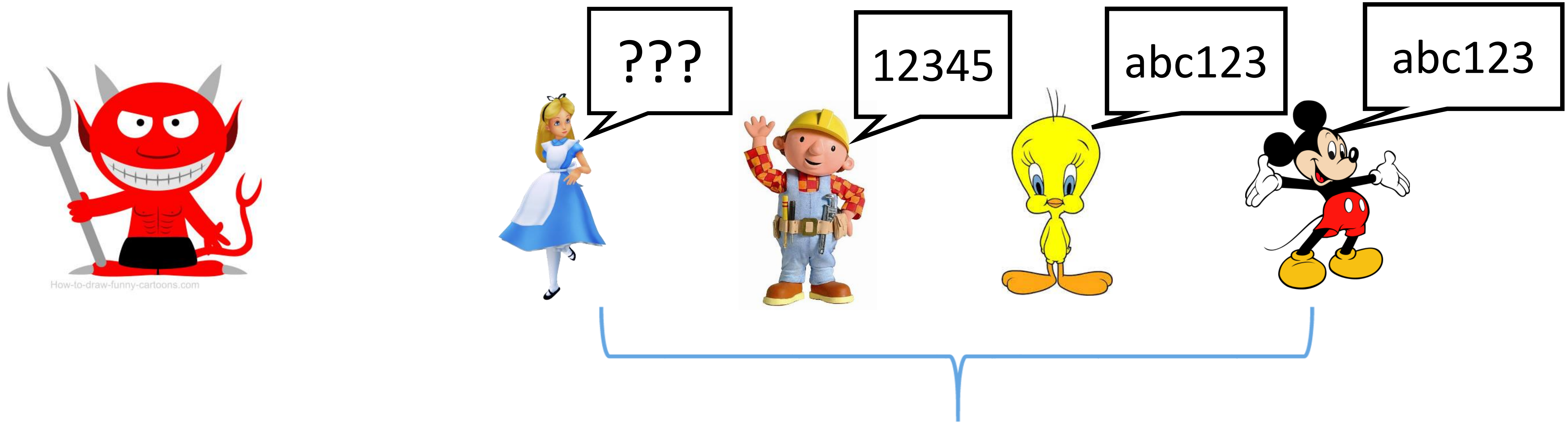
- Heuristic Approaches to Data Privacy often break down when the adversary has background knowledge
  - Massachusetts Group Insurance Medical Encounter Database [SS98]
    - Background Knowledge: Voter Registration Record
  - Netflix Prize Dataset[NS08]
    - Background Knowledge: IMDB

Why not just publish the original frequency lists?

- Heuristic Approaches to Data Privacy often break down when the adversary has background knowledge
  - Netflix Prize Dataset[NS08]
    - Background Knowledge: IMDB
  - Massachusetts Group Insurance Medical Encounter Database [SS98]
    - Background Knowledge: Voter Registration Record
  - Many other attacks [BDK07,…]

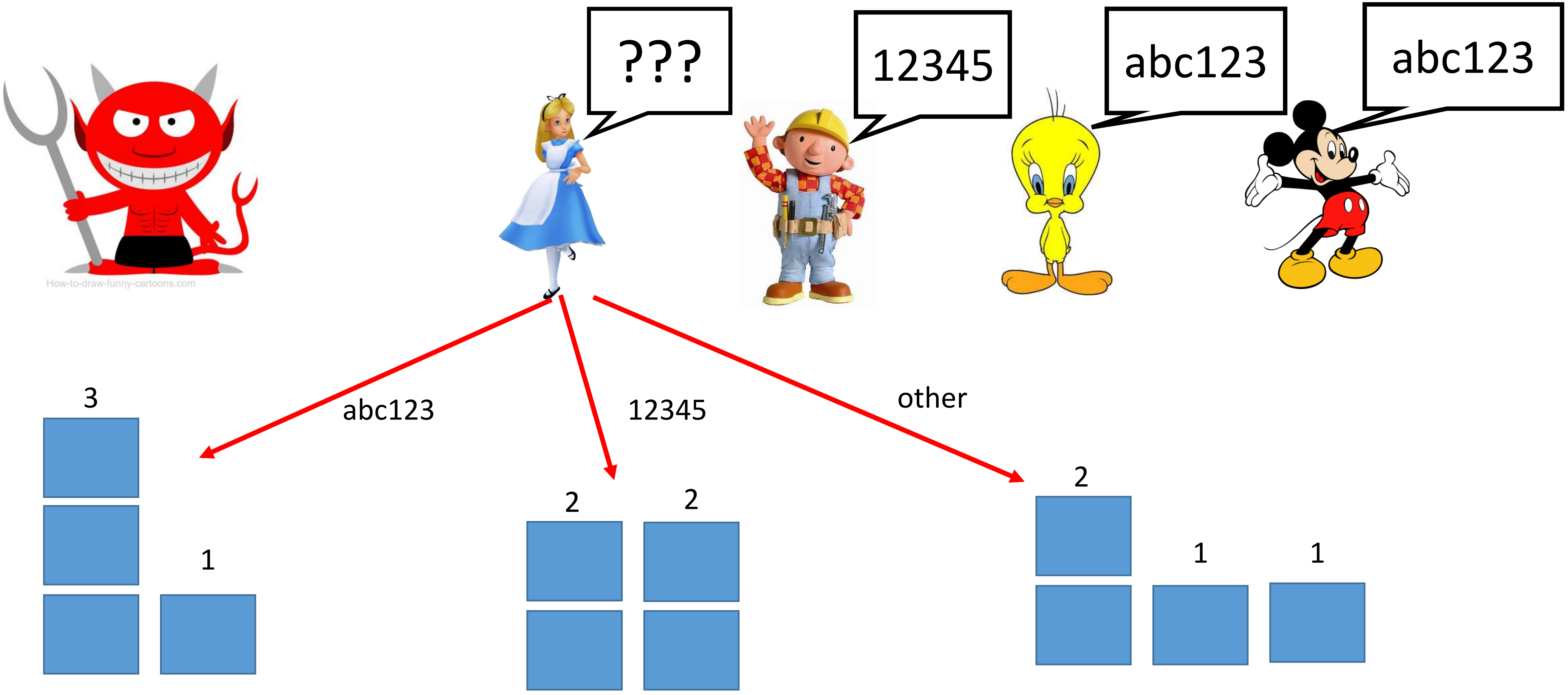
- In the absence of provable privacy guarantees Yahoo! was understandably reluctant to release these password frequency lists.

Security Risks (Example)



Adversary Background Knowledge

Security Risks (Example)



Differential Privacy (Dwork et al)

Definition: An (randomized) algorithm A preserves (ε, δ)-differential privacy if for *any* subset S$\subseteq Range(A)$ of possible outcomes and *any* we have

$$\Pr[A(f) \in S] \le e^{\varepsilon}\Pr[A(f') \in S] + \delta$$

for any pair of adjacent password frequency lists f and f',

$$\|f - f'\|_1 = 1.$$

$$\|f - f'\|_1 \stackrel{\text{def}}{=} \sum_i |f_i - f_i{}'|$$

Differential Privacy (Dwork et al)

Definition: An (randomized) algorithm A preserves (ε, δ)-differential privacy if for *any* subset S$\subseteq$ $Range(A)$ of possible outcomes and *any* we have

$$\Pr[A(f) \in S] \le e^{\varepsilon} \Pr[A(f') \in S] + \delta$$

for any pair of adjacent password frequency lists f and f',

$$\|f - f'\|_1 = 1.$$

f – original password frequency list

f' – remove Alice's password from dataset

Differential Privacy (Dwork et al)

Definition: An (randomized) algorithm A preserves (ε, δ)-differential privacy if for *any* subset S⊆ $Range(A)$ of possible outcomes and *any* we have

$$\Pr[A(f) \in \mathrm{S}] \le e^{\varepsilon} \Pr[A(f') \in \mathrm{S}] + \delta$$

for any pair of adjacent password frequency lists f and f',

$$\|f - f'\|_1 = 1.$$

Small Constant (e.g., $\varepsilon = 0.5$)

f – original password frequency list

f' – remove Alice's password from dataset

Differential Privacy (Dwork et al)

Definition: An (randomized) algorithm A preserves (ε, δ)-differential privacy if for *any* subset S⊆ $Range(A)$ of possible outcomes and *any* we have

$$\Pr[A(f) \in S] \le e^{\varepsilon}\Pr[A(f') \in S] + \delta$$

for any pair of adjacent password frequency lists f and f',

$$\|f - f'\|_1 = 1.$$

Small Constant (e.g., $\varepsilon = 0.5$)

Negligibly Small Value (e.g., $\delta = 2^{-100}$)

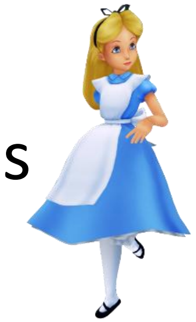
f – original password frequency list

f' – remove Alice's password from dataset

Differential Privacy (Example)

2 2

minus



=

2

1

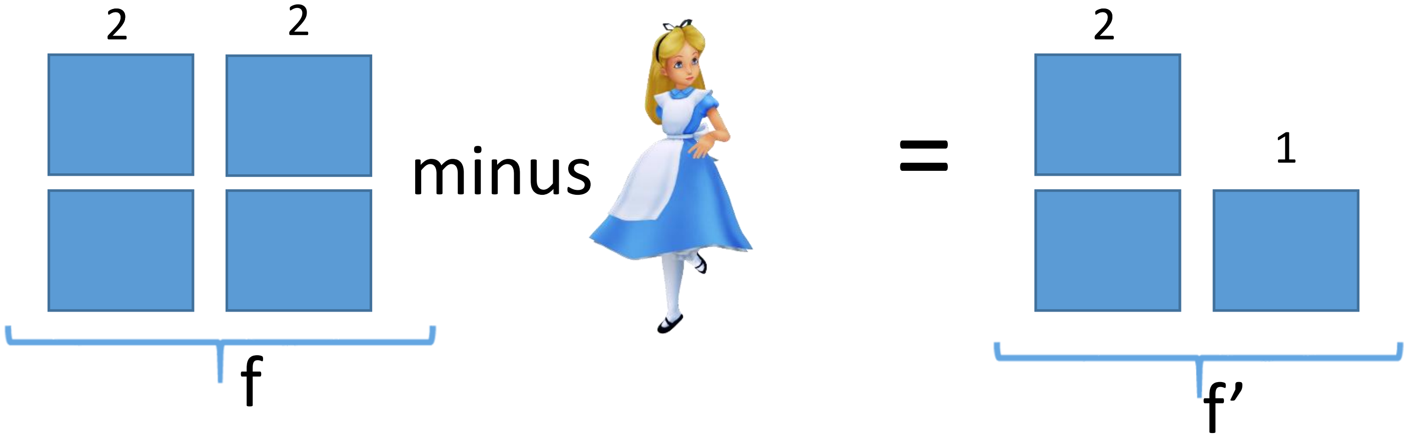
f

f'

Subset S of all potentially harmful outcomes to Alice

Outcomes

Differential Privacy (Example)



Subset S of all potentially harmful outcomes to Alice

$\mathbf{Pr}\big[A(f) \in$  $\big] \leq e^{\varepsilon}\,\mathbf{Pr}\big[A(f') \in$  $\big] + \delta$

Differential Privacy (Example)

Intuition: Alice will not harmed because her password was included in the dataset.

$\mathbf{Pr}\Big[A(f) \in$ 

$\Big] \leq e^{\varepsilon}\,\mathbf{Pr}\Big[A(f') \in$ 

$\Big] + \delta$

Main Technical Result

Theorem: There is a computationally efficient algorithm $\tilde{f} \leftarrow A(f)$ such that A preserves (ε, δ)-differential privacy and, except with probability δ, outputs $\tilde{f}$ s.t.

$$\frac{\|f - \tilde{f}\|_1}{N} \leq O\left(\frac{1}{\varepsilon\sqrt{N}} + \frac{\ln\left({}^{1}/{}_{\delta}\right)}{\varepsilon N}\right).$$

Main Tool: Exponential Mechanism [MT07]

Input: f

Output: $\Pr\left[\mathcal{E}^{\varepsilon}(f) = \tilde{f}\right] \propto e^{-\frac{\|f-\tilde{f}\|_1}{2\varepsilon}}$

Assigns very small probability to inaccurate outcomes.

Main Tool: Exponential Mechanism [MT07]

Input: f

Output: $\Pr\left[\mathcal{E}^{\varepsilon}(f) = \tilde{f}\right] \propto e^{-\frac{\|f-\tilde{f}\|_1}{2\varepsilon}}$

Assigns very small probability to inaccurate outcomes.

Theorem [MT07]: The exponential mechanism preserves $(\varepsilon, 0)$-differential privacy.

Analysis: Exponential Mechanism

Input: f

Output: $\Pr\left[\mathcal{E}^{\varepsilon}(f) = \tilde{f}\right] \propto e^{-\frac{\|f-\tilde{f}\|_1}{2\varepsilon}}$

Assigns very small probability to inaccurate outcomes.

Theorem [HR18]: There are $e^{O\left(\sqrt{N}\right)}$ partitions of the integer N.

Analysis: Exponential Mechanism

Input: f

Output: $\Pr\left[\mathcal{E}^{\varepsilon}(f) = \tilde{f}\right] \propto e^{-\frac{\|f-\tilde{f}\|_1}{2\varepsilon}}$

Assigns very small probability to inaccurate outcomes.

Theorem [HR18]: There are $e^{O(\sqrt{N})}$ partitions of the integer N.

Union Bound → $\|f - \tilde{f}\|_1 \leq O\left(\frac{\sqrt{N}}{\varepsilon}\right)$ with high probability.

Analysis: Exponential Mechanism

Input: f

Output: $\Pr\left[\mathcal{E}^{\varepsilon}(f) = \tilde{f}\right] \propto e^{-\frac{\|f-\tilde{f}\|_1}{2\varepsilon}}$ → **Assigns very small probability to inaccurate outcomes.**

Theorem: $\frac{\|f-\tilde{f}\|_1}{N} \leq O\left(\frac{1}{\varepsilon\sqrt{N}}\right)$ with high probability.

Theorem [MT07]: The exponential mechanism preserves $(\varepsilon, 0)$-differential privacy.

The Challenge --- Efficiency

Naïve Implementation: Exponential time (distribution assigns weights to infinitely many integer partitions)

Strong Evidence: Sampling from the exponential mechanism is computationally intractable in general (e.g., [U13]).

Good News

Theorem: There is an efficient algorithm A to sample from a distribution that is δ–close to the exponential mechanism $\mathcal{E}$ over integer partitions. The algorithm uses time and space

$$O\left(\frac{N\sqrt{N} + \mathrm{N}\ln\left(\frac{1}{\delta}\right)}{\varepsilon}\right)$$

Good News

Theorem: There is an efficient algorithm A to sample from a distribution that is δ–close to the exponential mechanism $\mathcal{E}$ over integer partitions. The algorithm uses time and space

$$O\left(\frac{N\sqrt{N} + \mathrm{N}\ln\left(\frac{1}{\delta}\right)}{\varepsilon}\right)$$

Key Idea 1: Novel dynamic programming algorithm to compute weights $\mathrm{W}_{\mathrm{i,k}}$ such that

$$\mathbf{Pr}\left[\tilde{f}_i = k \,\middle|\, \tilde{f}_{i-1}\right] = \frac{\mathrm{W}_{\mathrm{i},k}}{\sum_{\mathrm{t}=0}^{\tilde{f}_{i-1}} \mathrm{W}_{\mathrm{i,t}}}.$$

Good News

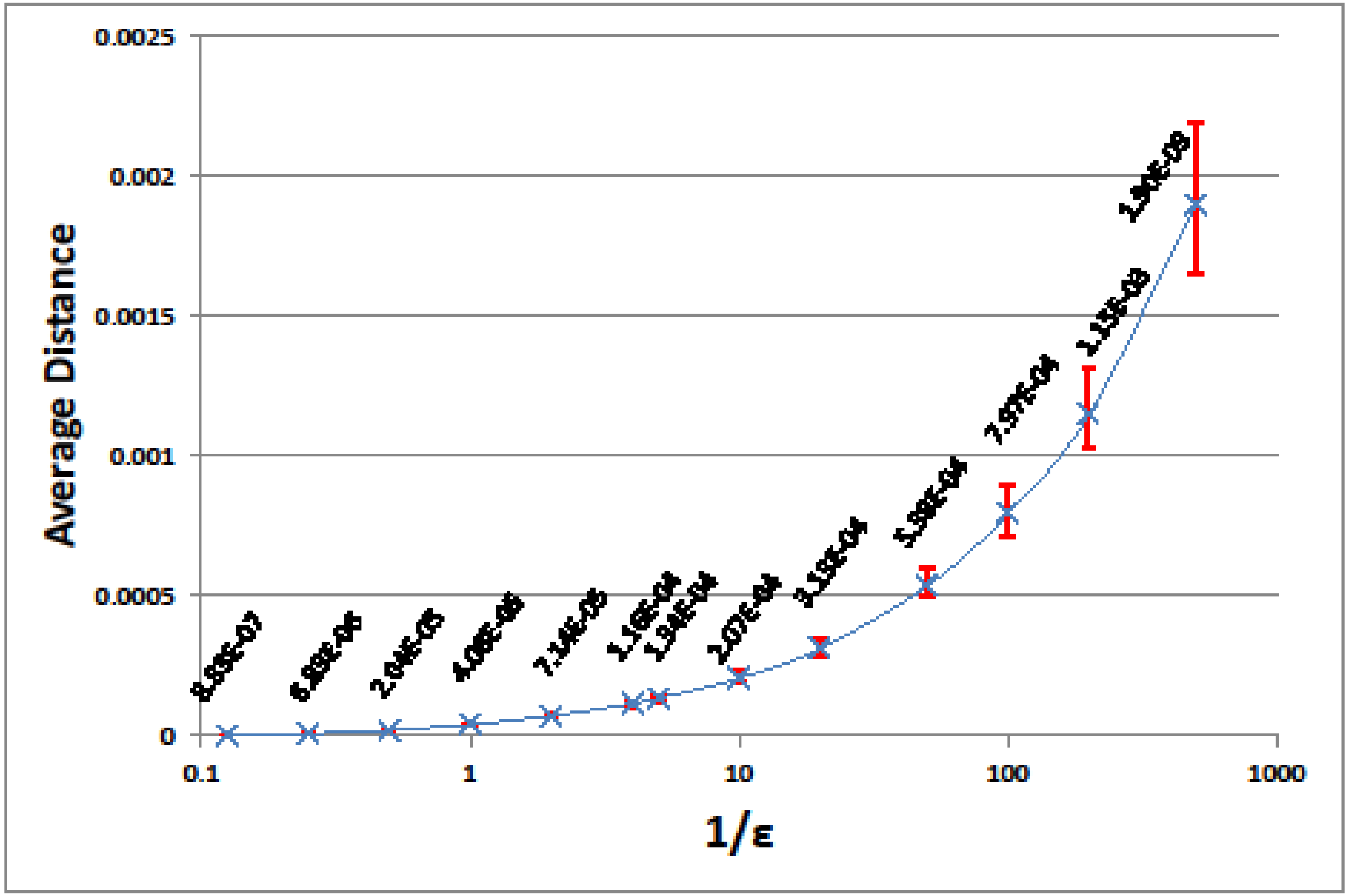
Theorem: There is an efficient algorithm A to sample from a distribution that is δ–close to the exponential mechanism Ɛ over integer partitions. The algorithm uses time and space

$$O\left(\frac{N\sqrt{N} + \mathrm{N}\ln\left(\frac{1}{\delta}\right)}{\varepsilon}\right)$$

Key Idea 1: Novel dynamic programming algorithm to compute weights $W_{i,t}$

Key Idea 2: Allow A to ignore a partition $\tilde{f}$ if $\left\|f - \tilde{f}\right\|_1$ very large.

RockYou Experiments



Yahoo! Results

	Original Data				Sanitized Data			
	N	$\log_2\left(\frac{N}{\lambda_1}\right)$	$\log_2\left(\frac{N}{\lambda_{100}}\right)$	$\log_2(G_{0.5})$	$\tilde{N}$	$\log_2\left(\frac{\tilde{N}}{\tilde{\lambda}_1}\right)$	$\log_2\left(\frac{\tilde{N}}{\tilde{\lambda}_{100}}\right)$	$\log_2(G_{0.5})$
All	69,301,337	6.5	11.4	21.6	69,299,074	6.5	11.4	21.6
gender (self-reported)								
Female	30,545,765	6.9	11.5	21.1	30,545,765	6.9	11.5	21.1
Male	38,624,554	6.3	11.3	21.8	38,624,554	6.3	11.3	21.8
…	…	…	…	…	…	…	…	…
language preference								
Chinese	1,564,364	6.5	11.1	22.0	1,571,348	6.5	11.1	*21.8*
…	…	…	…	…	…	…	…	…

Yahoo! Results

	Original Data				Sanitized Data			
	N	$\log_2\left(\frac{N}{\lambda_1}\right)$	$\log_2\left(\frac{N}{\lambda_{100}}\right)$	$\log_2(G_{0.5})$	$\tilde{N}$	$\log_2\left(\frac{\tilde{N}}{\tilde{\lambda}_1}\right)$	$\log_2\left(\frac{\tilde{N}}{\tilde{\lambda}_{100}}\right)$	$\log_2(G_{0.5})$
All	69,301,337	6.5	11.4	21.6	69,299,074	6.5	11.4	21.6
gender (self-reported)								
Female	30,545,765	6.9	11.5	21.1	30,545,765	6.9	11.5	21.1
Male	38,624,554	6.3	11.3	21.8	38,624,554	6.3	11.3	21.8
…	…	…	…	…	…	…	…	…
language preference								
Chinese	1,564,364	6.5	11.1	22.0	1,571,348	6.5	11.1	*21.8*
…	…	…	…	…	…	…	…	…

Yahoo! Results (Selecting Epsilon)

	Original Data				Sanitized Data			
	N	$\log_2\left(\frac{N}{\lambda_1}\right)$	$\log_2\left(\frac{N}{\lambda_{100}}\right)$	$\log_2(G_{0.5})$	$\tilde{N}$	$\log_2\left(\frac{\tilde{N}}{\tilde{\lambda}_1}\right)$	$\log_2\left(\frac{\tilde{N}}{\tilde{\lambda}_{100}}\right)$	$\log_2(G_{0.5})$
All	69,301,337	6.5	11.4	21.6	69,299,074	6.5	11.4	21.6
[illegible]								
Female	30,545,765	[illegible]	11.5	[illegible]	[illegible]	6.9	11.5	21.1
Male	38,624,554	6.3	11.3	21.8	38,624,554	6.3	11.3	21.8
…	…	…	…	…	…	…	…	…
language preference								
Chinese	1,564,364	6.5	11.1	22.0	1,571,348	6.5	11.1	*21.8*
…	…	…	…	…	…	…	…	…

Any individual participates in at most 23 groups (including All)

$$\varepsilon = \varepsilon_{all} + 22\varepsilon'$$

Yahoo! Results (Selecting Epsilon)

	Original Data				Sanitized Data			
	N	$\log_2\left(\frac{N}{\lambda_1}\right)$	$\log_2\left(\frac{N}{\lambda_{100}}\right)$	$\log_2(G_{0.5})$	$\tilde{N}$	$\log_2\left(\frac{\tilde{N}}{\tilde{\lambda}_1}\right)$	$\log_2\left(\frac{\tilde{N}}{\tilde{\lambda}_{100}}\right)$	$\log_2(G_{0.5})$
All	69,301,337		11.4		69,295,074	6.5	11.4	21.6
gender (self-reported)								
Female	30,545,765		11.5	21.1	30,545,765	6.9	11.5	21.1
Male	38,624,554	6.3	11.3	21.8	38,624,554	6.3	11.3	21.8
…	…							…
language preference								
Chinese	1,564,364	6.5	11.1	22.0	1,571,348	6.5	11.1	*21.8*
…	…	…	…	…	…	…	…	…

$$\varepsilon_{all} = 0.25$$

$$\varepsilon' = \frac{\varepsilon_{all}}{22}$$

$$\varepsilon = \varepsilon_{all} + 22\varepsilon'$$

Yahoo! Results (Selecting Epsilon)

	Original Data				Sanitized Data			
	N	$\log_2\left(\frac{N}{\lambda_1}\right)$	$\log_2\left(\frac{N}{\lambda_{100}}\right)$	$\log_2(G_{0.5})$	$\tilde{N}$	$\log_2\left(\frac{\tilde{N}}{\tilde{\lambda}_1}\right)$	$\log_2\left(\frac{\tilde{N}}{\tilde{\lambda}_{100}}\right)$	$\log_2(G_{0.5})$
All	69,301,337	6.5	11.4	21.6	69,299,074	6.5	11.4	21.6
gender (self-reported)								
Female	30,545,765	6.9	11.5	21.1	30,545,765	6.9	11.5	21.1
Male	38,624,554	6.3	11.3	21.8	38,624,554	6.3	11.3	21.8
…	**…**	**…**	**…**	**…**	**…**	**…**	**…**	**…**
language preference								
Chinese	1,564,364	6.5	11.1	22.0	1,571,348	6.5	11.1	*21.8*
…	**…**	**…**	**…**	**…**	**…**	**…**	**…**	**…**

$$\varepsilon = 0.5$$

Yahoo! Results (Selecting Epsilon)

	Original Data				Sanitized Data			
	N	$\log_2\left(\frac{N}{\lambda_1}\right)$	$\log_2\left(\frac{N}{\lambda_{100}}\right)$	$\log_2(G_{0.5})$	$\tilde{N}$	$\log_2\left(\frac{\tilde{N}}{\tilde{\lambda}_1}\right)$	$\log_2\left(\frac{\tilde{N}}{\tilde{\lambda}_{100}}\right)$	$\log_2(G_{0.5})$
All	69,301,337	6.5	11.4	21.6	69,299,074	6.5	11.4	21.6
gender (self-reported)								
Female	30,545,765	6.9	11.5	21.1	30,545,765	6.9	11.5	21.1
Male	38,624,554	6.3	11.3	21.8	38,624,554	6.3	11.3	21.8
…	…	…	…	…	…	…	…	…
language preference								
Chinese	1,564,364	6.5	11.1	22.0	1,571,348	6.5	11.1	*21.8*
…	…	…	…	…	…	…	…	…

$$\varepsilon = 0.5, \qquad \delta = 2^{-100}$$

Conclusions

- Novel differentially private algorithm for integer partitions
    - Password Frequency Lists
    - Degree Distribution in a Social Network?
    - Other applications?
- The Yahoo! Frequency data is now available
    - Search: “Yahoo! Password Frequency Corpus”
    - What exciting things can we do with it?
- Hope for other organizations to imitate Yahoo!