Distributed Differential Privacy in Multi-Armed Bandits

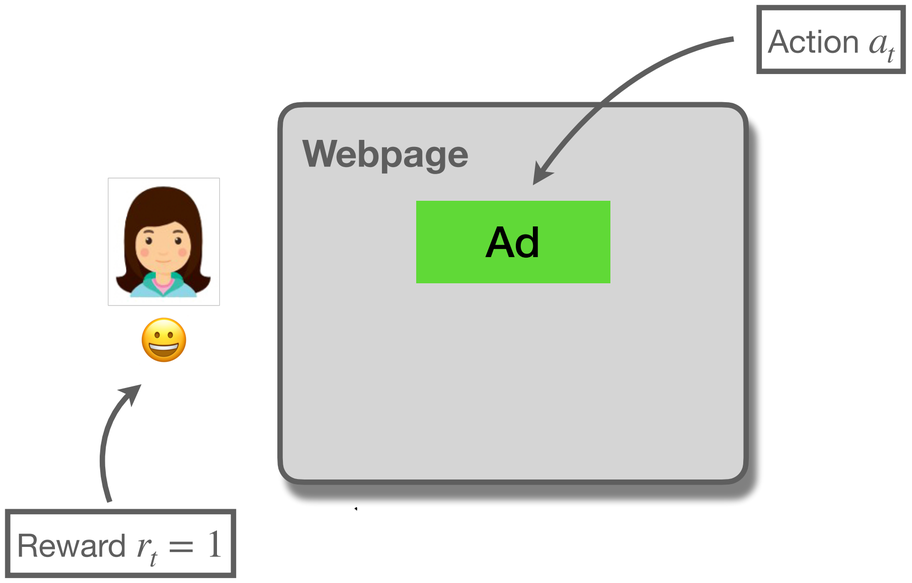
Xingyu Zhou*, Sayak Ray Chowdhury*

Wayne State University, Microsoft Research, India

* Equal Contributions

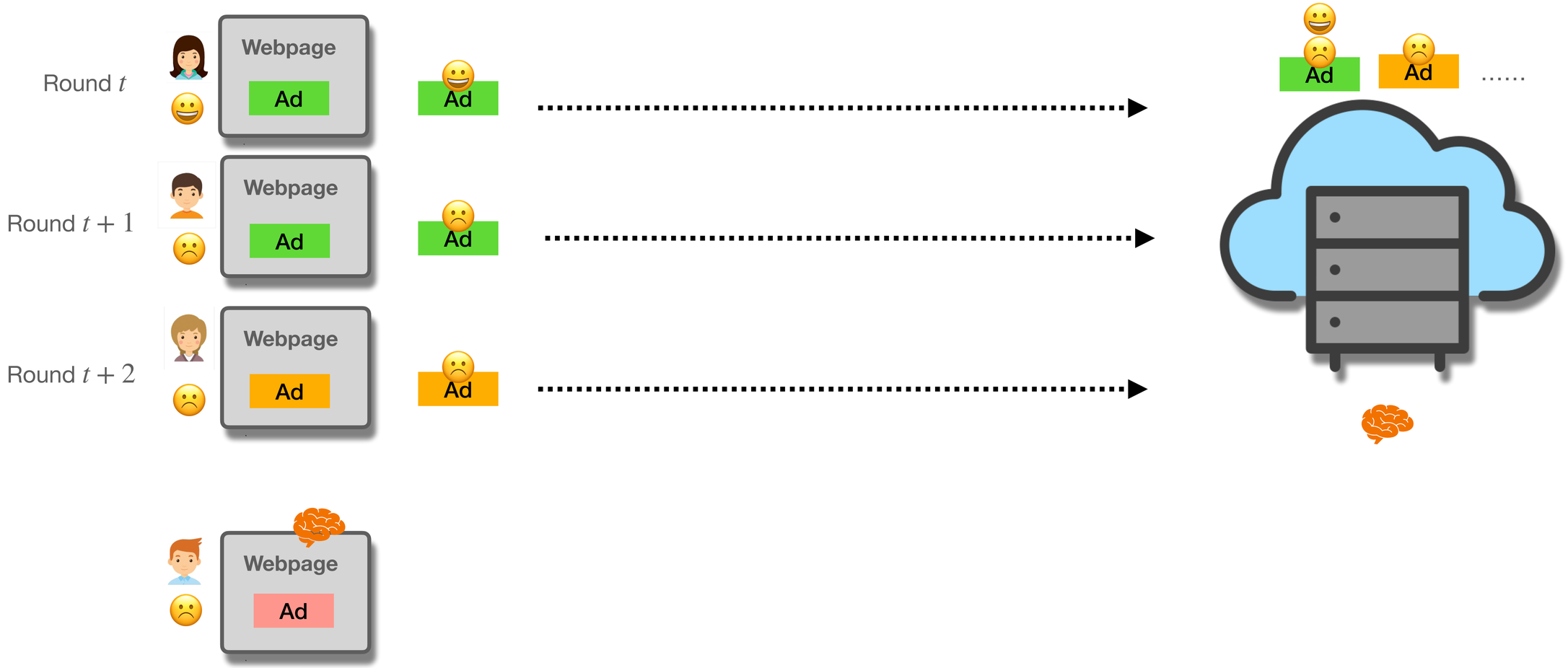
Multi-Armed Bandits

Ads recommendation example



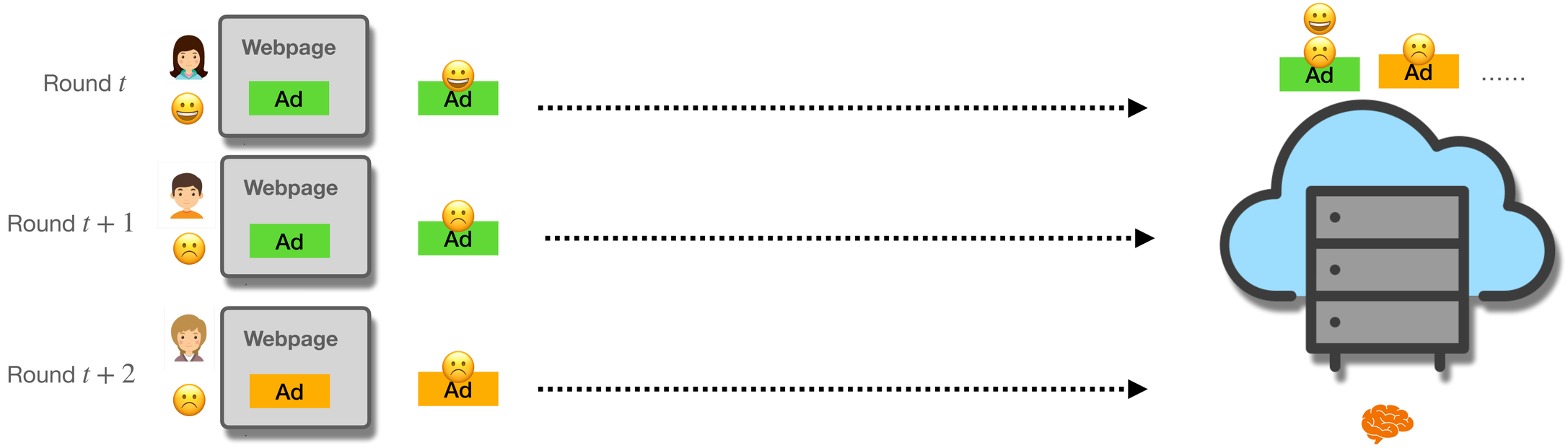
Multi-Armed Bandits

Ads recommendation example



Multi-Armed Bandits

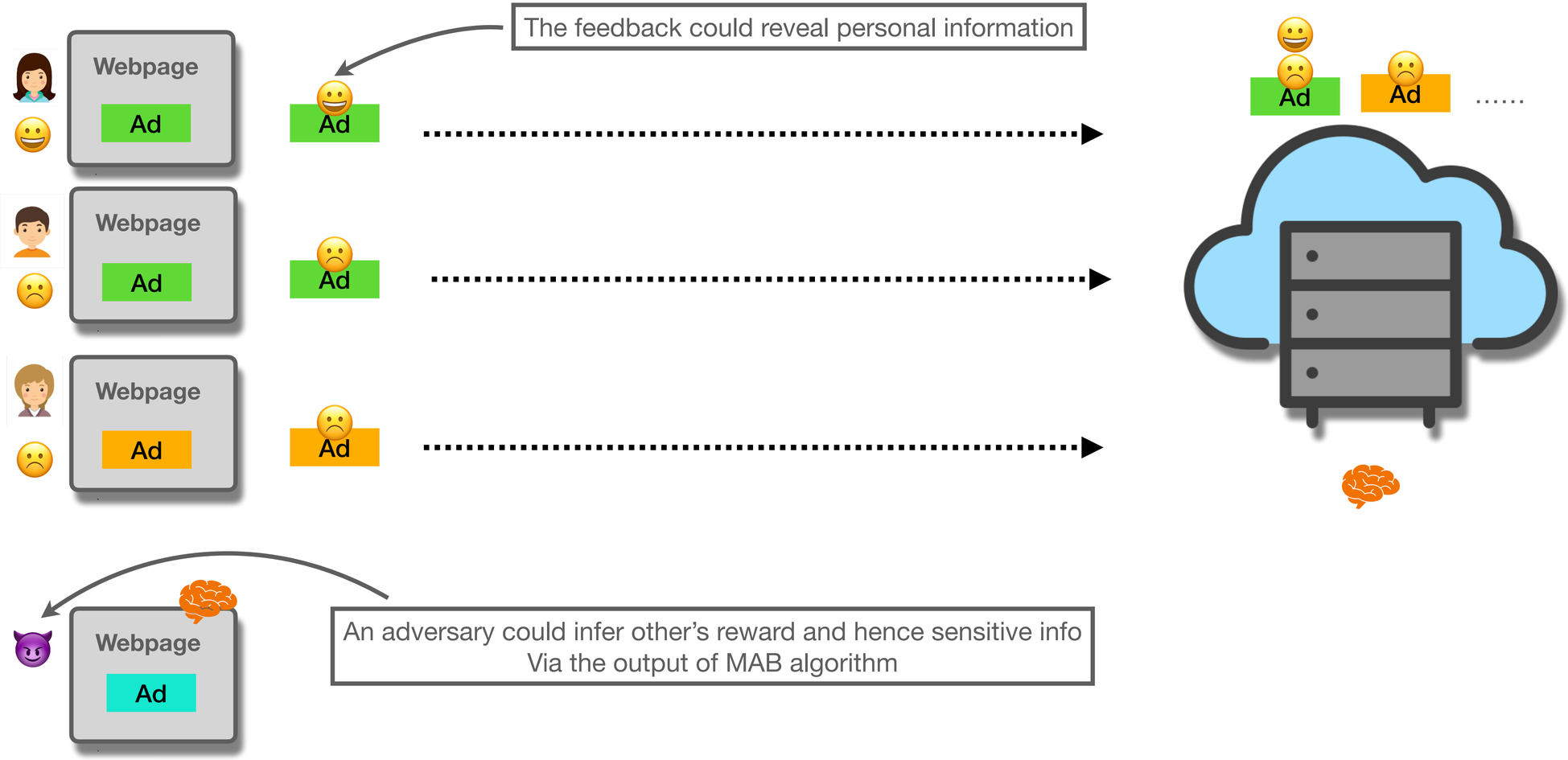
Regret minimization



$$\mathbb{E}[\text{Reg}(T)] = T \cdot \mu_{a^*} - \mathbb{E}\left[\sum_{t=1}^{T} \mu_{a_t}\right]$$

Privacy Concern

Reward is sensitive



Differential Privacy

Central model

Differential Privacy

For any two neighboring datasets D and D', and any outcome E

$$\mathbb{P}(M(D) \in E) \leq e^{\epsilon}\mathbb{P}(M(D') \in E) + \delta$$

Differential Privacy

Central model

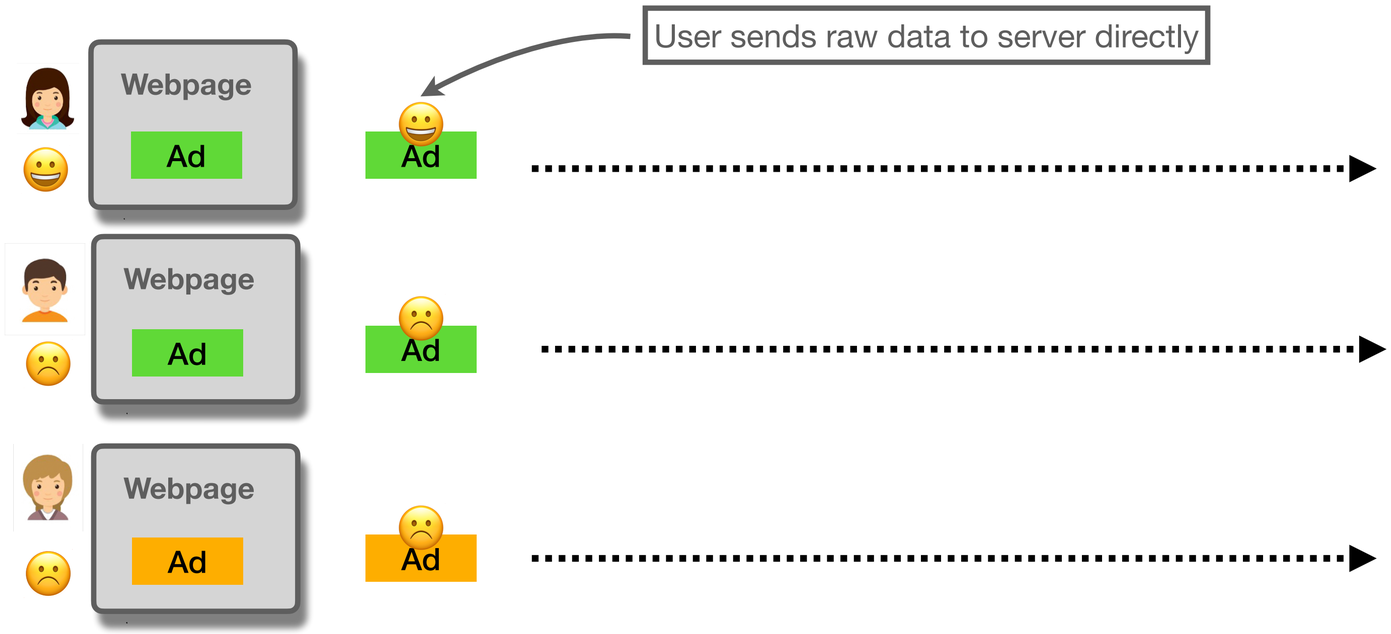
Differential Privacy

For any two neighboring datasets D and D', and any outcome E

$$\mathbb{P}(M(D) \in E) \leq e^{\epsilon}\mathbb{P}(M(D') \in E) + \delta$$

Differential Privacy

Central model



Differential Privacy

For any two neighboring datasets D and D', and any outcome E

$$\mathbb{P}(M(D) \in E) \leq e^{\epsilon}\mathbb{P}(M(D') \in E) + \delta$$

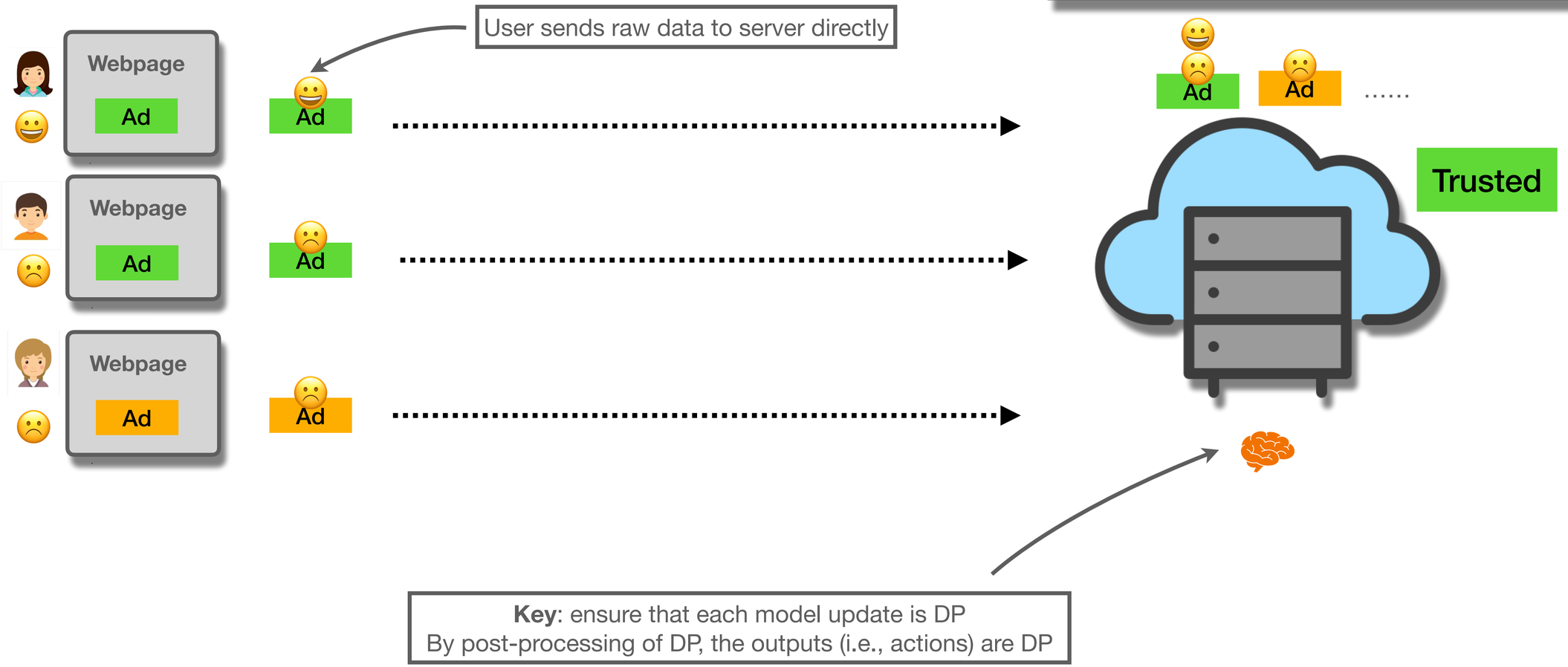
Differential Privacy

Central model

Differential Privacy

For any two neighboring datasets D and D', and any outcome E

$$\mathbb{P}(M(D) \in E) \leq e^{\epsilon}\mathbb{P}(M(D') \in E) + \delta$$



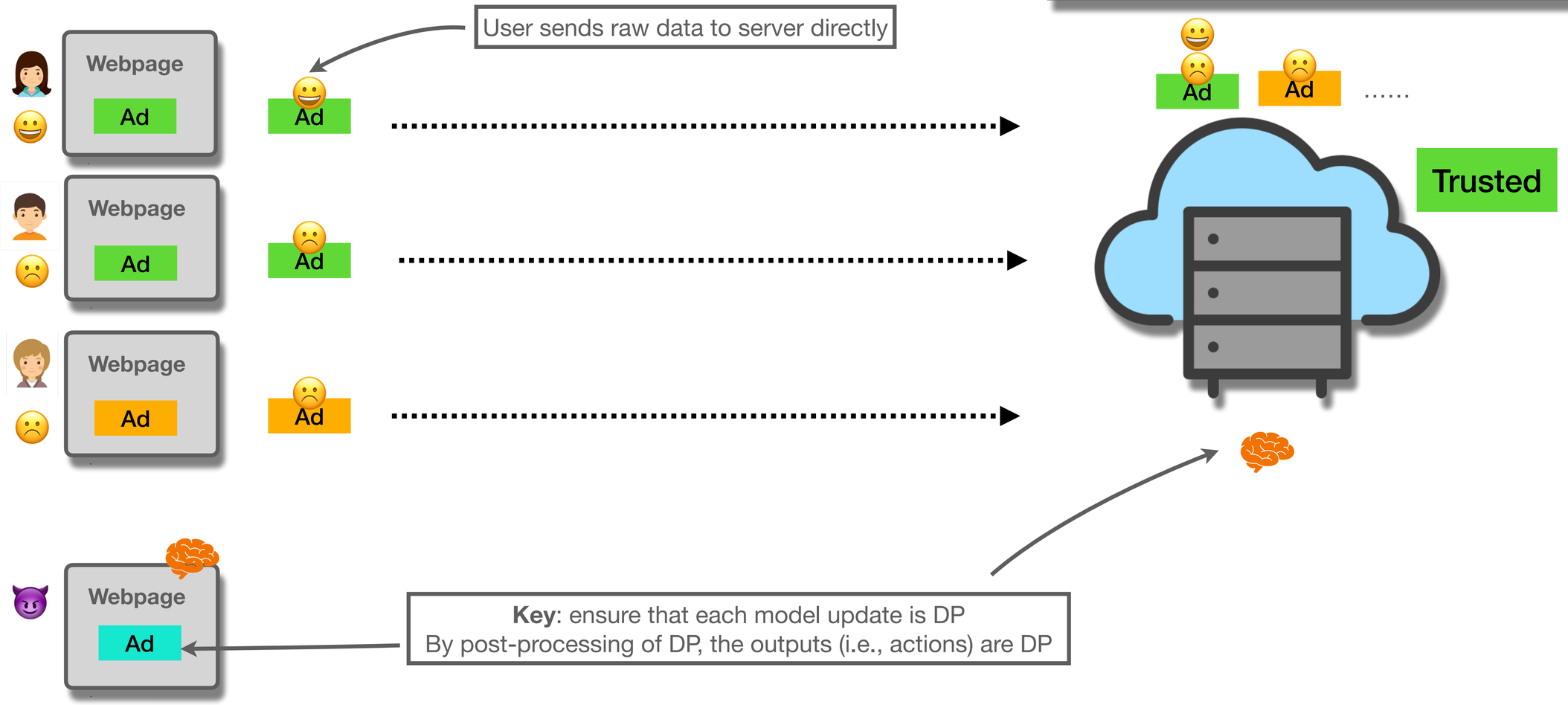
Differential Privacy

Central model

Differential Privacy

For any two neighboring datasets D and D', and any outcome E

$$\mathbb{P}(M(D) \in E) \leq e^{\epsilon}\mathbb{P}(M(D') \in E) + \delta$$



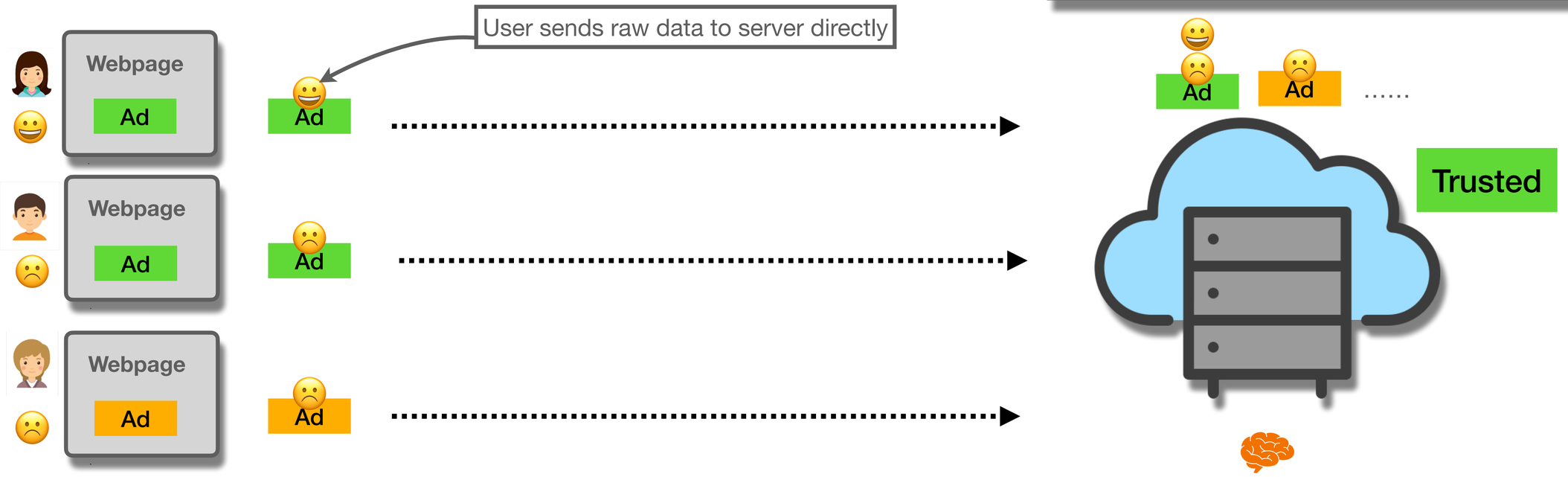
Optimal Regret in MAB

Central model

Differential Privacy (ϵ, δ)

For any two neighboring datasets D and D', and any outcome E

$$\mathbb{P}(M(D) \in E) \leq e^{\epsilon}\mathbb{P}(M(D') \in E) + \delta$$



[Sajed & Sheffet'19]

Algorithm: successive arm elimination + batching + Laplace noise

Privacy : pure DP , i.e., $(\epsilon,0)$-DP

Regret: optimal non-private regret + $\Theta\left(\frac{K \log T}{\epsilon}\right)$

Only an additive privacy cost
And it is optimal!

Differential Privacy

Local model

Differential Privacy

For any two neighboring datasets D and D', and any outcome E

$$\mathbb{P}(M(D) \in E) \leq e^{\epsilon}\mathbb{P}(M(D') \in E) + \delta$$

Untrusted

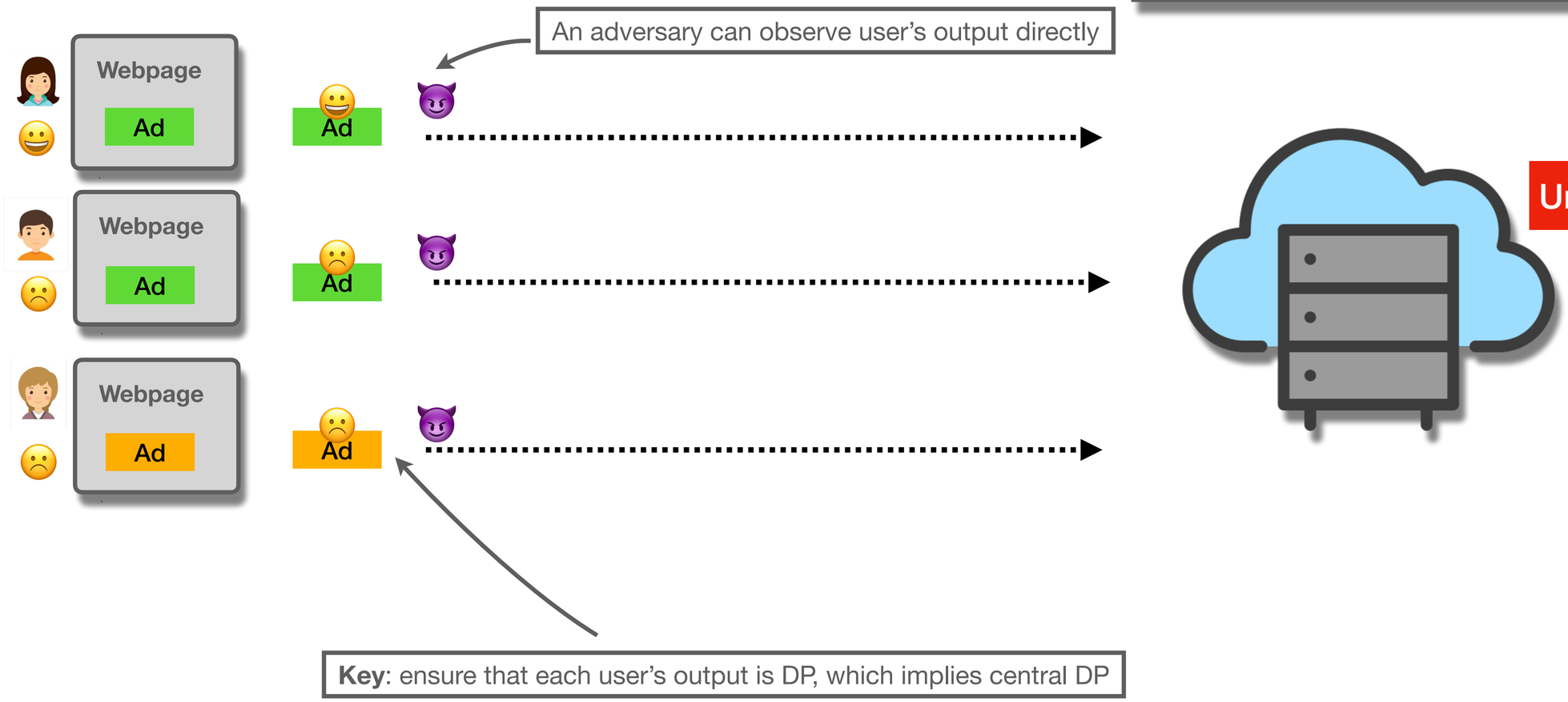
Differential Privacy

Local model

Differential Privacy

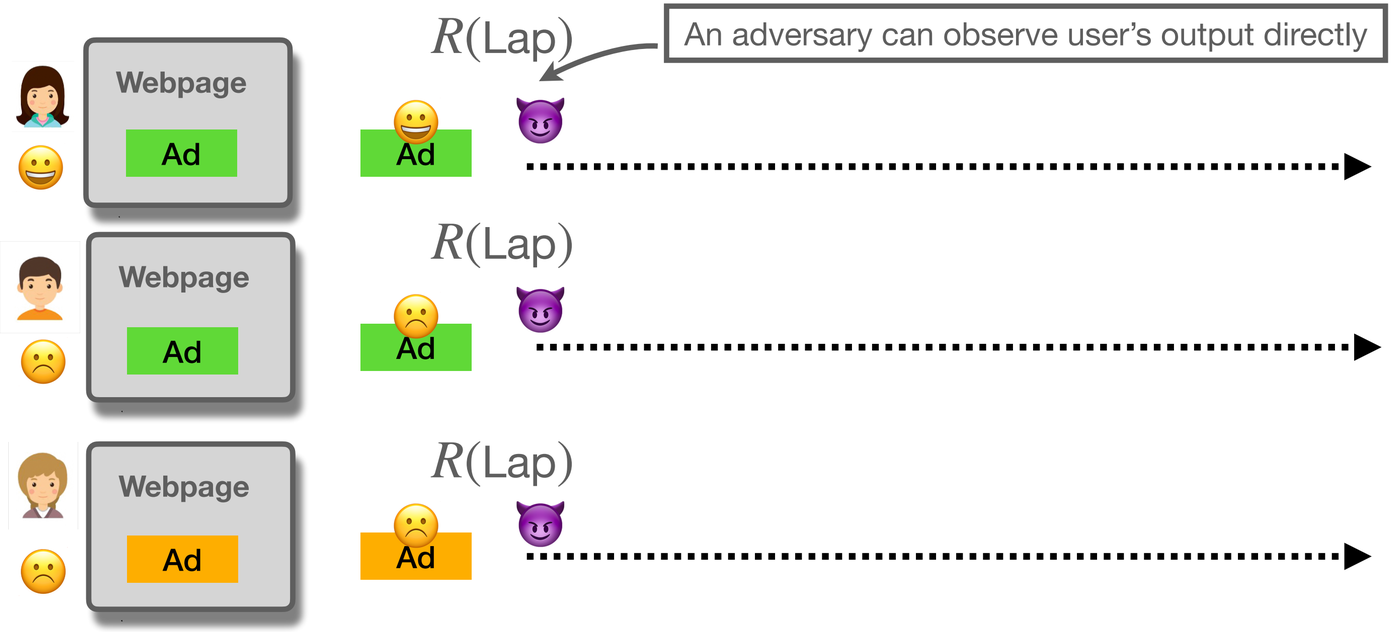
For any two neighboring datasets D and D', and any outcome E

$$\mathbb{P}(M(D) \in E) \leq e^{\epsilon}\mathbb{P}(M(D') \in E) + \delta$$



Optimal Regret in MAB

Local model



Differential Privacy

For any two neighboring datasets D and D', and any outcome E

$$\mathbb{P}(M(D) \in E) \leq e^{\epsilon}\mathbb{P}(M(D') \in E) + \delta$$

[Ren, Zhou, Liu, Shroff'20]

Algorithm: Laplace noise at local side + UCB

Privacy : pure DP , i.e., $(\epsilon,0)$-DP

Regret: $\approx \frac{1}{\epsilon^2}$ optimal non-private regret

Due to a large amount of noise
Privacy cost is now multiplicative

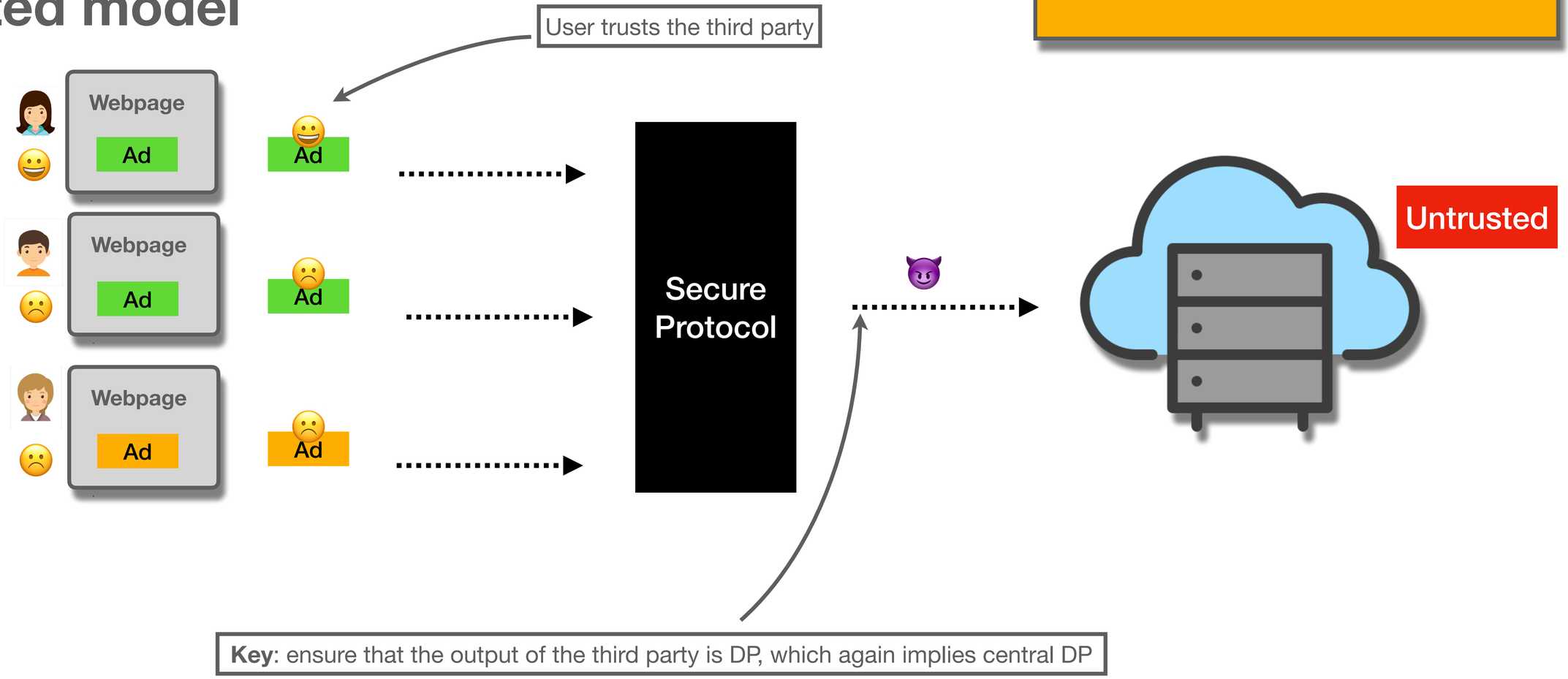
Differential Privacy

Distributed model

Differential Privacy

For any two neighboring datasets D and D', and any outcome E

$$\mathbb{P}(M(D) \in E) \leq e^{\epsilon}\mathbb{P}(M(D') \in E) + \delta$$



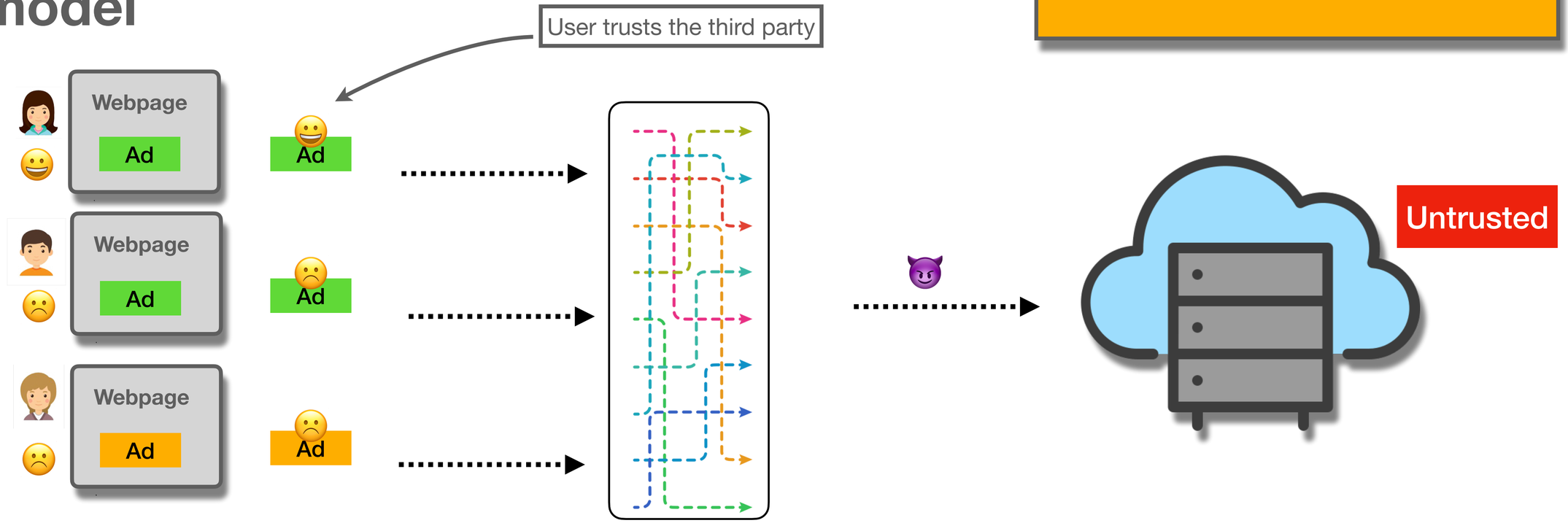
Distributed DP in MAB

Shuffle model

Differential Privacy

For any two neighboring datasets D and D', and any outcome E

$$\mathbb{P}(M(D) \in E) \leq e^{\epsilon}\mathbb{P}(M(D') \in E) + \delta$$



[Tenenbaum, Kaplan, Mansou, Stemme'21]

Algorithm: successive arm elimination + batching + shuffle protocol

Privacy: *approximate* DP , i.e., (ϵ, δ)-DP

Regret: optimal non-private regret + $O\left(\frac{K \log T\sqrt{\log(1/\delta)}}{\epsilon}\right)$

[Limitations]

Privacy: only approximate DP rather than pure DP

Regret: not optimal, additional log factors

Communication: current scheme only works for binary reward
* adapt to other scheme incurs extensive communication

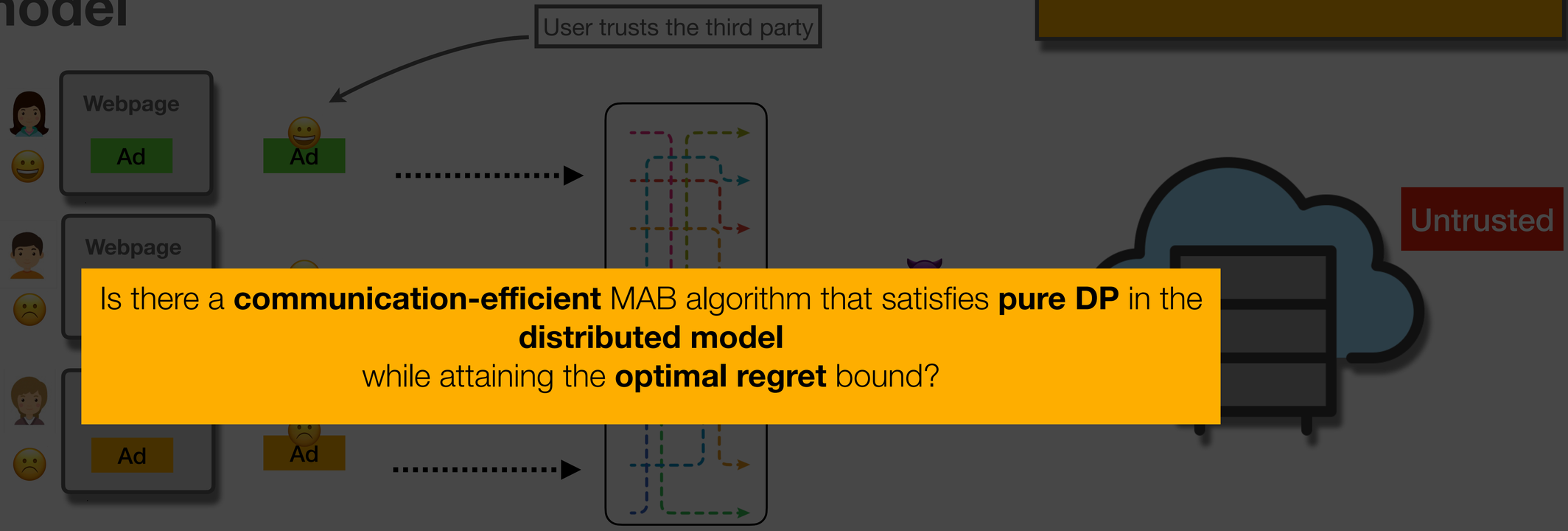
Distributed DP in MAB

Shuffle model

Differential Privacy

For any two neighboring datasets D and D', and any outcome E

$$\mathbb{P}(M(D) \in E) \leq e^{\epsilon}\mathbb{P}(M(D') \in E) + \delta$$



Is there a **communication-efficient** MAB algorithm that satisfies **pure DP** in the **distributed model**
while attaining the **optimal regret** bound?

[Tenenbaum, Kaplan, Mansou, Stemme'21]

Algorithm: successive arm elimination + batching + shuffle protocol

Privacy: *approximate* DP , i.e., (ϵ, δ)-DP

Regret: optimal non-private regret + $O\left(\frac{K \log T\sqrt{\log(1/\delta)}}{\epsilon}\right)$

[Limitations]

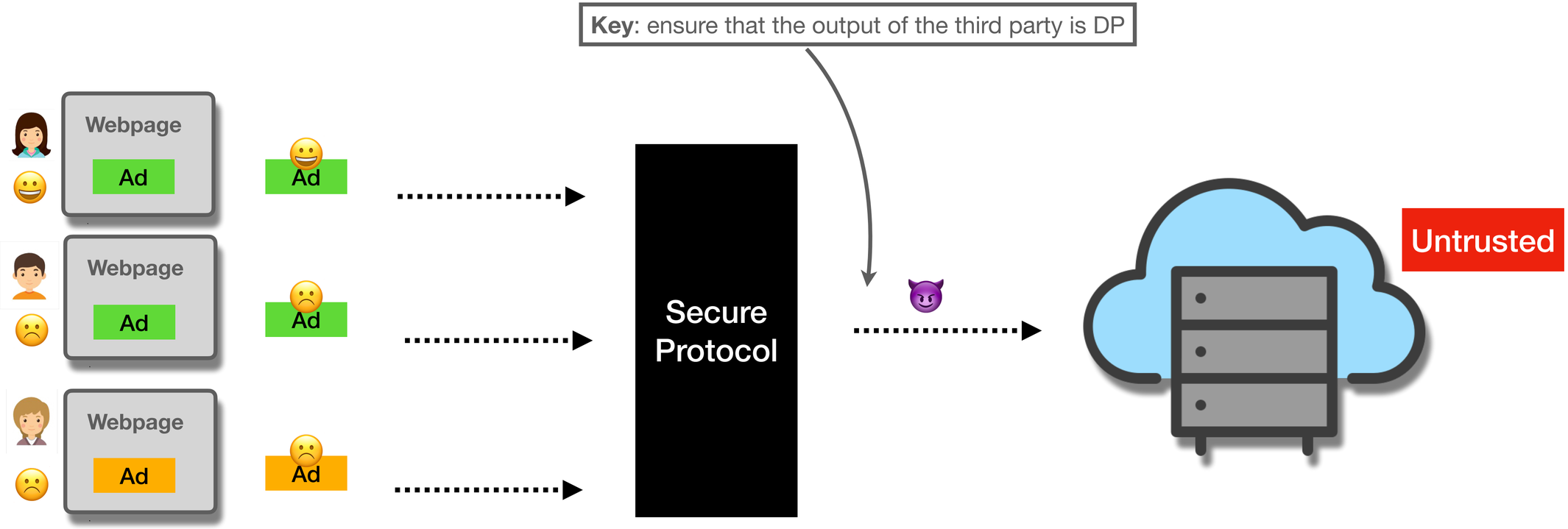
Privacy: only approximate DP rather than pure DP

Regret: not optimal, additional log factors

Communication: current scheme only works for binary reward
* adapt to other scheme incurs extensive communication

Contribution

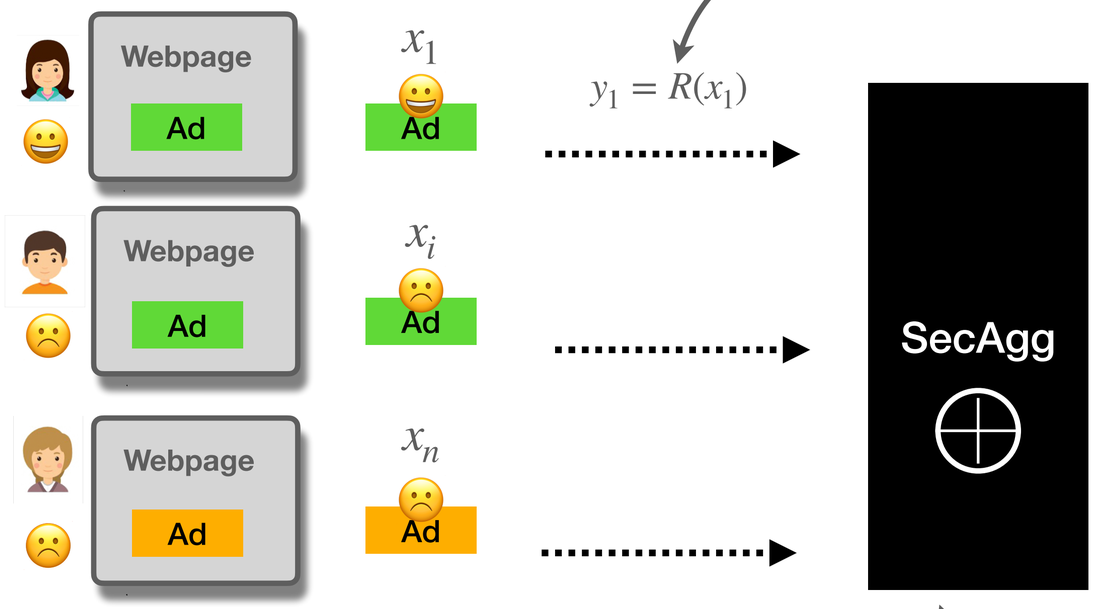
Main Results



1. The first algorithm to achieve optimal regret with pure DP in distributed model
2. The first algorithm to achieve RDP using only discrete privacy noise
3. A unified algorithmic framework for achieving optimal regret under central, local, distributed model
4. Extensive simulations and experiments to validate our theoretical results

Our Algorithm

Each local randomizer adds privacy noise on reward
Due to SecAgg: (i) only **discrete** privacy noise (ii) **modular** clipping

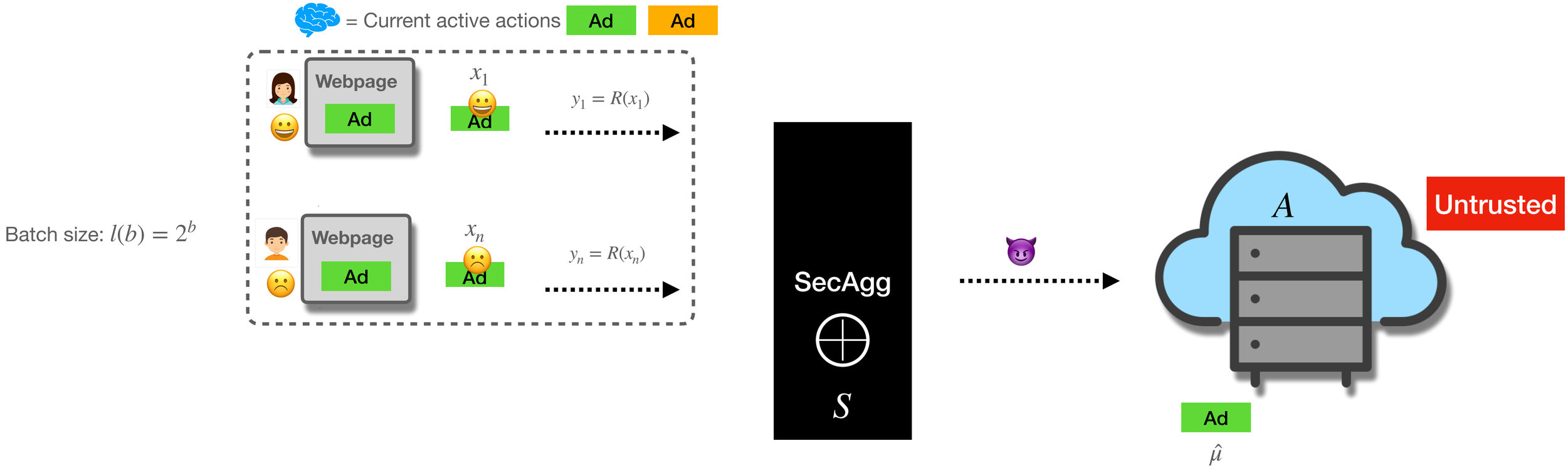


Given output from SecAgg, $\hat{y}$, try to calculate $z = A(\hat{y}) \approx \sum_{i=1}^{n} x_i$

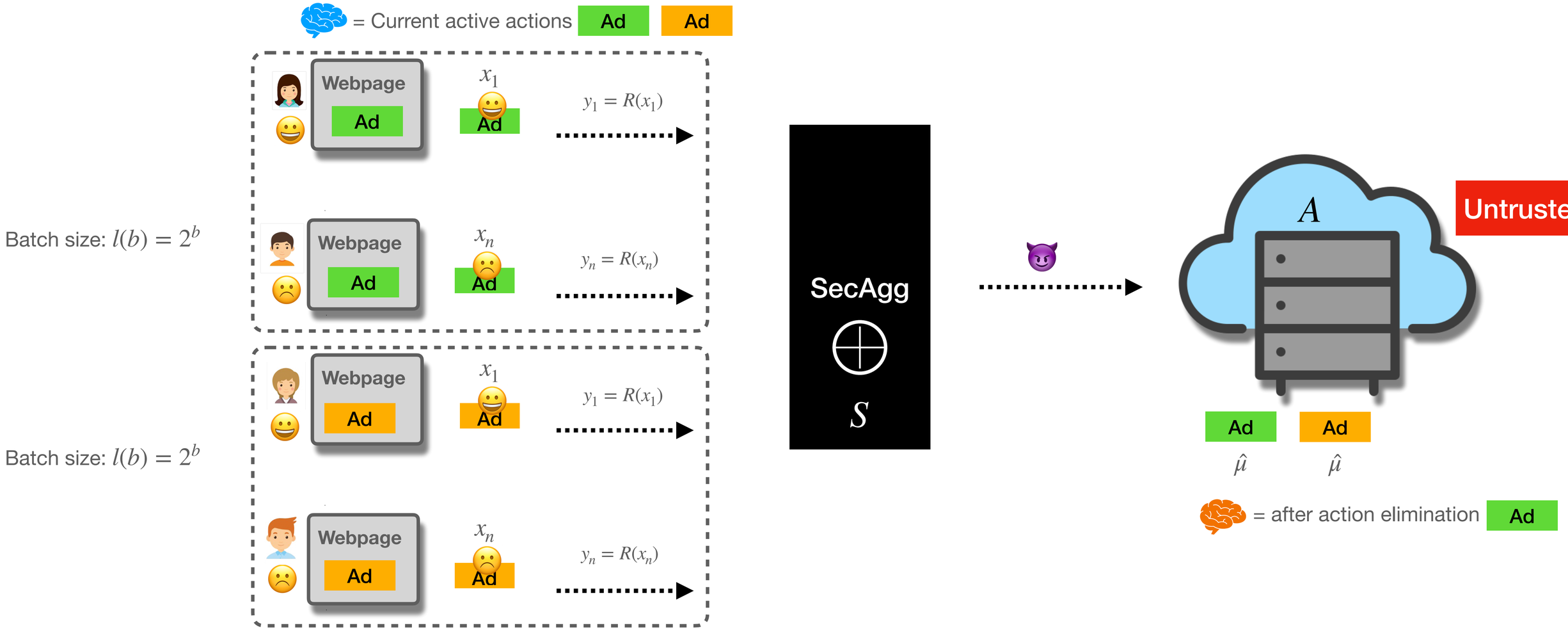
Given $y_1, y_2, \ldots, y_n$, output $\hat{y} = \left(\sum_{i=1}^{n} y_i \right) \mod m$

This alone does not provide formal privacy

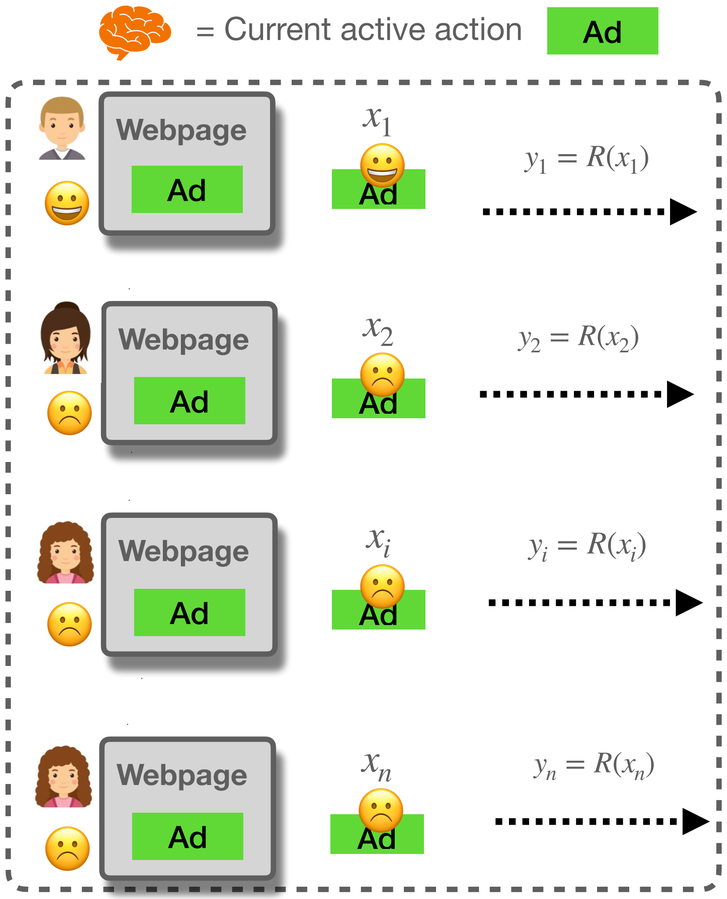
Our Algorithm



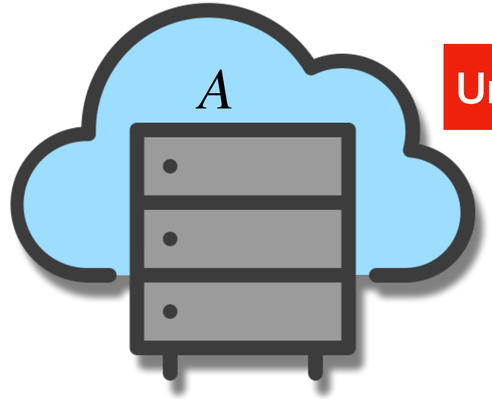
Our Algorithm



Our Algorithm



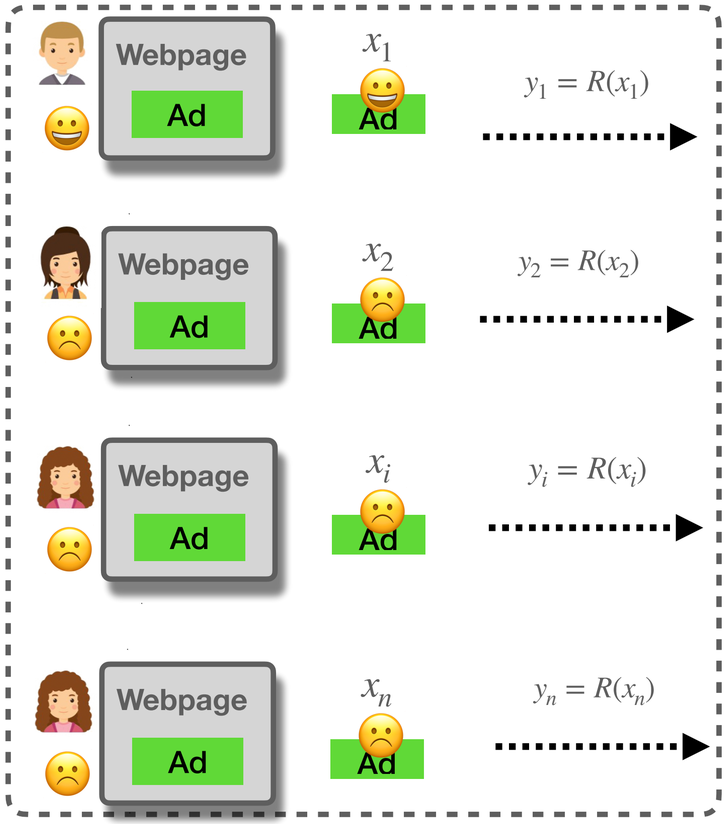
Batch size: $l(b) = 2^{b+1}$



Our Algorithm

= Current active action **Ad**

Batch size: $l(b) = 2^{b+1}$



Analyzer A

[Balle et al' 20, Cheu & Yan' 21]

Input $\hat{y}$, Output z

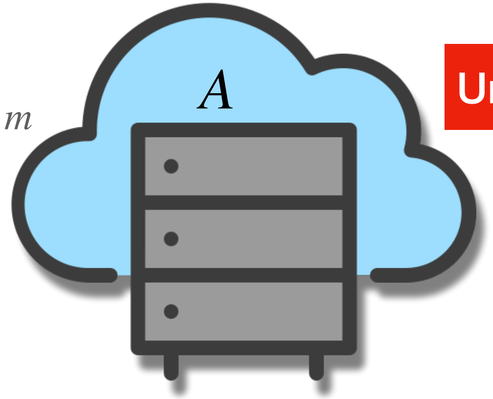
If $\hat{y} > ng + \tau$: $z = (\hat{y} - m)/g$; else $z = \hat{y}/g$

SecAgg

$\oplus$

S

$$\hat{y} = \left(\sum_{i=1}^{n} y_i \right) \mod m$$



Untrusted

How to choose privacy noise?

Local Randomizer R

[Balle et al' 20, Cheu & Yan' 21]

Input $x_i \in [0,1]$, Output y_i

$$x_i \xrightarrow{\lfloor x_i g \rfloor + \text{Ber}(x_i g - \lfloor x_i g \rfloor)} \hat{x}_i \xrightarrow{(\hat{x}_i + \eta_i) \mod m} y_i$$

Achieving Pure DP

Simulate discrete Laplace using Polya noise

Theorem 1 (Pure-DP via SecAgg)

Fix $\epsilon > 0$ and T. For each batch b, the noise $\eta_i = \gamma_i^+ - \gamma_i^-$, where $\gamma_i^+, \gamma_i^- \sim^{i.i.d} Poyla(1/n, e^{-\epsilon/g})$.

There exist proper choices of g, m, τ such that

Privacy: pure DP in the distributed model

Regret: optimal non-private regret + $\Theta\left(\frac{K \log T}{\epsilon}\right)$

Communication: bits scales logarithmically with the batch size

Local Randomizer R

[Balle et al' 20, Cheu & Yan' 21]

Input $x_i \in [0,1]$, Output y_i

$$x_i \xrightarrow{\lfloor x_i g \rfloor + \text{Ber}(x_i g - \lfloor x_i g \rfloor)} \hat{x}_i \xrightarrow{(\hat{x}_i + \eta_i) \mod m} y_i$$

Achieving Pure DP

Simulate discrete Laplace using Polya noise

Theorem 1 (Pure-DP via SecAgg)

Fix $\epsilon > 0$ and T. For each batch b, the noise $\eta_i = \gamma_i^+ - \gamma_i^-$, where $\gamma_i^+, \gamma_i^- \sim^{i.i.d} Poyla(1/n, e^{-\epsilon/g})$.

There exist proper choices of g, m, τ such that

Privacy: pure DP in the distributed model

Regret: optimal non-private regret + $\Theta\left(\frac{K \log T}{\epsilon}\right)$

Communication: bits scale logarithmically with the batch size

Remark on privacy

- First result on pure DP in distribute model for MABs
- Also achieve pure DP using advanced shuffle protocol

Remark on regret

- Match the optimal regret under central model
- Only use discrete privacy noise, w/o finite precision approx

Remark on communication

- Only communicate bits
- Previous works scale polynomially

Achieving RDP

Skellam noise

$(\alpha, \epsilon(\alpha))$-RDP

For any two neighboring datasets D and D', and any outcome E, $D_\alpha(M(D), M(D')) \leq \epsilon(\alpha)$

Theorem 2 (RDP via SecAgg)

Fix $\epsilon > 0$ and T. For each batch b, let the noise be $\eta_i \sim SK\left(0, \frac{g^2}{n\epsilon^2}\right)$.

There exist proper choices of g, m, τ such that

Privacy: $\approx (\alpha, \frac{\alpha\epsilon^2}{2})$- RDP

Regret: $\approx$ optimal non-private regret + $\Theta\left(\frac{K\sqrt{\log T}}{\epsilon}\right)$

Communication: scales logarithmically with the batch size

Achieving RDP

Skellam noise

Proposition (Tail bound of Skellam noise)

Skellam random variable has a sub-exponential tail

Theorem 2 (RDP via SecAgg)

Fix $\epsilon > 0$ and T. For each batch b, let the noise be $\eta_i \sim SK\left(0, \frac{g^2}{n\epsilon^2}\right)$.

There exist proper choices of g, m, τ such that

Privacy: $\approx (\alpha, \frac{\alpha\epsilon^2}{2})$- RDP

Regret: $\approx$ optimal non-private regret + $\Theta\left(\frac{K\sqrt{\log T}}{\epsilon}\right)$

Communication: scales logarithmically with the batch size

Achieving RDP

Skellam noise

Theorem 2 (RDP via SecAgg)

Fix $\epsilon > 0$ and T. For each batch b, let the noise be $\eta_i \sim SK\left(0, \frac{g^2}{n\epsilon^2}\right)$.

There exist proper choices of g, m, τ such that

Privacy: $\approx (\alpha, \frac{\alpha\epsilon^2}{2})$- RDP

Regret: $\approx$ optimal non-private regret + $\Theta\left(\frac{K\sqrt{\log T}}{\epsilon}\right)$

Communication: scales logarithmically with the batch size

Remark 1 (Conversion to approximate DP)

- After conversion, it is $\sqrt{\log T}$ tighter than SOTA

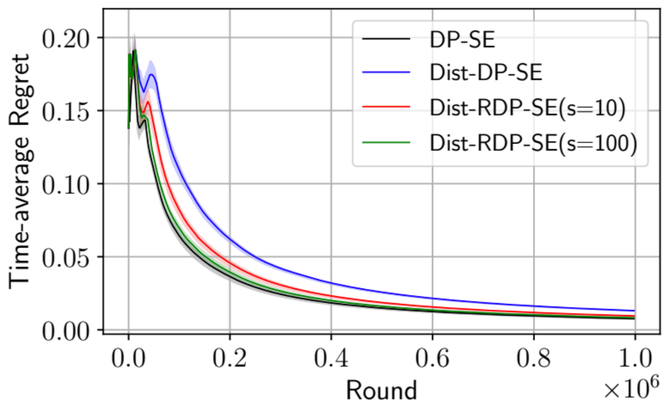
Remark 2 (Tight privacy accounting)

- Useful for analyzing returning users

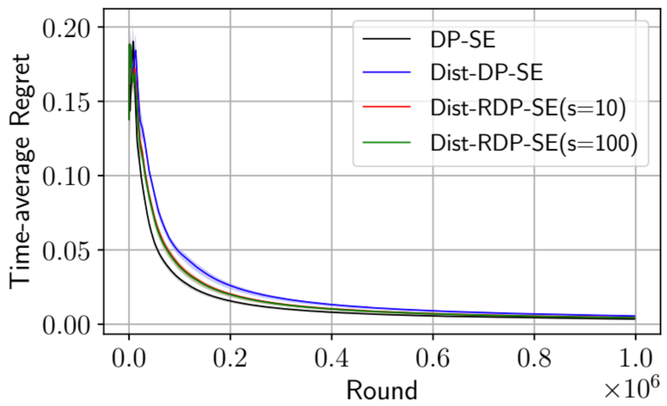
Remark 3 (RDP in other models)

- Our algorithm can be adapted to central and local
- Hence, the first result of RDP guarantees

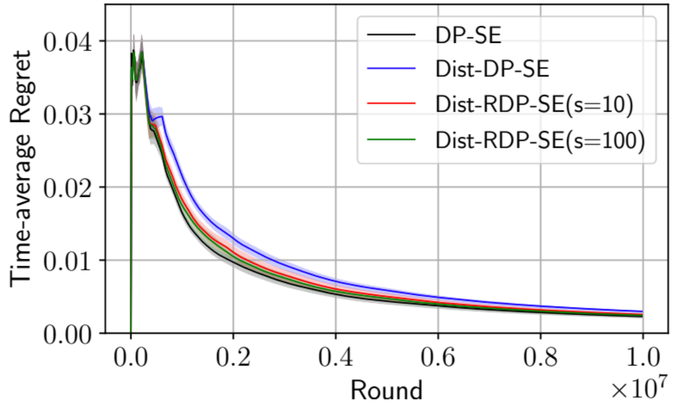
Simulations



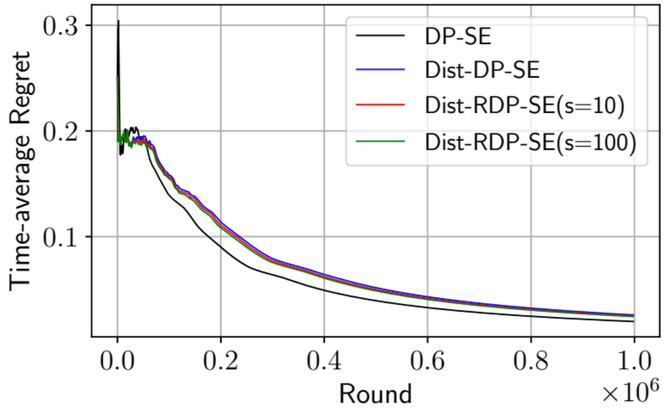
(a) $\varepsilon = 0.1, K = 10$



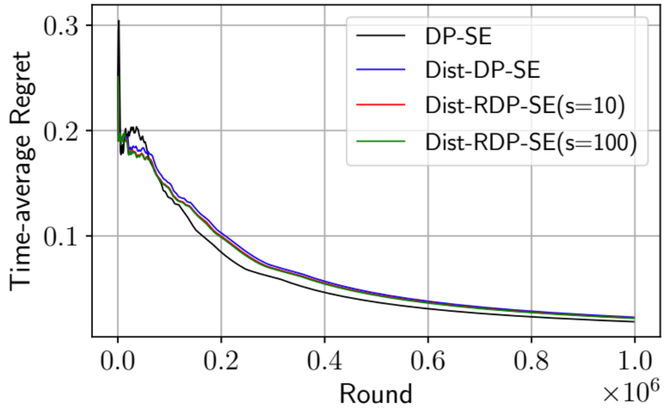
(b) $\varepsilon = 0.5, K = 10$



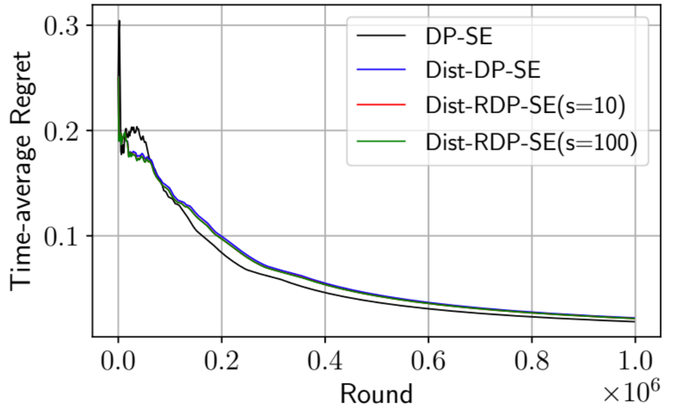
(c) $\varepsilon = 0.1, K = 10$



(d) $\varepsilon = 1, K = 50$



(e) $\varepsilon = 5, K = 50$



(f) $\varepsilon = 10, K = 50$

Thank you!