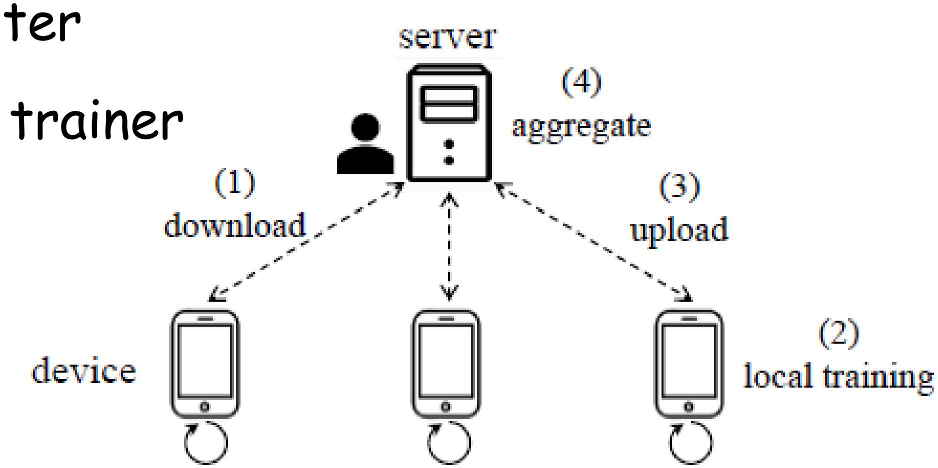
A Reward Response Game in the Federated Learning System

Suhan Jiang and Jie Wu
Dept. of Computer and Information Sciences
Temple University, USA

Mobile-crowd Federated Learning

- Classic federated learning system
  - A central server: model requester
  - A set of mobile devices: model trainer
    - Each device has its own data
- Model construction
  - A global training round
    1. server sends current global model to all devices
    2. each device trains its model using the local data
    3. all devices upload their updated models to server
    4. server aggregates all local models into a new global model.
  - Repeat forever or until meeting certain requirements



Economic Model

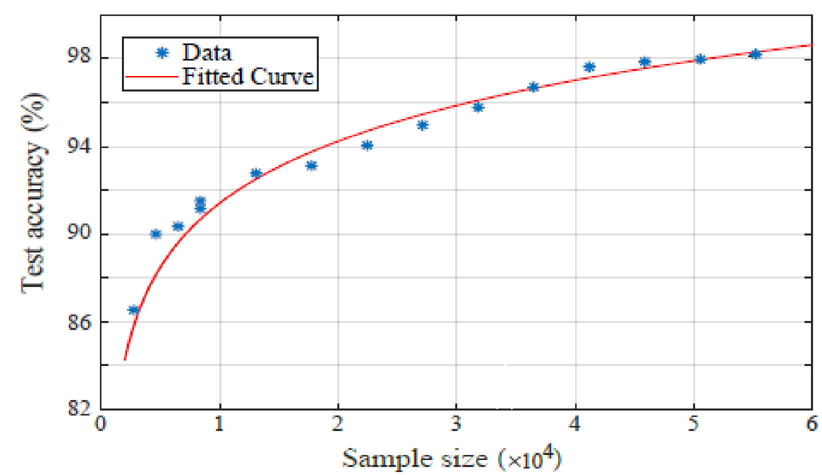
- Server
  - Monetary incentive: motivate devices with enough rewards
  - Model accuracy:
    - positively related to the size of trained data
    - diminishing marginal return
  - Trade-off: reward Vs. accuracy
- Device
  - Resource consumption: training speed
  - Contribution measurement:
    - Size of trained data
    - local model accuracy
  - Upload time

Device Side Problem

- Each device i determines its training time t_i to

$$
\begin{aligned}
&\textit{maximize} && u_i(t_i, \boldsymbol{t}_{-i}) = R\frac{\alpha_i}{\alpha} - c_i t_i, \\
&\textit{where} && \alpha = \sum\nolimits_{j=1}^{N} \alpha_j, \\
&\textit{subject to} && t_i + \tau_i \leq T.
\end{aligned}
$$

- Individual contribution: $a_i = \alpha_i(t_i)$
  - Size-based: $\alpha_i = \beta_i t_i$
  - Accuracy-based: $\alpha_i = \theta \log\left(1 + \lambda \beta_i t_i\right)$



Service Side Problem

- Central server determines its reward R to

$$maximize \qquad V = \theta \log\left(1 + \lambda \sum\nolimits_{i=1}^{N} \beta_i t_i\right) - R$$

- Stackelberg Equilibrium:
  - Neither leader nor followers have incentive to deviate
  - Find follower subgame perfect Nash equilibrium (NE) first
  - Apply induction to achieve the leader side equilibrium

Equilibrium in Stackelberg Game

- Analysis method: backward induction
- Theorem 1. A Nash equilibrium exists among all devices
- Theorem 2. A Stackelberg equilibrium exists among all devices and the server.

Robust Price of Anarchy

- Price of Anarchy (PoA)
  - Decentralized game: devices selfishly determine strategies
  - Centralized control: devices follow the server's instruction
  - Measurement of efficiency loss:
    - decentralized game solution Vs optimal centralized control
- Device non-cooperative game
  - A valid monotone utility game
  - Lower bound on the PoA: 0.5

Unstable Communication Channel

- Modeling stochastic upload time
  - Device i's upload time follows a normal distribution $\mathcal{N}(\mu_i, \sigma_i^2)$
  - PDF of i's upload time $F_i(\tau_i) = \int_{-\infty}^{\tau_i} \frac{1}{\sigma_i\sqrt{2\pi}} \exp\left\{-\frac{(x-\mu_i)^2}{2\sigma_i^2}\right\} \mathrm{d}x$
- Device side problem

$$\begin{aligned} &\textit{maximize} \quad && u_i(t_i, \boldsymbol{t}_{-i}) = R\frac{\alpha_i F_i(T-t_i)}{\hat{\alpha}_{-i}+\alpha_i} - c_i t_i, \\ &\textit{subject to} \quad && 0 \le t_i < T. \end{aligned}$$

- Follower subgame Nash equilibrium
  - An N-player non-zero-sum game
  - Nash equilibrium exists among all devices

Experiment

- Part 1
  - Device-side equilibrium analysis
  - Server-side equilibrium analysis
- Part 2
  - game-driven market equilibrium Vs optimal social welfare
- Part 3
  - Unstable upload channel

Follower Subgame Nash Equilibrium

- Parameters from the server side
  - T, R, and reward policies
  - Setting: 5 homogeneous devices
  - Results:
    - size-based policy leads devices to train for a longer time
    - an accuracy measurement function with a higher diminishing return will motivate devices for longer training time.

T \ R	200	400	600	800	1000
100	32	64	90	90	90
120	32	64	96	110	110
160	32	64	96	128	150

(a) Size-based policy.

T \ R	200	400	600	800	1000
100	31	61	90	90	90
120	31	61	90	110	110
160	31	61	90	118	144

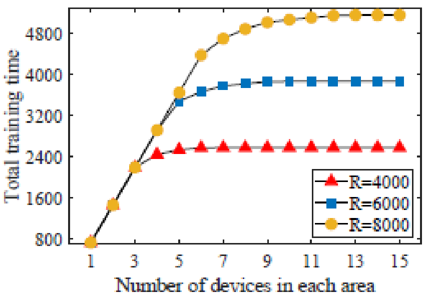
(b) Accuracy-based policy ($\theta = 10$, $\lambda = 8 \times 10^{-6}$).

T \ R	200	400	600	800	1000
100	29	54	76	90	90
120	29	54	76	97	110
160	29	54	76	97	117

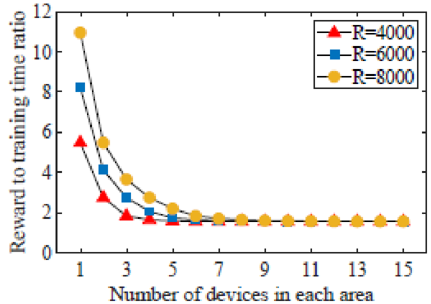
(c) Accuracy-based policy ($\theta = 10$, $\lambda = 4 \times 10^{-5}$).

Follower Subgame Nash Equilibrium

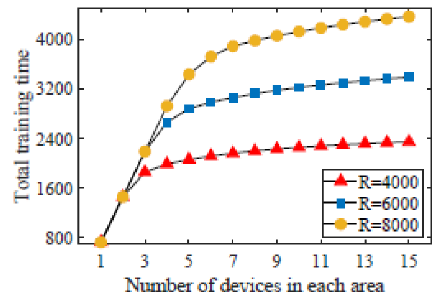
- Number of Participating Devices



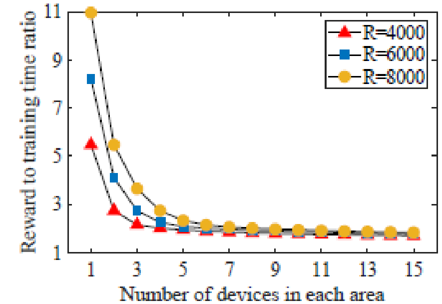
(a) Size-based policy.



(b) Size-based policy.



(c) Accuracy-based policy.



(d) Accuracy-based policy.

- Device Parameters (β, c, τ)

t \ R	200	400	600	800	1000
t_1	45.8	91.5	110	110	110
t_2	29.1	58.3	91.4	110	110
t_3	29.1	58.3	91.4	95	95
t_4	12.5	25	43.4	80	110
t_5	12.5	25	43.4	80	105
sum	**129**	**258.1**	**379.6**	**475**	**530**

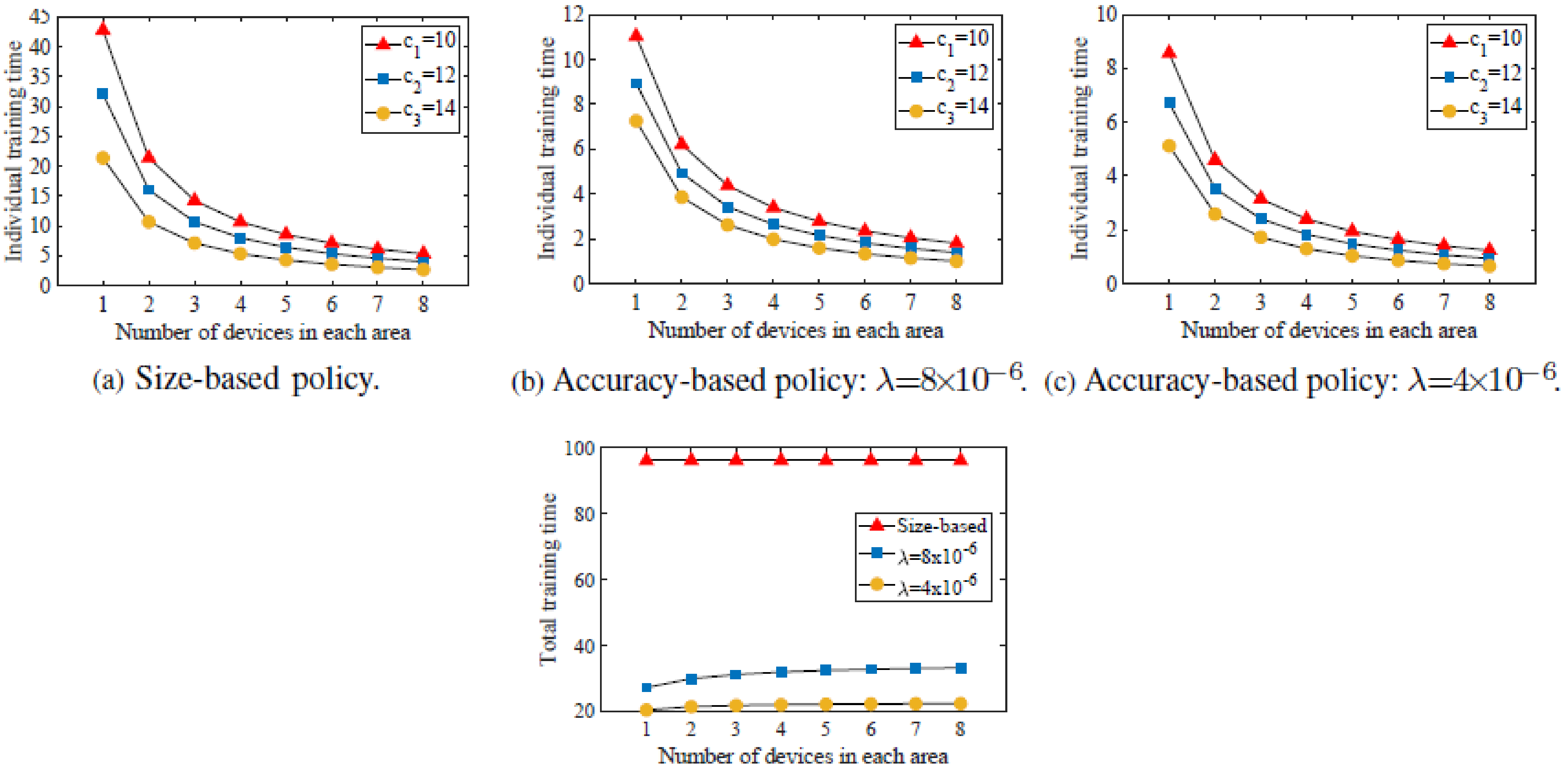
(a) Size-based policy.

t \ R	200	400	600	800	1000
t_1	36.7	64.9	89.6	110	110
t_2	25.5	47.2	66.9	85.5	106
t_3	25.5	47.2	66.9	85.5	95
t_4	16.1	33.1	49.3	64.8	82.5
t_5	16.1	33.1	49.3	64.8	82.5
sum	**119.9**	**225.5**	**322**	**410.6**	**476**

(b) Accuracy-based policy($\theta = 10$, $\lambda = 4 \times 10^{-5}$)..

Leader-Follower Stackelberg Equilibrium

- Optimal strategy on the server side



(a) Size-based policy.

(b) Accuracy-based policy: $\lambda=8\times10^{-6}$.

(c) Accuracy-based policy: $\lambda=4\times10^{-6}$.

(d) Size vs Accuracy.

Part 2 & 3

- Price of Anarchy
  - Social welfare: difference between the global model satisfaction and the total cost on the device side.

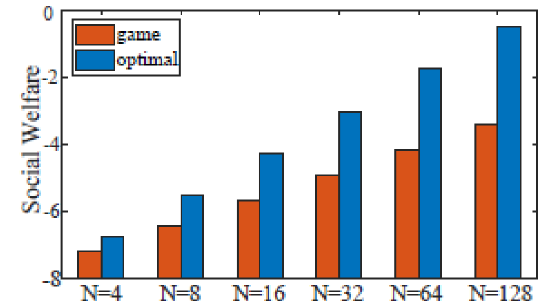


Fig. 5: Social welfare: game-driven vs optimal.

- Uncertainty in upload time

σ \ R	100	200	300	400	800
0	16	32	48	64	125
1	23	34.5	48.4	64	128
10	27	38	50	64.1	128

(a) Size-based policy.

σ \ R	100	200	300	400	800
0	15.8	31.2	46.3	61.1	117.6
1	17.8	31.3	46.3	61.1	117.6
10	24.6	35.9	46.8	61.1	117.6

(b) Accuracy-based policy ($\theta = 10$, $\lambda = 8 \times 10^{-6}$).

σ \ R	100	200	300	400	800
0	15.1	28.8	41.7	53.8	97.3
1	17.7	28.9	41.7	53.8	97.3
10	23.9	34.5	43.1	53.9	97.3

(c) Accuracy-based policy ($\theta = 10$, $\lambda = 4 \times 10^{-5}$).

5. Conclusion

- A Stackelberg game with two subgames
  - server-side deadline and device-side upload time
- Two different reward policies
  - size-based Vs accuracy-based
- Price of Anarchy
- Upload time
  - fixed Vs variable

Thank you

Q & A