Conference Paper Title*

*Note: Sub-titles are not captured in Xplore and should not be used

1st Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

2nd Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

3rd Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

4th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

5th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

6th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

***Abstract*—This document is a model and instructions for LaTeX. This and the IEEEtran.cls file define the components of your paper [title, text, heads, etc.]. *CRITICAL: Do Not Use Symbols, Special Characters, Footnotes, or Math in Paper Title or Abstract.**

***Index Terms*—component, formatting, style, styling, insert**

I. Introduction

This document is a model and instructions for LaTeX. Please observe the conference page limits.

II. System Model

A. Marginal Upper Bound

We denote $F^t(w(t)) \triangleq F^t(w(t); D(t))$. Then the marginal upper bound can be expressed as follows:

$$E\left[F^t(w(t))\right] - F^* \leq q^\tau \left(E\left[F^t(w(t-1))\right] - F^*\right) \tag{1}$$

$$+ \frac{1-q^\tau}{1-q} \cdot \frac{\beta\eta^2}{2D(t)^2} \sum_{k=1}^{N} \frac{M_k D_k(t)^2}{s_k} + \rho h(\tau)^2 \tag{2}$$

Proof 1:

$$E[F(w^t) - F(w^{t-1})] \tag{3}$$

$$\leq \underbrace{E[\langle \nabla F(w^{t-1}), w^t - w^{t-1} \rangle]}_{A} + \underbrace{\frac{\rho}{2} E[\|w^t - w^{t-1}\|_2^2]}_{B} \tag{4}$$

We focus on bounding A at first:

$$E[\langle \nabla F(w^{t-1}), w^t - w^{t-1} \rangle] \tag{5}$$

$$= E[\langle \nabla F(w^{t-1}), (-\eta)\nabla F(w^{t-1}) + n^t \rangle] \tag{6}$$

$$= E[\langle \nabla F(w^{t-1}), (-\eta)\nabla F(w^{t-1}) \rangle] + E[\langle \nabla F(w^{t-1}), n^t \rangle] \tag{7}$$

$$= (-\eta) E\|\nabla F(w^{t-1})\|_2^2 \tag{8}$$

Then, we focus on bounding B:

$$\frac{\rho}{2} E\|w^t - w^{t-1}\|_2^2 \tag{9}$$

$$= \frac{\rho}{2} E\|(-\eta)\nabla F(w^{t-1}) + n^t\|_2^2 \tag{10}$$

$$\leq \rho E\|(-\eta)\nabla F(w^{t-1})\|_2^2 + \rho E\|n^t\|_2^2 \tag{11}$$

$$\leq \rho\eta^2 E\|\nabla F(w^{t-1})\|_2^2 + \rho d \sum_{k=1}^{N} p_k^t {\sigma_k^t}^2 \tag{12}$$

$$\tag{13}$$

According to Polyak Lojasiewicz condition in 4) in assumption,

$$E[F(w^{t-1} - F(w^*))] \leq \frac{1}{2\mu} E\|\nabla F(w^{t-1})\|_2^2 \tag{14}$$

We set $\eta < \frac{1}{\rho}$, then $\rho\eta^2 - \eta < 0$

Combining A and B, we have

$$E[F(w^t) - F(w^{t-1})] \tag{15}$$

$$\leq (\rho\eta^2 - \eta) E\|\nabla F(w^{t-1})\|_2^2 + \rho d \sum_{k=1}^{N} p_k^t {\sigma_k^t}^2 \tag{16}$$

$$\tag{17}$$

B. the Utility Function of Center Server

$$\min_{L(t)} \sum_{t=0}^{T-1} E\left[F^t(w(t))\right] - F^* + R \tag{18}$$

$$\text{s.t.} \sum_{t=0}^{T-1} L(t) \leq \theta \tag{19}$$

$$\tag{20}$$

Identify applicable funding agency here. If none, delete this.

C. the Utility Function of Clients

$$\max_{v_k(t),\varepsilon_k(t)} \sum_{t=0}^{T-1} \frac{D_k(t)}{\sum_{i=1}^{N} D_i(t)} R - \alpha_k v_k(t)^2 - \beta_k \xi_k \frac{{c_k}^3 D_k(t)^3}{(L(t)-L_{com})^2} \tag{21}$$

$$\text{s.t.} D_k(t+1) = \gamma D_k(t) + v_k(t)(L(t) - L_{com}) \tag{22}$$

$$\sum_{t=0}^{T-1} \varepsilon_k(t) <= E_k, \forall k \tag{23}$$

III. Algorithm

A. Optimal Decision in Stage II

First of all, it makes sense for the following assumptions:

$$f(B_k^t, \varepsilon_k^t) = \log(B_k^t, \varepsilon_k^t) \tag{24}$$

$$g(B_k^t) = {B_k^t}^2 \tag{25}$$

$$h(\varepsilon_k^t) = {\varepsilon_k^t}^2 \tag{26}$$

There are some key challenges while finding the optimal decision. First, imcomplete infomation. Besides, long term constraint.

1) Mean Field Approach for Imcomplete Information: Besides, to overcome the problem of imcomplete information about other clients during the decision of control variable, we will introduce a mean-field term to estimate the global information. More specifically, we define $\phi^t = \sum_{i=1}^{N} \log(B_i^t \varepsilon_i^t)$ and replace the corresponding part in the objective function with it. From the prespective of mathematic, ϕ^t is a given function and have no relationship with other clients' decision, so the decoupling is accomplished. In the next section, we will also provide a fixed point algorithm to secure the value of ϕ^t. By introducing the mean-field term, we have the following problem:

$$\max_{B_k^t,\varepsilon_k^t} \sum_{t=1}^{T} \frac{\log(B_k^t \varepsilon_k^t)}{\phi^t} R^t - \alpha_k {B_k^t}^2 - \beta_k {\varepsilon_k^t}^2 \tag{27}$$

$$\text{s.t.} \sum_{t=1}^{T} \xi_k c_k f_k^2 B_k^t \le n_k \tag{28}$$

$$\sum_{t=1}^{T} \varepsilon_k^t \le m_k \tag{29}$$

$$B_k^t \in [0, B_k^{max}] \tag{30}$$

2) Lyapunov-Based Optimization for Long-Term Constraint: Under the long-term computing resource constraint and privacy budget constraint, it is difficult for a client to decide the optimal solution in current time slot when keeping unknown about the future infomation because mutual effect. Mechanisms like Lyapunov Drift and Penalty Algorithm can be adopted in this problem to solve the long term dependency. We give the recursive definition of two virtual queues following LDP Algorithm:

$$Q_k(t+1) = \max(Q_k(t) + \xi_k c_k f_k^2 B_k^t - \frac{n_k}{T}, 0), Q_k(0) = W \tag{31}$$

$$Z_k(t+1) = \max(Z_k(t) + \varepsilon_k^t - \frac{m_k}{T}, 0), Z_k(0) = W \tag{32}$$

$Q_k(t)$ and $Z_k(t)$ captures long term constraint about computing resource and privacy budget respectively. The more stable the queues are, the better the problem can meet the constraint conditions. By introducing the virtual queue, a client's decision among time slots are decoupled and the original problem can be divided into single slot optimization problem. Take the k-th clients for instance, the problem at the t-th slot is as follows:

$$\min_{B_k^t,\varepsilon_k^t} V(\alpha_k {B_k^t}^2 + \beta_k {\varepsilon_k^t}^2 - \frac{\log(B_k^t \varepsilon_k^t)}{\phi^t} R^t) + Q_k(t)(\xi_k c_k {f_k}^2 B_k^t - \frac{n_k}{T}) + Z_k(t)(\varepsilon_k^t - \frac{m_k}{T}) \tag{33}$$

$$\text{s.t.} B_k^t \in [0, B_k^{max}] \tag{34}$$

where the aim is to minimize the client's cost and queues' stability simultaneously. V is a weight factor and a larger V means paying more attention to minimizing client k's cost.

3) The Optimal Decison for Clients: We denote a client's cost function as $F(B_k^t, \varepsilon_k^t)$ and take the first order derivative of B_k^t and ε_k^t:

$$\frac{\partial F}{\partial B_k^t} = 2V\alpha_k B_k^t - V\frac{R^t}{\phi^t B_k^t} + Q_k(t)\xi_k c_k {f_k}^2 = 0 \tag{35}$$

$$\frac{\partial F}{\partial \varepsilon_k^t} = 2V\beta_k \varepsilon_k^t - V\frac{R^t}{\phi^t \varepsilon_k^t} + Z_k(t) = 0 \tag{36}$$

So the optimal solution for the k-th client at the t-th slot is

$$ {B_k^t}^* = \sqrt{\frac{R^t}{2\alpha_k \phi^t} + \left(\frac{Q_k(t)\xi_k c_k f_k^2}{4V\alpha_k}\right)^2} - \frac{Q_k(t)\xi_k c_k f_k^2}{4V\alpha_k} \tag{37}$$

$$ {\varepsilon_k^t}^* = \sqrt{\frac{R^t}{2\beta_k \phi^t} + \left(\frac{Z_k(t)}{4V\beta_k}\right)^2} - \frac{Z_k(t)}{4V\beta_k} \tag{38}$$

By observing the above formulas, ${B_k^t}^*$ and ${\varepsilon_k^t}^*$ are related to four variables: central server's reward R^t, mean field term ϕ^t, virutal queues $Q_k(t)$ and $Z_k(t)$. ${R^t}^*$ will be disscussed in the subsequent section. Algorithms to find ϕ^t and update virtual queues are also to be provided later.

B. Optimal Decision for Central Server in Stage I

In this section, we hope to derive central server's optimal reward ${R^t}^*$ at t-th slot that can minimize the central server's cost when clients' decision ${B_k^t}^*$ and ${\varepsilon_k^t}^*$ in Stage II are given. However, by observing (1), it is difficult to derive the close-form solution of ${R^t}^*$ directly. So here we try to prove ${R^t}^*$'s existence at first and then provide a heuristic algorithm to search it.

1) The Existence of R^{t^*}: Considering the t-th slot, We denote the server's cost function in (1) as $H(B_k^t, \varepsilon_k^t)$ and rewrite it as follows:

$$
\begin{aligned}
H(B_k^t, \varepsilon_k^t) =& \underbrace{\kappa^t \sum_{k=1}^{N} \frac{1}{(\sum_{i=1}^{N} B_k^t)^2 \varepsilon_k^t}}_{A} \\
&+ \underbrace{\sum_{k=1}^{N} \frac{\frac{c_k B_k}{f_k}}{\sum_i^N \frac{c_i B_i^t}{f_i}} log \frac{\frac{c_k B_k}{f_k}}{\sum_i^N \frac{c_i B_i^t}{f_i}}}_{B} + \underbrace{\gamma^{t-1} R^t}_{C} \quad (39)
\end{aligned}
\tag{40}
$$

where $\kappa^t = (1 + \mu(\rho\eta^2 - \eta)(Q_1 + Q_2))^{T-t} \rho d \beta^2 \eta^2 C^2$. We have *Theorem 1* as follows:

Theorem 1: When $R^t \in [0, \infty]$, there exists R^{t^*} in the t-th slot that minimize $H(B^{t^*}, \varepsilon_k^{t^*})$

Proof 2: For ease of reading, we define $X_k^t = \frac{Q_k(t)\xi_k c_k f_k^2}{2V\alpha_k}$ and $Y_k^t = \frac{Z_k(t)}{2V\beta_k}$. Then we have

$$
B_k^{t^*} = \sqrt{\frac{R^t}{2\alpha_k \phi^t} + \frac{{X_k^t}^2}{4}} - \frac{X_k^t}{2} \tag{41}
$$

$$
\varepsilon_k^{t^*} = \sqrt{\frac{R^t}{2\beta_k \phi^t} + \frac{{Y_k^t}^2}{4}} - \frac{Y_k^t}{2} \tag{42}
$$

So, we have

$$
\begin{aligned}
\frac{\partial A}{\partial R^t} &= \sum_{k=1}^{N} \kappa^t \left((B^t)^{-2} (\varepsilon_k^t)^{-2}\right)' \\
&= \sum_{k=1}^{N} \kappa^t (-2) \left[\underbrace{\left((B^t)^{-3} (\varepsilon_k^t)^{-2} \frac{\partial B^t}{\partial R^t}\right)}_{A1} + \underbrace{\left((B^t)^{-2} (\varepsilon_k^t)^{-3} \frac{\partial \varepsilon_k^t}{\partial R^t}\right)}_{A2} \right] \\
&= \sum_{k=1}^{N} \kappa^t (-2) \left[\left(\sum_{k=1}^{N} \sqrt{\frac{R^t}{2\alpha_k \phi^t} + \frac{{X_k^t}^2}{4}} - \frac{X_k^t}{2} \right)^{-3} \right. \\
&\cdot \left(\sqrt{\frac{R^t}{2\beta_k \phi^t} + \frac{{Y_k^t}^2}{4}} - \frac{Y_k^t}{2} \right)^{-2} \cdot \sum_{k=1}^{N} \frac{1}{4\alpha_k \phi^t} \left(\frac{R^t}{2\alpha_k \phi^t} + \frac{{X_k^t}^2}{4} \right)^{-\frac{1}{2}} \Bigg] \\
&+ \left[\left(\sum_{k=1}^{N} \sqrt{\frac{R^t}{2\alpha_k \phi^t} + \frac{X_k^t 2}{4}} - \frac{X_k^t}{2} \right)^{-2} \right. \\
&\left. \cdot \left(\sqrt{\frac{R^t}{2\beta_k \phi^t} + \frac{{Y_k^t}^2}{4}} - \frac{Y_k^t}{2} \right)^{-3} \cdot \frac{1}{4\beta_k \phi^t} \left(\frac{R^t}{2\beta_k \phi^t} + \frac{{Y_k^t}^2}{4} \right)^{-\frac{1}{2}} \right]
\end{aligned}
\tag{43}
$$

$$
\begin{aligned}
\frac{\partial B}{\partial R^t} &= \sum_{k=1}^{N} \underbrace{\left[1 + \log \frac{\frac{c_k B_k^t}{f_k}}{\sum_{i=1}^{N} \frac{c_i B_i^t}{f_i}} \right]}_{B1} \\
&\cdot \underbrace{\left[\frac{\frac{c_k {B_k^t}'}{f_k} \sum_{i=1}^{N} \frac{c_i B_i^t}{f_i} - \frac{c_k B_k^t}{f_k} \sum_{i=1}^{N} \frac{c_i {B_i^t}'}{f_i}}{\left(\sum_{i=1}^{N} \frac{c_i B_i^t}{f_i}\right)^2} \right]}_{B2} \\
&= \sum_{k=1}^{N} \left[1 + \log \frac{\frac{c_k}{f_k} \left(\sqrt{\frac{R^t}{2\alpha_k \phi^t} + \frac{{X_k^t}^2}{4}} - \frac{X_k^t}{2} \right)}{\sum_{i=1}^{N} \frac{c_i}{f_i} \left(\sqrt{\frac{R^t}{2\alpha_i \phi^t} + \frac{{X_i^t}^2}{4}} - \frac{X_i^t}{2} \right)} \right] \\
&\cdot \left[\frac{\frac{c_k}{f_k} \frac{1}{4\alpha_k \phi^t} \left(\frac{R^t}{2\alpha_k \phi^t} + \frac{{X_k^t}^2}{4} \right)^{-\frac{1}{2}} \sum_{i=1}^{N} \frac{c_i}{f_i} \left(\sqrt{\frac{R^t}{2\alpha_i \phi^t} + \frac{{X_i^t}^2}{4}} - \frac{X_i^t}{2} \right)}{\left(\sum_{i=1}^{N} \frac{c_i}{f_i} \left(\sqrt{\frac{R^t}{2\alpha_i \phi^t} + \frac{{X_i^t}^2}{4}} - \frac{X_i^t}{2} \right) \right)^2} \right. \\
&\left. - \frac{\frac{c_k}{f_k} \left(\sqrt{\frac{R^t}{2\alpha_k \phi^t} + \frac{{X_k^t}^2}{4}} - \frac{X_k^t}{2} \right) \sum_{i=1}^{N} \frac{c_i}{f_i} \frac{1}{4\alpha_i \phi^t} \left(\frac{R^t}{2\alpha_i \phi^t} + \frac{{X_i^t}^2}{4} \right)^{-\frac{1}{2}}}{\left(\sum_{i=1}^{N} \frac{c_i}{f_i} \left(\sqrt{\frac{R^t}{2\alpha_i \phi^t} + \frac{{X_i^t}^2}{4}} - \frac{X_i^t}{2} \right) \right)^2} \right]
\end{aligned}
\tag{44}
$$

$$
\frac{\partial C}{\partial R^t} = \gamma^{t-1} \tag{45}
$$

By analysing the above result, we can get TABLE I and TABLE II as follows:

Based on TABLE I and TABLE II, we can derive that $\lim_{R^t \to 0} \frac{\partial H^t}{\partial R^t} = -\infty$ and $\lim_{R^t \to \infty} \frac{\partial H^t}{\partial R^t} = \gamma^{t-1}$. Meanwhile,

	$\frac{\partial A1}{\partial R^t}$	$\frac{\partial A2}{\partial R^t}$	$\frac{\partial B1}{\partial R^t}$	$\frac{\partial B2}{\partial R^t}$
$\lim_{R^t \to 0}$	∞	∞	1	0
$\lim_{R^t \to \infty}$	0	0	1	0

TABLE I
*

	$\frac{\partial A}{\partial R^t}$	$\frac{\partial B}{\partial R^t}$	$\frac{\partial C}{\partial R^t}$
$\lim_{R^t \to 0}$	$-\infty$	0	γ^{t-1}
$\lim_{R^t \to \infty}$	0	0	γ^{t-1}

TABLE II
*

it's obvious that H^t is a continuous function in the interval of $R^t \in [0, \infty]$. So there exists R^{t^*} that minimizes H^t. The proof of *Theorem 1* ends. ∎

2) Algorithm to Find R^{t^*}: Heuristic algorithms such as Genetic Algorithm(GA) and Paritial Swarm Optimization(PSO) have been adopted widely in tackling complex optimization problem without analytical solution. In this article, we adopt the PSO to find the approximately optimal solution R^{t^*}. The detailed procedure will be exhibited in the later section.

C. Update of Mean Field Term

Recall the whole derivative process: the mean field term ϕ^t determines R^{t^*} in Stage I and further determines both $B_k^{t^*}$ and $\varepsilon_k^{t^*}$ in Stage II. Meanwhile, according to the definition of ϕ^t, it is in turn affected by $B_k^{t^*}$ and $\varepsilon_k^{t^*}$, which forms a close loop. So we have the following theorem:

Theorem 2: ϕ^t can be approximately estimated by adopting the fix point algorithm.

IV. Experiment