Accelerated Federated Learning with Dynamic Model Partition for H-IoT Online Proof

Theorem 1. For any $V > 0$, the long-term average delay and long-term average power consumption satisfy [1]:

$$\lim_{\tau \to T} \frac{1}{\tau} \sum_{t=0}^{\tau} \mathbb{E}\{C(\delta_t, t)\} \le \frac{K}{V} + c^*, \tag{15}$$

$$\lim_{\tau \to T} \frac{1}{\tau} \sum_{t=0}^{\tau} \mathbb{E}\{Q(\tau)\} \le \frac{K + V(c^* - c_{\min})}{\varepsilon}, \tag{16}$$

where K is a constant, and $K < \overline{E} - \sigma B / T$, c^* is the optimal solution to the FL local training delay in problem (8). It can be seen from (15) that when V is large enough, the delay obtained by the proposed algorithm can be infinitely closed to the optimal solution of the original problem, while maintaining the stability of the queue. However, the infinite approximation of the gap comes at the cost of a longer queue length.

Proof. Energy consumption can be reduced if we minimize the right side of (16) while ensuring the stability of the queue in (11). Given the observed $Q(\tau)$, we can get

$$\begin{aligned}
&\Delta Q(\tau) + V \cdot \mathbb{E}(C(\delta_\tau, \tau) \mid Q(\tau)) \\
&\le K + \mathbb{E}(VC(\delta_\tau, \tau) \\
&\qquad + Q(\tau)(E(\delta_\tau, \tau) - \sigma B / T) \mid Q(\tau)) \\
&= K + V \cdot \mathbb{E}(C(\delta_\tau^*, \tau)) \\
&\qquad + \mathbb{E}((E(\delta_\tau^*, \tau) - \sigma B / T) \mid Q(\tau)) \\
&\le K + V \cdot c^*
\end{aligned}, \tag{17}$$

where δ_τ^* is the optimal partition point for time slot τ. According to the conditions (18) and (19) satisfied by the optimal w-only strategy [1], it can be seen that (17) holds.

$$C(\delta_\tau^*, \tau) = c^*, \tag{18}$$

$$E(\delta_\tau^*, \tau) - \sigma B / T \le 0. \tag{19}$$

By accumulating the conclusion of (17) for the time slots, we can get

$$
\begin{aligned}
&(K+V\cdot c^*)\tau \ge \lim_{\tau\to T}\sum_{t=0}^{\tau}\mathbb{E}\{\Delta Q(\tau)+V\cdot C(\delta_\tau,\tau)\,|\,Q(\tau)\}\\
&=\mathbb{E}\left[L(Q(T))\right]+V\cdot \lim_{\tau\to T}\sum_{t=0}^{\tau}\mathbb{E}\left[C(\delta_\tau,\tau)\,|\,Q(\tau)\right] \quad .\\
&\ge V\cdot \lim_{\tau\to T}\sum_{t=0}^{\tau}\mathbb{E}\left[C(\delta_\tau,\tau)\,|\,Q(\tau)\right]
\end{aligned} \tag{20}
$$

Therefore, we can obtain (15) by deforming (20). For the proof of (20), there is a w-only optimal policy that satisfies $\exists\varepsilon>0$, such that $\mathbb{E}((E(\delta_\tau,\tau)-\sigma B/T))\le-\varepsilon$. Furthermore, $C(\delta_\tau,\tau)$ is a bounded function, which makes $c_{\min}\le C(\delta_\tau,\tau)\le c^*$ holds true. We can get

$$
\begin{aligned}
&\Delta Q(\tau)+V\cdot\mathbb{E}(C(\delta_\tau,\tau)\,|\,Q(\tau))\\
&\le K+\mathbb{E}(VC(\delta_\tau,\tau)+\\
&\qquad Q(\tau)(E(\delta_\tau,\tau)-\sigma B/T)\,|\,Q(\tau)) \quad \cdot\\
&\le K+V\cdot\mathbb{E}(C(\delta_\tau^{\bullet},\tau))\\
&\qquad +\mathbb{E}((E(\delta_\tau^{\bullet},\tau)-\sigma B/T)\,|\,Q(\tau))
\end{aligned} \tag{21}
$$

By scaling (21), we can get

$$
\begin{aligned}
&\Delta Q(\tau)+V\cdot c_{\min}\\
&\qquad \le V\cdot c^*+\mathbb{E}((E(\delta_\tau^{\bullet},\tau)-\sigma B/T)\,|\,Q(\tau)) \,.
\end{aligned} \tag{22}
$$

By calculating the expectation on both sides of (22), we can get

$$
\begin{aligned}
&\mathbb{E}(\Delta Q(\tau)\,|\,Q(\tau))+V\cdot c_{\min}\\
&=V\cdot c^*+\mathbb{E}((E(\delta_\tau^{\bullet},\tau)-\sigma B/T)\,|\,Q(\tau))\mathbb{E}(Q(\tau))\,.\\
&\le V\cdot c^*-\varepsilon\mathbb{E}(Q(\tau))
\end{aligned} \tag{23}
$$

Finally, we accumulate all the time slots of (23) to obtain (16). ■