Appendix

Zuguang Li

I. Proof of Theorem I

According to the max function in (16), we have

$$Q_{m,i}^{t+1} \geq Q_{m,i}^{t} + E_{m,i}^{t} - E_{th}^{B}. \tag{25}$$

Summing both sides of (25) over time slots $t = 0, 1, ..., T-1$, and then dividing the result by T, we can get

$$\frac{Q_{m,i}^{T} - Q_{m,i}^{0}}{T} \geq \frac{1}{T}\sum_{t=0}^{T-1} E_{m,i}^{t} - E_{th}^{B}. \tag{26}$$

The initial energy consumption in this queue is set to $E_{m,0}^{0} = 0,\ \forall m \in \mathcal{M}$. By taking expectations of both sides of (25) and then substituting $Q_{m,i}^{0} = 0$ and $\lim_{T\to\infty} Q_{m,i}^{T}/T = 0$ into the result, we can get the first constraint in (15). So Theorem I is proved.

II. Proof of Theorem II

With the definitions of the model training delay and energy consumption in (7) and (13), the objective function in (21) can be rewritten as

$$\begin{aligned} f\left(c_{m,i}^{t}\right) &= VD_{m,i}^{t} + Q_{m,i}^{t}E_{m,i}^{t} \\ &= V\left(\frac{\omega_1}{c_{m,i}^{t}} + \omega_2\right) + Q_{m,i}^{t}\left(\omega_3 + \omega_4 c_{m,i}^{t}\right) \\ &= Q_{m,i}^{t}\omega_4 c_{m,i}^{t} + \frac{V\omega_1}{c_{m,i}^{t}} + V\omega_2 + Q_{m,i}^{t}\omega_3, \end{aligned} \tag{27}$$

where

$$\begin{aligned} \omega_1 &= \frac{\eta - \eta_D\left(s_{m,i}^{t}\right)}{F_B\delta^B\sigma^B}, \\ \omega_2 &= d_{m,i}^{B,D} + d_{m,i}^{D,D} + d_{m,i}^{D,C} + d_{m,i}^{D,S} + d_{m,i}^{D,D}, \\ \omega_3 &= E_{m,i}^{D,T} + E_{m,i}^{D,C} + E_{m,i}^{D,X} + E_{m,i}^{B,T} + E_{m,i}^{B,X}, \\ \omega_4 &= \kappa\delta^B\sigma^B\left(F_B\right)^2\left(\eta - \eta_D\left(s_{m,i}^{t}\right)\right), \end{aligned} \tag{28}$$

are constants once the split point selection decision is determined.

The first and second derivative of the objective function in (27) with respect to $c_{m,i}^{t}$ are given by

$$\frac{\partial f\left(c_{m,i}^{t}\right)}{\partial c_{m,i}^{t}}=Q_{m,i}^{t}\omega_{4}-\frac{V\omega_{1}}{\left(c_{m,i}^{t}\right)^{2}},\tag{29}$$

and

$$\frac{\partial^{2} f\left(c_{m,i}^{t}\right)}{\partial\left(c_{m,i}^{t}\right)^{2}}=\frac{2V\omega_{1}}{\left(c_{m,i}^{t}\right)^{3}}.\tag{30}$$

Since $\omega_1>0$, we have $\partial^2 f\left(c_{m,i}^{t}\right)/\partial\left(c_{m,i}^{t}\right)^{2}>0$, indicating the objective function is convex. In addition, the inequality constraint in (21) is linear. As such, problem $\mathbf{P}_3$ is a convex optimization problem. The optimal edge computing resource allocation decision can obtained via setting the objective function's first-order derivative to zero, and then we can get the optimal solution for problem $\mathbf{P}_3$ given by (22).