Appendix

A.1 Proof of Theorem 1

Proof. First, we analyze how the difference between $\mathbb{E}[F(\omega_t)]$ and $F(\omega^*)$ (abbreviated as F^* for brevity) changes in each round. Since the loss function is β-smooth and there is $\nabla F(\omega^*)=0$, we have

$$\mathbb{E}[F(\omega_t)]-F^*\leq\frac{\beta}{2}\mathbb{E}\|\omega_t-\omega^*\|^2. \tag{A.1}$$

For simplicity, we let $\Omega_t=\mathbb{E}\|\omega_t-\omega^*\|^2$. According to (1) and (2), we have

$$\begin{aligned}\Omega_t&=\mathbb{E}\|\sum\nolimits_{i=1}^N\frac{n_i}{n}\omega_t^i-\omega_{t-1}+\omega_{t-1}-\omega^*\|^2\\&=\Omega_{t-1}+\mathbb{E}\|\sum\nolimits_{i=1}^N\tfrac{n_i}{n}(\omega_t^i-\omega_{t-1})\|^2\\&\quad+2\mathbb{E}\left\langle\omega_{t-1}-\omega^*,\sum\nolimits_{i=1}^N\tfrac{n_i}{n}(\omega_t^i-\omega_{t-1})\right\rangle.\end{aligned} \tag{A.2}$$

For ease of description, we use $A_1\triangleq\sum_{i=1}^N\frac{n_i}{n}(\omega_t^i-\omega_{t-1})$ and $A_2\triangleq\left\langle\omega_{t-1}-\omega^*,\sum_{i=1}^N\frac{n_i}{n}(\omega_t^i-\omega_{t-1})\right\rangle$ to describe the second and third terms in (A.2), respectively. Then, we need to derive the upper bounds of $\mathbb{E}[A_1]$ and $\mathbb{E}[A_2]$ successively. More specifically, for A_1, we can bound it by using the AM-GM inequality and the Cauchy-Schwarz inequality:

$$\begin{aligned}A_1&=\|-\sum\nolimits_{i=1}^N\tfrac{n_i\eta_t}{n}\sum\nolimits_{k=0}^{\tau-1}\nabla F_{t,i}(\omega_t^{i,k};\xi_t^i)\|^2\\&\leq N\tau\sum\nolimits_{i=1}^N\tfrac{n_i^2\eta_t^2}{n^2}\sum\nolimits_{k=0}^{\tau-1}\|\nabla F_{t,i}(\omega_t^{i,k};\xi_t^i)\|^2.\end{aligned}$$

According to Assumption 4, we further derive the bound of $\mathbb{E}[A_1]$, which is shown as follows:

$$\mathbb{E}[A_1]\leq N\tau^2\eta_t^2\sum\nolimits_{i=1}^N\tfrac{n_i^2}{n^2}\left[G_i^2+\Delta_i(t)\sigma_i^2\right]. \tag{A.3}$$

For A_2, we have the following equation:

$$\begin{aligned}A_2=&\left\langle\omega_{t-1}-\omega^*,-\sum\nolimits_{i=1}^N\tfrac{n_i\eta_t}{n}\nabla F_{t,i}(\omega_{t-1};\xi_t^i)\right\rangle\\&+\left\langle\omega_{t-1}-\omega^*,-\sum\nolimits_{i=1}^N\tfrac{n_i\eta_t}{n}\sum\nolimits_{k=1}^{\tau-1}\nabla F_{t,i}(\omega_t^{i,k};\xi_t^i)\right\rangle.\end{aligned} \tag{A.4}$$

Similarly, we make use of B_1 and B_2 to denote the first and second terms in (A.4), respectively. Next, we bound $\mathbb{E}[B_1]$ and $\mathbb{E}[B_2]$, respectively. Because $F_{t,i}(\cdot)$ is μ-strongly convex and there is $F_{t,i}^*\leq F_{t,i}(\omega_{t-1})$, we can bound $\mathbb{E}[B_1]$, which is presented as follows:

$$\begin{aligned}\mathbb{E}[B_1]&=-\sum\nolimits_{i=1}^N\tfrac{n_i\eta_t}{n}\left\langle\omega_{t-1}-\omega^*,\nabla F_{t,i}(\omega_{t-1})\right\rangle\\&\leq\sum\nolimits_{i=1}^N\frac{n_i\eta_t}{n}\left(F_{t,i}(\omega^*)-F_{t,i}(\omega_{t-1})-\frac{\mu}{2}\mathbb{E}\|\omega_{t-1}-\omega^*\|^2\right)\\&\leq\sum\nolimits_{i=1}^N\frac{n_i\eta_t}{n}\left[F_{t,i}(\omega^*)-F_{t,i}^*\right]-\frac{\mu\eta_t}{2}\Omega_{t-1}\\&\leq\sum\nolimits_{i=1}^N\frac{n_i\eta_t}{2n\mu}\|\nabla F_{t,i}(\omega^*)\|^2-\frac{\mu\eta_t}{2}\Omega_{t-1}\\&\leq\sum\nolimits_{i=1}^N\tfrac{n_i\eta_t}{2n\mu}\left[G_i^2+\Delta_i(t)\sigma_i^2\right]-\tfrac{\mu\eta_t}{2}\Omega_{t-1}.\end{aligned} \tag{A.5}$$

Next, we continue to derive the bound of $\mathbb{E}[B_2]$. By utilizing the properties of Assumptions 2 and 4, we can obtain the following inequality:

$$\begin{aligned}\mathbb{E}[B_2]&\leq\frac{\mu\eta_t}{4}\mathbb{E}\|\omega_{t-1}-\omega^*\|^2\\&\quad+\frac{1}{\mu\eta_t}\mathbb{E}\|\sum\nolimits_{i=1}^N\tfrac{n_i\eta_t}{n}\sum\nolimits_{k=1}^{\tau-1}\nabla F_{t,i}(\omega_t^{i,k};\xi_t^i)\|^2\\&\leq\frac{\mu\eta_t}{4}\Omega_{t-1}+\frac{N\eta_t(\tau-1)^2}{\mu}\sum\nolimits_{i=1}^N\tfrac{n_i^2}{n^2}\left[G_i^2+\Delta_i(t)\sigma_i^2\right].\end{aligned} \tag{A.6}$$

After substituting (A.3), (A.5), and (A.6) into (A.2), we can obtain the following inequality:

$$\begin{aligned}\Omega_t&\leq(1-\frac{\mu\eta_t}{2})\Omega_{t-1}+\frac{\eta_t}{\mu}\sum\nolimits_{i=1}^N\tfrac{n_i}{n}\left[G_i^2+\Delta_i(t)\sigma_i^2\right]\\&\quad+N\eta_t\frac{\tau^2\eta_t\mu+2(\tau-1)^2}{\mu}\sum\nolimits_{i=1}^N\tfrac{n_i^2}{n^2}\left[G_i^2+\Delta_i(t)\sigma_i^2\right]\\&\leq(1-\frac{\mu\bar{\eta}}{2})\Omega_{t-1}+\sum\nolimits_{i=1}^N\alpha_i\left[G_i^2+\Delta_i(t)\sigma_i^2\right],\end{aligned}$$

where $\alpha_i=\frac{\widetilde{\eta}n_i}{\mu n}+N\widetilde{\eta}\left(\tau^2\widetilde{\eta}+\frac{2(\tau-1)^2}{\mu}\frac{n_i^2}{n^2}\right)$. Clearly, the coefficient of Ω_{t-1} is a constant, so that we can directly derive Ω_T by induction, i.e.,

$$\Omega_T\leq(1-\frac{\mu\bar{\eta}}{2})^T\Omega_0+\sum\nolimits_{t=1}^T\sum\nolimits_{i=1}^N\alpha_i\left[G_i^2+\Delta_i(t)\sigma_i^2\right].$$

Finally, we substitute the above inequality into (A.1) and have the following bound:

$$\begin{aligned}\mathbb{E}[F(\omega_T)]-F^*&\leq\frac{\beta}{2}(1-\frac{\mu\bar{\eta}}{2})^T\|\omega_0-\omega^*\|^2\\&\quad+\sum\nolimits_{t=1}^T\sum\nolimits_{i=1}^N\tfrac{\alpha_i\beta}{2}\left[G_i^2+\Delta_i(t)\sigma_i^2\right].\end{aligned}$$

Now, we complete the proof of Theorem 1. □

A.2 Proof of Theorem 2

Proof. With the four components of MDP formulation described, we first present Bellman equations and the differential cost-to-go function. Consider the decoupled model with the AoI state Δ_i and the control variable a_i. Then, Bellman equations are given by $S(0)=0$ and the following equation:

$$S(\Delta_i)+\zeta=\min_{a_i\in\{0,1\}}\{\frac{\phi_i}{p_i}\Delta_i+S(\Delta_i+1),\frac{\phi_i}{p_i}\Delta_i+\lambda\}, \tag{A.7}$$

where $\Delta_i\in\{0,1,...\}$ and $S(\Delta_i)$ is the differential cost-to-go function. According to (24), we can find that the first term of (A.7) corresponds to $a_i=0$, while the second part is associated with the case of $a_i=1$.

In fact, any selection strategy can be regarded as a threshold strategy. Therefore, we start the proof by assuming that the optimal strategy π^* is a threshold strategy that selects client i when $0 \leq \Delta_i(t) \leq H-1$ and does not select client i when $\Delta_i(t) \geq H$ for a given value of $H \in \{1, 2, ...\}$. Under this assumption, we can solve the Bellman equations (i.e., (A.7)). For convenience, we rewrite Bellman equations as below.

$$S(\Delta_i) = S(\Delta_i+1) - \zeta + \frac{\phi_i}{p_i}\Delta_i + \min_{a\in\{0,1\}}\{0, \lambda - S(\Delta_i+1)\}.$$

First, we analyze the case: $\Delta_i \geq H$. According to (A.7), we can easily obtain the condition for the strategy π to select client i with the state $\Delta_i \geq H$, which is shown as follows:

$$S(\Delta_i+1) > \lambda\ ,\ S(\Delta_i) = \lambda - \zeta + \frac{\phi_i\Delta_i}{p_i}. \quad \text{(A.8)}$$

Next, we analyze the case: $0 \leq \Delta_i \leq H-1$. Similarly, the condition for the strategy π that does not select client i under this case is

$$S(\Delta_i+1) < \lambda\ ,\ S(\Delta_i) = S(\Delta_i+1) - \zeta + \frac{\phi_i\Delta_i}{p_i}. \quad \text{(A.9)}$$

After iterating the above equation, we have

$$S(\Delta_i) = S(H) - (H-\Delta_i)\left[\zeta - \frac{\phi_i}{p_i}\cdot\frac{H+\Delta_i-1}{2}\right]. \quad \text{(A.10)}$$

Merging the conditions in (A.8) and (A.9) with the appropriate values of Δ_i yields $S(H) < \lambda < S(H+1)$. Owing to the monotonicity of $S(\Delta_i)$ in (A.8), we can get that there exists a constant $\gamma \in (0,1)$ that satisfies the following equation:

$$S(H+\gamma) = \lambda. \quad \text{(A.11)}$$

By substituting (A.8) into (A.11), we obtain the value of ζ as follows:

$$\zeta = \frac{\phi_i}{p_i}(H+\gamma). \quad \text{(A.12)}$$

Next, we substitute the initial value $S(0) = 0$ into (A.10), we have the value of $S(H)$, i.e.,

$$S(H) = H\left[\zeta - \frac{\phi_i}{p_i}\cdot\frac{H-1}{2}\right].$$

According to (A.8), we can also know $S(H)$. Thus, there exists an equation by combining the above value, which is shown as follows:

$$\lambda - \zeta + \frac{\phi_i}{p_i}H = H\left[\zeta - \frac{\phi_i}{p_i}\cdot\frac{H-1}{2}\right]. \quad \text{(A.13)}$$

Furthermore, we substitute (A.12) into (A.13) and obtain the threshold H:

$$H = -\frac{1}{2} - \gamma + \sqrt{\frac{2p_i\lambda}{\phi_i} + (\gamma - \frac{1}{2})^2}.$$

It is easy to see that H is monotonically decreasing with γ. Hence, due to the value of γ ranges from 0 to 1, the value of H decreases from

$$H(\gamma=0) = -\frac{1}{2} + \sqrt{\frac{2p_i\lambda}{\phi_i} + \frac{1}{4}}$$

to

$$H(\gamma=1) = -\frac{3}{2} + \sqrt{\frac{2p_i\lambda}{\phi_i} + \frac{1}{4}}.$$

Since we have $H(\gamma=0) - H(\gamma=1) = 1$, there exists a unique $\gamma^* \in (0,1)$ such that $H(\gamma^*)$ is integer-valued and the expression for the threshold H is given by

$$H = \left\lfloor -\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{2\lambda p_i}{\phi_i}} \right\rfloor.$$

Now, we can derive Theorem 2 according to [1]. □

A.3 Proof Sketch of Theorem 3

Proof. Due to limited space, we borrow the basic idea in [1, 2] to present our proof sketch. The works in [1, 2] investigated the AoI minimization problem with single client selection. In this paper, we extend it to the case of multiple client selection under the FL scenario. Specifically, we denote the lower bound of the performance of Problem P2 as L_B and use U_B^{WI} to represent the upper bound of the performance of Problem P2 under WICS. Then, we analyze these bounds.

First, U_B^{WI} will be smaller than the upper bound in [1] since selecting more clients in each time slot will acquire a smaller average AoI value. Therefore, we can directly employ the existing analysis result and have:

$$U_B^{WI} \leq (9 - 1/N)\sum_{i=1}^N \phi_i.$$

In order to derive the bound L_B, we adopt the same method in [1], which will lead to a smaller optimal performance of Problem P2. When $T \to \infty$, we obtain the following inequalities based on the Fatou's lemma, i.e.,

$$\begin{aligned}
&\lim_{T\to\infty}\frac{1}{TN}\sum_{t=1}^T\sum_{i=1}^N \phi_i\mathbb{E}[\Delta_i(t)] \\
&\geq \lim_{T\to\infty}\frac{1}{2N}\sum_{i=1}^N \phi_i\frac{T}{\sum_{t=1}^T a_i^\pi(t)} - \frac{1}{2N}\sum_{i=1}^N \phi_i \\
&\geq \lim_{T\to\infty}\frac{T}{2N}\sum_{i=1}^N\frac{1}{\sum_{t=1}^T a_i^\pi(t)/\phi_i} - \frac{1}{2N}\sum_{i=1}^N \phi_i \\
&\geq \frac{T}{2}\cdot\frac{N\phi_{min}}{\sum_{i=1}^N\sum_{t=1}^T a_i^\pi(t)} - \frac{1}{2N}\sum_{i=1}^N \phi_i \\
&\geq \frac{N\phi_{min}}{2M} - \frac{1}{2N}\sum_{i=1}^N \phi_i,
\end{aligned}$$

where $M = \lfloor\frac{B}{p_{min}}\rfloor$ and $\phi_{min} = \min\{\phi_i | i \in \mathcal{N}\}$. The last inequality holds since there is $\sum_{i=1}^{N}\sum_{t=1}^{T} a_i^{\pi}(t) \leq mT$. Then, we can directly know

$$L_B \geq \frac{N\phi_{min}}{2M} - \frac{1}{2N}\sum_{i=1}^{N}\phi_i.$$

Based on the above upper bound and lower bound, we can derive the bound of the ratio as follows:

$$\rho^{WI} = \frac{U_B^{WI}}{L_B} < \frac{(9-1/N)\sum_{i=1}^{N}\phi_i}{N\phi_{min}/2M - \sum_{i=1}^{N}\phi_i/2N} < \frac{2M(9N-1)\sum_{i=1}^{N}\phi_i}{N^2\phi_{min} - M\sum_{i=1}^{N}\phi_i}.$$

Therefore, the theorem holds. □

References

[1] Kadota I, Sinha A, Uysal-Biyikoglu E, Singh R, Modiano E H. Scheduling policies for minimizing age of information in broadcast wireless networks. *IEEE/ACM Trans. Netw.*, 2018, 26(6):2637–2650. DOI: 10.1109/TNET.2018.2873606.

[2] Kadota I, Modiano E H. Age of information in random access networks with stochastic arrivals. In *Proc. the 40th IEEE Conference on Computer Communications (INFOCOM)*, 2021, pp. 1–10. DOI: 10.1109/INFOCOM42981.2021.9488897.