Appendix A
proof of Lemma 1

According to the definition of $Q'_a(t) = \max\{Q_m(t), m \in \mathcal{M}_a\}$, it is can be derived that

$$Pr\left(Q_m(t) \le Q'_a(t)\right) = 1, \forall m \in \mathcal{M}_a, t, \tag{22}$$

which can also be denoted by $Q_m(t) \le Q'_a(t), \forall m \in \mathcal{M}_a, t$. We take expectations of both sides. According to the monotonicity of expectation, there is

$$\mathbb{E}\{Q_m(t)\} \le \mathbb{E}\{Q'_a(t)\}, \forall m \in \mathcal{M}_a, t. \tag{23}$$

Similarly, the inequality can also be extended to

$$\lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T} \mathbb{E}\{Q_m(t)\} \le \lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} \mathbb{E}\{Q'_a(t)\}, \forall m \in \mathcal{M}_a. \tag{24}$$

Thus, if the constraint $\lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}\{Q'_a(t)\} < \infty, \forall a \in \mathcal{A}$ is satisfied, there must be

$$\begin{aligned}&\lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} \mathbb{E}\{Q_m(t)\} \le \lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} \mathbb{E}\{Q'_a(t)\} \le \infty,\\ &\forall a \in \mathcal{A}, m \in \mathcal{M}_a.\end{aligned} \tag{25}$$

This proved that if the constraint $\overline{Q'_a(t)} \triangleq \lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}\{Q'_a(t)\} < \infty, \forall a \in \mathcal{A}$ is satisfied, the constraint $\overline{Q_m(t)} \triangleq \lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}\{Q_m(t)\} < \infty, \forall m \in \mathcal{M}$ satisfies.

This completes the proof.

Appendix B
proof of Lemma 2

Recalling the equation (8), we can denote the $\Delta\left(\mathbf{\Theta(t)}\right)$ by

$$\begin{aligned}\Delta\left(\mathbf{\Theta(t)}\right) &= \mathbb{E}\left\{\frac{1}{2}\sum_{a=1}^{A} Q'_a(t+1)^2 - \frac{1}{2}\sum_{a=1}^{A} Q'_a(t)^2\right\}\\ &= \mathbb{E}\left\{\frac{1}{2}\sum_{a=1}^{A}\left(y'_a(t)+Q'_a(t)\right)^2 - \frac{1}{2}\sum_{a=1}^{A} Q'_a(t)^2\right\}\\ &= \mathbb{E}\left\{\frac{1}{2}\sum_{a=1}^{A}\left(y'_a(t)\right)^2 + \sum_{a=1}^{A} Q'_a(t)y'_a(t)\right\}.\end{aligned} \tag{26}$$

According to the definition of $y'_a(t)$ in (8), we have

$$\begin{aligned}y'_a(t) &= Q'_a(t+1) - Q'_a(t)\\ &= \max_{m\in\mathcal{M}_a} Q_m(t+1) - \max_{m\in\mathcal{M}_a} Q_m(t)\end{aligned} \tag{27}$$

According to (4), considering that each $m \in m \in \mathcal{M}_a$, there is $y_m(t) \le \max_{m\in\mathcal{M}_a} y_m(t)$, and

$$Q_m(t{+}1) = Q_m(t){+}y_m(t) \le Q_m(t){+}\max_{m\in\mathcal{M}_a} y_m(t), \forall m \in \mathcal{M}_a. \tag{28}$$

Based on (28), we have

$$\begin{aligned}\max_{m\in\mathcal{M}_a} Q_m(t+1) &\le \max_{m\in\mathcal{M}_a}\left(Q_m(t) + \max_{m\in\mathcal{M}_a} y_m(t)\right)\\ &\le \max_{m\in\mathcal{M}_a} Q_m(t) + \max_{m\in\mathcal{M}_a} y_m(t).\end{aligned} \tag{29}$$

Recalling (27), there is

$$\begin{aligned}y'_a(t) &= \max_{m\in\mathcal{M}_a} Q_m(t+1) - \max_{m\in\mathcal{M}_a} Q_m(t)\\ &\le \max_{m\in\mathcal{M}_a} y_m(t).\end{aligned} \tag{30}$$

According to (3), considering $x_m(t) \le Q_m(t)$ in (5, there is $Q_m(t) + A_m(t) - x_m(t) \ge 0$ and $Q_m(t+1) = Q_m(t) + A_m(t) - x_m(t)$. Thus, there is

$$y_m(t) = Q_m(t+1) - Q_m(t) = A_m(t) - x_m(t). \tag{31}$$

Associated with (1) and (5), we have

$$\begin{aligned}y_m(t) &= Q_m(t+1) - Q_m(t) = A_m(t) - x_m(t)\\ &\le D_m, \forall m \in \mathcal{M}.\end{aligned} \tag{32}$$

Therefore, the (30) can be rewritten as

$$y'_a(t) \le \max_{m\in\mathcal{M}_a} y_m(t) \le \max_{m\in\mathcal{M}_a} D_m, \forall a \in \mathcal{A}. \tag{33}$$

and also

$$\frac{1}{2}\sum_{a=1}^{A}\left(y'_a(t)\right)^2 \le \frac{1}{2}\sum_{a=1}^{A}\left(\max_{m\in\mathcal{M}_a} D_m\right)^2 = B. \tag{34}$$

Back to (26), we have

$$\Delta\left(\mathbf{\Theta(t)}\right) \le \mathbb{E}\left\{B + \sum_{a=1}^{A} Q'_a(t)y'_a(t)|\mathbf{\Theta(t)}\right\}. \tag{35}$$

and

$$\Delta_V\left(\mathbf{\Theta(t)}\right) \le B{+}\mathbb{E}\left\{\sum_{a=1}^{A} Q'_a(t)y'_a(t)|\mathbf{\Theta(t)}\right\}{+}V{\cdot}\mathbb{E}\{P(t)|\mathbf{\Theta(t)}\}\}, \tag{36}$$

This completes the proof.

Appendix C
proof of the conversion from Problem 2 to Problem 3

Due to at the beginning of time slot t, we have obtained the values of $\mathbf{\Theta(t)}$, $\{Q_m(t)\}$, $\{A_m(t)\}$, $\mathbf{b(t)}$. We can re-express the objective function as

$$\begin{aligned}&\mathbb{E}\left\{\sum_{a=1}^{A} Q'_a(t)y'_a(t)|\mathbf{\Theta(t)}\right\} + V\mathbb{E}\{P(t)|\mathbf{\Theta(t)}\}\\ &= \sum_{a=1}^{A} Q'_a(t)\mathbb{E}\left\{y'_a(t)\right\} + VP(t)\\ &\overset{(8)}{=} \sum_{a=1}^{A} Q'_a(t)\mathbb{E}\left\{Q'_a(t+1) - Q'_a(t)\right\} + V\frac{1}{A}\sum_{a\in\mathcal{A}} T'_a(t)\\ &\overset{(9)}{=} \sum_{a=1}^{A}\left(Q'_a(t)\mathbb{E}\left\{Q'_a(t+1)\right\} - (Q'_a(t))^2\right)\\ &\quad + V\frac{1}{A}\sum_{a\in\mathcal{A}}\max_{m\in\mathcal{M}_a} T_m(t)\end{aligned} \tag{37}$$

Since the $(Q'_a(t))^2$ is a constant item, we can ignore it and continue rewriting as

$$
\begin{aligned}
&\sum_{a=1}^{A} Q'_a(t)\mathbb{E}\left\{Q'_a(t+1)\right\} + V\frac{1}{A}\sum_{a\in\mathcal{A}}\max_{m\in\mathcal{M}_a} T_m(t) \\
&\overset{(6)(7)}{=} \sum_{a=1}^{A} Q'_a(t)\mathbb{E}\left\{\max_{m\in\mathcal{M}_a} Q_m(t+1)\right\} \\
&\quad + V\frac{1}{A}\sum_{a\in\mathcal{A}}\max_{m\in\mathcal{M}_a} \mathbf{1}(Q_m(t)+A_m(t)-x_m(t)>0) \\
&\overset{5}{=} \sum_{a=1}^{A} Q'_a(t)\mathbb{E}\left\{\max_{m\in\mathcal{M}_a} (Q_m(t)+A_m(t)-x_m(t))\right\} \\
&\quad + V\frac{1}{A}\sum_{a\in\mathcal{A}}\max_{m\in\mathcal{M}_a} \mathbf{1}(Q_m(t)+A_m(t)-x_m(t)>0) \\
&= \sum_{a=1}^{A} Q'_a(t)\max_{m\in\mathcal{M}_a} (Q_m(t)+A_m(t)-x_m(t)) \\
&\quad + V\frac{1}{A}\sum_{a\in\mathcal{A}}\max_{m\in\mathcal{M}_a} \mathbf{1}(Q_m(t)+A_m(t)-x_m(t)>0),
\end{aligned}
\tag{38}
$$

which is the objective function in Problem 3

APPENDIX D
PROOF OF THEOREM 1

[5]Here, due to our constraint $x_m(t) \leq Q_m(t)$ in (5), we have $Q_m(t) + A_m(t) - x_m(t) \geq 0$. And we can rewrite (3) as $Q_m(t+1) = Q_m(t) + A_m(t) - x_m(t)$.