A Proof

A.1 Proof of Theorem 1

Let $A = \{a_1, a_2, ..., a_M\}$ be the application placement decision space of all edge servers. At each iteration round, we randomly choose the edge server k and itsapplication placement decision from A. Following the iterations of algorithm, the application placement decisions $\mathcal{X}$ evolves as a L-dimension Markov chain in which the ith dimension corresponds to the ith edge server application placement decision. For the convenience of proving, let $L = 2$, so it is a two-dimensional Markov chain and denoted as S_{x_1,x_2}. In each iteration, we change the application placement decision to $S_{x_1^*,x_2}$ with the following probability:

$$\begin{aligned} \Pr(S_{x_1^*,x_2}|S_{x_1,x_2}) &= \frac{1}{2M} \times \frac{1}{1 + e^{(\vartheta(S_{x_1^*,x_2}) - \vartheta(S_{x_1,x_2}))/\omega}} \\ &= \frac{1}{2M} \times \frac{e^{-\vartheta(S_{x_1^*,x_2})/\omega}}{e^{-\vartheta(S_{x_1^*,x_2})/\omega} + e^{-\vartheta(S_{x_1,x_2})/\omega}} \end{aligned} \tag{1}$$

Let the stationary distribution be $\Pr^*$, according to the detailed balance condition,

$$\begin{aligned} &\Pr{}^*(S_{a_1,a_1}) \Pr(S_{a_1,a_m}|S_{a_1,a_1}) \\ &= \Pr{}^*(S_{a_1,a_m}) \Pr(S_{a_1,a_1}|S_{a_1,a_m}) \end{aligned} \tag{2}$$

it can be derived that:

$$\begin{aligned} &\Pr{}^*(S_{a_1,a_1}) \times \frac{1}{2M} \times \frac{e^{-\vartheta(S_{a_1,a_m})/\omega}}{e^{-\vartheta(S_{a_1,a_m})/\omega} + e^{-\vartheta(S_{a_1,a_1})/\omega}} \\ &= \Pr{}^*(S_{a_1,a_m}) \times \frac{1}{2M} \times \frac{e^{-\vartheta(S_{a_1,a_1})/\omega}}{e^{-\vartheta(S_{a_1,a_m})/\omega} + e^{-\vartheta(S_{a_1,a_1})/\omega}} \end{aligned} \tag{3}$$

By observation the above equation, we can find Eq. (3) is symmetric and can be balanced if the stationary joint distribution $\Pr^*(\tilde{S}) = \gamma e^{-\vartheta(\tilde{S})/\omega}$ for arbitrary state $\tilde{S}$ in the strategy space Ω, where γ is a constant. Therefore, based on the probability conservation law, the stationary distribution can be expressed as

$$\Pr{}^*(S_{x_1,x_2}) = \frac{e^{-\vartheta(S_{x_1,x_2})/\omega}}{\sum\limits_{S_{\tilde{x}_1,\tilde{x}_2} \in \Omega} e^{-\vartheta(S_{\tilde{x}_1,\tilde{x}_2})/\omega}} \tag{4}$$

Let $S_{x_1^*,x_2^*}$ be the globally optimal service placement decisions, thus $\vartheta(S_{x_1^*,x_2^*}) \leq \vartheta(S_{\tilde{x}_1,\tilde{x}_2})$ has always been established for any $S_{\tilde{x}_1,\tilde{x}_2} \in \Omega$. From Eq. (4), we observe that $\Pr^*(S_{x_1^*,x_2^*})$ will increase with the decrease of ω and $\lim_{\omega \to 0} \Pr^*(S_{x_1^*,x_2^*}) = 1$, which proves that the gibbs sample algorithm will converges to the optimal state in probability.

The above analysis can be directly extended to the L-dimensional Markov chain.

A.2 Proof of Theorem 2

According to the convex optimization theory, If the Hessian is a positive definite matrix, the optimization objective function is a convex function?. By analysis, the optimization objective of problem **P2** can be transformed into:

$$f(\mathcal{X}^*, \mathcal{Y}, \mathcal{Z}) = \sum_{s=1}^{S} \left[(1 - \sum_{i=1}^{L} z_{i,s}) n_s \phi_{c,s} \right] + \tag{5}$$
$$\sum_{i=1}^{L} \sum_{s \in \theta_i} \left[\frac{z_{i,s} n_s}{y_{i,s} F_i / c_s - z_{i,s} n_s / \Delta t} + max\{z_{i,s} n_s - n_{i,s}, 0\} \cdot \phi_s \right]$$

The Hessian of $f(y_{i,s}, z_{i,s})$ is

$$\mathbf{H} = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 f(y_{i,s}, z_{i,s})}{\partial^2 y_{i,s}} & \frac{\partial^2 f(y_{i,s}, z_{i,s})}{\partial y_{i,s} \partial z_{i,s}} \\ \frac{\partial^2 f(y_{i,s}, z_{i,s})}{\partial z_{i,s} \partial y_{i,s}} & \frac{\partial^2 f(y_{i,s}, z_{i,s})}{\partial^2 z_{i,s}} \end{bmatrix} \tag{6}$$

$$\triangle_1 = H_{11} = \frac{2 z_{i,s} n_s c_s F_i^2 \Delta t^3}{(y_{i,s} F_i \Delta t - z_{i,s} n_s c_s)^3} > 0 \tag{7}$$

$$\triangle_2 = \frac{-n_s^2 c_s^2 F_i^2 \Delta t^4 (y_{i,s} F_i \Delta t - z_{i,s} n_s c_s)^2}{(y_{i,s} F_i \Delta t - z_{i,s} n_s c_s)^6} < 0 \tag{8}$$

where $\triangle_1 = H_{11}$ and $\triangle_2 = H_{11} H_{22} - H_{12} H_{21}$. We can prove that the Hessian in Eq.(6) is indefinite by proving that the leading principal minors of $\mathbf{H}$ have both positive and negative,which are given by Eq. (7) and Eq. (8). Therefore, the problem **P2** is a non-convex optimization problem.

A.3 Proof of Theorem 3

The optimization objective of problem **P3** can be denoted as $f(\mathcal{X}^*, \mathcal{Z}^*, \mathcal{Y})$, then the second-order derivative of $f(\mathcal{X}^*, \mathcal{Z}^*, \mathcal{Y})$ is:

$$\frac{\partial^2 f(\mathcal{X}^*, \mathcal{Z}^*, \mathcal{Y})}{\partial y_{i,s} \partial y_{j,s}} = \begin{cases} \frac{2 z^*_{i,s} n_s c_s F_i^2 \Delta t^3}{(y_{i,s} F_i \Delta t - z^*_{i,s} n_s c_s)^3} & \text{if } i == j \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

The parameters in Eq.(9) are all positive, so the Hessian matrix of the objective function $f(\mathcal{X}^*, \mathcal{Z}^*, \mathcal{Y})$ is positive definite. And problem **P3** can be defined as a convex optimization problem since its constraints are linear.

The optimization objective of problem **P4** can be denoted as $f(\mathcal{X}^*, \mathcal{Y}^*, \mathcal{Z})$, then the second-order derivative of $f(\mathcal{X}^*, \mathcal{Y}^*, \mathcal{Z})$ is:

$$\frac{\partial^2 f(\mathcal{X}^*, \mathcal{Y}^*, \mathcal{Z})}{\partial z_{i,s} \partial z_{j,s}} = \begin{cases} \frac{2y_{i,s}^* n_s^2 c_s^2 F_i \Delta t^2}{(y_{i,s}^* F_i \Delta t - z_{i,s}^* n_s c_s)^3} & \text{if } i == j \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

The parameters in Eq.(10) are all positive, so the Hessian matrix of the objective function $f(\mathcal{X}^*, \mathcal{Y}^*, \mathcal{Z})$ is positive definite. And problem **P4** can be defined as a convex optimization problem since its constraints are linear.

A.4 Proof of Theorem 4

By using the KKT conditions, we can obtain that:

$$\begin{cases} \dfrac{\partial L(f(\mathcal{Y}), \lambda, \mu)}{\partial y_{i,s}} = 0 \\ \lambda_i (\sum\limits_{s \in \theta_i} y_{i,s} - 1) = 0 \\ \mu_i \left[\sum\limits_{s \in \theta_i} \left(\dfrac{m_s}{M_i} P_i^m + y_{i,s} P_i^f \right) - P_i^{budget} \right] = 0 \\ \lambda_i, \mu_i \geq 0 \end{cases} \tag{11}$$

Based on the above formulas, we can derive that:

$$\begin{cases} y_{i,s}^* = \dfrac{\sqrt{\frac{z_{i,s}^* n_s c_s F_i \Delta t^2}{\lambda_i + \mu_i P_i^f}} + z_{i,s}^* n_s c_s}{F_i \Delta t} \\ sum_{s \in \theta_i} y_{i,s} \leq \ 1 \quad or \quad \sum\limits_{s \in \theta_i} y_{i,s} \leq \ \Gamma_i \end{cases} \tag{12}$$

where $\Gamma_i = \left(p_i^{bud} - \sum_{s \in \theta_i} \frac{m_s}{M_i} P_i^m \right) / P_i^f$

Case 1: if $\Gamma_i < 1$, it means that $\sum_{s \in \theta_i} y_{i,s} \leq \ \Gamma_i$ and $\lambda_i = 0$.

$$\mu_i = \frac{(\sum_{s \in \theta_i} \sqrt{z_{i,s}^* n_s c_s F_i \Delta t^2})^2}{(\Gamma_i F_i \Delta t - \sum_{s \in \theta_i} z_{i,s}^* n_s c_s)^2 P_i^f} \tag{13}$$

$$y_{i,s}^* = \frac{\sqrt{z_{i,s}^* n_s c_s} (\Gamma_i F_i \Delta t - \sum_{s \in \theta_i} z_{i,s}^* n_s c_s)}{\sum_{s \in \theta_i} \sqrt{z_{i,s}^* n_s c_s} F_i \Delta t} + \frac{z_{i,s}^* n_s c_s}{F_i \Delta t} \tag{14}$$

Case 2: if $\Gamma_i \geq 1$, it means that $\sum_{s \in \theta_i} y_{i,s} \leq \ 1$ and $\mu_i = 0$.

$$\lambda_i = \left(\frac{\sum_{s\in\theta_i} \sqrt{z^*_{i,s} n_s c_s F_i \Delta t^2}}{F_i \Delta t - \sum_{s\in\theta_i} z^*_{i,s} n_s c_s} \right)^2 \tag{15}$$

$$y^*_{i,s} = \frac{\sqrt{z^*_{i,s} n_s c_s}(F_i \Delta t - \sum_{s\in\theta_i} z^*_{i,s} n_s c_s)}{\sum_{s\in\theta_i} \sqrt{z^*_{i,s} n_s c_s} F_i \Delta t} + \frac{z^*_{i,s} n_s c_s}{F_i \Delta t} \tag{16}$$

This ends the proof.

A.5 Proof of Theorem 5

In the follows, we will prove the optimality and convergence of the subgradient algorithm in **Theorem 5**. The final iteration of the algorithm converges to the optimal solution defined as $z^*_{i,s}$. Before analyzing convergence, we first introduce the following inequality:

$$\begin{aligned}
&\left\| z^{(n+1)}_{i,s} - z^*_{i,s} \right\|_2^2 \\
&= \left\| z^{(n)}_{i,s} - \alpha_n g^{(n)} - z^*_{i,s} \right\|_2^2 && (17)\\
&= \left\| z^{(n)}_{i,s} - z^*_{i,s} \right\|_2^2 - 2\alpha_n g^{(n)T}(z^{(n)}_{i,s} - z^*_{i,s}) + \alpha_n^2 \left\| g^{(n)} \right\|_2^2 && (18)\\
&\leq \left\| z^{(n)}_{i,s} - z^*_{i,s} \right\|_2^2 - 2\alpha_n \left(f(z^{(n)}_{i,s}) - f(z^*_{i,s}) \right) + \alpha_n^2 \left\| g^{(n)} \right\|_2^2 && (19)\\
&\leq \left\| z^{(0)}_{i,s} - z^*_{i,s} \right\|_2^2 - 2\sum_{k=0}^{n} \alpha_k \left(f(z^{(k)}_{i,s}) - f(z^*_{i,s}) \right) \\
&\quad + \sum_{k=0}^{n} \alpha_k^2 \left\| g^{(k)} \right\|_2^2 && (20)
\end{aligned}$$

where the inequality (19) is based on the convexity of $f(\mathcal{Z})$ proved by appendix C, as $f(z^*_{i,s}) \geq f(z^{(n)}_{i,s}) + g^{(n)T}(z^*_{i,s} - z^{(n)}_{i,s})$. The stepwise expansion of inequality (19) can derive to inequality (22).

Then let us defined that

$$f^{(n)}_{best} = \min\{f(z^{(0)}_{i,s}, ..., f(z^{(k)}_{i,s}), ..., f(z^{(n)}_{i,s})\} \tag{21}$$

we have:

$$\begin{aligned}
&2\sum_{k=0}^{n} \alpha_k \left(f^{(n)}_{best} - f(z^*_{i,s}) \right) \\
&\leq \left\| z^{(0)}_{i,s} - z^*_{i,s} \right\|_2^2 - \left\| z^{(n+1)}_{i,s} - z^*_{i,s} \right\|_2^2 + \sum_{k=0}^{n} \alpha_k^2 \left\| g^{(k)} \right\|_2^2 \\
&\leq \left\| z^{(0)}_{i,s} - z^*_{i,s} \right\|_2^2 + \sum_{k=0}^{n} \alpha_k^2 \left\| g^{(k)} \right\|_2^2 && (22)
\end{aligned}$$

Based on the inequality (22) above, it can be further deduced that

$$f_{best}^{(n)} - f(z_{i,s}^*) \leq \frac{\left\| z_{i,s}^{(0)} - z_{i,s}^* \right\|_2^2 + \sum_{k=0}^{n} \alpha_k^2 \left\| g^{(k)} \right\|_2^2}{2 \sum_{k=0}^{n} \alpha_k} \tag{23}$$

$$\leq \frac{R^2 + G^2 \sum_{k=0}^{n} \alpha_k^2}{2 \sum_{k=0}^{n} \alpha_k} \tag{24}$$

where $R = \|z_{i,s}^{(0)} - z_{i,s}^*\|_2$, $G = max_{0 \leq k \leq n} \|g^{(k)}\|_2$

For the step sizes, if choosing a fixed step size $\alpha_k = \alpha$, then we have:

$$f_{best}^{(n)} - f(z_{i,s}^*) \leq \frac{R^2}{2n\alpha} + \frac{G^2\alpha}{2} \tag{25}$$

$$\lim_{n \to \infty} f_{best}^{(n)} \leq f(z_{i,s}^*) + \frac{G^2\alpha}{2} \tag{26}$$

If the required accuracy is $\epsilon > 0$, from $\frac{R^2}{2n\alpha} \leq \epsilon$ and $\frac{G^2\alpha}{2} \leq \epsilon$ it can be derived that the required number of iterations $n \geq \frac{R^2 G^2}{4\epsilon^2}$.

If the step size satisfied that $\sum_{k=1}^{\infty} \alpha_k = \infty$ and $\sum_{k=1}^{\infty} \alpha_k^2 < \infty$, then we have:

$$\lim_{n \to \infty} f_{best}^{(n)} = f(z_{i,s}^*) \tag{27}$$

This ends the proof.