Analysis of Server Provisioning for Distributed Interactive Applications (supplementary material)

Hanying Zheng and Xueyan Tang

Appendix A
Derivation of Inequality (5) in the proof to Theorem 2

If we apply the k-median server placement S_M to the DIA server provisioning problem, its total interaction path length is given by

$$\begin{aligned} T_M = & \sum_{i=1}^{n}\sum_{j=1}^{n}\big(d\left(c_i, m_i\right) + d\left(m_i, m_j\right) + d\left(m_j, c_j\right)\big) \\ = & \, 2n\sum_{i=1}^{n} d\left(c_i, m_i\right) + \sum_{i=1}^{n}\sum_{j=1}^{n} d\left(m_i, m_j\right). \end{aligned}$$

By the α-triangle inequality, for any two clients c_i and c_j, we have

$$\begin{aligned} d(m_i, m_j) \leq & \ \alpha\big(d(m_i, p_i) + d(p_i, m_j)\big) \\ \leq & \ \alpha\big(\alpha\big(d(m_i, c_i) + d(c_i, p_i)\big) + d(p_i, m_j)\big). \quad (1) \end{aligned}$$

By the α-triangle inequality, we also have

$$\begin{aligned} d(p_i, m_j) &\leq \alpha\big(d(p_i, p_j) + d(p_j, m_j)\big) \\ &\leq \alpha\Big(d(p_i, p_j) + \alpha\big(d(p_j, c_j) + d(c_j, m_j)\big)\Big) \\ &= \alpha^2 \cdot d(p_j, c_j) + \alpha \cdot \big(d(p_i, p_j) + \alpha \cdot d(c_j, m_j)\big), \end{aligned}$$

and

$$\begin{aligned} d(p_i, m_j) &\leq \alpha\big(d(p_i, c_j) + d(c_j, m_j)\big) \\ &\leq \alpha\Big(\alpha\big(d(p_i, p_j) + d(p_j, c_j)\big) + d(c_j, m_j)\Big) \\ &= \alpha^2 \cdot d(p_j, c_j) + \alpha \cdot \big(\alpha \cdot d(p_i, p_j) + d(c_j, m_j)\big). \end{aligned}$$

Therefore,

$$\begin{aligned} d(p_i, m_j) \leq & \ \alpha^2 \cdot d(p_j, c_j) + \alpha \cdot \min\Big\{d(p_i, p_j) + \alpha \cdot d(c_j, m_j), \\ & \alpha \cdot d(p_i, p_j) + d(c_j, m_j)\Big\}. \quad (2) \end{aligned}$$

It follows from (1) and (2) that,

$$\begin{aligned} T_M \leq & \ 2n\sum_{i=1}^{n} d(c_i, m_i) + \alpha\sum_{i=1}^{n}\sum_{j=1}^{n}\Big(\alpha\big(d(m_i, c_i) + d(c_i, p_i)\big) \\ & + d(p_i, m_j)\Big) \\ \leq & \ 2n\sum_{i=1}^{n} d(c_i, m_i) \\ & + \alpha\sum_{i=1}^{n}\sum_{j=1}^{n}\Big(\alpha \cdot d(m_i, c_i) + \alpha \cdot d(c_i, p_i) \\ & + \alpha^2 \cdot d(p_j, c_j) + \alpha \cdot \min\Big\{d(p_i, p_j) + \alpha \cdot d(c_j, m_j), \\ & \alpha \cdot d(p_i, p_j) + d(c_j, m_j)\Big\}\Big) \\ = & \ 2n\sum_{i=1}^{n} d(c_i, m_i) + \alpha^2 n\sum_{i=1}^{n} d(m_i, c_i) \\ & + \alpha^2 n\sum_{i=1}^{n} d(c_i, p_i) + \alpha^3 n\sum_{j=1}^{n} d(c_j, p_j) \\ & + \alpha^2 \cdot \sum_{i=1}^{n}\sum_{j=1}^{n}\min\Big\{d(p_i, p_j) + \alpha \cdot d(c_j, m_j), \\ & \alpha \cdot d(p_i, p_j) + d(c_j, m_j)\Big\} \\ \leq & \ (2n + \alpha^2 n)\sum_{i=1}^{n} d(c_i, m_i) + (\alpha^2 n + \alpha^3 n)\sum_{i=1}^{n} d(c_i, p_i) \\ & + \alpha^2 \cdot \min\Big\{\sum_{i=1}^{n}\sum_{j=1}^{n} d(p_i, p_j) + \alpha n\sum_{j=1}^{n} d(c_j, m_j), \\ & \alpha\sum_{i=1}^{n}\sum_{j=1}^{n} d(p_i, p_j) + n\sum_{i=1}^{n} d(c_i, m_i)\Big\}. \end{aligned}$$

The definition of the k-median placement implies that $\sum_{i=1}^{n} d(c_i, m_i)$ is the minimum achievable total distance from all the clients to their nearest servers among all possible placements of up to k servers. Thus, we have

$$\sum_{i=1}^{n} d\left(c_i, m_i\right) \leq \sum_{i=1}^{n} d\left(c_i, p_i\right).$$

As a result,

$$\begin{aligned} T_M \leq & \ n(2 + 2\alpha^2 + \alpha^3)\sum_{i=1}^{n} d(c_i, p_i) \\ & + \alpha^2 \cdot \min\Big\{\sum_{i=1}^{n}\sum_{j=1}^{n} d(p_i, p_j) + \alpha n\sum_{i=1}^{n} d(c_i, p_i), \\ & \alpha\sum_{i=1}^{n}\sum_{j=1}^{n} d(p_i, p_j) + n\sum_{i=1}^{n} d(c_i, p_i)\Big\}. \end{aligned}$$

Define

$$x = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} d(p_i, p_j)}{n \cdot \sum_{i=1}^{n} d(c_i, p_i)}.$$

Then, we have

$$T_M \leq n \sum_{i=1}^{n} d(c_i, p_i) \cdot \Big(2 + 2\alpha^2 + \alpha^3 + \alpha^2 \cdot \min\Big\{x + \alpha,\ \alpha x + 1\Big\}\Big).$$

Hence, inequality (5) in the proof to Theorem 2 is proven.

APPENDIX B
RUNNING TIMES OF DIFFERENT ALGORITHMS

Table I reports the average running times of the algorithms for different numbers of candidate server locations in the experiments of Section VI.A of the main paper. They are measured on a machine with Intel Xeon 3.2GHz CPU and 16GB RAM. In general, our GREEDY algorithm has similar running time to the k-median and k-center placements, and is much faster than the k-favourable algorithm. The running time of the GREEDY algorithm is very acceptable in that server provisioning is often planned on mid- to long-term basis since deploying new servers may involve amendment to hardware infrastructures or lease agreements with third-party service providers.

TABLE I
AVERAGE RUNNING TIMES OF DIFFERENT ALGORITHMS (IN SECONDS)

Number of candidate server locations	GREEDY	k-median	k-center	k-favourable
75	0.285s	0.424s	0.398s	0.295s
150	0.475s	0.712s	0.663s	0.774s
300	0.846s	1.222s	1.086s	3.962s
600	1.945s	2.109s	2.094s	10.876s
900	3.413s	2.856s	2.846s	18.807s

As discussed in Section V of the main paper, the computational complexity of our proposed GREEDY algorithm is $O(k|Z|(|C| + k^2))$, where k is the number of iterations, $|Z|$ is the number of candidate server locations and $|C|$ is the number of clients. This indicates that the running time of our proposed algorithm grows linearly with the number of clients. Moreover, in our GREEDY algorithm, the evaluation of each unselected candidate server location in an iteration (lines 6 to 17 in Algorithm 1) can be carried out in parallel. Therefore, the running time of our algorithm could be further improved by utilizing parallel computing technology. Parallelizing the evaluations in each iteration would not affect the performance results of interactivity and the analysis on the approximation ratio of our proposed algorithm.