A PROOF FOR APPROXIMATION GUARANTEES

We provide the proof for approximation guarantees of SIMPLE on the achieved profits of dynamic data placement (ISPT problem) in this appendix. The detailed proof is presented in the following procedures:

- Lemma 1 indicates lower bound of the optimal solution for uniform size ISPT problem.
- Theorem 1 presents the approximation factor by SIMPLE for uniform size ISPT problem.
- Theorem 2 eventually shows the approximation factor by SIMPLE for ISPT problem of the general case.

Lemma 1. *For the uniform size ISPT problem (single type-d intervals), denote the number budget (how many intervals can overlap in every timespan) and its upper bound as $\overline{K}_d$ and $\overline{K}_d^*$. Denote the result by any feasible solution under $\overline{K}_d$ as $\Phi_x^d(\overline{K}_d)$, we have:*

$$\Phi_{OPT}^d(\overline{K}_d) \geq vol(d)\frac{\overline{K}_d}{2\overline{K}_d^*}, \tag{1}$$

where $vol(d) = \Phi_{OPT}^d(\overline{K}_d^*)$. It represents the maximal sum of profits of all type-d intervals (the maximal profits that can be gained via dynamic data placement for type-d input data).

Proof. We prove Lemma 1 by contradiction. Given two number budgets $\overline{K}_d(i)$ and $\overline{K}_d(j)$ ($\overline{K}_d(i) \leq \overline{K}_d(j)$, and $\overline{K}_d(i), \overline{K}_d(j) \in N^+$), if $\frac{\overline{K}_d(j)}{\overline{K}_d(i)} = c$ and $c \in N^+$, assume:

$$\frac{\Phi_{OPT}^d(\overline{K}_d(j))}{\overline{K}_d(j)} > \frac{\Phi_{OPT}^d(\overline{K}_d(i))}{\overline{K}_d(i)} \tag{2}$$

then we have:

$$\frac{\Phi_{OPT}(\overline{K}_d(j))}{c} > \Phi_{OPT}(\overline{K}_d(i)). \tag{3}$$

Note that $\Phi_{OPT}(\overline{K}_d(j))$ can be decomposed into several feasible solutions for ISPT problem with budget smaller than $\overline{K}_d(j)$, we can get:

$$\Phi_{OPT}^d(\overline{K}_d(j)) = c\Phi_x^d(\overline{K}_d(i)), \tag{4}$$

where x denotes any feasible solution. Based on Eq. 3 and the pigeonhole principle, we can know that there exists at least one $\Phi_x^d(\overline{K}_d(i))$ that $\Phi_x^d(\overline{K}_d(i)) > \Phi_{OPT}^d(\overline{K}_d(i))$, which contradicts to the fact that $\Phi_{OPT}^d(\overline{K}_d(i)) = max(\Phi_x^d(\overline{K}_d(i)))$. Thus the initial assumption is false when $c \in N^+$. Let $K_d(j) = K_d^*$, and we get:

$$\Phi_{OPT}^d(\overline{K}_d) \geq vol(d)\frac{\overline{K}_d}{\overline{K}_d^*}, \tag{5}$$

we can eventually obtain Eq. 5 through simple rounding operations under the general case ($c \in R^+$). □

Theorem 1. *For the uniform size ISPT problem (single type-d intervals), denote the number budget and its upper bound as K_d and $\overline{K}_d$, denote the solution by SIMPLE as ALG, then we have:*

$$\frac{\Phi_{ALG}^d(\overline{K}_d)}{\Phi_{OPT}^d(\overline{K}_d)} \geq \frac{1}{T}, \tag{6}$$

where T is the total timespan.

Proof. Under the number budget $\overline{K}_d$, *ALG* can at least choose the type-*d* intervals with top-$\overline{K}_d$ temporal net profit. Therefore, we have:

$$\Phi_{ALG}^d(\overline{K}_d) \geq \sum_{i=1}^{\overline{K}_d} f_i^*, \tag{7}$$

where f_i^* denotes the i-th largest temporal net profit. For the result of optimal solution (*OPT*), note that it can be decomposed into $\overline{K}_d$ groups of mutually non-overlapping intervals. Based on the largest temporal net profit contained in each of these $\overline{K}_d$ groups, we can sort these K_d groups in decreasing order and get the sorted groups denoted as $Q_1, Q_2, \cdots, Q_{\overline{K}_d}$. Further denote the largest temporal net profit contained in Q_i as q_i, then we can naturally have $q_i \leq f_i^*, \ \ \forall i = 1, 2, \cdots, K_d$. Note that the sum of makespans for all the intervals contained in each group is upper bounded by the maximal makespan, which is the total timespan T. Therefore, the sum of profit by Q_i is upper bounded by $f_i^* T$, and we can derive:

$$\Phi_{OPT}^d(\overline{K}_d) \leq T\sum_{i=1}^{\overline{K}_d} f_i^*. \tag{8}$$

Combining Eq. 7 and Eq. 8 ends the proof. □

THEOREM 2. *On ISPT problem with two types of intervals (the general case under the T-GNN training setting), denote the size budget as K, and the result by any feasible solution as $\Phi_x(K)$. For the solution by SIMPLE denoted as ALG, we have:*

$$\frac{\Phi_{ALG}(K)}{\Phi_{OPT}(K)} \geq \frac{K}{4MT}, \tag{9}$$

where M is the size sum of all intervals (total memory footprint of the input data in T-GNN training), and T represents the total timespan.

PROOF. Denote the node-related interval size, and the number budget for node-related intervals by SIMPLE as h_V and $\overline{K}_V$. The same definitions for edge-related intervals are represented by h_E and $\overline{K}_E$. Denote M as the total memory footprint, so that $M = \overline{K}_V^* h_V + \overline{K}_E^* h_E$. Then combine Lemma 1, Theorem 1, the budget allocation principle by SIMPLE, and the fact that $\Phi_{OPT}(K) \leq vol(V) + vol(E)$, we have:

$$\begin{aligned}
\frac{\Phi_{ALG}(K)}{\Phi_{OPT}(K)} &\geq \frac{\frac{vol(V)\overline{K}_V}{2T\overline{K}_V^*} + \frac{vol(E)\overline{K}_E}{2T\overline{K}_E^*}}{vol(V) + vol(E)} &(10)\\
&\geq \frac{vol(V)^2 K}{2Th_V\overline{K}_V(vol(V) + vol(E))^2} + \frac{vol(E)^2 K}{2Th_E\overline{K}_E(vol(V) + vol(E))^2} &(11)\\
&\geq \frac{K}{2MT}\frac{vol(V)^2 + vol(E)^2}{(vol(V) + vol(E))^2} &(12)\\
&\geq \frac{K}{4MT}, &(13)
\end{aligned}$$

which ends the proof. □