Appendix of The Paper Entitled "Joint Scheduling of Participants, Local Iterations, and Radio Resources for Fair Federated Learning over Mobile Edge Networks"

Appendix

A. Proof of Theorem I

Before proving Theorem I, we first prove five lemmas to facilitate the proof.

Lemma 1. *For $n \in \{0,1,...,N_i\}$, $||\boldsymbol{\omega}_m(t_{i-1}+nL\tau_0) - \boldsymbol{v}_i(t_{i-1}+nL\tau_0)|| \leq g_{m,i}(n)$, where $g_{m,i}(n) = \frac{\delta_{m,i}}{\beta}[(\eta\beta+1)^n - 1]$.*

Proof: We prove Lemma 1 by induction. First, when $n=0$, $g_{m,i}(0)=0$, and $||\boldsymbol{\omega}_m(t_{i-1}) - \boldsymbol{v}_i(t_{i-1})|| = 0$ since both $\boldsymbol{v}_i(t_{i-1})$ and $\boldsymbol{\omega}_{m,i}(t_{i-1})$ are the global model obtained in the previous aggrgation. Thus, $||\boldsymbol{\omega}_m(t_{i-1}) - \boldsymbol{v}_i(t_{i-1})|| = g_{m,i}(0)$.

By assuming that

$$||\boldsymbol{\omega}_m(t_{i-1}+(n-1)L\tau_0) - \boldsymbol{v}_i(t_{i-1}+(n-1)L\tau_0)|| \leq g_{m,i}(n-1),$$

we then have

$$\begin{aligned}
&||\boldsymbol{\omega}_m(t_{i-1}+nL\tau_0) - \boldsymbol{v}_i(t_{i-1}+nL\tau_0)|| \\
=&||\boldsymbol{\omega}_m(t_{i-1}+(n-1)L\tau_0) \\
&- \eta\nabla F_m(\boldsymbol{\omega}_m(t_{i-1}+(n-1)L\tau_0)) \\
&- \boldsymbol{v}_i(t_{i-1}+(n-1)L\tau_0) \\
&+ \eta\nabla F_i(\boldsymbol{v}_i(t_{i-1}+(n-1)L\tau_0))|| \\
=&||\boldsymbol{\omega}_m(t_{i-1}+(n-1)L\tau_0) - \boldsymbol{v}_i(t_{i-1}+(n-1)L\tau_0) \\
&- \eta[\nabla F_m(\boldsymbol{\omega}_m(t_{i-1}+(n-1)L\tau_0)) \\
&- \nabla F_m(\boldsymbol{v}_i(t_{i-1}+(n-1)L\tau_0)) \\
&+ \nabla F_m(\boldsymbol{v}_i(t_{i-1}+(n-1)L\tau_0)) \\
&- \nabla F_i(\boldsymbol{v}_i(t_{i-1}+(n-1)L\tau_0))]|| \\
\leq&||\boldsymbol{\omega}_m(t_{i-1}+(n-1)L\tau_0) - \boldsymbol{v}_i(t_{i-1}+(n-1)L\tau_0)|| \\
&+ \eta||\nabla F_m(\boldsymbol{\omega}_m(t_{i-1}+(n-1)L\tau_0)) \\
&- \nabla F_m(\boldsymbol{v}_i(t_{i-1}+(n-1)L\tau_0))|| \\
&+ \eta||\nabla F_m(\boldsymbol{v}_i(t_{i-1}+(n-1)L\tau_0)) \\
&- \nabla F_i(\boldsymbol{v}_i(t_{i-1}+(n-1)L\tau_0))|| \\
\leq&(\eta\beta+1)||\boldsymbol{\omega}_m(t_{i-1}+(n-1)L\tau_0) \\
&- \boldsymbol{v}_i(t_{i-1}+(n-1)L\tau_0)|| + \eta\delta_{m,i} \\
\leq&(\eta\beta+1)\frac{\delta_{m,i}}{\beta}[(\eta\beta+1)^{n-1} - 1] + \eta\delta_{m,i} \\
=&g_{m,i}(n).
\end{aligned} \tag{1}$$

The first equality in equation (1) comes from the following update rule

$$\begin{aligned}
&\boldsymbol{\omega}_m(t_{i-1}+nL\tau_0) \\
=&\boldsymbol{\omega}_m(t_{i-1}+(n-1)L\tau_0) - \eta\nabla F_m(\boldsymbol{\omega}_m(t_{i-1}+(n-1)L\tau_0)),
\end{aligned}$$

and

$$\begin{aligned}
&\boldsymbol{v}_i(t_{i-1}+nL\tau_0) \\
=&\boldsymbol{v}_i(t_{i-1}+(n-1)L\tau_0) - \eta\nabla F_i(\boldsymbol{v}_i(t_{i-1}+(n-1)L\tau_0)).
\end{aligned}$$

The first inequality in equation (1) comes from the triangular inequality. The second inequality in equation (1) comes from

$$\begin{aligned}
&||\nabla F_m(\boldsymbol{\omega}_m(t_{i-1}+(n-1)L\tau_0)) \\
&- \nabla F_m(\boldsymbol{v}_i(t_{i-1}+(n-1)L\tau_0))|| \\
\leq&\beta||\boldsymbol{\omega}_m(t_{i-1}+(n-1)L\tau_0) - \boldsymbol{v}_i(t_{i-1}+(n-1)L\tau_0)||,
\end{aligned}$$

and

$$\begin{aligned}
&||\nabla F_m(\boldsymbol{v}_i(t_{i-1}+(n-1)L\tau_0)) \\
&- \nabla F_i(\boldsymbol{v}_i(t_{i-1}+(n-1)L\tau_0))|| \leq \delta_{m,i};
\end{aligned}$$

The third inequality in equation (1) comes from the assumption that

$$\begin{aligned}
&||\boldsymbol{\omega}_m(t_{i-1}+(n-1)L\tau_0) - \boldsymbol{v}_i(t_{i-1}+(n-1)L\tau_0)|| \\
\leq&g_{m,i}(n-1) \\
=&\frac{\delta_{m,i}}{\beta}[(\eta\beta+1)^{n-1} - 1].
\end{aligned}$$

Thus, $||\boldsymbol{\omega}_m(t_{i-1}+nL\tau_0) - \boldsymbol{v}_i(t_{i-1}+nL\tau_0)|| \leq g_{m,i}(n)$. □

Lemma 2. *For $n \in \{0,1,...,N_i\}$, $||F(\boldsymbol{\omega}_i(t_{i-1}+nL\tau_0)) - F(\boldsymbol{v}_i(t_{i-1}+nL\tau_0))|| \leq \rho\psi_i(n)$, where $\boldsymbol{\omega}_i(t_{i-1}+nL\tau_0)$ is the auxiliary parameters of the global model that is obtained by the aggregation conducted after the n-th local iteration, and $\psi_i(n) = \frac{\delta_i}{\beta}[(\eta\beta+1)^n - 1] - \eta\delta_i n$.*

Proof: For interval i, we have

$$
\begin{aligned}
&||\boldsymbol{\omega}_i(t_{i-1}+nL\tau_0)-\boldsymbol{v}_i(t_{i-1}+nL\tau_0)|| \\
=&||\boldsymbol{\omega}_i(t_{i-1}+(n-1)L\tau_0)-\boldsymbol{v}_i(t_{i-1}+(n-1)L\tau_0) \\
&-\frac{\eta\sum_{m\in\mathcal{M}}|\mathcal{D}_m|s_{m,i}\nabla F_m(\boldsymbol{\omega}_m(t_{i-1}+(n-1)L\tau_0)}{\sum_{m\in\mathcal{M}}|\mathcal{D}_m|s_{m,i}} \\
&+\frac{\eta\sum_{m\in\mathcal{M}}|\mathcal{D}_m|s_{m,i}\nabla F_m(\boldsymbol{v}_i(t_{i-1}+(n-1)L\tau_0))}{\sum_{m\in\mathcal{M}}|\mathcal{D}_m|s_{m,i}}|| \\
\le&||\boldsymbol{\omega}_i(t_{i-1}+(n-1)L\tau_0)-\boldsymbol{v}_i(t_{i-1}+(n-1)L\tau_0)|| \\
&+||\frac{\eta\sum_{m\in\mathcal{M}}|\mathcal{D}_m|s_{m,i}\nabla F_m(\boldsymbol{\omega}_m(t_{i-1}+(n-1)L\tau_0)}{\sum_{m\in\mathcal{M}}|\mathcal{D}_m|s_{m,i}} \\
&-\frac{\eta\sum_{m\in\mathcal{M}}|\mathcal{D}_m|s_{m,i}\nabla F_m(\boldsymbol{v}_i(t_{i-1}+(n-1)L\tau_0))}{\sum_{m\in\mathcal{M}}|\mathcal{D}_m|s_{m,i}}||.
\end{aligned}
$$

Based on β-*smooth* and Lemma 1, we have

$$
\begin{aligned}
&||\boldsymbol{\omega}_i(t_{i-1}+nL\tau_0)-\boldsymbol{v}_i(t_{i-1}+nL\tau_0)|| \\
\le&||\boldsymbol{\omega}_i(t_{i-1}+(n-1)L\tau_0)-\boldsymbol{v}_i(t_{i-1}+(n-1)L\tau_0)|| \\
&+\frac{\eta\beta\sum_{m\in\mathcal{M}}|\mathcal{D}_m|s_{m,i}g_{m,i}(n-i)}{\sum_{m\in\mathcal{M}}|\mathcal{D}_m|s_{m,i}} \\
=&||\boldsymbol{\omega}_i(t_{i-1}+(n-1)L\tau_0)-\boldsymbol{v}_i(t_{i-1}+(n-1)L\tau_0)|| \\
&+\eta[(\eta\beta+1)^{n-1}-1]\frac{\sum_{m\in\mathcal{M}}|\mathcal{D}_m|s_{m,i}\delta_{m,i}}{\sum_{m\in\mathcal{M}}|\mathcal{D}_m|s_{m,i}} \\
=&||\boldsymbol{\omega}_i(t_{i-1}+(n-1)L\tau_0)-\boldsymbol{v}_i(t_{i-1}+(n-1)L\tau_0)|| \\
&+\eta\delta_i[(\eta\beta+1)^{n-1}-1].
\end{aligned}
$$

Since $||\boldsymbol{\omega}_i(t_{i-1})-\boldsymbol{v}_i(t_{i-1})||=0$,

$$
\begin{aligned}
&||\boldsymbol{\omega}_i(t_{i-1}+nL\tau_0)-\boldsymbol{v}_i(t_{i-1}+nL\tau_0)|| \\
=&\sum_{x=1}^{n}[||\boldsymbol{\omega}_i(t_{i-1}+xL\tau_0)-\boldsymbol{v}_i(t_{i-1}+xL\tau_0)|| \\
&-||\boldsymbol{\omega}_i(t_{i-1}+(x-1)L\tau_0)-\boldsymbol{v}_i(t_{i-1}+(x-1)L\tau_0)||] \\
\le&\eta\delta_i\sum_{x=1}^{n}[(\eta\beta+1)^{x-1}-1] \\
=&\frac{\delta_i}{\beta}[(\eta\beta+1)^n-1]-\eta\delta_i n \\
=&\psi_i(n).
\end{aligned}
\tag{2}
$$

Based on ρ-*Lipschitz*, we have $||F(\boldsymbol{\omega}_i(t_{i-1}+nL\tau_0))-F(\boldsymbol{v}_i(t_{i-1}+nL\tau_0))||\le\rho\psi_i(n)$. □

Lemma 3. *If the learning rate $\eta<\frac{1}{\beta}$, the the gap between $\boldsymbol{v}_i$ and the optimal parameter vector $\boldsymbol{\omega}^*$ does not increase with time, i.e., $||\boldsymbol{v}_i(t_{i-1}+(n+1)L\tau_0)-\boldsymbol{\omega}^*||\le||\boldsymbol{v}_i(t_{i-1}+nL\tau_0)-\boldsymbol{\omega}^*||$ for $n\in\{0,1,...,N_i\}$.*

Proof:

$$
\begin{aligned}
&||\boldsymbol{v}_i(t_{i-1}+(n+1)L\tau_0)-\boldsymbol{\omega}^*||^2 \\
=&||\boldsymbol{v}_i(t_{i-1}+nL\tau_0)-\eta\nabla F_i(\boldsymbol{v}_i(t_{i-1}+nL\tau_0))-\boldsymbol{\omega}^*||^2 \\
=&||\boldsymbol{v}_i(t_{i-1}+nL\tau_0)-\boldsymbol{\omega}^*||^2 \\
&-2\eta\nabla F_i(\boldsymbol{v}_i(t_{i-1}+nL\tau_0))^T(\boldsymbol{v}_i(t_{i-1}+nL\tau_0) \\
&-\boldsymbol{\omega}^*)+\eta^2||\nabla F_i(\boldsymbol{v}_i(t_{i-1}+nL\tau_0))||^2 \\
\le&||\boldsymbol{v}_i(t_{i-1}+nL\tau_0)-\boldsymbol{\omega}^*||^2-\frac{\eta}{\beta}||\nabla F_i(\boldsymbol{v}_i(t_{i-1}+nL\tau_0))||^2 \\
&+\eta^2||\nabla F_i(\boldsymbol{v}_i(t_{i-1}+nL\tau_0))||^2 \\
=&||\boldsymbol{v}_i(t_{i-1}+nL\tau_0)-\boldsymbol{\omega}^*||^2 \\
&-\eta(\frac{1}{\beta}-\eta)||\nabla F_i(\boldsymbol{v}_i(t_{i-1}+nL\tau_0))||^2.
\end{aligned}
\tag{3}
$$

Since $\eta(\frac{1}{\beta}-\eta)||\nabla F_i(\boldsymbol{v}_i(t_{i-1}+nL\tau_0))||^2\ge 0$, it can be obtained that $||\boldsymbol{v}_i(t_{i-1}+(n+1)L\tau_0)-\boldsymbol{\omega}^*||\le||\boldsymbol{v}_i(t_{i-1}+nL\tau_0)-\boldsymbol{\omega}^*||$. □

Lemma 4. *If the learning rate $\eta<\frac{1}{\beta}$, the the gap between $F(\boldsymbol{v}_i(t_{i-1}+(n+1)L\tau_0))$ and $F(\boldsymbol{v}_i(t_{i-1}+nL\tau_0)))$ is bounded by $\frac{\xi-1}{2\beta}||\nabla F(\boldsymbol{v}_i(t_{i-1}+nL\tau_0))||^2$ for $n\in\{0,1,...,N_i\}$, where $\xi\in(0,1)$ satisfies $||\nabla F(\boldsymbol{\omega})-\eta\beta\nabla F_i(\boldsymbol{\omega})||\le\xi||\nabla F(\boldsymbol{\omega})||$.*

Proof:

$$
\begin{aligned}
&F(\boldsymbol{v}_i(t_{i-1}+(n+1)L\tau_0))-F(\boldsymbol{v}_i(t_{i-1}+nL\tau_0)) \\
\le&\nabla F(\boldsymbol{v}_i(t_{i-1}+nL\tau_0))^T(\boldsymbol{v}_i(t_{i-1}+(n+1)L\tau_0)) \\
&-\boldsymbol{v}_i(t_{i-1}+nL\tau_0)) \\
&+\frac{\beta}{2}||\boldsymbol{v}_i(t_{i-1}+(n+1)L\tau_0))-\boldsymbol{v}_i(t_{i-1}+nL\tau_0)||^2 \\
=&-\eta\nabla F(\boldsymbol{v}_i(t_{i-1}+nL\tau_0))^T\nabla F_i(\boldsymbol{v}_i(t_{i-1}+nL\tau_0)) \\
&+\frac{\beta\eta^2}{2}||\nabla F_i(\boldsymbol{v}_i(t_{i-1}+nL\tau_0))||^2 \\
=&\frac{1}{2\beta}||\nabla F(\boldsymbol{v}_i(t_{i-1}+nL\tau_0))||^2 \\
&-\eta\nabla F(\boldsymbol{v}_i(t_{i-1}+nL\tau_0))^T\nabla F_i(\boldsymbol{v}_i(t_{i-1}+nL\tau_0)) \\
&+\frac{\beta\eta^2}{2}||\nabla F_i(\boldsymbol{v}_i(t_{i-1}+nL\tau_0))||^2 \\
&-\frac{1}{2\beta}||\nabla F(\boldsymbol{v}_i(t_{i-1}+nL\tau_0))||^2 \\
=&\frac{1}{2\beta}||\nabla F(\boldsymbol{v}_i(t_{i-1}+nL\tau_0))-\eta\beta\nabla F_i(\boldsymbol{v}_i(t_{i-1}+nL\tau_0))||^2 \\
&-\frac{1}{2\beta}||\nabla F(\boldsymbol{v}_i(t_{i-1}+nL\tau_0))||^2 \\
\le&\frac{\xi-1}{2\beta}||\nabla F(\boldsymbol{v}_i(t_{i-1}+nL\tau_0))||^2.
\end{aligned}
\tag{4}
$$

□

Lemma 5. *Define $\theta_i(n)=F(\boldsymbol{v}_i(t_{i-1}+nL\tau_0))-F(\boldsymbol{\omega}^*)$ and $\theta_i(n)>0)$ according to [1]. We then have $\frac{1}{\theta_i(n+1)}-\frac{1}{\theta_i(n)}\ge-\frac{\kappa(\xi-1)}{2\beta}$, where $\kappa=min_i\frac{1}{||\boldsymbol{v}_i(t_{i-1})-\boldsymbol{\omega}^*||^2}$.*

Proof: According to Lemma 4, $\theta_i(n+1)-\theta_i(n)\le\frac{\xi-1}{2\beta}||\nabla F(\boldsymbol{v}_i(t_{i-1}+nL\tau_0))||^2$. In addition, due to the convexity of $F(\cdot)$, we have

$$\begin{aligned}\theta_i(n) =&F(\boldsymbol{v}_i(t_{i-1}+nL\tau_0))-F(\boldsymbol{\omega}^*)\\ \le&\nabla F(\boldsymbol{v}_i(t_{i-1}+nL\tau_0))^T(\boldsymbol{v}_i(t_{i-1}+nL\tau_0)-\boldsymbol{\omega}^*)\\ \le&||\nabla F(\boldsymbol{v}_i(t_{i-1}+nL\tau_0))||||\boldsymbol{v}_i(t_{i-1}+nL\tau_0)-\boldsymbol{\omega}^*||.\end{aligned} \tag{5}$$

Combining the above two items together yields

$$\begin{aligned}\theta_i(n+1) &\le \theta_i(n)+\frac{\xi-1}{2\beta}||\nabla F(\boldsymbol{v}_i(t_{i-1}+nL\tau_0))||^2\\ &\le \theta_i(n)+\frac{\frac{\xi-1}{2\beta}\theta_i(n)^2}{||\boldsymbol{v}_i(t_{i-1}+nL\tau_0)-\boldsymbol{\omega}^*||^2}\\ &\le \theta_i(n)+\frac{\kappa(\xi-1)}{2\beta}\theta_i(n)^2.\end{aligned} \tag{6}$$

By dividing both sides with $\theta_i(n)\theta_i(n+1)$, we have

$$\frac{1}{\theta_i(n)} \le \frac{1}{\theta_i(n+1)}+\frac{\kappa(\xi-1)}{2\beta}\frac{\theta_i(n)}{\theta_i(n+1)}. \tag{7}$$

Since $\frac{\theta_i(n)}{\theta_i(n+1)}>1$, we can show $\frac{1}{\theta_i(n+1)}-\frac{1}{\theta_i(n)}\ge-\frac{\kappa(\xi-1)}{2\beta}$. □

Theorem I can then be proved as follows:

we first have $\frac{1}{\theta_i(N_i)}-\frac{1}{\theta_i(0)}=\sum_{n=0}^{N_i}(\frac{1}{\theta_i(n+1)}-\frac{1}{\theta_i(n)})\ge\frac{N_i\kappa(1-\xi)}{2\beta}$. By summing up over all aggregation intervals, we have

$$\begin{aligned}&\sum_{i\in\mathbf{I}}\frac{1}{\theta_i(N_i)}-\frac{1}{\theta_i(0)}\\ =&\frac{1}{\theta_{|\mathbf{I}|}(l_{|\mathbf{I}|})}-\frac{1}{\theta_1(0)}-\sum_{i=1}^{|\mathbf{I}|-1}(\frac{1}{\theta_{i+1}(0)}-\frac{1}{\theta_i(N_i)})\\ \ge&\sum_{i\in\mathbf{I}}\frac{N_i\kappa(1-\xi)}{2\beta},\end{aligned} \tag{8}$$

where

$$\begin{aligned}&\frac{1}{\theta_{i+1}(0)}-\frac{1}{\theta_i(N_i)}\\ =&\frac{\theta_i(N_i)-\theta_{i+1}(0)}{\theta_i(N_i)\theta_{i+1}(0)}=\frac{F(\boldsymbol{v}_i(t_i))-F(\boldsymbol{v}_{i+1}(t_i))}{\theta_i(N_i)\theta_{i+1}(0)}\\ \ge&\frac{-\rho\psi_i(N_i)}{\theta_i(N_i)\theta_{i+1}(0)}\ge\frac{-\rho\psi_i(N_i)}{\zeta^2},\end{aligned} \tag{9}$$

and $\zeta=|F[\boldsymbol{\omega}(T)]-F(\boldsymbol{\omega}^*)|$.

Thus, we have

$$\frac{1}{\theta_{|\mathbf{I}|}(l_{|\mathbf{I}|})}-\frac{1}{\theta_1(0)}\ge\sum_{i\in\mathbf{I}}\frac{N_i\kappa(1-\xi)}{2\beta}-\sum_{i=1}^{|\mathbf{I}|-1}\frac{-\rho\psi_i(N_i)}{\zeta^2}. \tag{10}$$

Furthermore, we have $\frac{-1}{(F(\boldsymbol{\omega}(T))-F(\boldsymbol{\omega}^*))\theta_{|\mathbf{I}|}(l_{|\mathbf{I}|})}\ge\frac{-1}{\zeta^2}$, then

$$\begin{aligned}&\frac{1}{F(\boldsymbol{\omega}(T))-F(\boldsymbol{\omega}^*)}-\frac{1}{\theta_{|\mathbf{I}|}(l_{|\mathbf{I}|})}\\ =&\frac{\theta_{|\mathbf{I}|}(l_{|\mathbf{I}|})-F(\boldsymbol{\omega}(T))+F(\boldsymbol{\omega}^*)}{(F(\boldsymbol{\omega}(T))-F(\boldsymbol{\omega}^*))\theta_{|\mathbf{I}|}(l_{|\mathbf{I}|})}\\ =&\frac{F(\boldsymbol{v}_I(T))-F(\boldsymbol{\omega}(T))}{(F(\boldsymbol{\omega}(T))-F(\boldsymbol{\omega}^*))\theta_{|\mathbf{I}|}(l_{|\mathbf{I}|})}\\ \ge&\frac{-\rho\psi_I(l_I)}{(F(\boldsymbol{\omega}(T))-F(\boldsymbol{\omega}^*))\theta_{|\mathbf{I}|}(l_{|\mathbf{I}|})}\ge\frac{-\rho\psi_I(l_I)}{\zeta^2}.\end{aligned} \tag{11}$$

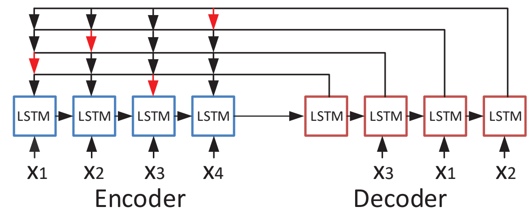

Fig. 1. The architecture of the pointer network.

Summing up the above two items, it is obvious that

$$\begin{aligned}&\frac{1}{F(\boldsymbol{\omega}(T))-F(\boldsymbol{\omega}^*)}-\frac{1}{\theta_1(0)}\\ &\ge\sum_{i\in\mathbf{I}}\frac{N_i\kappa(1-\xi)}{2\beta}-\sum_{i\in\mathbf{I}}\frac{\rho\psi_i(N_i)}{\zeta^2}.\end{aligned} \tag{12}$$

Since $\theta_1(0)>0$, we have

$$F(\boldsymbol{\omega}(T))-F(\boldsymbol{\omega}^*)\le\frac{1}{\sum_{i\in\mathbf{I}}(\frac{N_i\kappa(1-\xi)}{2\beta}-\frac{\rho\psi_i(N_i)}{\zeta^2})}. \tag{13}$$

□

B. Proof of Theorem II

The maximum density knapsack problem can be reduced to Ω. Given a set of items x_i with different weights w_i and price p_i and a knapsack with the capacity of W, the maximum density knapsack problem is to choose the items that can maximize their total price normalized by their weight, under the constraint that the total weight of items is within the knapsack's capacity. Consider a simplified version of Ω with only an interval, with a given number of local iterations, and without the constraints of radio resources, the problem to minimize training time is equivalent to maximize the reciprocal of the sum transmission time of each participant. The constraint of convergence rate can be viewed as that the sum weight of selected participants is less than a constraint, i.e., by taking minus of both sides of E.q. (23). Hence, the solution to the simplified Ω is also the optimal solution to the maximum density knapsack problem. Since the maximum density knapsack problem is NP-hard, Ω is also NP-hard.

C. Preliminary of pointer network

The pointer network was first proposed in [2] to solve combinatorial optimization problems like the traveling salesman problem. Inheriting from the sequence-to-sequence model [3], the pointer network has two recurrent neural networks (RNNs) constructed with long short term memory (LSTM) [4] units, i.e., the encoder and the decoder, as shown in Fig. 1. The target of the pointer network can be explained as follows. Given a sequence pair with the input sequence $\mathbf{X}=\{x_1,x_2,...,x_n\}$ and the output sequence $\mathbf{Y}=\{y_1,y_2,...,y_m\}$, the probability $p(\mathbf{Y}|\mathbf{X};\boldsymbol{\omega})$ needs to be maximized, where $\boldsymbol{\omega}$ is the model parameters. The probability can be further expressed via the probabilistic chain rule as $p(\mathbf{Y}|\mathbf{X};\boldsymbol{\omega})=\prod_{j=1}^m p(y_j|y_1,...,y_{j-1},\mathbf{X};\boldsymbol{\omega})$.

To determine such an output sequence $\mathbf{Y}$, the pointer network adopts a modified attention mechanism: After $\mathbf{X}$ has been fed into the encoder one element by one element, each decoder unit generates an output element y_j to form the output sequence $\mathbf{Y}$. y_j is determined by finding the maximized $p(y_j|y_1,...,y_{j-1},\mathbf{X};\boldsymbol{\omega}) = \text{softmax}(\mathbf{u}^j)$, in which $\mathbf{u}^j = \{u_1^j, u_2^j, ..., u_n^j\}$ and $u_i^j = v^T \text{thah}(W_1 e_i + W_2 d_j)$; e_i and d_j are the hidden states of the encoder and the decoder, respectively; v, W_1, and W_2 are learnable parameters. From the above illustration, u_i^j is actually a pointer to an input element, which is the origin of the "pointer network" and the reason for the ability to adapt to variable input data size.

The application of the pointer network, however, suffers from the lack of labeled data for training, since the optimal solutions to NP-hard problems are hard to obtain. Fortunately, RL is justified feasible to train the pointer network [5]. With the combination of RL and the pointer network, various combinatorial optimization problems have successfully been solved recently, including the 3D bin packing problem [6], the binary quadratic programming problem [7], the maximum cut and the minimum vertex cover problems [8], etc.

References

[1] S. Bubeck, "Convex optimization: Algorithms and complexity," *arXiv:1405.4980*, 2014.

[2] O. Vinyals, M. Fortunato, and N. Jaitly, "Pointer networks," in *Advances in Neural Inf. Process. Syst. (NeurIPS)*, 2015, pp. 2692–2700.

[3] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in *Advances in Neural Inf. Process. Syst. (NeurIPS)*, 2014, pp. 3104–3112.

[4] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[5] I. Bello, H. Pham, Q. Le, M. Norouzi, and S. Bengio, "Neural combinatorial optimization with reinforcement learning," in *Int. Conf. Learn. Representations (ICLR) Workshop*, 2017.

[6] H. Hu, X. Zhang, X. Yan, L. Wang, and Y. Xu, "Solving a new 3d bin packing problem with deep reinforcement learning method," *arXiv:1708.05930*, 2017.

[7] S. Gu, T. Hao, and H. Yao, "A pointer network based deep learning algorithm for unconstrained binary quadratic programming problem," *Neurocomputing*, 2020.

[8] T. Guo, C. Han, S. Tang, and M. Ding, "Solving combinatorial problems with machine learning methods," in *Nonlinear Combinatorial Optim.* Springer, 2019, pp. 207–229.