Supplementary Materials For SMC

I. Problem Formulation

The DRCESPP can be formulated as a mixed integer linear programming (MILP) as follows:

$$min \sum_{(i,j)\in E} w_{ij}^{1} x_{ij} \tag{1}$$

s.t.

$$\sum_{(i,j)\in E} x_{ij} - \sum_{(j,i)\in E} x_{ji} = \begin{cases} 1 & if \;\; v_i = v_s \\ 0 & if \;\; v_i \in V \backslash \{v_s, v_t\}, \;\; \forall v_i \in V \\ -1 & if \;\; v_i = v_t \end{cases} \tag{2}$$

$$L_k \leq \sum_{(i,j)\in E} w_{ij}^{k} x_{ij} \leq U_k, \;\; k = 2, ..., K \tag{3}$$

$$u_i - u_j + n x_{ij} \leq n - 1, \;\; \forall (i,j) \in E \tag{4}$$

$$x_{ij} \in \{0, 1\}, \;\; u_i \in \mathbb{R} \tag{5}$$

where the zero-one decision variable x_{ij} denotes whether the optimal path contains arc (i, j). L_k and U_k represent the lower and upper limits on the $k-1$ th resource consumed on the path respectively and $0 \leq L_k \leq U_k$. The shortest path problem can be formulated by (1), (2), and (5). The addition of double sided constraints (3) makes the SPP a DRCSPP. Further, subtour elimination constraints (4) guarantee that the optimal path is elementary, and accordingly the problem is transformed into the DRCESPP.

II. Preprocessing Techniques

The relevant content is in our other work, which is currently under review. We will provide an update here once the work has been accepted.

III. Improved Pulse Algorithm

The relevant content is in our other work, which is currently under review. We will provide an update here once the work has been accepted.

IV. NOTATION

G	a directed cyclic graph
G'	the reverse graph of G (obtained by reversing the direction of all arcs of G)
V	set of nodes
E	set of edges
v_s	source node
v_t	target node
$\pi = (\pi_0 = v_s, \pi_1, \pi_2, \ldots, \pi_n)$	a path
n	length of a path
$V(\pi)$	set of nodes of π
$E(\pi)$	set of edges of π
$\lvert E(\pi)\rvert$ or $\lvert\pi\rvert$	length of π
K	number of features (cost or resource) associated with each arc
w_{ij}^k	the k-th feature of edge (i, j)
L_k	lower limit of resource k
U_k	upper limit of resource k
$W^k(\pi)$	total quantity of the k-th weight feature (cost or resource) accumulated along π
$\overrightarrow{\pi}_i^k$	the shortest path form v_s to node v_i on the k-th feature network
$\overleftarrow{\pi}_i^k$	the shortest path form node v_i to v_t on the k-th feature network
UB	cost bounds (the cost of the best-so-far path from preprocessing: cost_min)
d_v	dimension of node representation
d_e	dimension of edge representation
M	number of heads of multi-head attention
s or s_i	a instance
$\overleftarrow{s_i}$ or $\overleftarrow{s}_i$(discarded)	a reverse instance
π_i^j	the j-th sampled path from v_s in the original graph for instance s_i
$\overleftarrow{\pi}_i^j$	the j-th sampled path from v_t in the reverse graph for instance s_i (i.e., $\overleftarrow{s_i}$)
B	size of mini-batch
t	time step
N	number of sampled paths in the forward (backward) direction for a insatnce
$\mathcal{N}_i$	set of the first-order outgoing neighbors of node i. $\mathcal{N}_i = \{v_j \in V \mid (i, j) \in E\}$
$\hat{\mathcal{N}}_\pi$	set of unpruned successor nodes of partial path π after the pruning strategies