DISTRIBUTED LOAD BALANCING FREIGHT ROUTING

WITH A CO-SIMULATION OPTIMIZATION METHOD

Pengfei Chen
Petros Ioannou, Maged Dessouky
University of Southern California
Global Transhipment

Jean-Paul Rodrigue (2017), Maritime Transportation, New York: Routledge
Los Angeles Metropolitan Area

US Federal Gov, Map of Los Angeles, California
Research Problem

To assign a number of demands from origins to destinations with considerations of several factors: *time window, road network traffic condition, emissions*, etc.

- **Time window**: pick up time window, delivery time window
- **Road network traffic condition**: Vehicle Travel Time
- **Emissions**: $CO_2, PM_{2.5}, NO_x$
Formulation

\[
\min TC(X) = \sum_{k \in K} \sum_{i \in I} \sum_{j \in J} \sum_{r \in R_{i,j}} S^r_{i,j}(k)X^r_{i,j}(k)
\]

\[= \sum_{k \in K} \sum_{i \in I} \sum_{j \in J} \sum_{r \in R_{i,j}} \left( C^r_{i,j}(k) + \theta^r T^r_{i,j}(k) \right)X^r_{i,j}(k) \]

subject to:

\[\sum_{k \in K} \sum_{r \in R_{i,j}} X^r_{i,j}(k) = d_{i,j}, \forall i \in I, \forall j \in J \quad (1)\]

\[\sum_{i \in I} \sum_{j \in J} \sum_{r \in R_{i,j}} \sum_{\tau \leq k} X^r_{i,j}(\tau) \delta^r_{i,j,\tau,k} = x_l(k), \forall l \in L, \forall k \in K \quad (2)\]

\[0 \leq x_l(k) \leq u_l v_l(k), \forall l \in L, \forall k \in K \quad (3)\]

\[X^r_{i,j}(k) \geq 0, \forall i \in I, \forall j \in J, \forall k \in K \quad (4)\]

\[d_{i,j} : \text{The total demand from an origin } i \text{ to a destination } j; \]

\[X^r_{i,j}(k) : \text{The freight demand in units of containers from an origin } i \text{ to a destination } j \text{ using a route } r \text{ with a departure time } k; \]

\[x_l(k) : \text{The number of containers using edge } l \text{ at time } k; \]

\[u_l : \text{The edge capacity in units of vehicles for edge } l; \]

\[v_l(k) : \text{The vehicle capacity in units of containers per freight vehicle for edge } l; \]

\[S^r_{i,j}(k) : \text{combination of the non-travel time vehicle cost } C^r_{i,j}(k) \text{ and the cost of the route travel time } T^r_{i,j}(k); \]
Methodology: co-simulation optimization

- Demand: 1. O/D, 2. Volume
- Truck Model & Distribution
- Emission Model
- Predetermined parameters

Service Graph Optimization:
1. Simulator traffic data processing
2. Route generation
3. Marginal cost calculation
4. Descent direction & step size generation with marginal cost

Traffic Simulator:
Route flow vector $X$,
Traffic network status ($\hat{w}$, $\hat{z}$, $\hat{v}$, $\hat{u}$, ...)

if satisfying stop criteria
freight assignment decision
Methodology:

• Step 1: Initialize iteration number, cost coefficients, route collections.
• Step 2: Check if the stop criteria is satisfied. If not, go to Step 3; otherwise, terminate the algorithm and output the result.
• Step 3: Input the route flow vector into the traffic simulator and obtain the marginal cost of each segment.
• Step 4: Update the marginal cost of each segment as well as routes for each O/D pair and check whether there is a new minimal marginal cost route. If there is, then add it into the route collection.
• Step 5: Construct an augmented route flow vector by assigning the trucks onto the minimum cost route in the current traffic status in an all-or-nothing manner.
• Step 6: Set the route flow vector for the next iteration as a linear combination of the previous route flow vector and the augmented route flow vector. Go to Step 2.
Methodology: marginal cost

- Marginal cost: the change in total cost if we add one truck at time instance $k$ on service link $l$

\[
MCP_{l'}^{p'}(k') \approx c_{l'}^{p'} + \sigma_{l'}^{p'} t_{l'}^{p'}(k) \\
+ \sum_{n_{p'}=1}^{N_{p'}} \left( \frac{1}{v_{l'}(e_{a_{p'},n_{p'}},(k'))} \frac{\partial w_{a_{p'},n_{p'}},(k')}{\partial z_{a_{p'},n_{p'}},(e_{a_{p'},n_{p'}},(k'))} \right)
\]

- $e_{a_{p'},n_{p'}},(k')$: entering time at arc $a_{p'},n_{p'}$ for a freight vehicle using path $p'$ with a departure time of $k'$ from the origin.
- $w_{a_{p'},n_{p'}},(k')$: the travel time of arc $a_{p'},n_{p'}$ at time $k'$
- $z_{a_{p'},n_{p'}},(k')$: the traffic volume on road network arc $a_{p'},n_{p'}$ at time $k'$
Distributed co-simulation optimization method

- Limitation of centralized co-simulation optimization method: computational complexity when applied to a large-scale road network.

Abstracted service network

service subnetworks

road network in simulator
Distributed co-simulation optimization method

- Step 1: Initialize iteration number, cost coefficients, route collections, etc.
- Step 2: Check if the stop criteria is satisfied. If not, go to Step 3; otherwise, terminate the algorithm and output the result.
- Step 3: Input the abstract service network route flow vector into each service subnetwork.
- Step 4: For each service subnetwork and road subnetwork pair, perform a load balancing scheme as in the centralized co-simulation method.
- Step 5: Update the marginal cost for the abstracted service network and check whether there is a new minimal marginal cost route. If there is, add it to the associated route collection and created an augmented route flow vector.
- Step 6: Set the route flow vector for the next outer loop iteration as a linear combination of the previous route flow vector and the augmented route flow vector. Go to Step 2.
Numerical Results

• Long Beach Network

- Total number of demand: 3514
- Configure the scenarios by decomposing the network into 2, 3, 4 subnetworks. Apply the distributed co-simulation optimization method for each scenario and gain the total cost value and computation time of each scenario.
- The centralized load balancing method is also applied on the network as a base for comparison.
Numerical Results

- Metropolitan Network
  - Total number of demand: 13600
  - Configure the scenarios by decomposing the network into 3, 4, 5 subnetworks. Apply the distributed co-simulation optimization method for each scenario and gain the total cost value and computation time of each scenario.
  - The centralized load balancing method can not be applied on the network since the capability of the simulator and exponentially growing variables.
Numerical Results

- Long Beach Network
- Performance pattern under different number of subnetworks

<table>
<thead>
<tr>
<th># of Subnetworks</th>
<th>0</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Cost</td>
<td>$C_1$</td>
<td>$1.053C_1$</td>
<td>$1.091C_1$</td>
<td>$1.147C_1$</td>
</tr>
<tr>
<td>Computation Time (second)</td>
<td>4792</td>
<td>3558</td>
<td>3102</td>
<td>3907</td>
</tr>
</tbody>
</table>

$c_1 = $342189.00

- Observation:
With the increase of the number of subnetworks, the total cost gained from the distributed load balancing method is moving away from the optimal one gained from centralized load balancing method. However, the computational time is gained through dividing the network.
Numerical Results

- Long Beach Network
- Pattern under different number of demands

<table>
<thead>
<tr>
<th>Demands</th>
<th># of Subnetworks</th>
<th>0</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td># of demands = 3514</td>
<td>Total Cost</td>
<td>$c_1$</td>
<td>1.053$c_1$</td>
<td>1.091$c_1$</td>
<td>1.147$c_1$</td>
</tr>
<tr>
<td></td>
<td>Computation Time (second)</td>
<td>4792</td>
<td>3558</td>
<td>3102</td>
<td>3907</td>
</tr>
<tr>
<td># of demands = 7028</td>
<td>Total Cost</td>
<td>$c_2$</td>
<td>1.059$c_2$</td>
<td>1.096$c_2$</td>
<td>1.151$c_2$</td>
</tr>
<tr>
<td></td>
<td>Computation Time (second)</td>
<td>$T_2$</td>
<td>0.784$T_2$</td>
<td>0.664$T_2$</td>
<td>0.827$T_2$</td>
</tr>
<tr>
<td># of demands = 14056</td>
<td>Total Cost</td>
<td>$c_3$</td>
<td>1.073$c_3$</td>
<td>1.117$c_3$</td>
<td>1.194$c_3$</td>
</tr>
<tr>
<td></td>
<td>Computation Time (second)</td>
<td>$T_3$</td>
<td>0.806$T_3$</td>
<td>0.675$T_3$</td>
<td>0.862$T_3$</td>
</tr>
</tbody>
</table>

$c_1 = $342189.00  
$c_2 = $728162.00  
$c_3 = $1708326.00  
$T_2 = 5238$ s  
$T_3 = 7965$ s
Numerical Results

- Long Beach Network
- Performance pattern under different number of boundary nodes

<table>
<thead>
<tr>
<th># of demands = 3514</th>
<th># of Boundary Nodes</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Cost</td>
<td>1.024C</td>
<td>1.012C</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>Computation Time</td>
<td>T</td>
<td>1.037T</td>
<td>1.083T</td>
<td></td>
</tr>
</tbody>
</table>

$C = 355880.00$
$T = 3558$ s

- Observation:
  With the decrease of the number of boundary nodes, we gain benefits on computation time, while lose some optimality.
### Numerical Results

- **Metropolitan Network**
- **Performance pattern under different number of subnetworks**

<table>
<thead>
<tr>
<th>Demands</th>
<th># of Subnetworks</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td># of demands = 13600</td>
<td>Total Cost</td>
<td>$C_4$</td>
<td>$1.096C_4$</td>
<td>$1.155C_4$</td>
</tr>
<tr>
<td></td>
<td>Computation Time (second)</td>
<td>1.648$T_4$</td>
<td>1.217$T_4$</td>
<td>$T_4$</td>
</tr>
<tr>
<td># of demands = 27200</td>
<td>Total Cost</td>
<td>$C_5$</td>
<td>1.114$C_5$</td>
<td>1.189$C_5$</td>
</tr>
<tr>
<td></td>
<td>Computation Time (second)</td>
<td>1.918$T_5$</td>
<td>1.39$T_5$</td>
<td>$T_5$</td>
</tr>
</tbody>
</table>

- **Observation:**
With the increase of the number of subnetworks, the total cost gained from the distributed load balancing method is moving away from the optimal one gained from centralized load balancing method. However, the computational time is gained through dividing the network.
Conclusions

- The distributed co-simulation optimization method is tested under two networks: Long Beach network and metropolitan network and tested can address the scalability issue encountered by centralized co-simulation method.
- For Long Beach network, the distributed co-simulation optimization method is tested based on different number of subnetworks, boundary nodes and demand. By decreasing the number of boundary nodes, we can achieve less computational time with some loss on optimality. By increasing the number of subnetworks, we can achieve saving a large amount of computational time with a relatively small loss on the optimality. However, a proper decomposition is needed since if the network is decomposed too much, the interactions between subnetworks will compromise the computational time gained from decomposition.
- Under various number of demand, the pattern of computational time VS optimality sustains.
- For metropolitan network, similar relation between performance and number of subnetworks is revealed.
Thank you!