JOINT POLICY OPTIMIZATION FOR n STORES AND 1 DC

1. Network

We consider n stores indexed by $i = 1, 2, \ldots, n$. The n stores are served inventory from a common DC. In general, we will index a DC variable with 0 when refering to the same variable that we might use for stores. The DC is served inventory from a vendor with infinite inventory. The lead time for inventory to move from DC to Store i is L_i $(1 \leq i \leq n)$ and the lead time for inventory to move from Vendor to DC is L_0.

2. Costs

- The holding costs at the $n+1$ nodes (n stores and 1 DC) are $h_i > 0$. h_i is cost per unit of on-hand inventory before receiving inventory in an epoch.
- The stockout cost (cost of missed sales) at store i $(1 \leq i \leq n)$ is $p_i > h_i$. p_i is cost per unit of missed sales in an epoch.
- The presentation-minimum at store i $(1 \leq i \leq n)$ is $PM_i \geq 0$. The cost of violation of presentation-mimimum at store i $(1 \leq i \leq n)$ is $v_i > h_i$ (also, $v_i < p_i$). v_i is cost per unit of inventory below PM_i before receiving inventory in an epoch.
- The maximum number of units that node i can carry is M_i, for all $0 \leq i \leq n$. The throwout cost at node i is c_i for all $0 \leq i \leq n$. c_i is cost per unit of inventory above M_i after receiving inventory in an epoch.
- The cost of initiating x units of inventory from parent of node i to node i $(0 \leq i \leq n)$ is $\mathbb{I}_{x>0} \cdot (K_i + J_i \cdot x)$

3. Inventory

- Denote on-hand inventory (a.k.a. Inventory Level) at node i at the start of epoch t as: $IL_{t,i} \geq 0$
- Denote on-order inventory arriving at node i in k epochs $(1 \leq k \leq L_i)$ at the start of epoch t as $OO_{t,i,k} \geq 0$

4. Inventory Movements

Denote quantity of inventory movement initiated in epoch t from the parent of node i to node i $(0 \leq i \leq n)$ as $q_{t,i} \geq 0$. Node i will receive that inventory of $q_{t,i}$ in epoch $t + L_i$. Denote $R_{t,i}$ as the inventory received in epoch t at node i. Following the epoch of initiation of inventory movement t and until the epoch of inventory receipt $t + L_i$, this quantity q_i will appear as on-order $OO_{t+j,i,L_i-j+1}, 1 \leq j \leq L_i$.

For the special case where $L_i = 0$, $R_{t,i} = q_{t,i}$ (Seqence of Events below illustrates that within an epoch, receipt of inventory happens after initiation of movement).

Demand at store i $(1 \leq i \leq n)$ in epoch t is denoted by random variable $D_{t,i}$.

5. States and Actions

State S_t in epoch t is defined by the vector:

$$[IL_{t,0}, OO_{t,0,1}, \ldots OO_{t,0,L_0}, IL_{t,1}, OO_{t,1,1}, \ldots, OO_{t,1,L_1}, \ldots, IL_{t,n}, OO_{t,n,1}, \ldots, OO_{t,n,L_n}]$$

Action A_t in epoch t is defined by the vector:

$$[q_{t,0}, q_{t,1}, \ldots, q_{t,n}]$$

6. Sequence of events in an epoch

(1) Observe *State* (simultaneous observation at all $n+1$ nodes of the inventory levels and at all $n+1$ arcs of the on-orders).
(2) Perform *Action* (simultaneous initiation of movements of inventory at all $n+1$ arcs).
(3) Calculate movement costs and holding costs for all nodes based on inventory levels at this point of the epoch
(4) Receipt of inventory (simultaneous receipt of inventory at all $n+1$ nodes).
(5) Calculate throwout cost at all nodes based on inventory levels at this point of the epoch.
(6) Occurrence of demand at stores (including missed sales, i.e., stockouts at the stores).
(7) Calculate stockout costs and presentation violation costs for all stores based on inventory levels at this point of the epoch.

7. Equations defining Inventory Flow

The following equations define the inventory flow in any epoch t:

$$R_{t,i} = \begin{cases} OO_{t,i,1} & \text{if } L_i > 0 \\ q_{t,i} & \text{if } L_i = 0 \end{cases} \text{ for all } t, \text{ for all } 0 \leq i \leq n$$

$$IL_{t+1,0} = \min(M_0, IL_{t,0} - \sum_{i=1}^{n} q_{t,i} + R_{t,0})$$

$$IL_{t+1,i} = \max(0, \min(M_i, IL_{t,i} + R_{t,i}) - D_{t,i}) \text{ for all } 1 \leq i \leq n$$

$$OO_{t+1,i,k} = OO_{t,i,k+1} \text{ for all } 0 \leq i \leq n, \text{ for all } 1 \leq k < L_i$$

$$OO_{t+1,i,L_i} = q_{t,i} \text{ for all } 0 \leq i \leq n$$

$$\sum_{i=1}^{n} q_{t,i} \leq IL_{t,0}$$

8. Cost Equations

The *Reward* is defined by the following costs incurred in any epoch t:

- Movement Cost:
$$\sum_{i=0}^{n} \mathbb{I}_{q_{t,i}>0} \cdot (K_i + J_i \cdot q_{t,i})$$
- Holding Cost:
$$h_0 \cdot (IL_{t,0} - \sum_{i=1}^{n} q_{t,i}) + \sum_{i=1}^{n} h_i \cdot IL_{t,i}$$
- Throwout Cost:
$$c_0 \cdot \max(0, IL_{t,0} - \sum_{i=1}^{n} q_{t,i} + R_{t,0} - M_0) + \sum_{i=1}^{n} c_i \cdot \max(0, IL_{t,i} + R_{t,i} - M_i)$$
- Stockout Cost:
$$\sum_{i=1}^{n} p_i \cdot \max(0, D_{t,i} - \min(M_i, IL_{t,i} + R_{t,i}))$$
- Presentation-Violation Cost:
$$\sum_{i=1}^{n} v_i \cdot \max(0, PM_i - IL_{t+1,i})$$