A Submodular Representation for Hub Network Design Problems with Profits and Single Assignments

Maryam Hemmati [1]

[1]Concordia University and Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT), Montreal H3G 1M8, Canada

Scan to accss full article

Introduction

Hub Location Problems (HLPs) are NP-hard, involving decisions on hub placements and network design to optimize cost efficiency and service levels. In this research, we focus on the Hub Network Design Problem with Profits and Single Assignments (HNDPSA), aiming to maximize total profit using a polynomial approximation algorithm.

Contributions: We present a novel representation of HNDPSA as a submodular maximization problem. Additionally, we propose a polynomial-time greedy heuristic leveraging submodularity for strong approximation results. We then mathematically show an upper bound for the heuristic solution and make a bound tighter under certain circumstances.

Let N be a finite set and $z : 2^N \to \mathbb{R}$ be a real-valued function. The incremental value of adding an element j to set S is $\rho_j(S) = z(S \cup \{j\}) - z(S)$.

Definition:
(a) z is *submodular* if $\rho_j(S) \geq \rho_j(T)$ for all $S \subseteq T \subseteq N$ and $j \in N \setminus T$.
(b) z is *non-decreasing* if $\rho_j(S) \geq 0$ for all $S \subseteq N$ and $j \in N$.
According to [2] and [3], $f(S) = -\sum_{j \in S} d_j$ is submodular and a positive linear combination of submodular functions is submodular.

Problem Definition

In this part, we introduce a path-based representation for the HNDPSA.

Mathematical Notations:

- $G = (N, E)$ is a complete undirected graph.
- P represents paths between nodes of G.
- S, S', H, A, and R denote the sets of hub nodes, non-hub nodes, hub edges, access edges, and paths subset.
- $U = N \cup E \cup P$ denotes the finite ground set.
- R_k represents paths for commodity $k \in K$, where $R = \bigcup_{k=1}^{|K|} R_k$.
- $N_s(R)$, $N_{s'}(R)$, $E_h(R)$, and $E_a(R)$ represent sets of hub nodes, non-hub nodes, hub edges, and access edges in R.
- $f(S, S') = -\sum_{j \in S} f_j - \sum_{j \in Q} f'_j$: Setup cost for hub and non-hub nodes.
- $g(H) = -\sum_{e \in H} g_e - \sum_{e \in A} g'_e$: Setup cost for hub and access edges.
- $h_k(R_k) = \max_{p_k \in R_k} h_{p_k}$: Profit for transporting commodity k.
- $z(S, S', H, A, R) = h(R) + g(H, A) + f(S, S')$: Objective function.

The HNDPSA seeks to maximize $z(S, S', H, A, R)$, where $(S, S', H, A, R) \subseteq U$, subject to various constraints ensuring consistency and limiting cardinality:

$$\max_{(S,S',H,A,R) \subseteq U} \{ z(S, S', H, A, R) : |S| \leq p, |H| \leq q, N_s(R) = S, N_{s'}(R) = S', \\ E_h(R) = H, E_a(R) = A, S \cap S' = \emptyset, H \cap A = \emptyset, |\delta(\{i\})| = 1, \forall i \in S' \}. \tag{1}$$

The first two constraints enforce cardinality constraints on the set of hub nodes and hub edges. The following four constraints ensure that the selected hub nodes, non-hub nodes, hub edges, and access edges match exactly with the nodes and edges of the chosen paths R, respectively. The final set of constraints ensures that each non-hub node $i \in S'$ has exactly one incident access edge selected.

Defining Finite Set of Paths P on Graph G

Figure 1 shows feasible paths for routing commodity k. Black squares represent the set of hub nodes and black circles correspond to the set of non-hub nodes. The weight of hub arcs in the original graph will be multiplied by $0 \leq \alpha \leq 1$ for scaling the profit using the hub network.

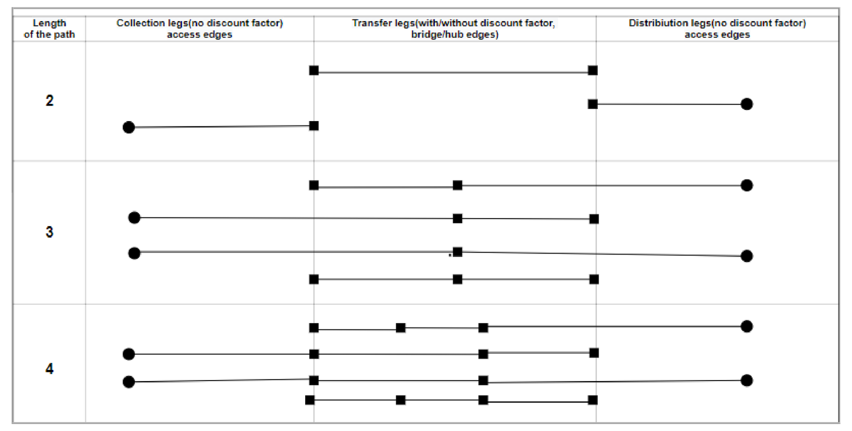

Figure 1. The set of possible paths for routing a commodity

Submodular Properties of HNDPSA

Proposition:(a) $h(R)$ is submodular and non-decreasing.
(b) $z(S, H, R)$ is submodular.

Proof: Given $(S, S', H, A, R) \subseteq (T, T', L, B, Q) \subseteq U$, $q \in T \setminus Q$ is a path with endpoints of an arbitrary commodity $k = \{i, j\}$. For $R_k \neq \emptyset$:

$$h_k(R_k \cup \{q\}) = \max_{p_k \in R_k \cup \{q\}} h_{p_k} - \max_{p_k \in R_k} h_{p_k} = \max\{0, h_q - \max_{p_k \in R_k} h_{p_k}\}$$
$$\geq \max\{0, h_q - \max_{p_k \in Q_k} h_{p_k}\} = \max_{p_k \in Q_k \cup \{q\}} h_{p_k} - \max_{p_k \in Q_k} h_{p_k} = h_k(Q_k \cup \{q\}) - h_k(Q_k)$$

Since $\max_{p_k \in R_k} h_{p_k} \leq \max_{p_k \in Q_k} h_{p_k}$, h_k is non-decreasing. Summing over all commodities, $h(R)$ is submodular and non-decreasing. Since f and g are submodular (see [3]).

Approximation Algorithm

We propose a greedy heuristic algorithm leveraging submodularity to solve the HNDPSA. The algorithm iteratively selects the most profitable feasible path until no further improvement can be made or all commodities are served. Let $q \in P$, and $M_q = q \cup N_s(q) \cup E_h(q) \cup N_{s'}(q) \cup E_a(q) \subseteq U$.
Algorithm Steps:

1. **Initialization:**
   - Set $(S, S', H, A, R)^0$ to empty, $U^0 = (N, E, P)$.
   - Initialize $S' \leftarrow \emptyset$, $A \leftarrow \emptyset$, $\rho_0 \leftarrow 0$, and $t \leftarrow 1$.
2. **Iteration:**
   - While $t \leq |K|$:
     - Select $q^t \subseteq P^{t-1}$ maximizing $\rho_{M_q{}^t}(S, S', H, A, R)^{t-1}$.
     - Update $\rho_{t-1} \leftarrow \rho_{M_q{}^t}$.
     - If $\rho_{t-1} \leq 0$, stop.
     - Update $S' \leftarrow N_{s'}(q^t)$, $A \leftarrow E_a(q^t)$, $S \leftarrow N_s(q^t)$, $H \leftarrow E_h(q^t)$.
     - Update $(S, S', H, A, R)^t \leftarrow (S, S', H, A, R)^{t-1} \cup M_q{}^t$.
     - Update $U^t \leftarrow U^{t-1} \setminus \{q^t\}$.
     - For each $i \in N(q^t)$ such that $i \in S'$:
     - For each $e \in E$ such that $i \in e$ and $e \notin E_a(q^t)$:
     - Update $U^t \leftarrow U^t \setminus \{e\}$.
     - Increment t by 1.
3. **Output:**
   - Output $(S, S', H, A, R)^t$ as the solution.

Finding Maximum Increment using Features of DAGs

The set $M_q{}^t$ at iteration t is efficiently found using longest-path algorithms on an auxiliary directed acyclic graph (DAG) for each unserved commodity $k = \{i, j\} \in K$. We construct an auxiliary DAG $G_k(V_k, E_k)$, a multi-layer graph representing $U^{t-1} = (N^{t-1}, E^{t-1}, P^{t-1})$ at iteration t. Nodes from the original graph G not in S'^{t-1} are duplicated as intermediate hub nodes between i and j to account for potential routing through multiple hubs (see Figure 2). Considerations include cardinality and single assignment constraints, with adjustments made for hub arc discounts.

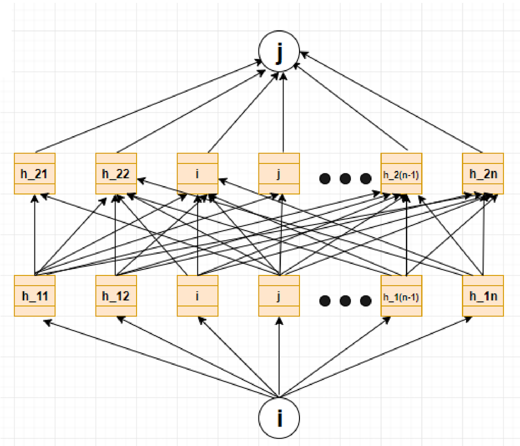

Figure 2. $G_k(V_k, E_k)$ designed for serving commodity $k = \{i, j\}$

Computational Complexity and Theoretical Bound

Proposition: The complexity of Algorithm is $O(|K|^2|N|^2)$.

We could prove there is an upper bound for the solution generated by the algorithm. z^G is the solution coming from our algorithm and z^* is the optimal solution.

Proposition: If the greedy heuristic terminates after t^* iterations:

$$\frac{z^* - z^G}{z^* - z(\emptyset)} \leq \frac{|K| - 1}{|K|} \tag{2}$$

and if z is non-decreasing (considering there is no setup cost),

$$\frac{z^G}{z^*} \leq 1 - \left(\frac{|K| - 1}{|K|}\right)^{|K|} \leq \frac{1 - e}{e} \tag{3}$$

Acknowledgments

We thank Concordia University for supporting this research. Many thanks to Dr. Camilo Ortiz, for his guidance throughout this project.

References

[1] Ivan Contreras and Elena Fernández. Hub location as the minimization of a supermodular set function. *Operations Research*, 62(3):557–570, 2014.

[2] Gilbert LAPORTE. Multi-level facility location as the maximization of a submodular set function. *European Journal of Operational Research*, 247(3), 2015.

[3] George L Nemhauser and Laurence A Wolsey. Analysis of linear production programming. *Journal of the Society for Industrial and Applied Mathematics*, 26(2):418–444, 1978.

[4] Diego Ortiz and Ivan Contreras. Multi-level hub location problems: Formulations and exact algorithms. *Transportation Research Part B: Methodological*, 121:113–129, 2019.