Math Models of OR: Multicommodity Network Flow Problems

John E. Mitchell

Department of Mathematical Sciences
RPI, Troy, NY 12180 USA

November 2018
Outline

1. Introduction
2. Arc-based formulation
3. A path-based formulation
Multiple commodities

A network may be shared by many different services or commodities. For example:

- telephone network: each call is a commodity with its own source and sink
- transportation network: each trip is its own commodity
- supply chain: different goods move through the network
Multiple commodities

A network may be shared by many different services or commodities. For example:

- telephone network: each call is a commodity with its own source and sink
- transportation network: each trip is its own commodity
- supply chain: different goods move through the network
Multiple commodities

A network may be shared by many different services or commodities. For example:
- telephone network: each call is a commodity with its own source and sink
- transportation network: each trip is its own commodity
- supply chain: different goods move through the network
Example
Example

Commodity A
Example
Example

Commodity A
Commodity B
Commodity C
Example
Node-arc incidence matrix

The network can be directed or undirected.

We’ll assume it is a directed network $D = (V, E)$ and let $A$ denote the node-arc incidence matrix. So $A$ has one row for each node, one column for each arc, and one $+1$ and one $-1$ in each column.

Node 1: Supply 10
Nodes 2, 3: Transshipment
Node 4: Demand 10

![Network Diagram]

\[
A = \begin{bmatrix}
    x_{12} & x_{13} & x_{23} & x_{24} & x_{34} \\
    1 & 1 & 0 & 0 & 0 \\
    -1 & 0 & 1 & 1 & 0 \\
    0 & -1 & -1 & 0 & 1 \\
    0 & 0 & 0 & -1 & -1 \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
    v_1 \\
    v_2 \\
    v_3 \\
    v_4 \\
\end{bmatrix}
\]

\[
Ax = \begin{bmatrix}
    10 \\
    0 \\
    0 \\
    -10 \\
\end{bmatrix}
\]
Parameters

We have $K$ commodities, each with its own sources and sinks. The vector $b^k$ gives the net demand at each node for commodity $k$. Transporting one unit of commodity $k$ on arc $(i, j) \in E$ costs $c^k_{ij}$. Each arc $(i, j) \in E$ has a capacity $u_{ij}$. The total amount shipped through arc $(i, j) \in E$ of all the commodities must be no larger than $u_{ij}$.

Without this shared bound, the problem would separate: we could solve separate min cost network flow problems for each commodity.
Formulations

We’ll give two formulations: an *arc-based* formulation and a *path-based* formulation.

Each formulation can be modified straightforwardly to handle the case that the network is undirected.

For more details, see for example Chapter 17 of the text by Ahuja, Magnanti, and Orlin, *Network Flows*, Prentice Hall, 1993.
Outline

1. Introduction

2. Arc-based formulation

3. A path-based formulation
An arc-based formulation

We let $x^k_{ij}$ denote the amount of commodity $k$ shipped on arc $(i,j) \in E$. The arc-based LP formulation for the multicommodity network flow problem is:

$$\min_{x \in \mathbb{R}^{|E| \times K}} \sum_{k=1}^{K} \left( \sum_{(i,j) \in E} c^k_{ij} x^k_{ij} \right)$$

subject to

$$Ax^k = b^k \quad \text{flow conservation}$$

$$\sum_{k=1}^{K} x^k_{ij} \leq u_{ij} \quad \forall (i,j) \in E \quad \text{arc capacity}$$

$$x \geq 0$$

Without the shared capacity constraint, we’d have separate problems for each commodity.

Even if all the data is integer, this problem may well have a fractional optimal solution, in contrast with the single commodity case.
An arc-based formulation

We let $x_{ij}^k$ denote the amount of commodity $k$ shipped on arc $(i, j) \in E$. The arc-based LP formulation for the multicommodity network flow problem is:

$$\min_{x \in \mathbb{R}^{|E| \times K}} \sum_{k=1}^{K} \left( \sum_{(i, j) \in E} c_{ij}^k x_{ij}^k \right)$$

subject to

$$Ax^k = b^k$$

$$\sum_{k=1}^{K} x_{ij}^k \leq u_{ij}, \quad \forall (i, j) \in E$$

$$x \geq 0$$

Without the shared capacity constraint, we'd have separate problems for each commodity.

Even if all the data is integer, this problem may well have a fractional optimal solution, in contrast with the single commodity case.
Let $n = |E|$ and $m = |V|$. The arc-based formulation has $nK$ variables and $n + mK$ constraints (excluding nonnegativity constraints).

For example, in a graph with 1000 nodes and 5000 arcs and with one commodity for each pair of nodes, we have on the order of $10^9$ variables and constraints.

In principle, the problem can be solved using an LP solver, but in practice the size of the problem can be daunting.
Outline

1. Introduction

2. Arc-based formulation

3. A path-based formulation
A path-based formulation

We assume all the costs $c_{ij}^k$ are nonnegative, so in particular there are no negative length cycles.

Any feasible flow in commodity $k$ can be represented as a sum of flows on paths.

We assume *each commodity has a single source $s^k$ and a single sink $t^k$*; this can be done without loss of generality, by introducing a supersource and/or a supersink for each commodity and constructing capacities and costs appropriately for the edges linking the supersource to the sources and the sinks to the supersink.
Paths to supply B:

- P1 - A - D - B
- P1 - A - D - C - B
- P3 - C - B
- P4 - E - F - B

Sum of flows on paths = demand at F
A set of paths for commodity $k$

A path-based formulation

Node $s^k$: Supply 10
Node $t^k$: Demand 10

$$c_{13} = 6, c_{32} = 5, c_{43} = 2, c_{45} = 6, c_{12} = 2, c_{24} = 1, c_{35} = 3$$
A set of paths for commodity $k$

Path $s^k_{24}t$ cost: $c^k_{s24t} = 2 + 1 + 6 = 9$
A set of paths for commodity $k$

Path $s3t$ cost: $c_{s3t}^k = 6 + 3 = 9$
A path-based formulation

A set of paths for commodity $k$

Path $s243t$ cost: $c^k_{s243t} = 2 + 1 + 2 + 3 = 8$
A set of paths for commodity $k$

Path $s^k 3 2 4 t$ cost: $c^k_{s^k 3 2 4 t} = 6 + 5 + 1 + 6 = 18$
A path-based formulation

A set of paths for commodity \( k \)

Path \( s24t \) cost: \( c_{s24t}^k = 2 + 1 + 6 = 9 \)
Path \( s3t \) cost: \( c_{s3t}^k = 6 + 3 = 9 \)
Path \( s243t \) cost: \( c_{s243t}^k = 2 + 1 + 2 + 3 = 8 \)
Path \( s324t \) cost: \( c_{s324t}^k = 6 + 5 + 1 + 6 = 18 \)

meet demand:
\[
f_{s24t}^k + f_{s3t}^k + f_{s243t}^k + f_{s324t}^k = 10
\]
A path-based formulation

A set of paths for commodity $k$

Path $s24t$ cost: $c^k_{s24t} = 2 + 1 + 6 = 9$
Path $s3t$ cost: $c^k_{s3t} = 6 + 3 = 9$
Path $s243t$ cost: $c^k_{s243t} = 2 + 1 + 2 + 3 = 8$
Path $s324t$ cost: $c^k_{s324t} = 6 + 5 + 1 + 6 = 18$

capacity constraint for $(2, 4)$:

$$f^k_{s24t} + f^k_{s243t} + f^k_{s324t} \leq u_{24}$$
A path-based formulation

A set of paths for commodity $k$

Path $s24t$ cost: $c_{s24t}^k = 2 + 1 + 6 = 9$
Path $s3t$ cost: $c_{s3t}^k = 6 + 3 = 9$
Path $s243t$ cost: $c_{s243t}^k = 2 + 1 + 2 + 3 = 8$
Path $s324t$ cost: $c_{s324t}^k = 6 + 5 + 1 + 6 = 18$

capacity constraint for $(2, 4)$:
$f_{s24t}^k + f_{s243t}^k + f_{s324t}^k \leq u_{24}$
also add in flows for other commodities
A path-based formulation

Each commodity has a set of possible paths

We let $d^k \in \mathbb{R}$ denote the (scalar) demand for commodity $k$ and we let $P^k$ denote the set of all paths from the source for commodity $k$ to the sink for commodity $k$. Each path $p \in P^k$ has a cost per unit shipped:

$$c^k_p = \sum_{(i,j) \in p} c^k_{ij}.$$ 

We define an indicator parameter to capture which edges are on a path $p$:

$$\delta_{ij}(p) = \begin{cases} 1 & \text{if } (i,j) \in p \\ 0 & \text{otherwise} \end{cases}$$

Our variables are $f^k_p$ for each $p \in P^k$, for each $k = 1, \ldots, K$, which gives the amount of commodity $k$ shipped on path $p$. 
The path-based formulation of the multicommodity problem is then:

$$\min_f \sum_{k=1}^K \left( \sum_{p \in P^k} c_p^k f_p^k \right)$$

subject to

$$\sum_{p \in P^k} f_p^k = d^k \quad \text{for } k = 1, \ldots, K \quad \text{demand}$$

$$\sum_{k=1}^K \sum_{p \in P^k} \delta_{ij}(p)f_p^k \leq u_{ij} \quad \forall (i, j) \in E \quad \text{capacity}$$

$$f_p^k \geq 0 \quad \forall p \in P^k, k = 1, \ldots, K$$

The path-based formulation has $n + K$ constraints (excluding nonnegativity constraints). For the earlier numbers of 1000 nodes and 5000 arcs and with one commodity for each pair of nodes, this results in on the order of $10^6$ constraints.
Variables

The drawback is that the path-based formulation has an exponential number of variables. Since, we have an exponential number of possible paths for any one commodity, in the worst case.

Thus, the paths are not all enumerated a priori; instead, they are generated as needed.

They can be found by solving shortest path problems.

To check optimality:

- Check every primal variable (that is, every path) has a nonnegative reduced cost.
- So check dual feasibility. Each dual constraint corresponds to a different path.
- Can find the most violated dual constraint for commodity $k$ by solving a shortest path problem: edge weights are the original weights, with an adjustment representing the congestion on the edges.
Variables

The drawback is that the path-based formulation has an exponential number of variables. Since, we have an exponential number of possible paths for any one commodity, in the worst case.

Thus, the paths are not all enumerated \textit{a priori}; instead, they are generated as needed.

They can be found by solving \textit{shortest path problems}.

To check optimality:

- Check every primal variable (that is, every path) has a nonnegative reduced cost.
- So check dual feasibility. Each dual constraint corresponds to a different path.
- Can find the most violated dual constraint for commodity $k$ by solving a shortest path problem: edge weights are the original weights, with an adjustment representing the congestion on the edges.
Variables

The drawback is that the path-based formulation has an exponential number of variables. Since, we have an exponential number of possible paths for any one commodity, in the worst case.

Thus, the paths are not all enumerated \textit{a priori}; instead, they are generated as needed.

They can be found by solving \textbf{shortest path problems}.

To check optimality:

- Check every primal variable (that is, every path) has a nonnegative reduced cost.
- So check dual feasibility. Each dual constraint corresponds to a different path.
- Can find the most violated dual constraint for commodity $k$ by solving a shortest path problem: edge weights are the original weights, with an adjustment representing the congestion on the edges.
Jun office hours this week:
Thursday 10-11.