Supply Chain Management: Logistics Network Design

Donglei Du
(ddu@unb.edu)

Faculty of Business Administration, University of New Brunswick, NB Canada Fredericton
E3B 9Y2

Table of contents I

Section 1

Introduction

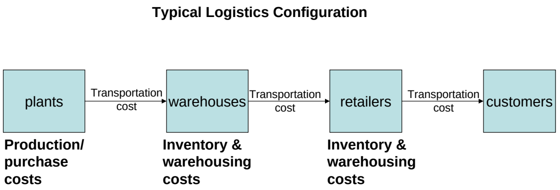
Subsection 1

The Logistics Network

The Logistics Network I

- The objective of this chapter is to present some issues involved in the design and configuration of the logistics network. Obviously these are strategic decisions because they have a long-standing effect on the firm.
- The Logistics Network consists of:
  - Facilities:
    - Plants/Vendors Ports
    - Warehouse
    - Retailers/Distribution Centers
    - Customers
  - Raw materials and finished products that flow between the facilities.

The Logistics Network II



Key Strategic Decisions in the Logistics Network I

- Assuming that plants and retailer locations are fixed, we concentrate on the following strategic decisions in terms of warehouses.
  - Pick the optimal number, location, and size of warehouses
  - Determine optimal sourcing strategy
    - Which plant/vendor should produce which product
  - Determine best distribution channels
    - Which warehouses should service which retailers
  - The objective is to design or reconfigure the logistics network so as to minimize annual system-wide costs, including
    - Production/ purchasing costs
    - Inventory carrying costs, and facility costs (handling and fixed costs)
    - Transportation costs

Key Strategic Decisions in the Logistics Network II

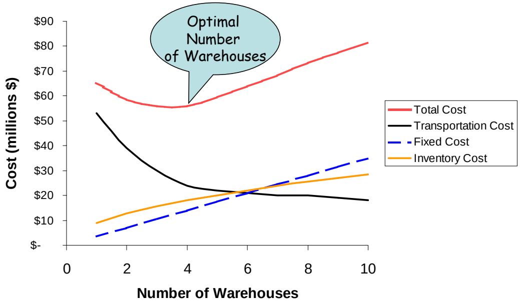
- That is, we would like to find a minimal-annual-cost configuration of the distribution network that satisfies product demands at specified customer service levels.

The trade-off in this problem I

- Increasing the number of warehouses yields
  - An improvement in service level due to the reduction in average travel time to the customers.
  - An increase in inventory costs due to increased safety stocks required to protect each warehouse against uncertainties in customer demands.
  - An increase in overhead and setup costs
  - A reduction in outbound transportation costs: transportation costs from the warehouse to the customers.
  - An increase in inbound transportation costs: transportation costs from the suppliers and/or manufacturers to the warehouse.
- In essence, the firm must balance the costs of opening new warehouses with the advantages of being close to the customer.

The trade-off in this problem II

- Thus warehouse location decisions are crucial determinants of whether the supply chain is an efficient channel for the distribution of products.



Subsection 2

Major Steps in Network Design

Major Steps in Network Design

Step 1. Data Collection

Step 2. Data Aggregation

Step 3. Data Validation and Model

Step 4. Optimization

Step 1. Data Collection I

A typical network configuration problem involves large amount of data, including information on

1. Location of customers, stocking points and sources—location theory
2. A listing of all products
3. Demand for each product by customer location–forecast technique
4. Transportation rates by mode—information system, like rating engine
5. Mileage estimation—GIS
6. Warehousing costs (handling and fixed)—inventory management
7. Service level requirement—probabilistic technique
8. Shipment sizes by product

Step 1. Data Collection II

9. Order patterns by frequency, size, season, content
10. Order processing costs
11. Customer service goals

Transportation Cost I

$$\text{Transportation costs} = \text{Transportation rate} \times \text{Distance}$$

1. Transportation rate: the cost per mile per SKU. An important characteristic of a class of rates for truck, rail, UPS and other trucking companies is that the *rates are quite linear with the distance but not with volume.* Usually there are two kinds of transportation rates:
   - Internal fleet (company-owned): It can be easily calculated from information like annual costs per truck, annual mileage per truck, annual amount delivered, and truck's effective capacity.

Transportation Cost II

- External fleet (third-part): More complex calculation is needed: There are rating engines available, such as the SMC^3 RateWare—www.smc3.com.

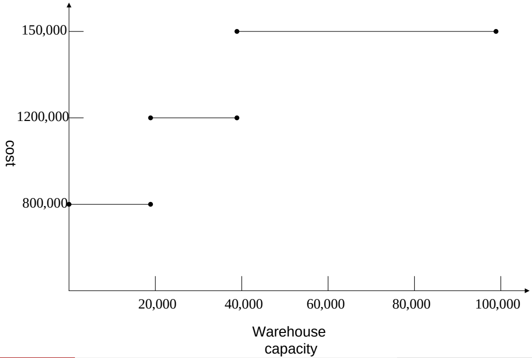
2. Mileage estimation: Once we know the transportation rates, which usually depends on the distance, we need to estimate the mileage between any two locations. Depending on your situation, you may want
    - Exact estimation: this usually can be obtained using GIS system, but the drawback is cost and speed—you may need to install GIS receiver and slow down the operation of a Decision-Support System (See Chapter 12 for more information)

Transportation Cost III

- Approximate estimation: For most of the applications, this will be sufficient.

Warehousing costs (handling, and fixed costs) I

- Handling cost: proportional to the amount of material the flows through the warehouse
- Fixed Cost: All costs that are not proportional to the amount of material the flows through the warehouse. It is typically proportional to *warehouse space size (or warehouse capacity)* but in a nonlinear way.



Warehousing costs (handling, and fixed costs) II

1. So we need to estimate warehouse capacity. Obviously the capacity is proportional to the peak inventory, not the average inventory or annual flow.
2. We introduce the concept of inventory turnover ratio, given by

$$\text{inventory turnover ratio} = \frac{\text{annual flow}}{\text{average inventory level}}$$

3. The warehouse capacity is given by

$$\begin{aligned} \text{Warehouse Capacity} &= 3 \times (2 \times \text{average inventory level}) \\ &= 6 \times \frac{\text{annual flow}}{\text{inventory turnover ratio}} \end{aligned}$$

Step 2. Data Aggregation

- Aggregate and clean the data because
  1. the data collected in Step 1 is usually overwhelming,
  2. the cost of obtaining and processing the real data is huge,
  3. the form in which data is available must be streamlined,
  4. the size of the resulting location model is huge, and
  5. the accuracy of forecast demand is improved.
  6. Of course, data aggregation only approximates the real data, so the impact on model's effectiveness must be addressed.

The impact of aggregate demand: $\sigma_1 + \sigma_2 \geq \sqrt{\sigma_1^2 + \sigma_2^2}$ I

- Consider the following example with two customers: Please do it yourself in class.
- Given historical demands for customers 1 and 2 in the following table:

Year	1992	1993	1994	1995	1996	1997	1998
Customer 1	22346	28549	19567	25457	31986	21897	19854
Customer 2	17835	21765	19875	24346	22876	14653	24987
Total 2	40181	50314	39442	49803	54862	36550	44841

The impact of aggregate demand: $\sigma_1 + \sigma_2 \geq \sqrt{\sigma_1^2 + \sigma_2^2}$ II

- Here is the summary of the historical data, we can see the average demand for the aggregated customer is the sum of the two averages. However, the variability, measured by standard deviation and coefficient of variation is smaller.

Statistics	Average	standard deviation	coefficient of variation
Customer 1	24237	4685	0.192
Customer 2	20905	3427	0.173
Total 2	45142	6757	0.150

A heuristic to aggregate data I

- Customer-based Clustering: Customers located in close proximity are aggregated using a grid network or clustering techniques. All customers within a single cell or a single cluster are replaced by a single customer located at the centroid of the cell or cluster. We refer to a cell or a cluster as a customer zone.
- Product type-based clustering: Place all SKU's into a source-group. A source group is a group of SKU's all sourced from the same place(s). Within each of the source-groups, aggregate the SKU's by similar logistics characteristics (Weight, Volume, Holding Cost).
- A rule of thumb for aggregate customers and product types is give by
  - Aggregate 150-200 customers or 20-50 product types points for each zone.

A heuristic to aggregate data II

- Make sure each zone has an approximate equal amount of total demand
- Place the aggregated point at the center of the zone.
- In this case, the error is typically no more than 1%

Step 3. Data Validation and Model

- Once the data are collected and cleaned, we need to ensure that the data and model accurately reflect the network design problem.
- This is typically done by reconstructing the existing network configuration using the model and collected data, and comparing the output of the model to existing data. The purpose is to answer the following questions:
  - Does the model make sense?
  - Are the data consistent?
  - Can the model results be fully explained?
  - Did you perform sensitivity analysis?

Step 4. Optimization

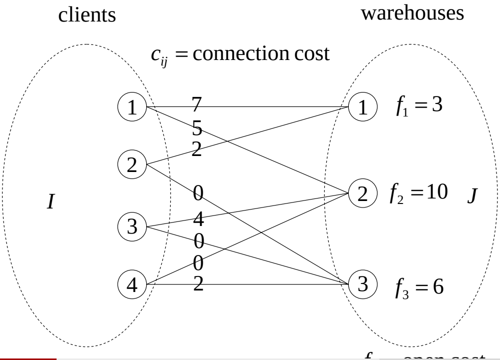
- Once the data are collected, cleaned, and verified, the next step is to optimize the configuration of the logistics networks. In practice, two techniques are employed:
    - Mathematical optimization techniques, including
        - Exact algorithms: find optimal solutions
        - Heuristics: find "good" solutions, not necessarily optimal
- Simulation models that provide a mechanism to evaluate specified design alternatives created by the designer.

Section 2

A Facility Location Problem

A Facility Location Problem I

- Let $J = \{1, 2, 3\}$ be a set of three potential sites for establishing new warehouses and $I = \{1, 2, 3, 4\}$ be a set of four clients (distribution centers or retailers). There is an open cost f_j for establishing a warehouse at site $j \in J$. There is a transportation cost c_{ij} of warehouse j serving client i. These data are summarized in the following graph.



A Facility Location Problem II

- Assume we assign clients according to the closest-site rule, i.e., assign a client to the open warehouse which has the minimal connection cost. The objective is to decide
    - the number of warehouses that should be established and their locations, so as to minimize the total cost, including connection and open costs.

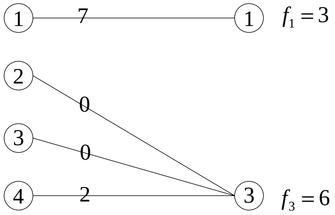
Solution—Enumeration approach I

- There are $2^3 - 1$ different combinations for open and closed warehouses (excluding the case where no warehouse is established at all).

Open warehouses	Open cost	Connection cost	Total cost
1	3	∞	∞
2	10	∞	∞
3	6	∞	∞
1,2	3+10	11	24
1,3	3+6	9	<u>18</u>
2,3	10+6	5	21
1,2,3	3+10+6	5	24

Solution—Enumeration approach II

- The optimal solution is given in the following graph.



$$\text{Total cost} = (3+6)+(7+0+0+2) = 18$$

Three steps in formulating (Integer) Linear Program

- Step 1. Define the decision variables
- Step 2. Write the objective in terms of theses decision variables
- Step 3. Write the constraints in terms of theses decision variables

Solution—Integer Linear Programming approach I

- Define variable for $i = 1, 2, 3, 4$, $j = 1, 2, 3$.

$$x_{ij} = \begin{cases} 1, & \text{If warehouse } j \text{ serves client } i \\ 0, & \text{Otherwise} \end{cases}$$

$$y_j = \begin{cases} 1, & \text{If warehouse } j \text{ is established} \\ 0, & \text{Otherwise} \end{cases}$$

Solution—Integer Linear Programming approach II

- The ILP is given by

$$
\begin{aligned}
\min \; & \sum_{j=1}^{3} f_j y_j + \sum_{i=1}^{4} \sum_{j=1}^{3} c_{ij} x_{ij} \\
& \sum_{j=1}^{3} x_{ij} = 1, i = 1, 2, 3, 4 \\
& x_{ij} \leq y_j, i = 1, 2, 3, 4, j = 1, 2, 3 \\
& x_{ij} = 1, 0, i = 1, 2, 3, 4, j = 1, 2, 3 \\
& y_j = 1, 0, j = 1, 2, 3
\end{aligned}
$$

- Solving ILP, we get the same optimal solution as in the previous graph.

Section 3

Optimial Distribution and outsourcing

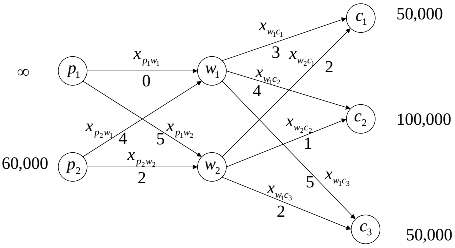
An Example of Optimizing Distribution and outsourcing Channels I

- Single product.
- Twp plants, referred to p_1 and p_2.
- Plant p_2 has an annual capacity of 60,000 units, while p_1 has an unlimited capacity.
- The two plants have the same production costs.
- Two existing warehouses, referred to w_1 and w_2, have identical warehouse handling costs.
- Three market areas, c_1 c_2 and c_3, with demands of 50,000, 100,000, and 50,000, respectively.

An Example of Optimizing Distribution and outsourcing Channels II

- The following table provides distribution cost per unit.

	p_1	p_1	c_1	c_2	c_3
w_1	0	4	3	4	5
w_1	5	2	2	1	2



An Example of Optimizing Distribution and outsourcing Channels III

- Our objective is to find a distribution strategy that specifies the flow of products from the suppliers through the warehouses to the market areas without violating the plant p_2 production capacity constraint, that satisfies market area demands, and that minimizes the total distribution costs.

Analysis and Solution I

- We formulate this problem as a linear programming:
- Define the following variables
  - $x_{p_i w_j}$ be the flow from plant p_i $(i = 1, 2)$ to warehouse w_j $(j = 1, 2)$.
  - $x_{w_j c_k}$ be the flow from warehouse w_j $(i = 1, 2)$ to market c_k $(k = 1, 2, 3)$.

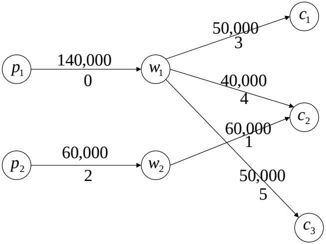
Analysis and Solution II

- Then the LP is given by

$$
\begin{aligned}
\min\ 0x_{p_1w_1} + 5x_{p_1w_2} + 4x_{p_2w_1} + 2x_{p_2w_2} + 3x_{w_1c_1} \\
+4x_{w_1c_2} + 5x_{w_1c_3} + 2x_{w_2c1} + 1x_{w_2c_2} + 2x_{w_2c_3} \\
x_{p_2w_1} + x_{p_2w_2} \leq 60,000 \\
x_{p_1w_1} + x_{p_2w_1} = x_{w_1c_1} + x_{w_1c_2} + x_{w_1c_3} \\
x_{p_1w_2} + x_{p_2w_2} = x_{w_2c_1} + x_{w_2c_2} + x_{w_2c_3} \\
x_{w_1c_1} + x_{w_2c_1} = 50,000 \\
x_{w_1c_2} + x_{w_2c_2} = 100,000 \\
x_{w_1c_3} + x_{w_2c_3} = 50,000 \\
x \geq 0
\end{aligned}
$$

Analysis and Solution III

- The optimal solution is given in the following figure



$$\text{total cost} = [140(0)+60(2)+50(3)+40(4)+60(1)+50(5)](1000) = 740{,}000$$

Section 4

Maximum flow problem

Maximum flow problem I

- Given a flow network $G(V, E, c)$, where V is a set of nodes, E is a set of arcs and c is capacity function such that each edge $(u, v) \in E$ has a flow capacity c_{uv}, we distinguish two vertices : a source s and a sink t.
- A *feasible flow* in G is a function $f : V \times V \to \mathbb{R}^+$ that satisfies two properties:
  1. Capacity constraint: for each arc $(u, v) \in E$, we require
  $$f_{uv} \leq c_{uv}.$$
  2. Flow conservation: for each $u \in V - \{s, t\}$, we require
  $$\sum_v f_{uv} = 0.$$
- The objective is to find a flow of maximum value from s to t.

Maximum flow problem: LP formulation I

- Let $\delta_i^- = \{j : (j,i) \in A\}$ and $\delta_i^+ = \{j : (i,j) \in A\}$ denote the sets of in-neighbors and out-neighbors of vertex $i \in V$, respectively. Let F be the total flow across the network
- The so called *arc-flow* formulation is given by

$$\max F \tag{1}$$

$$\sum_{j \in \delta_i^+} f_{ij} - \sum_{j \in \delta_i^-} f_{ji} = \begin{cases} F, & i = s \\ 0, & \forall i \neq s,\ t \\ -F, & i = t \end{cases} \tag{2}$$

$$0 \leq f_{ij} \leq u_{ij} \qquad \forall (i,j) \in A \tag{3}$$

An equivalent *path-flow* LP formulation

- The *path-flow* LP formulation. Let $\mathcal{P}$ denote the set of all m-routes, and let $\mathcal{P}_{ij}$ denote the set of s-t routes that contain a given arc $(i,j) \in A$. Define vector variable $x = \{x_p\}_{p\in\mathcal{P}} \in R_+^{|\mathcal{P}|}$, where each x_p is the weight assigned to s-t route $p \in \mathcal{P}$.

$$\max \sum_{p\in\mathcal{P}} x_p \tag{4}$$

$$\sum_{p\in\mathcal{P}_{ij}} x_p \leq u_{ij}, \qquad \forall (i,j) \in A \tag{5}$$

$$x_p \geq 0, \qquad \forall p \in \mathcal{P} \tag{6}$$

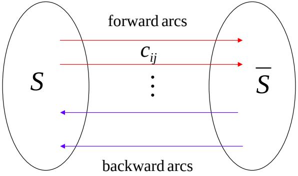
How to solve the maximum flow problem

- General-purpose method: LP solver, such as Simplex method or Interior point method.
- Special purpose method: we introduce the Ford-Fulkerson labeling method in this class.

The max-flow-min-cut theorem I

- A *cut* is a partition of node set V into two disjoint sets S and $\bar{S}$ such that $s \in S$ and $t \in \bar{S}$, denoted as $(S, \bar{S})$.
- The *cut capacity* is the sum of capacities across the cut; that is

$$C(S, \bar{S}) = \sum_{(i,j) \in (S, \bar{S})} c_{ij}$$



- The max-flow-min-cut theorem says that: the maximum flow value is always equal to the minimum cut capacity.

The Ford-Fulkerson labeling algorithm I

Initialization: Fix a feasible flow $f_{ij} = 0$ with value $F = 0$. Lable s with $(-, \infty)$ and let $S = \{s\}$ be the set of labled nodes and $\bar{S} = V - S$ the set of unlabeled nodes.

Label the nodes: Whenever $t \notin S$, choose a pair of nodes $i \in S$ and $j \in \bar{S}$ such that

- either $f_{ij} < c_{ij}$ for $(i, j) \in A$, called *forward* arc, in which case, label node j with

$$(i^+, \epsilon_j), \text{ where } \epsilon_j = \min\{\epsilon_i, c_{ij} - f_{ij}\}$$

- or $f_{ji} > c_{ji}$ for $(j, i) \in A$, called *backward* arc, in which case, label node j with

$$(i^-, \epsilon_j), \text{ where } \epsilon_j = \min\{\epsilon_i, f_{ij}\}$$

The Ford-Fulkerson labeling algorithm II

In both cases, update S and $\bar{S}$ as follows

$$S = S \cup \{j\}, \quad \bar{S} = \bar{S} - \{j\}.$$

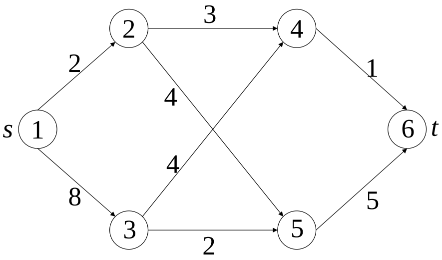
If no such pair of nodes i and j can be found. Stop, the current flow is optimal.

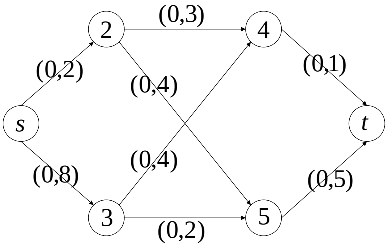
Locate an augmenting path: $t \in S$, we can locate an augmenting path by working backward from t.

Update the feasible flow: Let $\epsilon = \min_{j \in \text{augmenting path}} \epsilon_j$. Update the flow on the augmenting by

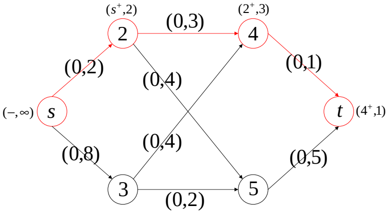
$$f_{ij} := \begin{cases} f_{ij} + \epsilon, & \text{if } (i,j) \text{ is a froward arc;} \\ f_{ij} - \epsilon, & \text{if } (i,j) \text{ is a backward arc.} \end{cases}$$

An example

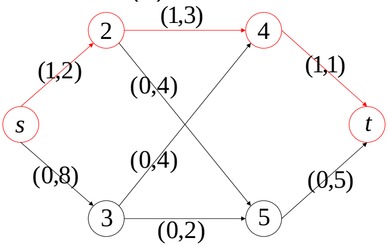




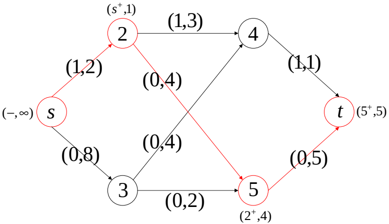
(a) $F = 0$



(b) Augmenting path

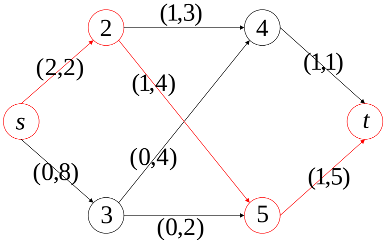


(c) $F = 1$

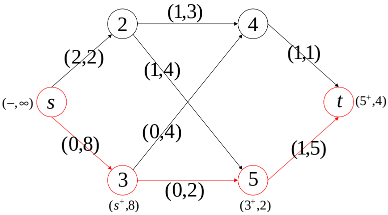


(d) Augmenting path

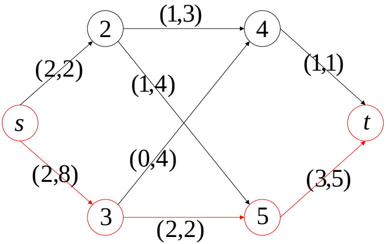
An example



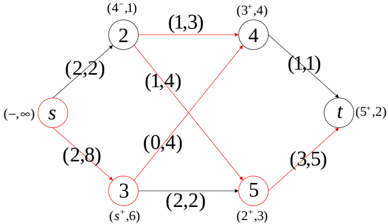
(e) $F = 2$



(f) Augmenting path

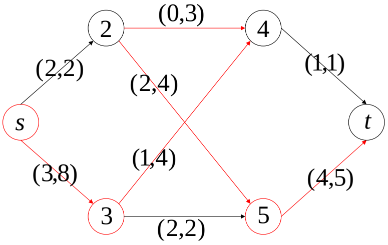


(g) $F = 4$

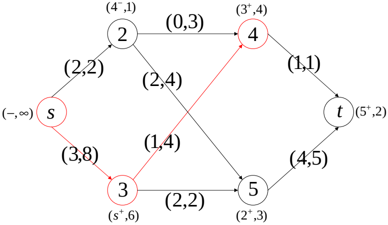


(h) Augmenting path

An example



(i) $F = 5$



(j) minimum cut

Figure: Labeling algorithm

Solving as an LP

- $$\begin{aligned}
\max x_{12} + x_{13} \\
x_{12} - x_{24} - x_{25} = 0 \\
x_{13} - x_{34} - x_{35} = 0 \\
x_{24} + x_{34} - x_{46} = 0 \\
x_{25} + x_{35} - x_{56} = 0 \\
0 \le x_{12} \le 2 \\
0 \le x_{13} \le 8 \\
0 \le x_{24} \le 3 \\
0 \le x_{25} \le 4 \\
0 \le x_{34} \le 4 \\
0 \le x_{35} \le 2 \\
0 \le x_{46} \le 1
\end{aligned}$$

Section 5

Network synthesis problem

Subsection 1

Offline version

Network synthesis problem: The offline version

- A symmetric requirement matrix $R = (r_{ij}) \in \mathbb{R}^{n\times n}$, where each entry r_{ij} indicates the minimum flow requirement between sites i and j. By default $r_{ii} = 0$, $\forall i \in N$.
- The goal is to construct an undirected, simple (no loops and parallel edges) network $G = [N, E, c]$ on site set N, with edge set E and edge capacities $\{c(e) \geq 0 : e \in E\}$, such that
  1. (i) all the minimum flow requirements are met one at a time (that is, for any $i, j \in N,\ i \neq j$, the maximum flow value in G between i and j is at least r_{ij}), and
  2. (ii) $\sum_{e\in E} c(e)$ is minimum.

Network synthesis problem: LP formulation I

- Here is the LP formulation for this problem:

$$\left[\min \sum_{(i,j)\in(S,\overline{S}} c_{ij}x_{ij} : \sum_{(i,j)\in(S,\overline{S}} x_{ij} \geq \max_{(i,j)\in(S,\overline{S}} r_{ij},\ x_{ij} \geq 0, \forall(i,j)\right]$$

- The offline NSP is strongly polynomially solvable as showed by Gomory and Hu.
    - The following closed-form expression for the optimal value of the offline NSP is due to Gomory and Hu, and Mayeda, independently:

$$\frac{1}{2}\sum_{i\in N} \pi_i,$$

where $\pi_i = \max_{j\in N} r_{ij}$ is the *potential* of site i, $\forall i \in N$.

Network synthesis problem: LP formulation II

- Here is an example to illustrate the algorithm of Gomory and Hu

R	1	2	3	4	5	π
1	0	2	1	3	4	4
2	2	0	3	6	7	7
3	1	3	0	8	2	8
4	3	6	8	0	5	8
5	4	7	2	5	0	7

$\Rightarrow$

Step 1. (Sorting)

$\pi_4 \geq \pi_3 = 8 \geq \pi_5 \geq \pi_2 = 7 \geq \pi_1 = 4$

Step 2. (Peeling)

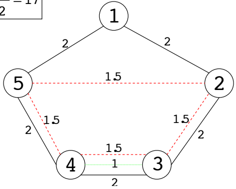
(1) peeling 4 : 4, 4, 3, 3, 0

(2) peeling 3 : 1, 1, 0, 0

(3) peeling 1 : 0, 0

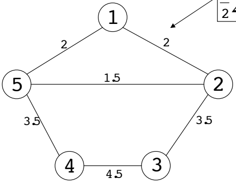
Step 3. (Superposing)

$\Downarrow$



$\Leftarrow$

$$\frac{1}{2}\sum \pi_i = \frac{34}{2} = 17$$



Subsection 2

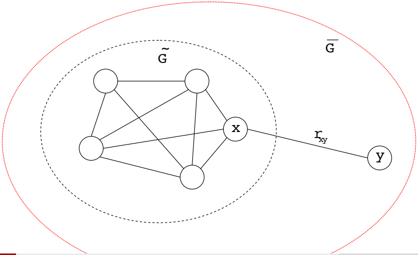
Online version

Network synthesis problem: The online version I

- Very often, in practical network designing, the source and destination and the minimum flow requirements only become known and/or are updated one by one in sequence and after all the previous requirements in the sequence have been served by installing necessary edge-capacities. Any installed edge-capacity cannot be decreased, but can only be increased in future.
  - At any point in time, a certain set $\{r_{ij} : (i,j) \in S\}$ of requirements between some set S of pairs of distinct sites, and through it the set $\widetilde{N} = \{i : (i,j) \in S \text{ for some } j\}$ of sites are known to us.
    - The on-line algorithm is required to have designed a network $\widetilde{G}$ on site set $\widetilde{N}$ that meets the revealed set of minimum flow requirements one at a time.

Network synthesis problem: The online version II

- The next piece of information revealed is some requirement r_{xy}. If some requirement between sites x and y was revealed before then the new value is greater than the previous and replaces the previous; else, r_{xy} is a new revealed requirement and in that case, the new revealed set of sites is $\overline{N} = \widetilde{N} \cup \{x, y\}$.
  - The on-line algorithm is required to update $\widetilde{G}$ to a network $\overline{G}$ on site set $\overline{N}$ (that includes at most two more sites) such that none of the previous edges-capacities in $\widetilde{G}$ is decreased and the new requirement r_{xy} is also satisfied.
- Here is a graphical illustration for Version 1:



Network synthesis problem: The online version III

Competitive ratio I

- We can analyze the performance of any online algorithm via *competitive ratio*: the worst-case ratio between the total edge-capacity of the on-line algorithm and the corresponding optimal (off-line) total edge-capacity over all instances of the problem.
- We present a best possible online algorithm for Version 1.
  - We shall show that the best competitive ratio for Version 1 is given by
$$\alpha_n = 2 - \frac{2}{n},\ n \geq 2.$$
  - There are two aspects in showing the claim above.

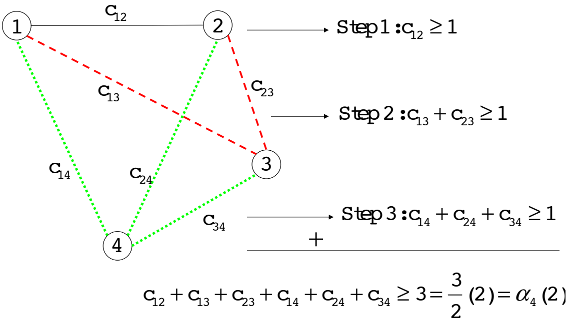
Competitive ratio II

1. Lower-bounding: No online algorithm can achieve better than α_n. for the instance below:

$$(r_{12} = 1, r_{23} = 1, \cdots, r_{n-1,n} = 1),$$

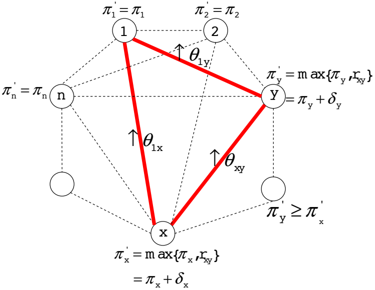
   revealed in this order.

2. We illustrate the argument using a small example of $n = 4$ $(\alpha_4 = \frac{3}{2})$.



Competitive ratio III

3. Upper-bounding: The competitive ratio of the algorithm proposed is no more than α_n.



$$\pi_1 \geq \cdots \geq \pi_y \geq \cdots \geq \pi_x \geq \cdots \geq \pi_n$$

If $\pi_y' < \pi_1'$

$$\begin{cases} \theta_{xy} = \frac{1}{2}\delta_y, \\ \theta_{1x} = \delta_x - \frac{1}{2}\delta_y, \\ \theta_{1y} = \frac{1}{2}\delta_y. \end{cases}$$

If $\pi_y' \geq \pi_1'$

$$\begin{cases} \theta_{xy} = \delta_x - \frac{1}{2}(\pi_1 - \pi_y), \\ \theta_{1x} = \frac{1}{2}(\pi_1 - \pi_y), \\ \theta_{1y} = \frac{1}{2}(\pi_1 - \pi_y). \end{cases}$$