Network Optimisation

Aim

To identify, formulate (using network models and LP models) and solve minimum cost flow and shortest path problems within the *Solver*/*Excel*. Identify shortest spanning tree problems, and solve them by the Prim algorithm.

Definition and Properties

A **graph** or **network** is an ordered pair $G=(V,A)$ where:

$V=\{1, 2, \dots, n\}$ is the set of **nodes** or **vertices**;

$A=\{a_1,a_2,\dots,a_m\}$ is the set of lines linking two nodes (a correspondence from *V* to *V*); $a_k=(i,j),\ i,j\in V$.

The directed elements of *A* are named **arcs**, and the non directed elements of *A* are **edges** (or **links** or **undirected arcs**).

An **arc/edge** $(i,j)\in A$ is **adjacent** to *i* and *j*.

$G=(V,A)$ is a **directed network** if all the elements of *A* are arcs.
$G=(V,A)$ is an **undirected network** if all the elements of *A* are edges.

If *A* has both arcs and edges it is a **mixed network**.

If $(i,j)\in A$ is an arc:

i is the **initial node** of (i,j); *j* is the **final node** of (i,j);
j is the **successor** of *i*; *i* is **predecessor** of *j*.

If $(i,j)\in A$ is an edge:

i and *j* are the **terminal nodes** or **extremities** of (i,j); *i* and *j* are **adjacent nodes**.

An **oriented path** from $x\in V$ to $y\in V$ is a sequence of distinct arcs, from *x* to *y*, represented by $C(x,y)=\{(x,x_1),(x_1,x_2),\dots,(x_{k-1},x_k),(x_k,y)\}$, where the final node of an arc is the initial node of the following one; and with only different nodes, with the possible exception of *x* and *y*. If $x=y$, the path is a ***circuit***.

If the direction of the links is not considered the concepts of **undirected path** (or **chain**) and of **cycle** (or **undirected circuit**) arise.

A **connected network** is a network such that any two nodes are linked by at least one chain.

A **tree** is a connected network without cycles. Given a network, $G=(V,A)$, a tree with the same set of nodes, *V*, and arcs/edges belonging to *A*, is a **spanning tree** of *G*.

The parameters defined, to each node and link, are problem dependent; for example, costs, distances, times, capacities; supplies, demands, weights, may be considered.

In network model three types of nodes may be considered:

- **source** (or a **supply node**, or an **origin**) is a node that is a net generator of flow, i.e., the flow out of the node exceeds the flow into it;
- **destination** (or a **demand node**, or a **sink**) is a node that is a net absorber of flow, i.e., the flow into the node exceeds the flow out of it;
- **transhipment node** (or an **intermediate node**) satisfies conservation of flow, so flow in equals flow out.

Minimum Cost Flow Problem (MCFP)

Given a directed and connected network, with at least a source and at least a sink, determine how to send the available supply through the network, to satisfy the demand, respecting arc capacities, at minimum cost.

Problem Data:

$G=(V,A)$ directed and connected network;

For each node, $i \in V$, b_i is its net flow, and:

if $b_i > 0$, i is a source with supply equal to b_i;

if $b_i < 0$, i is a destination with demand $-b_i$;

se $b_i = 0$, i is a transhipment node;

For each link $(i, j) \in A$:

c_{ij} is the cost for unit of flow through arc (i, j);

$u_{ij}>0$ is the **arc capacity**, i.e., the maximum value of flow that may traverse it.

Hypothesis: Arc capacities are compatible with supplies and demands; and the problem is balanced, i.e., the total supply equal the total demand: $\sum_{i \in V} b_i = 0$.

To formulate the MCFP as a network model, the network must be identified, $G=(V,A)$, as well as all the problem parameters (c_{ij}, u_{ij}, b_i) and the problem to solve (identifying the sources and the sink nodes).

Defining x_{ij} as the amount of flow through arc $(i, j) \in A$, the formulation of the MCFP with an LP model is:

$$Min\ Z = \sum_{(i,j) \in A} c_{ij} x_{ij}$$

$$\text{subject to: } \begin{cases} \sum_{j:(i,j)\in A} x_{ij} - \sum_{k:(k,i)\in A} x_{ki} = b_i & \forall i \in V \\ 0 \le x_{ij} \le u_{ij} \quad \forall (i, j) \in A \end{cases}.$$

In this model, the first summation represents the total flow out of node i, whereas the second summation represents the total flow into node i. So the difference is the net flow generated at this node (b_i). A set of restrictions is needed to guarantee that the flow on each arc does not exceed its capacity.

Properties of the MCFP

Property 1: The MCFP has, at least, a feasible solution.

Corollary: The MCFP has an optimal solution

Property 2: A MCFP where every b_i and u_{ij} are integer values, has, at least, one optimal solution with all variables assuming integer values.

Resolution of the MCFP with the *Solver/Excel*

Prototype example 1 – *Distribution Unlimited Co.* (HL[1], &3.4, pp. 58) – write down the problem to solve in an Excel sheet. Consider two tables: i) one with the network data for arcs: arc identifiers (B5-C11); flow values (D5:D11), initially equal to zero; arc capacities (F5:F11); unit costs (G5:G11); ii) a second table with all the data for the nodes: node identifiers (I5:I9); formulas representing the difference between the flow out of each node and the flow into that node (J5:J9); and the net flow at each node (L5:L9). The objective function formula must also be written (D13).

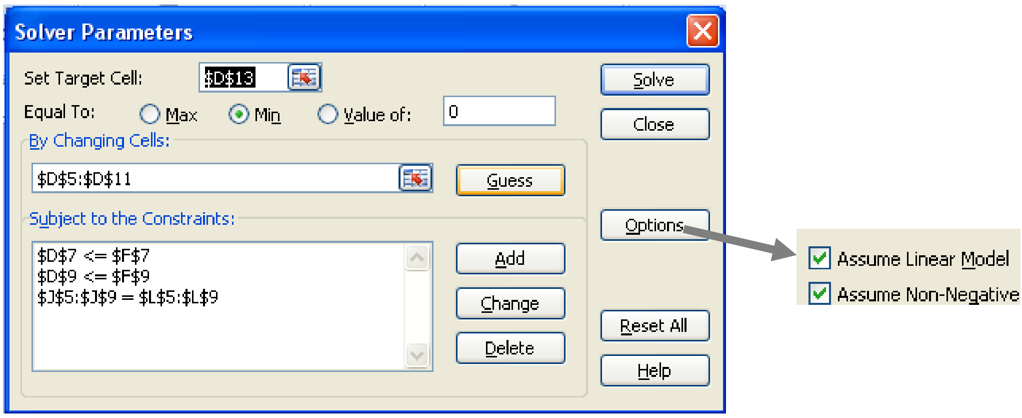
	A	B	C	D	E	F	G	H	I	J	K	L
1						data						
2	Distribution Unlimited Co.											
3							Unitary			Flow		Supply/
4		From	To	Flow		Capacity	Cost		Node	"in-out"		Demand
5		F1	W1	0	<=		$ 900		F1	0	=	50
6		F1	DC	0	<=		$ 400		F2	0	=	40
7		F1	F2	0	<=	10	$ 200		DC	0	=	0
8		F2	DC	0	<=		$ 300		W1	0	=	-30
9		DC	W2	0	<=	80	$ 100		W2	0	=	-60
10		W1	W2	0	<=		$ 300				Total	0
11		W2	W1	0	<=		$ 200					
12												
13			Total Cost	$ -								

	D
13	`=SUMPRODUCT(D5:D11;G5:G11)`

	I	J
5	F1	`=SUMIF($B$5:$B$11;"F1";$D$5:$D$11)-SUMIF($C$5:$C$11;"F1";$D$5:$D$11)`
6	F2	`=SUMIF($B$5:$B$11;"F2";$D$5:$D$11)-SUMIF($C$5:$C$11;"F2";$D$5:$D$11)`
7	DC	`=SUMIF($B$5:$B$11;"DC";$D$5:$D$11)-SUMIF($C$5:$C$11;"DC";$D$5:$D$11)`
8	W1	`=SUMIF($B$5:$B$11;"W1";$D$5:$D$11)-SUMIF($C$5:$C$11;"W1";$D$5:$D$11)`
9	W2	`=SUMIF($B$5:$B$11;"W2";$D$5:$D$11)-SUMIF($C$5:$C$11;"W2";$D$5:$D$11)`

[1] Hillier, Lieberman, "Introduction to Operations Research", 9th ed., McGraw-Hill 2010.

Solver – Identify the target cell (D13), the objective ("Minimize" or "Maximize") and the cells for the values of the variables (D5:D11). Define the functional constraints ("Add") of the model (J5:J9=L5:L9 and D5:D11<=F5:F11). The problem solution may then be found ("Solve").



Solution – Interpretation of the solution, using all the information provided by the *Excel* sheet or by the answer report.

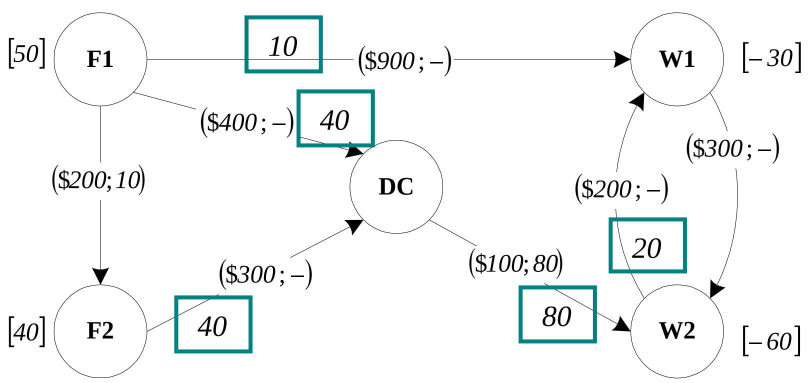
Distribution Unlimited Co.

From	To	Flow		Capacity	Unitary Cost		Node	Flow "in-out"		Supply/ Demand
F1	W1	10	<=		$ 900		F1	50	=	50
F1	DC	40	<=		$ 400		F2	40	=	40
F1	F2	0	<=	10	$ 200		DC	0	=	0
F2	DC	40	<=		$ 300		W1	-30	=	-30
DC	W2	80	<=	80	$ 100		W2	-60	=	-60
W1	W2	0	<=		$ 300				Total	0
W2	W1	20	<=		$ 200					

Total Cost $ 49.000

Answer: Send 10 units of flow from F1 directly to W1 and 40 to the distribution center DC. From F2 send 40 units to DC. To W2, send the 80 units of flow that were sent to DC. Then 20 of these units should be forward to W1, and the remaining 60 must stay at W2. The total cost is $ 49000.

The solution may also be presented using the network, where flows are depicted by: x_{ij}



Variants of the MCFP

(1) Total Supply > Total Demand: $\sum_{i \in V} b_i > 0$

The net flow at the sources is a maximum value that must be respected, and the constraints at the sources should be of type "≤".

(2) Total Supply < Total Demand: $\sum_{i \in V} b_i < 0$

The net flow at the sinks is a maximum value that must be respected, and the constraints at the sink nodes should be of type "≥".

Shortest-Path Problem (SPP)

Given a directed and connected network, with only one origin and only one destination node, determine a path with the minimum total distance from the origin to the destination.

Problem Data:

$G=(V,A)$ directed network; $s \in V$ origin; $t \in V$ destination

$c_{ij} > 0$ distance (cost, time, length) of arc $(i,j) \in A$.

To formulate the SPP within a network model, the network must be identified, $G=(V,A)$, as well as the origin, the destination and all the distances (c_{ij}) and the problem to solve.

Defining $x_{ij} = \begin{cases} 1 & \text{if } (i,j) \text{ is in the path} \\ 0 & \text{otherwise} \end{cases}$ $\forall (i,j) \in A$, the LP model associated to the SPP is:

$$Min\ Z = \sum_{(i,j) \in A} c_{ij} x_{ij}$$

$$\text{subject to: } \begin{cases} \sum_{j:(s,j) \in A} x_{sj} = 1 \\ \sum_{i:(i,t) \in A} x_{it} = 1 \\ \sum_{j:(i,j) \in A} x_{ij} - \sum_{k:(k,i) \in A} x_{ki} = 0 & \forall i \in V \setminus \{s,t\} \\ x_{ij} \in \{0,1\} \quad \forall (i,j) \in A \end{cases}.$$

In this formulation, the first two constraints are needed to ensure that the path starts at the origin, s, and ends at the destination, t. The third set of restrictions defines the remaining nodes as nodes that may be used to form the path.

It may be proved that the feasible range for variables values may be replaced by nonnegative constraints $x_{ij} \geq 0$. Additionally, if all the arcs in the network are assumed to have capacity one, the SPP is a particular MCFP with only one source with supply one, and one sink, with demand one.

Resolution of the SPP with the *Solver*/*Excel*

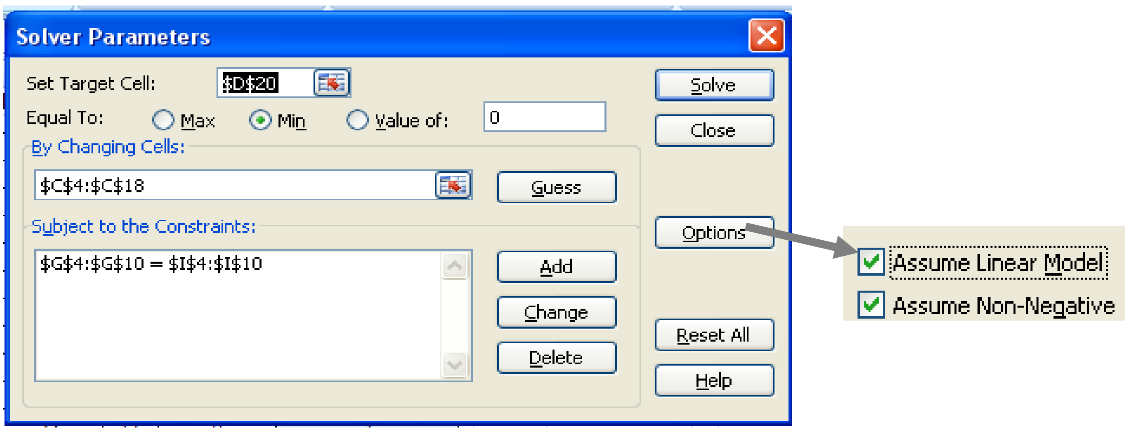
Prototype example 2 – *SEERVADA PARK*, 1st – write down the problem to solve in an Excel sheet. As in the MCFP, consider two tables: i) one with the network data for arcs: arc identifiers (A4:B18); solution (C4:C18), initially equal to zero; distances (D4:D18); ii) a second table with all the data for the nodes: node identifiers (F4:F10); formulas representing the difference between the flow out of each node and the flow into that node (G4:G10); and the value of one for the origin node (I4); -1 for the destination node (I10); and zero for remaining nodes (I5:I9). The objective function formula must also be written (D20).

	A	B	C	D	E	F	G	H	I
1	Seervada Park			data					
2									
3	From	To	Solution	Distance		Node	Total		Supply/Demand
4	O	A	0	2		O	0	=	1
5	O	B	0	5		A	0	=	0
6	O	C	0	4		B	0	=	0
7	A	B	0	2		C	0	=	0
8	A	D	0	7		D	0	=	0
9	B	A	0	2		E	0	=	0
10	B	C	0	1		T	0	=	-1
11	B	D	0	4					
12	B	E	0	3					
13	C	B	0	1					
14	C	E	0	4					
15	D	E	0	1					
16	D	T	0	5					
17	E	D	0	1					
18	E	T	0	7					
19									
20		Total Distance		0					

	D
20	=SUMPRODUCT(C4:C18;D4:D18)

	G
4	=SUMIF($A$4:$A$18;F4;$C$4:$C$18)-SUMIF($B$4:$C$18;F4;$C$4:$C$18)
5	=SUMIF($A$4:$A$18;F5;$C$4:$C$18)-SUMIF($B$4:$C$18;F5;$C$4:$C$18)
6	=SUMIF($A$4:$A$18;F6;$C$4:$C$18)-SUMIF($B$4:$C$18;F6;$C$4:$C$18)
7	=SUMIF($A$4:$A$18;F7;$C$4:$C$18)-SUMIF($B$4:$C$18;F7;$C$4:$C$18)
8	=SUMIF($A$4:$A$18;F8;$C$4:$C$18)-SUMIF($B$4:$C$18;F8;$C$4:$C$18)
9	=SUMIF($A$4:$A$18;F9;$C$4:$C$18)-SUMIF($B$4:$C$18;F9;$C$4:$C$18)
10	=SUMIF($A$4:$A$18;F10;$C$4:$C$18)-SUMIF($B$4:$C$18;F10;$C$4:$C$18)

Solver – Identify the target cell (D20), the objective ("Minimize" or "Maximize") and the cells for the values of the variables (C4:C18). Define the functional constraints ("Add") of the model (G4:G10=I4:I10).



Solution – Interpretation of the solution, using all the information provided by the *Excel* sheet or by the answer report.

	A	B	C	D	E	F	G	H	I
1	**Seervada Park**								
2									
3	From	To	Solution	Distance		Node	Total		Supply/Demand
4	O	A	1	2		O	1	=	1
5	O	B	0	5		A	0	=	0
6	O	C	0	4		B	0	=	0
7	A	B	1	2		C	0	=	0
8	A	D	0	7		D	0	=	0
9	B	A	0	2		E	0	=	0
10	B	C	0	1		T	-1	=	-1
11	B	D	1	4					
12	B	E	0	3					
13	C	B	0	1					
14	C	E	0	4					
15	D	E	0	1					
16	D	T	1	5					
17	E	D	0	1					
18	E	T	0	7					
19									
20		Total Distance		13					

Answer: The shortest path from *O* to *T* is (*O*,*A*,*B*,*D*,*T*), and the total distance of this path is *13*.

Minimum Spanning Tree Problem (MSTP)

Given an undirected and connected network, with lengths associated to the network edges, choose the set of edges that represent a spanning tree (a tree including all network nodes) with minimum total length.

Having an associated linear programming model with many constraints to guarantee the connectivity of the solution, this problem is only modeled within a network context. The solution may be easily found by Prim's algorithm.

Problem Data:

$G=(V,A)$ undirected network.

$c_{ij} > 0$ distance (cost, time, length) of edge $(i,j) \in A$.

To formulate the MSTP within a network model, the network must be identified, $G=(V,A)$, as well as all the distances (c_{ij}) and the problem to solve.

Property 3: A spanning tree of a network with *n* nodes has n nodes and $n-1$ edges (no cycles).

Prim Algorithm de Prim (1957) - determines the minimum spanning tree of *G*

Objective for iteration *k* - Select from the nodes not yet included in the tree the one that is closest to it. Link the node to the tree.

Repeat until all the nodes are in the tree.

Algorithm

0. Input: Undirected connected network with *n* nodes $G=(V,A)$;
Lengths of the edges;

1. Initialisation

Choose any node and the shortest edge incident on it;

Initialise the tree with the edge and respective nodes;

$k \leftarrow 2$;

2. Iteration *k*

If all the nodes are in the tree ($k = n$) **go to step 3.**

otherwise, select the shortest edge linking a node outside the tree to a node already in the tree;

Add the edge to the tree;

$k \leftarrow k+1$;

Go to step 2.

3. Draw the minimum spanning tree and determine the total length of the tree. **Stop.**

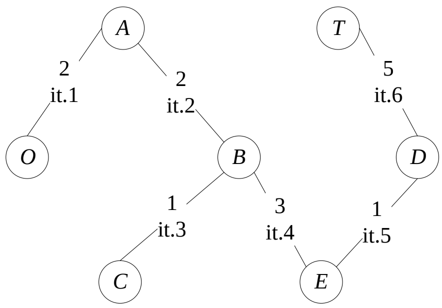
Resolution of the MSTP with the Prim Algorithm *Solver/Excel*

Prototype example – *SEERVADA PARK*, 2nd – the graph has $n=7$ nodes $\Rightarrow$ $n\text{-}1=6$ iterations.

Iteration	Node **in** the Tree	Closest and Adjacent Node ∉Tree	Edge Length	Edge to include in the Tree
1	*O*	*A*	2	(*O*,*A*)
2	*O*	*C*	4	(*A*,*B*)
	A	*B*	**2** ←	
3	*O*	*C*	4	(*B*,*C*)
	A	*D*	7	
	B	*C*	**1** ←	
4	*O*	–		(*B*,*E*)
	A	*D*	7	
	B	*E*	**3** ←	
	C	*E*	4	
5	*A*	*D*	7	(*E*,*D*)
	B	*D*	4	
	C	–		
	E	*D*	**1** ←	
6	*D*	*T*	**5** ←	(*D*,*T*)
	E	*T*	7	

Legend: "←" is the link with the shortest length.

Minimum Spanning Tree



The total MST length is 14