A 2-edge-connected spanning subgraph problem: Robert Carr, Ojas Parekh, Sandia Labs

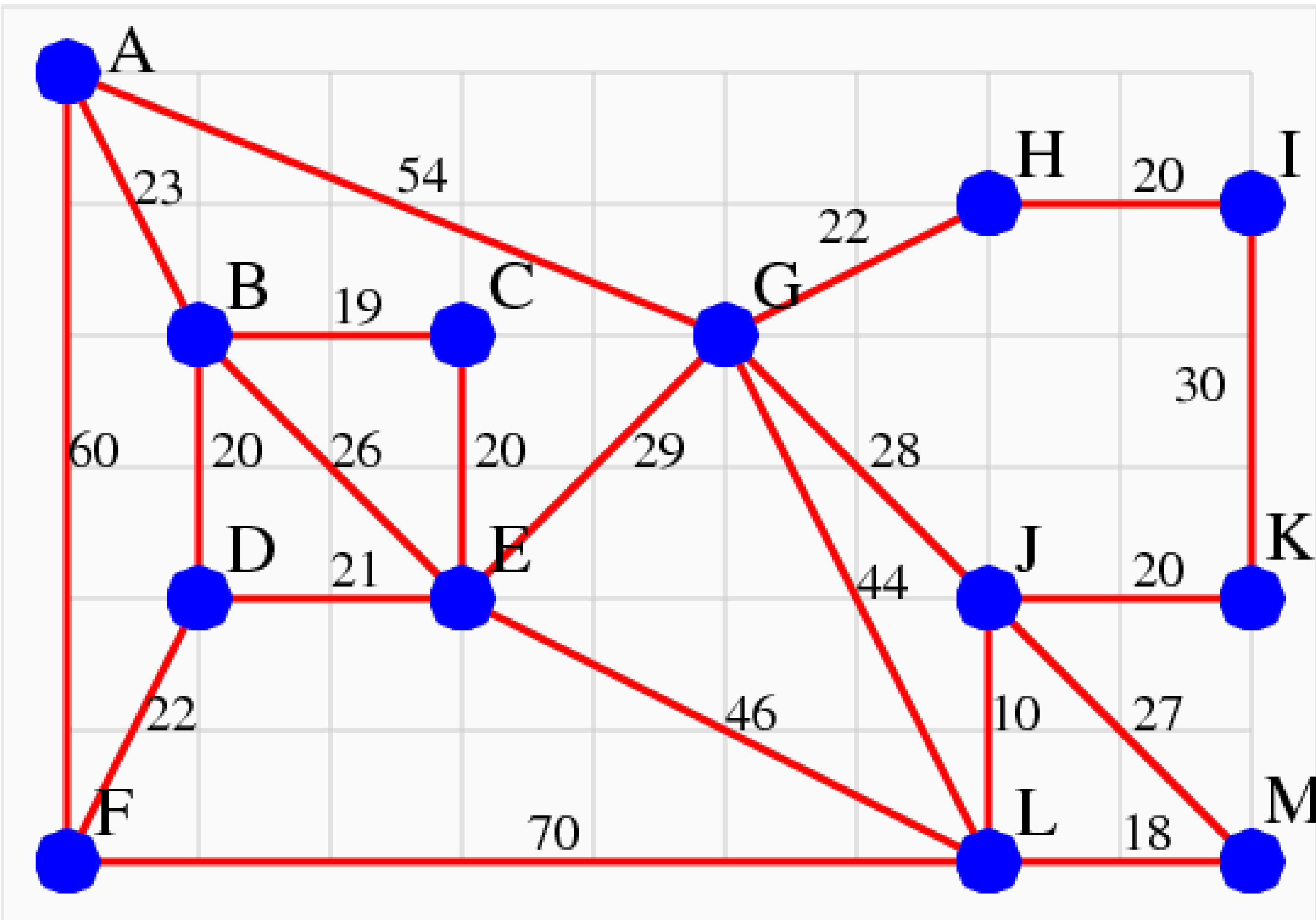


illustration from `http://people.sc.fsu.edu/~jburkardt/latex/asa_2011_graphs_homework/`

Integer programming solves the Labs difficult **discrete optimization problems**.

Water, Road **sensor** placement, subway, building **sensor** management, **Network** interdiction, **Scheduling** quantum EC, **Protein** structure, **Peptide** docking, **Meshing**, **Space-filling** curves, **Energy** systems, **Pantex** planning, **Vehicle routing**, **Conference** schedule.

Integer Program: Minimize a linear **cost function** subject to **linear inequality and integrality constraints.**

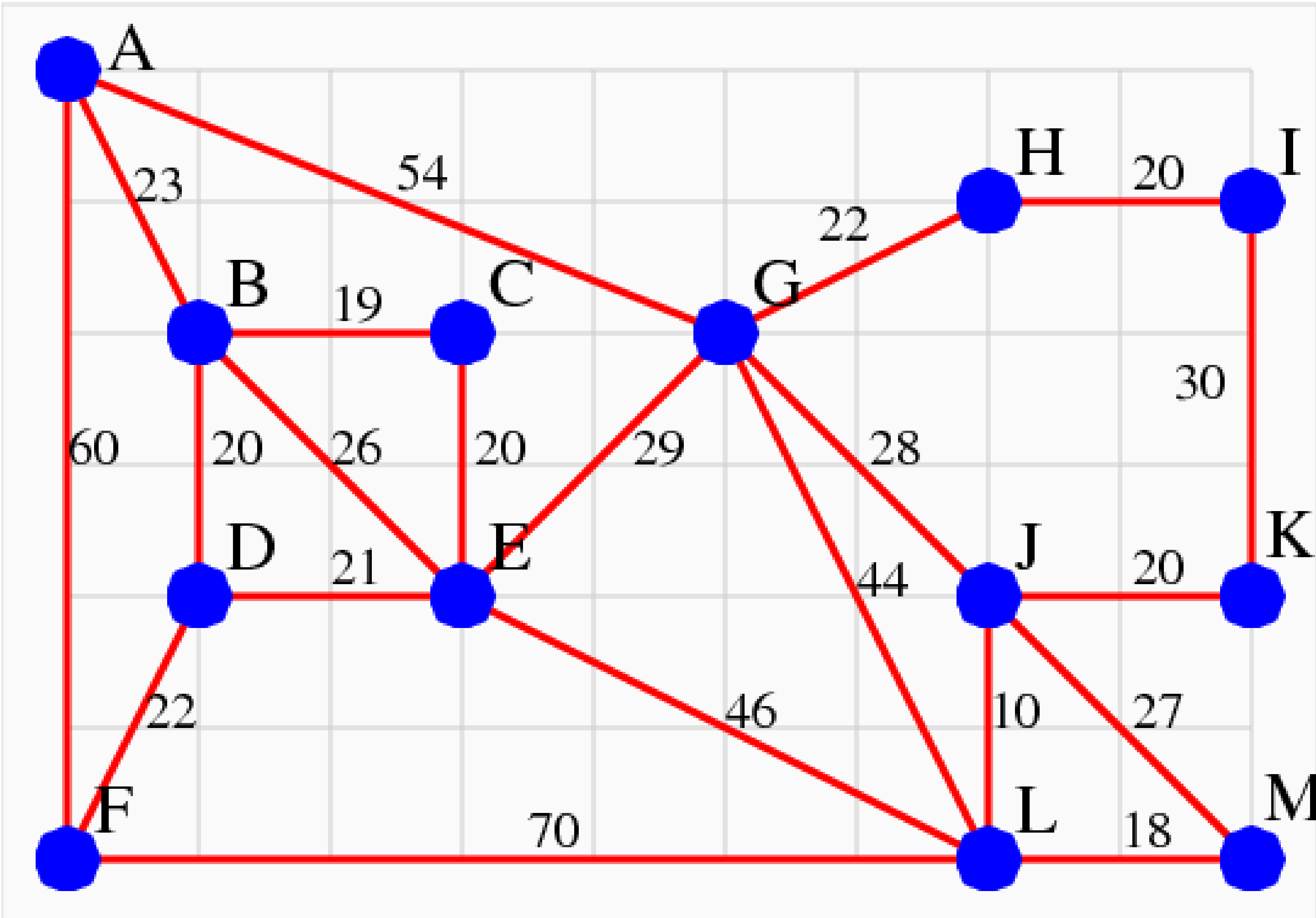
$$\begin{array}{rl} \text{minimize} & c \cdot x \\ \text{subject to} & \\ & A \cdot x \geq b \\ & x \in \mathbf{Z}^n. \end{array}$$

Recent focus on **creating formulations** of identified (tractible) problem structures.

Hardness arguments of modeling difficult structures.

Predict solution efficiency of a formulation.

Lockheed Martin Tech Refresh (Watson), **improved formulation** changed solution times **from days to minutes**.



Find minimum cost **2-edge connected** spanning subgraph.

Protect shipments **against single failure.**

Doubled edges are allowed and provided at a **discount**.

$$\begin{array}{rcl} \delta(S) & := & \{e = \{i,j\} \in E : |S \cap e| = 1\}, \\ E(S) & := & \{e = \{i,j\} \in E : |S \cap e| = 2\} \quad \forall S \subset V, \\ x(F) & := & \sum_{e \in F} x_e \qquad \forall F \subset E. \end{array}$$

A Classic 2-edge connected spanning subgraph problem.

$x_e \in \{0,1,2\}$ vars: Buy **edge** at **price** c_e.

$$\begin{array}{rll} \text{min} & c \cdot x & \\ \text{subj to} & & \\ & x(\delta(S)) \geq 2 & \forall S \subset V, \\ & 0 \leq x_e \leq 2 & \forall e \in E, \\ & x_e \in \mathbf{Z} & \forall e \in E. \end{array}$$

Drop integrality constraints to get LP relaxation.

Double-tree and Christofides heuristics

Select a minimum cost spanning tree $T = (V, E^T)$.

The edge incidence vector:
$\chi_e^T = 1$ iff $e \in E^T$ else $\chi_e^T = 0$.

Double each $e \in E^T$: $2\chi^T$ is the multi-edge incidence vector (has 2s) of our 2-edge connected graph.

Take the set T^{odd} of odd degree nodes of T.

A T^{odd}-join is a graph $M = (V, E^M)$ such that the degree of $v \in V$ is odd iff $v \in T^{odd}$.

Select a minimum cost T^{odd}-join M.

$\chi^T + \chi^M$ is the multi-edge incidence vector of a connected, Eulerian, hence 2-edge connected graph.

Double-tree approximation

Let x^* be optimal for LP relaxation.

$x^*(\delta(S)) \geq 2 \quad \forall S \subset V$ implies that x^* satisfies the partition inequalities for spanning trees.

Since x^* satisfies the partition inequalities, x^* dominates a convex combination of incidence vectors of spanning trees:
$x^* \geq \sum_i \lambda_i \chi^{T,i}$, $(\sum_i \lambda_i = 1)$.

Each tree can be doubled to get a 2-edge connected graph:
$2x^* \geq \sum_i \lambda_i (2\chi^{T,i})$.

By averaging argument, one 2-edge connected $2\chi^{T,i}$ costs at most that of $2x^*$.

Christofides approximation

x^* dominates a convex combination of tree vectors:
$x^* \geq \sum_i \lambda_i \chi^{T,i}$.

Let T_i be set of odd degree nodes for tree i.
$\frac{1}{2}x^*(\delta(S)) \geq 1 \quad \forall S \subset V$ implies that $\frac{1}{2}x^*$ satisfies the T_i-join inequalities for each i.

Since $\frac{1}{2}x^*$ satisfies the T_i-join inequalities, $\frac{1}{2}x^*$ dominates a convex combination of T_i-join vectors:
$\frac{1}{2}x^* \geq \sum_j \mu_{ij}\chi^{M,ij}$, $(\sum_j \mu_{ij} = 1)$.

For each i, j, $\chi^{T,i} + \chi^{M,ij}$ is 2-edge connected.

$\frac{3}{2}x^* \geq \sum_i \sum_j \lambda_i \mu_{ij}(\chi^{T,i} + \chi^{M,ij})$.

By averaging argument, one 2-edge connected $\chi^{T,i} + \chi^{M,ij}$ costs at most that of $\frac{3}{2}x^*$.

Our new 2-edge connected problem

$x_e \in \{0,1,2\}$ vars: Buy each **edge** at **price** c_e**,**
$y_e \in \{0,1\}$ Buy **doubled edge** at a **discount,**
$x \oplus y \in \mathbf{R}^E \times \mathbf{R}^E$, $c \oplus c' \in \mathbf{R}^E \times \mathbf{R}^E (c'_e \leq 2c_e)$.

$$
\begin{array}{rll}
\min & (c \oplus c') \cdot (x \oplus y) & \\
\text{subj to} & & \\
& x(\delta(S)) + 2y(\delta(S)) \geq 2 & \forall S \subset V \\
& 0 \leq x_e \leq 2, 0 \leq y_e \leq 1 & \forall e \in E \\
& x_e, y_e \in \mathbf{Z}. &
\end{array}
$$

Drop integrality constraints to get LP relaxation.

Integrality gap of 2: $y_e^* = \frac{1}{2}$ for edges of a Hamilton (n edge) cycle. But optimal integer solution is $y_e^{opt} = 1$ for all but one edge of cycle.

A better LP relaxation

Idea: $x_e + y_e$ dominates a spanning tree vector, denoted by z_e. That is, $x + y$ has enough mass ($n-1$ edges) to contain a spanning tree, and the tree (z) has acyclic structure.

$$\begin{aligned} \textbf{Add } x_e + y_e \geq z_e \; z(E(V)) &= n-1, \\ \forall S \subset V \;\; z(E(S)) &\leq |S|-1. \end{aligned}$$

Now $y_e^* = \frac{1}{2}$ on a Hamilton cycle no longer feasible.

Worst **gap** seen is **now** $\frac{3}{2}$ when horizontal edges of a square have $x_e^* = 1$ and a cost of 1 and vertical edges of that square have $y_e^* = \frac{1}{2}$ and a cost of 2.

Let $x^*\oplus y^*$ be an optimal extreme point solution to our LP.

To keep things simple, assume $x_e^* = 0$ or $y_e^* = 0$ for each $e \in E$.

$x_e^* + y_e^* \geq z_e^*$ and the spanning tree constaints on z^* imply that $x^* + y^*$ dominates a convex combination of incidence vectors $\chi^{T,i}$ of spanning trees $x^* + y^* \geq \sum_i \lambda_i \chi^{T,i}$.

For each spanning tree, break up its set of edges into a set of x-edges and a set of y-edges. Then its incidence vector $\chi^{T,i}$ is broken up into incidence vectors $\chi^{T,x,i}$ and $\chi^{T,y,i}$.

So, $\chi^{T,i} = \chi^{T,x,i} + \chi^{T,y,i}$. Thus,
$x^* + y^* \geq \sum_i \lambda_i (\chi^{T,x,i} + \chi^{T,y,i})$.

Finally, $x^* \oplus y^* \geq \sum_i \lambda_i (\chi^{T,x,i} \oplus \chi^{T,y,i})$.

Double-tree approximation

Let $x^* \oplus y^*$ be optimal for LP relaxation.

$x^*_e + y^*_e \geq z^*_e$ and the constaints on z^* imply that $x^* \oplus y^*$ dominates a convex combination of incidence vectors of spanning trees in $x \oplus y$ variable space:
$x^* \oplus y^* \geq \sum_i \lambda_i (\chi^{T,x,i} \oplus \chi^{T,y,i})$.

The x-part of each tree can be doubled to get a 2-edge connected graph:
$2x^* \oplus y^* \geq \sum_i \lambda_i (2\chi^{T,x,i} \oplus \chi^{T,y,i})$.

By averaging argument, one 2-edge connected $2\chi^{T,x,i} \oplus \chi^{T,y,i}$ costs at most that of $2x^* \oplus y^*$.

Christofides approximation

$x^* \oplus y^*$ dominates a convex combination of tree vectors: $x^* \oplus y^* \geq \sum_i \lambda_i(\chi^{T,x,i} \oplus \chi^{T,y,i})$.

Let T_i be set of odd degree nodes for tree i. $\frac{1}{2}x^*(\delta(S)) + y^*(\delta(S)) \geq 1 \quad \forall S \subset V$ implies that $\frac{1}{2}x^* + y^*$ satisfies T_i-join inequalities for each i.

Since $\frac{1}{2}x^* + y^*$ satisfies the T_i-join inequalities, $\frac{1}{2}x^* \oplus y^*$ dominates a convex combination of T_i-join vectors:
$\frac{1}{2}x^* \oplus y^* \geq \sum_j \mu_{ij}(\chi^{M,x,ij} \oplus \chi^{M,y,ij})$.

For each i, j, $(\chi^{T,x,i} + \chi^{M,x,ij}) \oplus (\chi^{T,y,i} + \chi^{M,y,ij})$ is 2-edge connected.

$\frac{3}{2}x^* \oplus 2y^* \geq \sum_i \sum_j \lambda_i \mu_{ij}(\chi^{T,x,i} + \chi^{M,x,ij}) \oplus (\chi^{T,y,i} + \chi^{M,y,ij})$.

By averaging argument, one 2-edge connected $(\chi^{T,x,i} + \chi^{M,x,ij}) \oplus (\chi^{T,y,i} + \chi^{M,y,ij})$ costs at most that of $\frac{3}{2}x^* \oplus 2y^*$.

The 5/3 approximation

From the Double-tree approximation, $2x^* \oplus y^*$ dominates a convex combination of 2-edge connected graphs G_i^1.

From the Christofides approximation, $\frac{3}{2}x^* \oplus 2y^*$ dominates a convex combination of 2-edge connected graphs G_i^2.

We can combine these as follows:

$$\begin{array}{rrll} & \frac{1}{3} & (\; 2x^* \oplus y^* & \geq \; \sum_i \lambda_i G_i^1) \\ + & \frac{2}{3} & (\; \frac{3}{2}x^* \oplus 2y^* & \geq \; \sum_i \lambda_i G_i^2) \\ \hline & & \frac{5}{3}x^* \oplus \frac{5}{3}y^* & \geq \; \sum_i \lambda_i G_i. \end{array}$$

The $\frac{5}{3}$ approximation and integrality gap follows since one of the G_is cost at most that of $\frac{5}{3}(x^* \oplus y^*)$ by our averaging argument.