Introduction to Operations Research

Matthew Galati

magh@lehigh.edu

Department of Industrial and Systems Engineering, Lehigh University

Service Parts Solutions, IBM Corporation

Outline

- What is Operations Research?
- Optimization
- Problems and Applications
- Personal Examples

What is Operations Reseach?

- Operations Research (OR) started just before World War II in Britain with the establishment of teams of scientists to study the strategic and tactical problems involved in military operations. The objective was to find the most effective utilization of limited military resources by the use of quantitative techniques.
- Following the war, numerous peacetime applications emerged, leading to the use of OR and management science in many industries and occupations.
- Occupational Outlook Handbook - http://stats.bls.gov/oco/ocos044.htm.

Definitions

- Model: A schematic description of a system,theory, or phenomenon that accounts for its known or inferred properties and may be used for further study of its characteristics.
- System: A functionally related group of elements, especially:
  - The human body regarded as a functional physiological unit.
  - An organism as a whole, especially with regard to its vital processes or functions.
  - A group of physiologically or anatomically complementary organs or parts: the nervous system; the skeletal system.
  - A group of interacting mechanical or electrical components.
  - A network of structures and channels, as for communication, travel, or distribution.
  - A network of related computer software, hardware, and data transmission devices.

Definitions

- Operations Research (OR) is the study of mathematical models for complex organizational systems.
- Optimization is a branch of OR which uses mathematical techniques such as linear and nonlinear programming to derive values for system variables that will optimize performance.

Models

- Linear Programming
  - Typically, a single objective function, representing either a profit to be maximized or a cost to be minimized, and a set of constraints that circumscribe the decision variables. The objective function and constraints all are linear functions of the decision variables.
  - Software has been developed that is capable of solving problems containing millions of variables and tens of thousands of constraints.
- Network Flow Programming
  - A special case of the more general linear program. Includes such problems as the transportation problem, the assignment problem, the shortest path problem, the maximum flow problem, and the minimum cost flow problem.
  - Very efficient algorithms exist which are many times more efficient than linear programming in the utilization of computer time and space resources.

Models

- Integer Programming
  - Some of the variables are required to take on discrete values.
  - NP-hard: Most problems of practical size are very difficult or impossible to solve.
- Nonlinear Programming
  - The objective and/or any constraint is nonlinear.
  - In general, much more difficult to solve than linear.
  - Most (if not all) real world applications require a nonlinear model. In order to be make the problems tractable, we often approximate using linear functions.

Models

- Dynamic Programming
  - A DP model describes a process in terms of states, decisions, transitions and returns. The process begins in some initial state where a decision is made. The decision causes a transition to a new state. Based on the starting state, ending state and decision a return is realized.
  - The process continues through a sequence of states until finally a final state is reached. The problem is to find the sequence that maximizes the total return.
  - Objectives with very general functional forms may be handled and a global optimal solution is always obtained.
  - "Curse of dimensionality" - the number of states grows exponentially with the number of dimensions of the problem.

Models

- Stochastic Processes
  - In many practical situations the attributes of a system randomly change over time.
  - Examples include the number of customers in a checkout line, congestion on a highway, the number of items in a warehouse, and the price of a financial security to name a few.
  - The model is described in part by enumerating the states in which the system can be found. The state is like a snapshot of the system at a point in time that describes the attributes of the system. Events occur that change the state of the system.
  - Consider an Automated Teller Machine (ATM) system. The state is the number of customers at or waiting for the machine. Time is the linear measure through which the system moves. Events are arrivals and departures.

Models

- Markov Chains
  - A stochastic process that can be observed at regular intervals such as every day or every week can be described by a matrix which gives the probabilities of moving to each state from every other state in one time interval.
  - Assuming this matrix is unchanging with time, the process is called a Markov Chain. Computational techniques are available to compute a variety of system measures that can be used to analyze and evaluate a Markov Chain model.
- Markov Processes
  - A continuous time stochastic process in which the duration of all state changing activities are exponentially distributed. Time is a continous parameter.
  - The process is entirely described by a matrix showing the rate of transition from each state to every other state. The rates are the parameters of the associated exponential distributions. The analytical results are very similar to those of a Markov Chain.

Models

- Simulation
  - It is often difficult to obtain a closed form expression for the behavior of a stochastic system.
  - Simulation is a very general technique for estimating statistical measures of complex systems.
  - A system is modeled as if the random variables were known. Then values for the variables are drawn randomly from their known probability distributions. Each replication gives one observation of the system response. By simulating a system in this fashion for many replications and recording the responses, one can compute statistics concerning the results. The statistics are used for evaluation and design.

Models

- Time Series and Forecasting
  - A time series is a sequence of observations of a periodic random variable.
  - Typically serve as input to OR decision models.
  - Example - inventory model requires estimates of future demands.
  - Example - a course scheduling and staffing model for the university department requires estimates of future student inflow.
  - Example - A model for providing warnings to the population in a river basin requires estimates of river flows for the immediate future.

Models

- Inventory Theory
  - Inventories are materials stored, waiting for processing, or experiencing processing.
  - When and how much raw material should be ordered?
  - When should a production order should be released to the plant?
  - What level of safety stock should be maintained at a retail outlet?
  - How is in-process inventory maintained in a production process?
- Reliability Theory
  - Attempts to assign numbers to the propensity of systems to fail.
  - Estimating reliability is essentially a problem in probability modeling.
  - Extremely important in the telecommunications and networking industry.

Mathematical Programming

- A mathematical model consists of:
  - Decision Variables, Constraints, Objective Function, Parameters and Data
- The general form of a math programming model is:

$$
\begin{array}{lll}
\mathit{min} \text{ or } \mathit{max} & f(x_1, \ldots, x_n) & \\
s.t. & g_i(x_1, \ldots, x_n) \left\{ \begin{array}{c} \leq \\ = \\ \geq \end{array} \right\} b_i \\
& x \in X &
\end{array}
$$

- Linear program (LP): all functions f and g_i are linear and X is continuous.
- Integer program (IP): X is discrete.

Mathematical Programming

- A solution is an assignment of values to variables.
- A feasible solution is an assignment of values to variables such that all the constraints are satisfied.
- The objective function value of a solution is obtained by evaluating the objective function at the given solution.
- An optimal solution (assuming minimization) is one whose objective function value is less than or equal to that of all other feasible solutions.

Outline

- What is Operations Research?
- Optimization
- Problems and Applications
- Personal Examples

LP Example: Two Crude Petroleum

- Two Crude Petroleum distills crude from two sources:
  - Saudi Arabia,Venezuela
- They have three main products:
  - Gasoline,Jet Fuel,Lubricants
- Yields

	Gasoline	Jet Fuel	Lubricants
Saudi Arabia	0.3 barrels	0.4 barrels	0.2 barrels
Venezuela	0.4 barrels	0.2 barrels	0.3 barrels

- Availability and cost

	Availability	Cost
Saudi Arabia	9000 barrels	$20/barrel
Venezuela	6000 barrels	$15/barrel

LP Example: Two Crude Petroleum

- Production Requirements (per day)

Gasoline	Jet fuel	Lubricants
2000 barrels	1500 barrels	500 barrels

- Objective: Minimize production cost.
- This yields the following LP formulation:

$$
\begin{aligned}
\min \quad & 20x_1 + 15x_2 \\
\text{s.t.} \quad & 0.3x_1 + 0.4x_2 \geq 2.0 \\
& 0.4x_1 + 0.2x_2 \geq 1.5 \\
& 0.2x_1 + 0.3x_2 \geq 0.5 \\
& 0 \leq x_1 \leq 9 \\
& 0 \leq x_2 \leq 6
\end{aligned}
$$

- Graphical Demo

Simplex Method

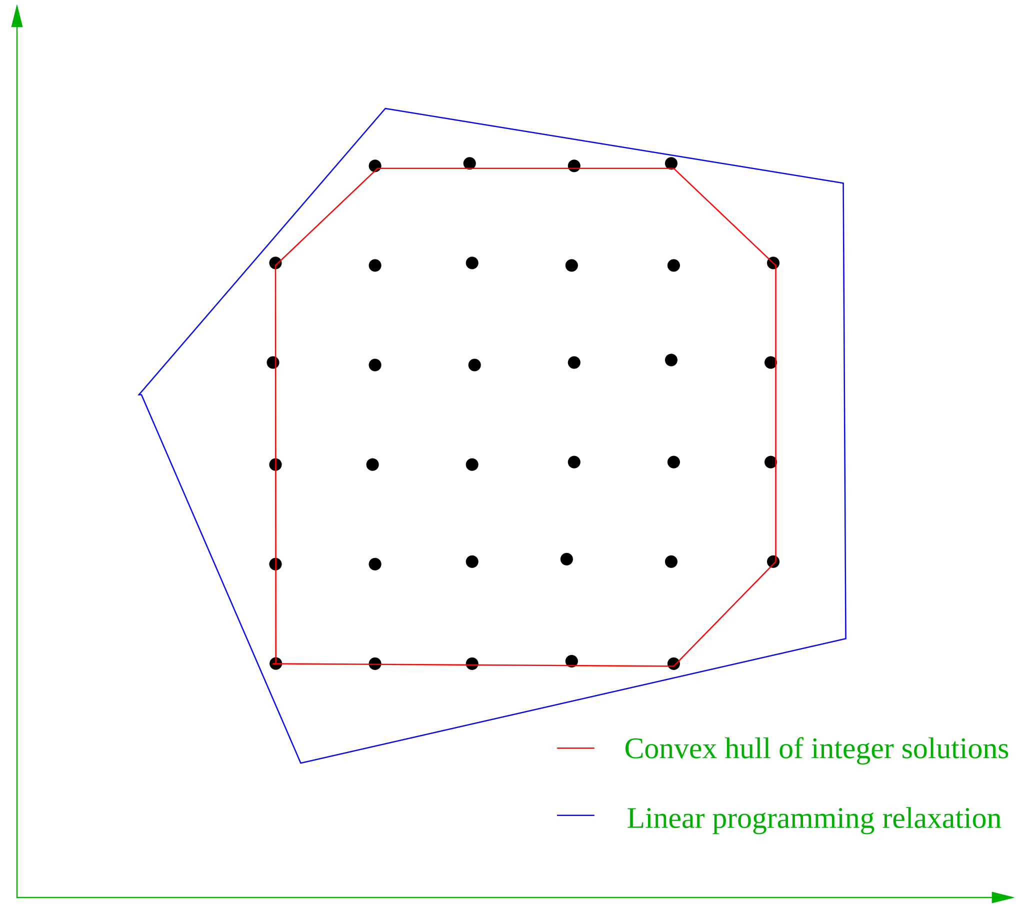
- The simplex method generates a sequence of feasible iterates by repeatedly moving from one vertex of the feasible set to an adjacent vertex with a lower value of the objective function . When it is not possible to find an adjoining vertex with a lower value, the current vertex must be optimal, and termination occurs.
- Exponential (worst case) run time; in practice, runs very fast.

Optimization - Linear Programming

- Simplex (Dantzig 1947)
- Ellipsoid (Khachian 1979) - the "first" polynomial-time algorithm
- Interior Point - the "first" practical polynomial-time algorithm
  - Projective Method (Karmarkar 1984)
  - Affine Method (Dikin 1967)
  - Logarithmic Barrier Method (Frisch 1955, Fiacco 1968, Gukk et. al 1986)

Integer Programming

- Let X be a discrete set (integers).



Integer Programming

- Why not just solve the LP and round?

$$
\begin{aligned}
\max \quad & 1.00x_1 + 0.64x_2 \\
\text{s.t.} \quad & 50x_1 + 31x_2 \leq 250 \\
& 3x_1 - 2x_2 \geq -4 \\
& x_1 \geq 0 \\
& x_2 \geq 0
\end{aligned}
$$

LP-based Branch and Bound

- Consider problem P:

$$\begin{aligned} &min\ c^T x \\ &s.t. \quad Ax \leq b \\ &\qquad\quad x_i \in \mathbf{Z}\ \forall\, i \in I \end{aligned}$$

  where $(A, b) \in \mathbf{R}^{m \times n+1}, c \in \mathbf{R}^n$.

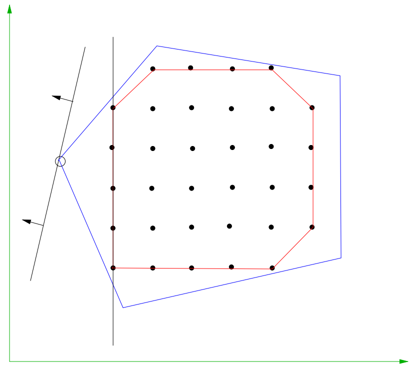
- Let $\mathcal{P} = conv\{x \in \mathbf{R}^n : Ax \leq b, x_i \in \mathbf{Z}\ \forall\, i \in I\}$.
- Basic Algorithmic Approach
  - Use LP relaxations to produce lower bounds.
  - Branch using hyperplanes.
- Basic Algorithmic Elements
  - A method for producing and tightening the LP relaxations.
  - A method for branching.

Branch, Cut, and Price

- Weyl-Minkowski
  - $\exists(\bar{A}, \bar{b}) \in \mathbf{R}^{\bar{m} \times n+1}$ s.t. $\mathcal{P} = \{x \in \mathbf{R}^n : \bar{A}x \leq \bar{b}\}$
  - We want the solution to $min\{c^T x : \bar{A}x \leq \bar{b}\}$.
  - Solving this LP isn't practical (or necessary).
- BCP Approach
  - Form LP relaxations using submatrices of $\bar{A}$.
  - The submatrices are defined by sets $\mathcal{V} \subseteq [1..n]$ and $\mathcal{C} \subseteq [1..\bar{m}]$.
  - *Forming/managing these relaxations efficiently is one of the primary challenge of BCP*.

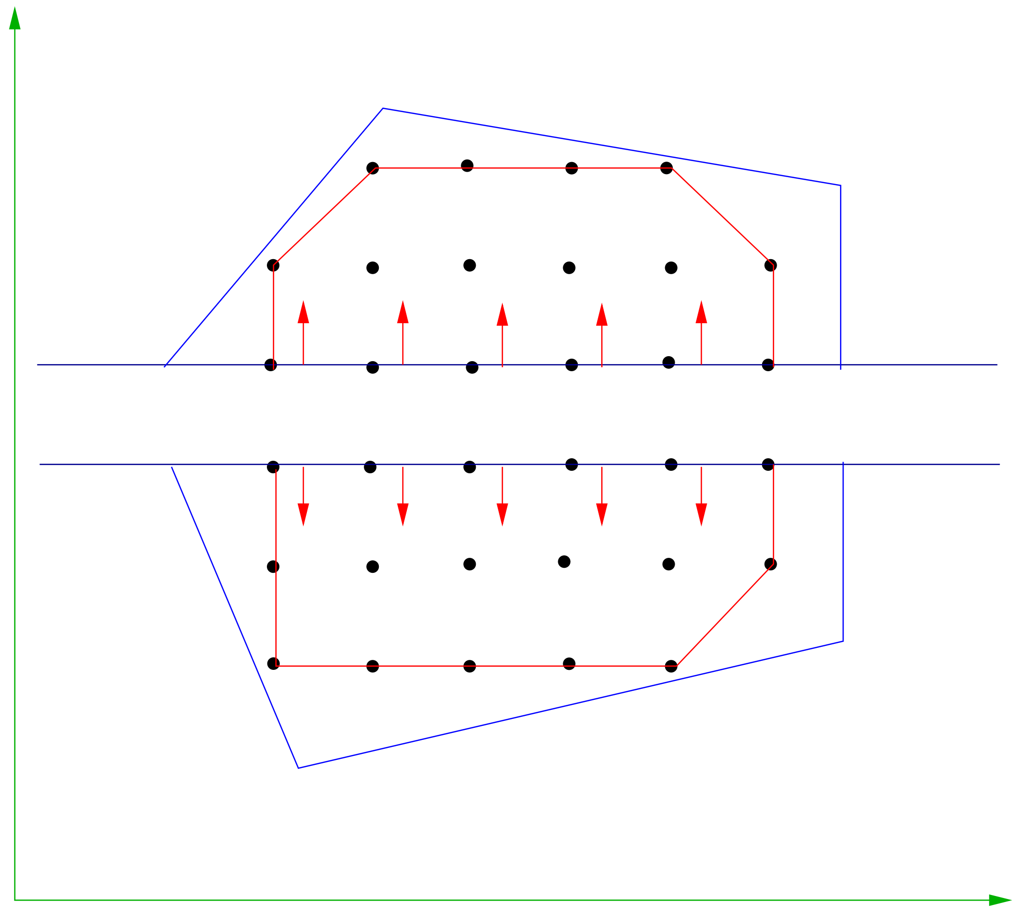
Cutting Plane Method

- Basic cutting plane algorithm
  - Relax the integrality constraints.
  - Solve the relaxation. Infeasible $\Rightarrow$ STOP.
  - If $\hat{x}$ integral $\Rightarrow$ STOP.
  - Separate $\hat{x}$ from $\mathcal{P}$.
  - No cutting planes $\Rightarrow$ algorithm fails.
- The key is good separation algorithms.



Branch and Cut Methods

- If the cutting plane approach fails, then we divide and conquer (branch).



The Challenge of BCP

- The efficiency of BCP depends heavily on the size (number of rows and columns) and tightness of the LP relaxations.
- Tradeoff
  - Small LP relaxations $\Rightarrow$ faster LP solution.
  - Big LP relaxations $\Rightarrow$ better bounds.
- The goal is to keep relaxations small while not sacrificing bound quality.
- We must be able to easily move constraints and variables in and out of the *active* set.
- This means dynamic generation and deletion.

Outline

- What is Operations Research?
- Optimization
- Problems and Applications
- Personal Examples

Common Problems in OR

- Set Covering, Packing, Partitioning
  - Let S be a set of objects and Ω a set of subsets of S. Let $a_{ij} = 1$, if $i \in \Omega_j$ and define variable $x_j = 1$, if the j^{th} member of Ω is used.

$$\begin{aligned}
\text{Set Covering} \quad &: \quad \min_{x\in\{0,1\}} \{\sum_{i\in\Omega} c_i x_i : \sum_j a_{ij} x_j \geq 1, \forall i \in S\} \\
\text{Set Packing} \quad &: \quad \max_{x\in\{0,1\}} \{\sum_{i\in\Omega} c_i x_i : \sum_j a_{ij} x_j \leq 1, \forall i \in S\} \\
\text{Set Partitioning} \quad &: \quad \min_{x\in\{0,1\}} \text{ or } \max_{x\in\{0,1\}} \{\sum_{i\in\Omega} c_i x_i : \sum_j a_{ij} x_j = 1, \forall i \in S\}
\end{aligned}$$

  - Air Crew Scheduling (Covering): Consider S to be a set of "legs" that the airline has to cover and the members of Ω are possible rotations involving particular combinations of flights.

Common Problems in OR

- Quadratic Assignment Problem
  - Given two sets of objects S and T where $|S| = |T|$, each member of S must be assigned to exactly one member of T, and vice versa. Costs are incurred for an assignment of $i \in S$ to $j \in T$ *and* $k \in S$ to $l \in T$.

$$\min_{x \in \{0,1\}} \{ \sum_{\substack{i,k \in S \\ k > i}} \sum_{j,l \in T} t_{ik} d_{jl} x_{ij} x_{kl} : \sum_{j \in T} x_{ij} = 1, \forall i \in S, \sum_{i \in S} x_{ij} = 1, \forall j \in T \}$$

  - Facility Location: Consider S to be a set of n factories and T to be a set of n cities. Locate one factory in each city and minimize the total communication cost between factories. Interpret t_{ik} as the frequency of communication between factories i and k and d_{jl} as the cost per unit of communication between cities j and l.
  - Circuit Design: Consider S to be a set of n electronic modules and T to be a set of n predetermined positions on a backplate. The modules have to be connected to each other by a series of wires. Interpret t_{ik} as the number of wires which must connect module i to the module k and d_{jl} as the distance between position j and position l on the backplate.

Common Problems in OR

- Routing problems. Finding a path or cycle in a network. An easy routing problem is the shortest path; a hard one is the travelling salesman problem. One prevalent class, with many variations, is vehicle routing.
    - Shortest path. In a graph or network, this is a path from one node to another whose total cost is the least among all such paths. The "cost" is usually the sum of the arc costs, but it could be another function (e.g., the product for a reliability problem, or max for a fuzzy measure of risk).
    - Vehicle routing problem (VRP). Find optimal delivery routes from one or more depots to a set of geographically scattered points (e.g., population centers). In its most complex form, the VRP is a generalization of the TSP, as it can include additional time and capacity constraints, precedence constraints, plus more.

Common Problems in OR

- Production Scheduling Problem
  - To determine levels of production over time. Constraints include demand requirements (possibly with backordering), capacity limits (including warehouse space for inventory), and resource limits. Define

    x_t = level of production in period t (before demand);
    y_t = level of inventory at the end of period t;
    U_t = production capacity in period t;
    W_t = warehouse capacity in period t;
    h_t = holding cost (per unit of inventory);
    p_t = production cost (per unit of production);
    D_t = demand at the end of period t.

$$\min_{x,y}\{\sum px + hy : y_{t+1} = y_t + x_t - D_t, \forall t, 0 \leq x \leq U, 0 \leq y \leq W\}$$

Common Problems in OR

- Portfolio Selection Problem.
  - The objective is to minimize the variance on returns. Let x_j be the percent of capital invested in the j^{th} opportunity (e.g., stock or bond), so x must satisfy $x \geq 0$ and $\sum_j x_j = 1$. Let v_j be the expected return per unit of investment in the j^{th} opportunity, so that vx is the sum total rate of return per unit of capital invested. It is required to have a lower limit on this rate: $vx \geq b$ (where b is between $\min v_j$ and $\max v_j$). Subject to this rate of return constraint, the objective is the quadratic form, $x^T Qx$, where Q is the variance-covariance matrix associated with the investments (i.e., if the actual return rate is V_j, then $Q_i j = E[(V_i - v_i)(V_j - v_j)]$
- Chemical Equilibrium Problem.
  - The problem is to $min_x cx + \sum_j x_j \log x_j : x > 0, \sum_j x_j = M, Ax = b$, where the objective is the Gibbs free energy function for x_j = number of moles of species j, b_i = atomic weight of atom of type i, and a_{ij} = number of atoms of type i in one mole of species j. The equation, $Ax = b$, is mass balance.

Outline

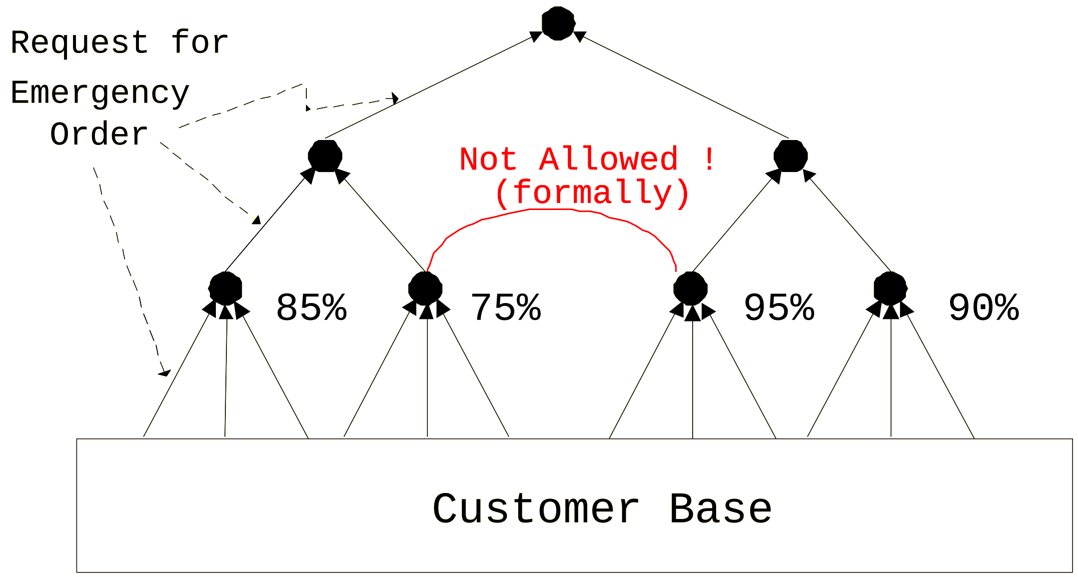
- What is Operations Research?
- Optimization
- Problems and Applications
- Personal Examples

TRL Trans Refrigerated Lines

- Dynamic Load Assignment Problem (Trucks to Loads)
- Parametric Multi-Criteria Objective
  - Early/Late Delivery, Early/Late Pickup
  - Reduction in Empty Travel, Vehicle Maintenance Schedule
  - Driver Vacation, Driver Hours Balancing (Union Labor)
- GUI Interface for Load Planners (Real Time Dispatch)
- Lesson in Industry - Systems Development Paradox
  - In order to create the system correctly we needed the knowledge base of the load planners.
  - Load planners are well aware that system would serve as their replacement.
  - Upper management was willing to fund the development due to the potential cost reduction.

IBM SPS - Industry Standard

- Rigid Structured Echelon Stocking/Order Routing Model
  - Stocking/order routing predetermined via pass-up schemes.
- Service Delivery Performance Based on Parts Availability
  - No time/distance component linked to "Point-of-Demand"
- Stock Network Locations Independently



IBM SPS - Neighborhood

- Regional Contained Stocking/Ordering
  - Neighborhoods are created around the time-based requirements of customers.
  - All locations may serve as emergency stocking points.
  - Stocking/order routing based on "pooled risk".
  - Pass-up converted to pass-along

IBM SPS - Neighborhood

$$P(\text{Part is required from location i}) = \beta_i \prod_{j \in N_1(i)} (1-\beta_j)$$

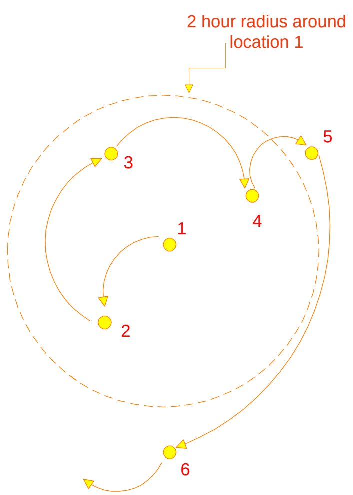
$$P(\text{Part is available in the neighborhood}) = 0.99994$$

IBM SPS - Neighborhood

- For a given stocking level at each location and demand made at location 1.

$$P(\text{Part is required from location } i) = \beta_i \prod_{j \in N_1(i)} (1 - \beta_j)$$

$$P(\text{Part is available in the neighborhood}) = 0.99994$$



Location	$P(NoParts) = 1 - \beta$	$P(PartRequired)$
1	0.10	0.90
2	0.08	0.092
3	0.09	0.0073
4	0.11	0.0064
5	0.10	0.0007
6	0.10	0.000007

IBM SPS - Neighborhood

- Mixed Integer Programming Model
- Variables
  - Demand flow rate from facilities to customers (continuous).
  - Stocking levels at facilities, distribution centers, and hubs (integral).
- Constraints
  - Satisfy demand to all customers (flow balance).
  - Probabilistically meet all service contract requirements (parts procurement).
- Size (United States)
  - Customers: over 2,000,000 individual across 40,000 zipcodes - aggregated to 4,000 regions. Facilities: inside 110, outside 3000. Distribution Centers: 3
  - Variables in MIP: 400,000 (Continuous), 122,000 (Integer)
  - Constraints in MIP: 100,000 (Original), 660,00 (Additional Valid Cuts)

IBM SPS - Neighborhood

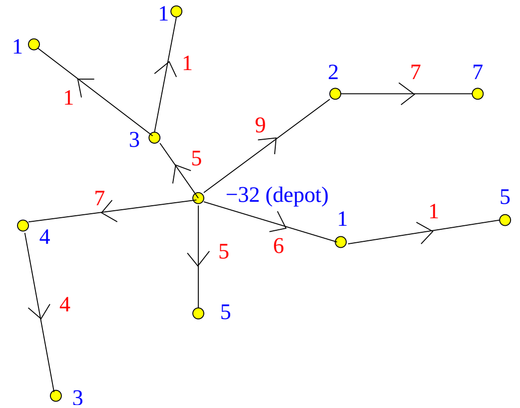
- Inventory Implications
  - Lowers inventory investment for given PPT criteria when compared to hierarchical structure
  - Locations not stocked independently of other locations in various vitality class neighborhoods
  - Churn : Regional view of available inventory lessens need to "instantly" reestablish stocking levels
  - Returns : Real-time inventory visibility
  - pickup/dropoff parts using scheduled runs
  - no need to return parts to original location due to regional visibility

IBM SPS - Neighborhood

- Service Implications
  - Customers are assigned to closest primary location
  - If no stock at primary location, pass-up strategy converts to pass-along
  - Demand satisfied directly from least cost/least critical secondary location contained in PPT class "neighborhood"
  - Regional orientation for providing customer service through "pooled risk"

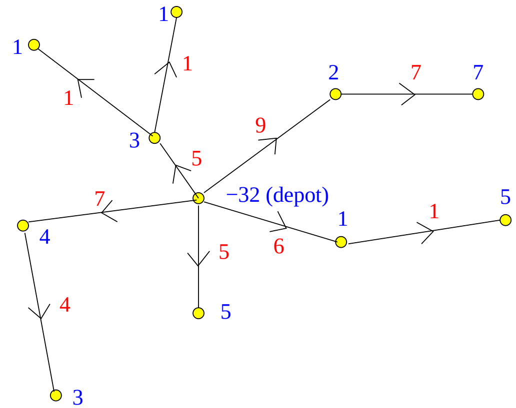
Node Routing

- We are given an undirected graph $G = (V, E)$.
  - The edges represent transportation arteries or communication links.
  - Each edge has an associated *cost* or *length*.
  - The nodes represent supply/demand points.
- Assume one supply point (the *depot*).
- A *node routing* is a directed subgraph G' of G satisfying the following properties:
  - G' is (weakly) connected.
  - The in-degree of each non-depot node is 1.



Optimal Node Routing

- Properties of a node routing.
  - It is a spanning arborescence plus (possibly) some edges returning to the depot.
  - There is a unique path from the depot to each demand point.
- We wish to construct a least cost routing.
- Cost Measures
  - Sum the lengths of all edges in G'.
  - Sum the length of all paths from the depot.
  - Some linear combination of these two.



IP Formulation

- IP formulation for this routing problem:

$$
\begin{aligned}
& Min \sum_{\{i,j\}\in E} \gamma\, c_{ij}x_{ij} + \tau\, c_{ij}y_{ij} \\
& s.t. \qquad x(\delta(V\setminus\{i\})) = 1 \qquad \forall i \in V^- \\
& y(\delta(V\setminus\{i\})) - y(\delta(\{i\})) = d_i \qquad \forall i \in V^- \\
& y_{ij} \le M x_{ij} \quad \forall \{i,j\} \in E \\
& y_{ij} \ge 0 \qquad \forall \{i,j\} \in E \\
& x_{ij} \in \{0,1\} \quad \forall \{i,j\} \in E
\end{aligned}
$$

where:

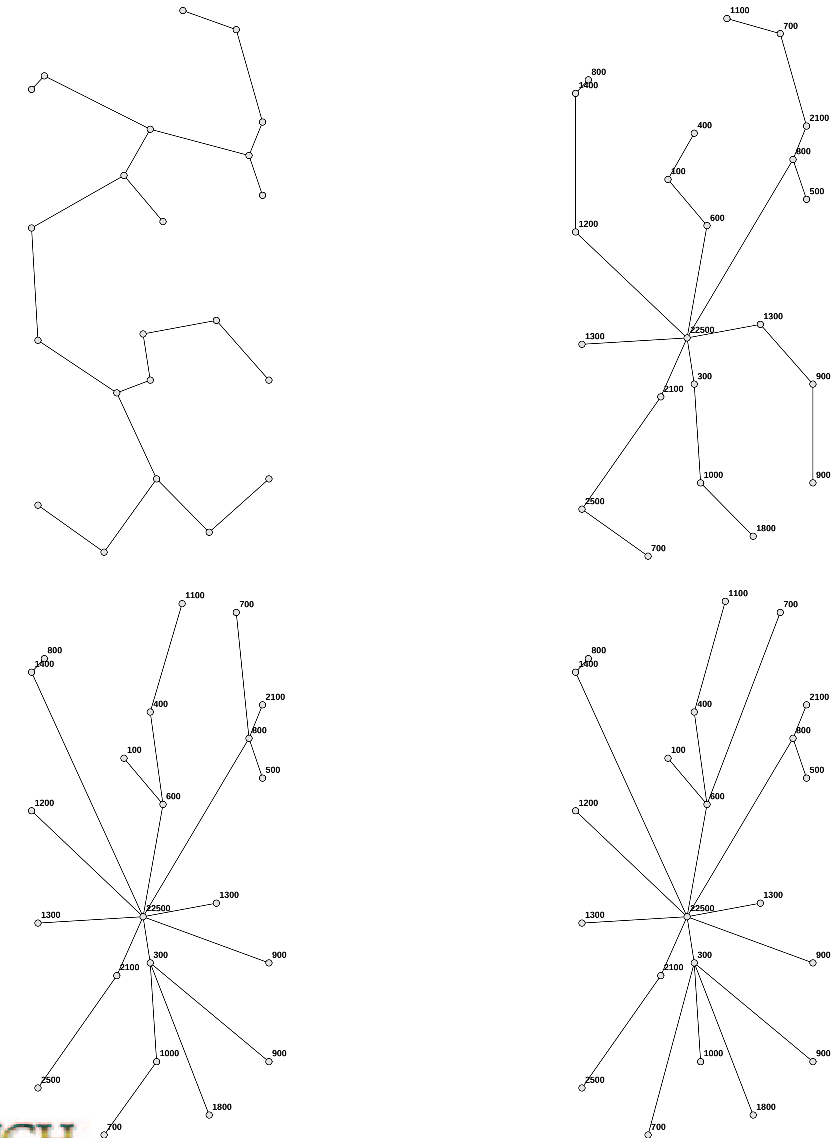
- $V^- = V \setminus \{0\}$.
- x_{ij}, x_{ji} (*fixed-charge variables*) indicate whether $\{i,j\}$ is in the routing *and its orientation*. and y_{ij} (*flow variable*) represents demand flow.

Complexity

- This node routing problem is NP-complete in general.
- Polynomially solvable special cases.
  - $\tau = 0 \Rightarrow$ Minimum Spanning Tree Problem.
  - $\gamma = 0 \Rightarrow$ Shortest Paths Tree Problem.
  - Note that demands are irrelevant.
- Other special cases.
  - $\tau, \gamma > 0 \Rightarrow$ Cable-Trench Problem (CTP).
  - $\tau = 0$ and $x(\delta(\{i\}) = 1 \Rightarrow$ Traveling Salesman Problem (TSP).
  - $\tau > 0$ and $x(\delta(\{i\}) = 1 \Rightarrow$ Variable Cost TSP (VCTSP).
  - $x(\delta(V \setminus \{0\}) = x(\delta(\{0\}) = k \Rightarrow$ k-TSP.

IBM LEHIGH University

Complexity


1100
700
800
1400
400
2100
100
800
500
600
1200
1300
22500
1300
300
900
2100
1000
900
2500
1800
700
1100
700
800
1400
2100
400
800
100
500
600
1200
1300
22500
1300
900
300
2100
1000
900
2500
1800
700
1100
700
800
1400
2100
400
800
100
500
600
1200
1300
22500
1300
900
300
2100
1000
900
2500
1800
700