Operations Research: Opportunities and Challenges

Willem-Jan van Hoeve

Associate Professor
Tepper School of Business
Carnegie Mellon University

vanhoeve@andrew.cmu.edu
@wjvanhoeve
Outline

• What is Operations Research?
  – examples

• Future opportunities
  – due to technological and economic changes

• Challenges
  – scale, uncertainty, algorithms

Acknowledgement: Thanks go to Michael Trick for an earlier version of this presentation
What is Operations Research?

What is Operations Research?
Role of Operations Research

• Data is not information; information is not improved decision making

• Operations research allows companies to transform data into better decision making

• New roles:
  – Competitive advantage
  – Business Opportunity
  – Unlocking the value of information
• Trends are moving towards more operations research
  – Increased Data
  – Faster Computers
  – Better Algorithms
  – Lower Fixed Cost for Optimization
  – Service Applications
Increased Data

Machine-generated versus authored data

- Storage online
- All medical imaging
- Medical data stored
- Surveillance bytes
- Surveillance for urban areas
- Personal multimedia
- In databases
- Static Web data
- Text data

Giga-bytes/US capita/year

Year

16 TFlops now costs US$3,000 (AMD 7970 desktop computer)
Algorithms are getting better also!

- Evolution of Linear Programming solver CPLEX

Table 5: PDS models—Solution times

<table>
<thead>
<tr>
<th>Instance</th>
<th>CPLEX 1.0</th>
<th>CPLEX 5.0</th>
<th>CPLEX 7.1</th>
<th>CPLEX 7.1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dual</td>
<td>Primal</td>
<td>Dual</td>
<td></td>
</tr>
<tr>
<td>pds100</td>
<td>–</td>
<td>50413.1</td>
<td>2414.8</td>
<td>256.3</td>
</tr>
<tr>
<td>pds90</td>
<td>–</td>
<td>59981.0</td>
<td>2452.2</td>
<td>320.3</td>
</tr>
<tr>
<td>pds80</td>
<td>–</td>
<td>42055.4</td>
<td>2201.5</td>
<td>304.4</td>
</tr>
<tr>
<td>pds70</td>
<td>335292.1</td>
<td>21120.4</td>
<td>1504.1</td>
<td>197.8</td>
</tr>
<tr>
<td>pds60</td>
<td>205798.3</td>
<td>7442.6</td>
<td>852.4</td>
<td>160.5</td>
</tr>
<tr>
<td>pds50</td>
<td>122195.9</td>
<td>8509.9</td>
<td>493.2</td>
<td>114.6</td>
</tr>
<tr>
<td>pds40</td>
<td>58920.3</td>
<td>2816.8</td>
<td>188.3</td>
<td>79.3</td>
</tr>
<tr>
<td>pds30</td>
<td>15891.9</td>
<td>1154.9</td>
<td>74.8</td>
<td>39.1</td>
</tr>
<tr>
<td>pds20</td>
<td>5168.8</td>
<td>232.6</td>
<td>27.9</td>
<td>20.9</td>
</tr>
<tr>
<td>pds10</td>
<td>208.9</td>
<td>13.0</td>
<td>3.7</td>
<td>2.6</td>
</tr>
<tr>
<td>pds06</td>
<td>26.4</td>
<td>2.4</td>
<td>1.4</td>
<td>0.9</td>
</tr>
<tr>
<td>pds02</td>
<td>0.4</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
</tbody>
</table>

546,469 variables
156,171 constraints

[Bixby, 2002]
• The combination of improved algorithms and faster computers mean that many problems can now be solved 1 million times faster than they could 10 years ago
  – solve in few seconds instead of days
• Huge increase in applicability of optimization methods
• Optimization methods can now often be applied in real-time
Lower Fixed Cost for Optimization

• Past. Big projects for big companies
  – Fighting World War II
  – Airline Crew Scheduling
  – Material Planning at Ford

• Current. Everywhere
  – Much more accessible
  – Optimization software on your computer
    (e.g., Solver in Excel)
New Areas of Application

• Traditional view of OR:
  – Manufacturing
  – Services limited to transportation and logistics

• New view
  – Operations Research everywhere
  – Services are a great opportunity
The Rise of the Service Economy

- United States
- Japan
- Germany
- China
- India
- Russia

Rise of service economy 1800-2010
Specific OR Opportunities

• Healthcare
  – cancer treatment planning, (influenza) vaccine composition, medical staff planning, in-home health care, ...

• Financial services
  – portfolio optimization (with side constraints such as transaction costs, shared budget, ...), balance risk and profit, credit score estimation, ...

• Energy market
  – smart grid, pricing, strategic inventory, ...

• Mass media
  – next generation advertising, product marketing, political campaigning, ...
Dollar slips, gold climbs in pre-Fed jockeying

The U.S. dollar eased and gold prices hit a three-month high on Tuesday as investors anticipated the new head of the U.S. Federal Reserve's outlook for the economy and policy would do nothing to rock the boat for markets. Fed Chair Janet Yellen gives her first testimony before the House Financial Services Committee at 1500 GMT, and will likely face questions on the state of the labor market and the future pace of tapering.
Why Now?

- Trends are moving towards more operations research
  - Increased Data
  - Faster Computers
  - Better Algorithms
  - Lower Fixed Cost for Optimization
  - Service Applications

...not present 10~15 years ago
Impact and Challenges

• Bright future for OR practitioners...
  – many job opportunities, new application areas

 Operations Research Analyst

#2 best business job

Mean Salaries Shown

<table>
<thead>
<tr>
<th>Year</th>
<th>Salary</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td>$50,000</td>
</tr>
<tr>
<td>2006</td>
<td>$58,000</td>
</tr>
<tr>
<td>2008</td>
<td>$66,000</td>
</tr>
<tr>
<td>2010</td>
<td>$74,000</td>
</tr>
<tr>
<td>2012</td>
<td>$82,000</td>
</tr>
</tbody>
</table>

Operations research analysts are high-level problem-solvers who use advanced techniques, such as optimization, data mining, statistical analysis and mathematical modeling, to develop solutions that help businesses and organizations operate more efficiently and cost-effectively.
Impact and Challenges

• Bright future for OR practitioners...
  – many job opportunities, new application areas

• ... and for researchers!

  ▪ new application areas
  ▪ large scale problems
  ▪ real-time application
  ▪ data driven
  ▪ handle uncertainty

  ▪ different requirements
  ▪ new methodologies
  ▪ new algorithms
  ▪ combine technologies
  ▪ more funding
Challenge 1: Scale

• Graph theory is one of the corner stones of operations research (e.g., networks)

• Up to recently, many theoretic results involved the relationship between specific subgraphs
  – Specify (family of) graphs by forbidden substructures
  – Example: the Petersen graph is the smallest bridgeless cubic graph with no three-edge-coloring

• Small graphs, can draw on piece of paper
Size of the internet:

• In 1998: 26 million unique URLs
• In 2008: 1 trillion unique URLs

(source: Google)
Large scale optimization

- Huge networks
  - internet-based applications, e.g., advertising via facebook or twitter network
- Different questions can be asked
  - importance of nodes in a network
  - connectedness; small-world phenomenon
- Different graph-theoretic approaches
  - ‘social networks’ very active research area
- Other large-scale applications
  - routing applications, client management for service industry; complex supply chains, ...
Challenge 2: Uncertainty

- New applications will demand better handling of uncertainty
- Example application: Vehicle routing
Many approaches in the literature assume

• One commodity needs to be picked up or delivered
• Client demand is deterministic
• Vehicles are uniform and 1-dimensional
• Distances are given and fixed

In practice, usually *none* of these assumptions applies

• Moreover, there are often side constraints (time windows, stacking conditions, ...)

*Vehicle Routing Approaches*
Next generation vehicle routing

• Need to integrate traditional optimization with data mining, statistics, forecasting, ...
  – Use historical data to build demand distributions
  – Cluster clients together such that trucks do not overload with $p=0.95$

• Assume uncertain distances (again, can be based on historic traffic data)

• Try to accommodate all side constraints

• We are still far from optimal solutions...
Challenge 3: Algorithms

- Previously known ‘good’ algorithms may no longer be applicable

- Example: *Minimum spanning tree* for graph $G = (V,E)$
  - Prim: $O(|E| + |V| \log |V|)$
  - Kruskal: $O(|E| \log |V|)$

- Graph on 1M nodes?

- Note: often used as subroutine
Traveling Salesman Problem

Find the shortest closed tour that visits each city exactly once

Applications:

- Truck routing
- Electronic circuit design
- Genome sequencing
- Parcel delivery services
- Robotic arm movement planning
- and many, many more
For $n$ locations, there are $(n-1)!$ possible routes.

Example:

\begin{align*}
n &= 5 & n! &= 120 \\
n &= 10 & n! &= 3\,628\,800 \\
n &= 20 & n! &= 2.43\times10^{18} \\
n &= 40 & n! &= 8.15\times10^{47} \\
n &= 60 & n! &= 8.32\times10^{81}
\end{align*}

This is more than the total number of atoms in the observable universe! (estimated to be around $10^{80}$)
71,009 Cities in China

http://www.tsp.gatech.edu
### Milestones

<table>
<thead>
<tr>
<th>Year</th>
<th>Research Team</th>
<th>Size of Instance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1954</td>
<td>G. Dantzig, R. Fulkerson, and S. Johnson</td>
<td>49 cities</td>
</tr>
<tr>
<td>1971</td>
<td>M. Held and R.M. Karp</td>
<td>64 cities</td>
</tr>
<tr>
<td>1975</td>
<td>P.M. Camerini, L. Fratta, and F. Maffioli</td>
<td>67 cities</td>
</tr>
<tr>
<td>1977</td>
<td>M. Grötschel</td>
<td>120 cities</td>
</tr>
<tr>
<td>1980</td>
<td>H. Crowder and M.W. Padberg</td>
<td>318 cities</td>
</tr>
<tr>
<td>1987</td>
<td>M. Padberg and G. Rinaldi</td>
<td>532 cities</td>
</tr>
<tr>
<td>1987</td>
<td>M. Grötschel and O. Holland</td>
<td>666 cities</td>
</tr>
<tr>
<td>1987</td>
<td>M. Padberg and G. Rinaldi</td>
<td>2,392 cities</td>
</tr>
<tr>
<td>1994</td>
<td>D. Applegate, R. Bixby, V. Chvátal, and W. Cook</td>
<td>7,397 cities</td>
</tr>
<tr>
<td>1998</td>
<td>D. Applegate, R. Bixby, V. Chvátal, and W. Cook</td>
<td>13,509 cities</td>
</tr>
<tr>
<td>2001</td>
<td>D. Applegate, R. Bixby, V. Chvátal, and W. Cook</td>
<td>15,112 cities</td>
</tr>
<tr>
<td>2004</td>
<td>D. Applegate, R. Bixby, V. Chvátal, W. Cook, and K. Helsgaun</td>
<td>24,978 cities</td>
</tr>
<tr>
<td>2005</td>
<td>Applegate et al.</td>
<td>85,900 cities</td>
</tr>
</tbody>
</table>

chip design application for AT&T/Bell Labs, solved to optimality in 136 CPU years (on a 250-node cluster this took around one year)
• Integer linear programming model
  – binary variable $x_e$ for each edge $e$ to represent tour
  – constraints to ensure that there are no sub-tours
  – minimize $\sum_e d_e x_e$

• Challenge:
  – huge number of edges and constraints

• Remedy: problem decomposition
  – variable generation based on marginal cost of edge
  – constraint generation for detected subtours, integrality, ...
  – heuristic solutions to find upper bounds
  – optimality is still guaranteed, with fraction of full model
China TSP revisited

Optimal tour [Hung Dinh Nguyen]
Summary

• What is Operations Research?
  – examples

• Future opportunities
  – due to technological and economic changes

• Challenges
  – scale, uncertainty, algorithms

https://www.informs.org/