Robust capacity expansion of a network under demand uncertainty: a bicriteria approach

H. Aissi and D. Vanderpooten

LAMSADE, Université Paris-Dauphine

Table of contents

1. The capacity expansion problem
2. Classical robustness methodology
3. Bi-objectif approach
4. Bi-objective flow on expansion graphs

The standard version

- Some commodity provided by a set I of suppliers must be dispatched to a set J of customers expressing demands.
- Assume that each demand d_j, for any $j \in J$, is a given integer.
- The existing supply/transportation capacities are not sufficient to meet all demands.
- Increase these capacities with minimum total cost in order to satisfy all demands.
- This problem is formulated as a minimum cost maximum flow problem defined over a network $N = (V, A)$ with vertex set $V = I \cup J \cup K \cup \{s, t\}$ where s is a source, t is a sink, and K is the set of transit nodes.

The standard version

- Decompose supply or transportation capability into two parallel arcs: *original* and *expansion* arcs.
- The arc set A is partitioned into $A = O \cup E \cup D$ where:
  - $O = \{(s,i), \text{ for all } i \in I\} \cup \{(a,b) \in A : a \in I \cup K, b \in K \cup J\}$ is the set of *original arcs* with lower bounds l_{ab} and upper bounds u_{ab} limiting the existing supply and transportation capacities,
  - $E = \{(s,i), \text{ for all } i \in I\} \cup \{(a,b) \in A : a \in I \cup K, b \in K \cup J\}$ is the set of *expansion arcs* corresponding to the possibility of expanding existing supply and transportation capacities up to U_{ab},
  - $D = \{(j,t), \text{ for all } j \in J\}$ is the set of *demand arcs* with upper bounds $u_{jt} = d_j$ and lower bounds $l_{jt} = 0$, $j \in J$.

Picture of the problem

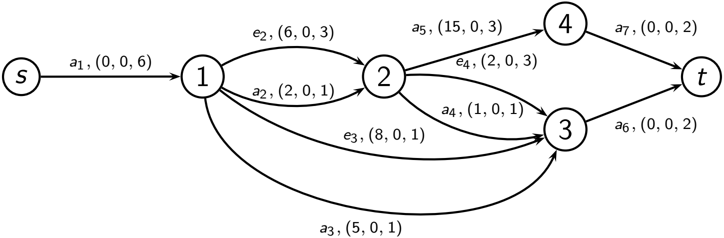

Figure: Network N: values on arcs are (c, lower bound, upper bound). $I = \{1\}$, $K = \{2\}$, and $J = \{3, 4\}$. The arc set A is partitioned into original arcs $O = \{a_1, a_2, a_3, a_4, a_5\}$, expansion arcs $E = \{e_2, e_3, e_4\}$, and demand arcs $D = \{a_6, a_7\}$. The demands are $d_3 = d_4 = 2$.

Uncertain demands

- In practice, network operators use statistical models, together with market surveys, to forecast the evolution of demand.
- Experience shows that forecasts are wrong and far from the observed reality
  - Traffic measurements on a backbone network of France Telecom, compared with the amounts forecasted one year before, have revealed gaps up to 25% on the global amount of traffic in the network depending on the year [BKOV'10].

Uncertain demands

Assumptions:

- Each demand d_j, for any $j \in J$, can take any integer value from interval $[\underline{d}_j, \bar{d}_j]$ independently of any other demand d_h, for $h \in J \setminus \{j\}$.
- No distribution assumption is made on the uncertain demands.
- The initial network can support the minimal demands $\underline{d}_j$, for all $j \in J$.

Two stages decision context

Decisions are actually to be made in two stages:

1. definition of a capacity expansion plan when the demand is still unknown,
   $\Rightarrow$ the expansion cost is known but supply and transportation cost is estimated,
2. definition of a supply and transportation policy once the demand is revealed,
   $\Rightarrow$ the supply and transportation cost is known.

Adjustable Robust Optimization (ARO) formulation

- Ben-Tal et al. [BGN'09,BGN'04,BN'98] introduced the ARO formulation in order to model this decision environment.
- The ARO model requires to define an uncertainty domain $\mathcal{U} \subset \mathcal{D} =$ hyperrectangle formed by all intervals and excluding the extreme combinations of demand values.
- The shape of $\mathcal{U}$ is usually ellipsoidal [BGN'09,BGN'04,BN'98] or polyhedral [BS'03,BS'04,TTE'09].
- The definition of $\mathcal{U}$ requires to specify a parameter Γ controlling the 'size' of $\mathcal{U}$.

Adjustable Robust Optimization (ARO) formulation

- An expansion plan is *robust* if it is able to cope with any demand in $\mathcal{U}$.
- Only expansion plans admitting a supply and transportation policy able to satisfy any demand $d \in \mathcal{U}$ are considered.
- The worst-case supply and transportation cost of any feasible expansion plan x^E is

$$\kappa(x^E, \mathcal{U}) = \max_{d \in \mathcal{U}} \min_{x \in \mathcal{F}(d, x^E)} c^T x$$

where c^T is the vector of the supply and transportation costs and $\mathcal{F}(d, x^E)$ is the set of feasible flows satisfying exactly demand $d \in \mathcal{U}$ and consistent with x^E.

First stage

- The goal of the first stage problem is to compute an expansion plan minimizing the worst-case total cost for all demands in $\mathcal{U}$.

$$\min_{x^E \in \mathcal{E}} c_{ARO}(x^E) = c^E x^E + \kappa(x^E, \mathcal{U})$$

where $\mathcal{E}$ denote the set of all feasible expansion plans and where c^E is the vector of expansion costs.

Second stage problem

- Once the demand $\widetilde{d}$ is revealed, the goal of the second stage problem is to compute a supply and transportation plan by solving the following minimum cost flow problem

$$\min_{x \in \mathcal{F}(\widetilde{d}, x^{*E})} c^T x$$

where x^{*E} denote an optimal solution of the ARO model.

Limits

- The definition of an uncertainty set $\mathcal{U} \subset \mathcal{D}$ is a non-trivial task.
- Setting a priori a specific value for Γ is uneasy.
  In [BS'04,TTE'09] it is suggested that the budget of uncertainty Γ should be of the order of $\sqrt{n}$ where n is the number of uncertain parameters.
- No a priori guarantee on the running time can be provided for the resulting integer, linear or conic, programs.

Main ideas

- The feasibility is not a binary state but indicates the ability to satisfy a certain level of demand.
  $\Rightarrow$ The *Quality of Service* (QoS) can be taken as a performance measure.
- If we impose to satisfy any possible demand, a *totally robust solution* may have a large cost.
- The decision maker may accept to reduce its requirements in terms of guaranteed QoS depending on the cost saving.

First stage

- In order to make a decision about an expansion plan x^E at the first stage, we need to have a worst case estimate of the total cost $\widehat{cost}(x^E)$ and the QoS $\widehat{QoS}(x^E)$.
- For any feasible flow $x = (x^O, x^E, x^D)$, where x^O are flow values on the original arcs, x^E are flow values on the expansion arcs, and x^D are flow values on demand arcs
  - the total cost of x is

$$c(x) = \sum_{(a,b)\in E} c^E_{ab} x^E_{ab} + \Big(\sum_{(a,b)\in O} c^T_{ab} x^O_{ab} + \sum_{(a,b)\in E} c^T_{ab} x^E_{ab} \Big)$$

  - the ability of x to satisfy the maximum demand $\bar{d}$ can be defined by a QoS measure

$$q(x) = \frac{1}{|J|} \sum_{j\in J} \frac{x^D_{jt}}{\bar{d}_j} \text{ or } \min_{j\in J} \frac{x^D_{jt}}{\bar{d}_j} \ (bottleneck)$$

First stage

- For any solution $x = (x^O, x^E, x^D)$, we show that $c(x)$ and $q(x)$ provide tight worst-case guarantees of $\widehat{cost}(x^E)$ and $\widehat{QoS}(x^E)$ for any demand $d \in \mathcal{D}$.
- The conflicting criteria $c(x)$ and $q(x)$ can be taken as criteria for evaluating *ex ante* the corresponding expansion plan x^E.

First stage

- Each solution x can be represented in the criterion space by a point $(c(x), q(x))$.

Definition

Since c is to be minimized and q is to be maximized, a solution x is *efficient* if and only if there is no solution x' such that $c(x') \leqslant c(x)$ and $q(x') \geqslant q(x)$, with at least one strict inequality.

First stage

- All the efficient points are generated and for each such point a corresponding efficient solution is provided.
  $\Rightarrow$ *potentially robust* solutions.
- The decision maker chooses a non-dominated point, associated with a potentially robust solution $\bar{x}$, offering the best tradeoff between guaranteed cost and QoS.
- This way he/she selects an expansion plan, characterized by variables $\bar{x}^E$, for which we guarantee that, for any demand $d \in \mathcal{D}$, the total cost will be at most $\widehat{cost}(\bar{x}^E) = c(\bar{x})$ and the QoS will be at least $\widehat{QoS}(\bar{x}^E) = q(\bar{x})$.

First stage

In this bi-objective formulation, the network N is completed by adding on each arc the estimate QoS criterion values.

Figure: Network N: values on arcs are (c, q, lower bound, upper bound).

Second stage

- Once the demand $\widetilde{d}$ is revealed, we exhibit a supply and transportation plan $y(\bar{x}^E, \widetilde{d})$ which is both consistent with the expansion plan $\bar{x}^E$ and the conflicting guarantees that its total cost will be at most $\widehat{cost}(\bar{x}^E)$ and its QoS will be at least $\widehat{QoS}(\bar{x}^E)$.
- For this purpose, we just need to solve a minimum cost maximum flow problem on a network N'

Figure: Network N': values on arcs are (cost, lower bound, upper bound).

Second stage

The total cost of solution $\bar{x}$ selected in the first stage is

$$cost(\bar{x}^E, \widetilde{d}) = \sum_{(a,b)\in E} c^E_{ab}\bar{x}^E_{ab} + C^T(y(\bar{x}^E, \widetilde{d})),$$

where $C^T(y(\bar{x}^E, \widetilde{d}))$ denote the optimal supply and transportation cost and its QoS is

$$QoS(\bar{x}^E, \widetilde{d}) = \frac{1}{|J|}\sum_{j\in J} \frac{y_{jt}(\bar{x}^E, \widetilde{d})}{\widetilde{d}_j}.$$

Properties of exact and approximate cost and QoS criteria

Proposition

Let $x = (x^E, x^O, x^D)$ be any feasible flow in N and $\tilde{d} \in \mathcal{D}$ be any revealed demand, we have:

(i) $cost(x^E, \tilde{d}) \leqslant \widehat{cost}(x^E)$,

(ii) *if x is efficient and $\tilde{d} \geqslant x^D$, then $\widehat{cost}(x^E) = cost(x^E, \tilde{d})$.*

(iii) $QoS(x^E, \tilde{d}) \geqslant \widehat{QoS}(x^E)$ *for any $\tilde{d} \in \mathcal{D}$,*

(iv) $\widehat{QoS}(x^E) = QoS(x^E, \bar{d})$.

Example

Consider the expansion graph $N = (V, A)$ with demands in $[1;3]$ ($\bar{d}_3 = \bar{d}_4 = 3$ and $\underline{d}_3 = \underline{d}_4 = 1$).

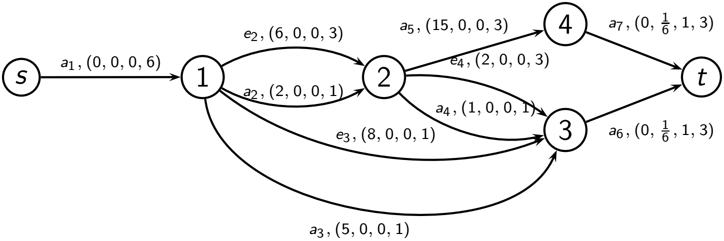

Figure: Values on arcs are (c, q, lower bound, upper bound).

Example

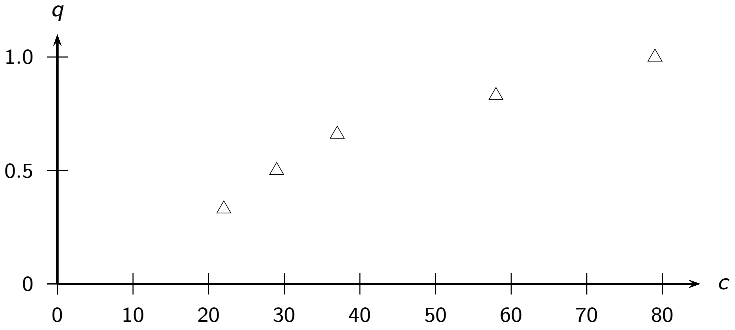

Figure: Non-dominated points.

Example

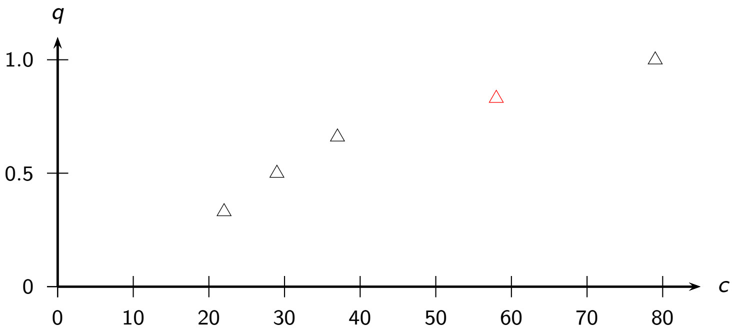

Figure: Non-dominated points.

Example

- Assume that he/she chooses, among these non-dominated points, a QoS of at least 0.83 and a total cost of at most 58.
- The corresponding solution $\bar{x}$ is defined as

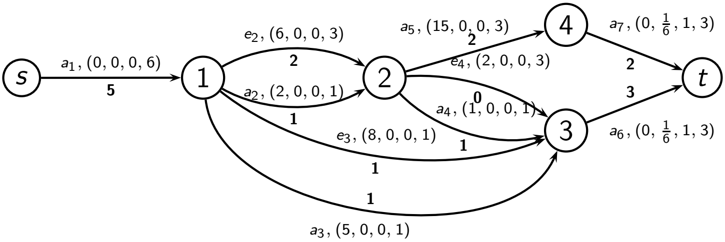

Figure: Network N: values on arcs are (c, q, lower bound, upper bound). The expansion cost is 11.

Example

- Suppose now that the real demand is $\widetilde{d}_3 = 2$ and $\widetilde{d}_4 = 3$.
- The second stage problem consists of solving the minimum cost maximum flow problem defined over graph N'

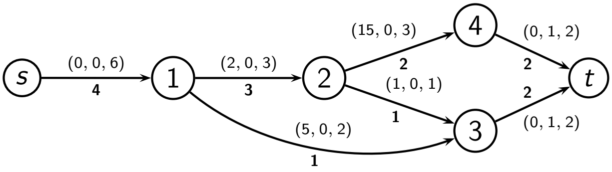

Figure: Values on arcs are (cost, lower bound, upper bound). The supply and transportation cost $C^T(y(\bar{x}^E, \widetilde{d})) = 42$, and the total cost is $cost(\bar{x}^E, \widetilde{d}) = 53$. The quality of service is $QoS(\bar{x}^E, \widetilde{d}) = \frac{1}{2}(\frac{2}{2} + \frac{2}{3}) = 0.83$.

Example

Demand		*cost*	QoS
$\widetilde{d}_3$	$\widetilde{d}_4$		
1	1	31	1
1	2	48	1
1	3	48	0.83
2	1	36	1
2	2	53	1
2	3	53	0.83
3	1	41	1
3	2	58	1
3	3	58	0.83

Table: Cost and QoS of solution $\bar{x}$ for all possible real demands.

Comparison with the standard ARO model

- The ARO model provides conclusions like ‘with a probability at least w, the proposed robust solution is feasible and has a cost at most x’.
- Our robust approach provides conclusions like ‘the proposed robust solution guarantees a cost at most y and a QoS at least z’.
- Our model computes all the interesting values of worst-case costs and QoS.

Comparison with the standard ARO model

- For any feasible expansion $x^E \in \mathcal{E}$, the ARO model provides a supply and transportation plan with a QoS 1 for all demands in $\mathcal{U}$. No guarantees are given for demands in $\mathcal{D} \setminus \mathcal{U}$ which may occur with a (small) nonzero probability.
- Our model computes for each expansion plan x^E a supply and transportation plan which may partially satisfy any demand in $\mathcal{U}$ or in $\mathcal{D} \setminus \mathcal{U}$.

Comparison with the standard ARO model

- Does an optimal solution of the ARO model corresponds to a non-dominated solution in our model?
- If any non-dominated solution is an optimal solution of the ARO model for some uncertainty set?

Interpreting the guarantees of the ARO model in our framework

An uncertainty set $\mathcal{U} \subseteq \mathcal{D}$ can be interpreted as imposing a minimum QoS level

$$q^*(\mathcal{U}) = \max_{d \in \mathcal{U}} \frac{1}{|J|} \sum_{j \in J} \frac{d_j}{\bar{d}_j}.$$

Proposition

For any feasible expansion plan $x^E \in \mathcal{E}$, our model outputs an efficient solution $\bar{x} = (\bar{x}^O, \bar{x}^E, \bar{x}^D)$ whose associated expansion plan $\bar{x}^E$ guarantees a worst case total cost $\widehat{cost}(\bar{x}^E) \leqslant c_{ARO}(x^E)$ and a worst case QoS $\widehat{QoS}(\bar{x}^E) \geqslant q^*(\mathcal{U})$.

Interpreting an optimal solution of the ARO model in our framework

Let $\mathcal{U}^* = \{d \in \mathcal{U} : \frac{1}{|J|}\sum_{j \in J} d_j/\bar{d}_j = q^*(\mathcal{U})\}$ and for any feasible expansion plan x^E of (11), let $\mathcal{W}(x^E, \mathcal{U})$ denote the set of worst-case demands $d \in \arg\max_{d \in \mathcal{U}} \min_{x \in \mathcal{F}(d, x^E)} c^T x$.

Proposition

If there exists an optimal solution x^{*E} of the ARO model such that $\mathcal{U}^* \setminus \mathcal{W}(x^{*E}, \mathcal{U}) \neq \emptyset$, our model outputs an efficient solution $\bar{x} = (\bar{x}^O, \bar{x}^E, \bar{x}^D)$ whose associated expansion plan $\bar{x}^E$ guarantees a worst case total cost $\widehat{cost}(\bar{x}^E) < c_{ARO}(x^E)$ and a worst case QoS $\widehat{QoS}(\bar{x}^E) \geqslant q^*(\mathcal{U})$.

Interpreting an efficient solution in the ARO model

Proposition

For any efficient solution $x = (x^O, x^E, x^D)$, there exists an uncertainty set $\mathcal{U}(x)$ such that the associated expansion plan x^E is an optimal solution of the ARO model.

In particular, $\mathcal{U}(x) = \{d \in \mathcal{D} : \underline{d}_j \leqslant d_j \leqslant x^D_{jt}\}$.

Example

- In the ARO model, suppose that $\mathcal{U}$ is included in the ellipse $(\frac{d_3 - 2}{0.5})^2 + (d_4 - 2)^2 = 1$.
- The optimal solution x^{*E} is (3,0,0) with an expansion cost 12 and a worst supply and transportation cost 59.

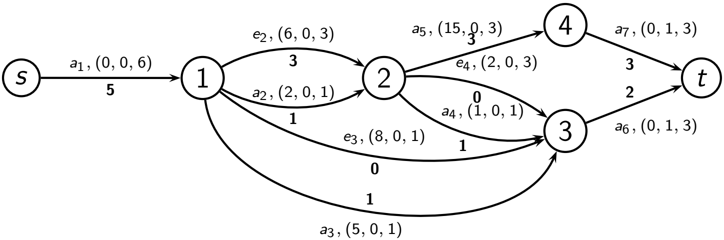

Figure: Network N: values on arcs are (c, lower bound, upper bound).

Example

- The uncertainty domain $\mathcal{U}$ can be interpreted in our model as guaranteeing a QoS level $q^*(\mathcal{U}) = 0.83$.
- Our model outputs an efficient solution $\bar{x} = (\bar{x}^O, \bar{x}^E, \bar{x}^D)$ offering a guaranteed QoS $q^*(\mathcal{U})$ but at much lower worst-case total cost 58 (19% cost saving) for any demand $d \in \mathcal{D}$.

Example

- Consider the domain $\mathcal{U}(\bar{x}) = \{d \in \mathcal{D} : 1 \leqslant d_3 \leqslant 3 \text{ and } 1 \leqslant d_4 \leqslant 2\}$.
- One can verify that $\bar{x}^E$ is the optimal solution of the ARO model defined with the uncertainty set $\mathcal{U}(\bar{x})$.

Abstract model

- The robust capacity expansion problem can be stated as a bi-objective flow problem on specific graphs.
- By construction, each arc has a criterion vector with at most one non-zero entry.
- Each path from the source s to the sink t is formed by a sequence of arcs $(a, b) \in O \cup E$ such that the QoS is zero, followed by only one arc $(j, t) \in D$ with zero cost and a positive QoS.
  $\Rightarrow$ Expansion graphs

Complexity results

	Bi-Objective Flow on Expansion Graphs	Bottleneck Bi-Objective Flow on Expansion Graphs
Complexity	NP-hard	Poly
Intractability	Intractable	Intractable
Exact algorithms	Pseudo-poly*	Pseudo-poly
Approximation	ptas, fptas*	fptas

The starred entries correspond to cases where the number of distinct values of demands $\bar{d}$.