Tutorial: New Challenges in Network Optimization

Dimitri Papadimitriou[1]
dimitri.papadimitriou@nokia-bell-labs.com

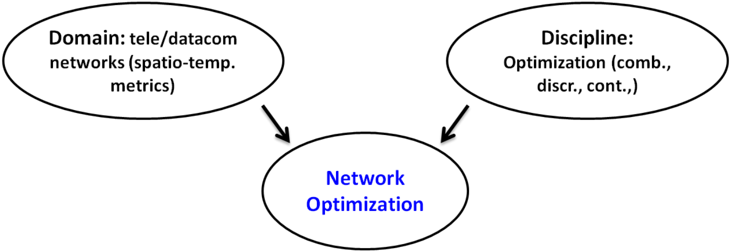
[1]**Bell Labs**
Antwerp, Belgium

IEEE HPSR - Yokohama
June 14-17, 2016

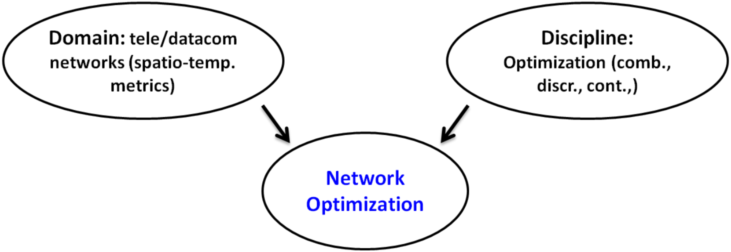
Outline

Outline

Introduction



Introduction



Evolution/Trends

- Wireless, circuit/optical and packet switching networks → Information exchange networks (switching, storage, processing)
- Node-centric (network as collection of nodes) → Network-centric (realize function)
- Physical resources (memory & CPU, link/node capacity) with fixed allocation → Logical resources (abstraction) with dynamic allocation
- Open-loop, static, centralized, and dependent → Closed-loop (feedback, adaptive, model-reference), dynamic, distributed/multi-agent and autonomous control

⇒ Specialized (network design(access/aggregation), multicommodity flow routing, placement/location, etc.) → Combined studies

Problem Classes

1. Positioning/Location and Dimensioning
2. Configuring and Provisioning
3. Planning and Scheduling

} Traditional applications

Problem Classes

1. Positioning/Location and Dimensioning
2. Configuring and Provisioning
3. Planning and Scheduling

} Traditional applications

... Not only about CAPEX and OPEX

4. Protocol and System Design (early phases)

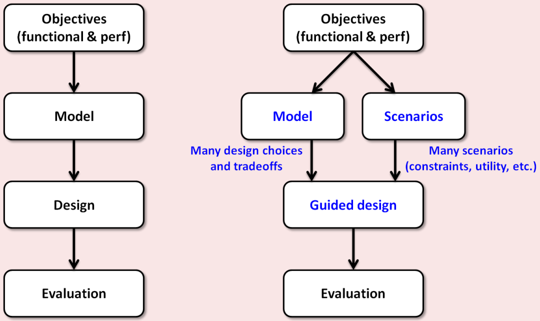
Problem Classes

1. Positioning/Location and Dimensioning
2. Configuring and Provisioning
3. Planning and Scheduling

} Traditional applications

. . . Not only about CAPEX and OPEX

4. Protocol and System Design (early phases)



Main trends: 3R

Reliability (time or space)

- Probabilistic parameters and model
- Invalidate independence property or balance scale-oriented decisions

Main trends: 3R

Reliability (time or space)

- Probabilistic parameters and model
- Invalidate independence property or balance scale-oriented decisions

Routing

- Distance functions/metrics beyond graph distance, e.g., load, delay
- Multi-level (partition), multi-period (dynamics, evolution), multi-layer (beyond overlays)
- Coupling constraints, e.g., network $\times$ routing decisions

Main trends: 3R

Reliability (time or space)

- Probabilistic parameters and model
- Invalidate independence property or balance scale-oriented decisions

Routing

- Distance functions/metrics beyond graph distance, e.g., load, delay
- Multi-level (partition), multi-period (dynamics, evolution), multi-layer (beyond overlays)
- Coupling constraints, e.g., network $\times$ routing decisions

Robustness $\rightarrow$ Robust optimization

- Parameter space (variability) $\rightarrow$ Construction (automatic) of uncertainty sets (machine/stat. learning)
- Computational complexity tradeoff

Original Problem	LP	MILP	QCQP	SOCP
Polyhedral Set	LP	MILP	MINLP	MINLP
Ellipsoidal Set	SOCP	MISOCP	SDP	SDP

Challenging Problems

1. (Reliable) capacitated Facility Location Problem (cFLP) $\times$ Multicommodity Flow Routing (MCF) $\rightarrow$ cFLRP

Challenging Problems

1. (Reliable) capacitated Facility Location Problem (cFLP) $\times$ Multicommodity Flow Routing (MCF) $\rightarrow$ cFLRP

   $\Rightarrow$ **Applicability**: information networks

Challenging Problems

1. (Reliable) capacitated Facility Location Problem (cFLP) $\times$ Multicommodity Flow Routing (MCF) $\rightarrow$ cFLRP
   - $\Rightarrow$ **Applicability**: information networks

2. Hub-Location Problem (HLP) $\times$ Location-Routing Problem (LRP) $\rightarrow$ HLRP

Challenging Problems

1. (Reliable) capacitated Facility Location Problem (cFLP) $\times$ Multicommodity Flow Routing (MCF) $\rightarrow$ cFLRP

   $\Rightarrow$ **Applicability**: information networks

2. Hub-Location Problem (HLP) $\times$ Location-Routing Problem (LRP) $\rightarrow$ HLRP

   $\Rightarrow$ **Applicability**: virtualization/cloud networks

Challenging Problems

1. (Reliable) capacitated Facility Location Problem (cFLP) $\times$ Multicommodity Flow Routing (MCF) $\rightarrow$ cFLRP
   $\Rightarrow$ **Applicability**: information networks
2. Hub-Location Problem (HLP) $\times$ Location-Routing Problem (LRP) $\rightarrow$ HLRP
   $\Rightarrow$ **Applicability**: virtualization/cloud networks
3. Mixed-Integer Programming Model for the Multi-Stage Hub Location Problem $\rightarrow$ mHLRP

Challenging Problems

1. (Reliable) capacitated Facility Location Problem (cFLP) $\times$ Multicommodity Flow Routing (MCF) $\rightarrow$ cFLRP
   $\Rightarrow$ **Applicability**: information networks

2. Hub-Location Problem (HLP) $\times$ Location-Routing Problem (LRP) $\rightarrow$ HLRP
   $\Rightarrow$ **Applicability**: virtualization/cloud networks

3. Mixed-Integer Programming Model for the Multi-Stage Hub Location Problem $\rightarrow$ mHLRP
   $\Rightarrow$ **Applicability**: multi-tenant virtualization/cloud networks

Challenging Problems

1. (Reliable) capacitated Facility Location Problem (cFLP) $\times$ Multicommodity Flow Routing (MCF) $\rightarrow$ cFLRP

   $\Rightarrow$ **Applicability**: information networks

2. Hub-Location Problem (HLP) $\times$ Location-Routing Problem (LRP) $\rightarrow$ HLRP

   $\Rightarrow$ **Applicability**: virtualization/cloud networks

3. Mixed-Integer Programming Model for the Multi-Stage Hub Location Problem $\rightarrow$ mHLRP

   $\Rightarrow$ **Applicability**: multi-tenant virtualization/cloud networks

4. Robust cFLP (variant of)

Challenging Problems

1. (Reliable) capacitated Facility Location Problem (cFLP) $\times$ Multicommodity Flow Routing (MCF) $\rightarrow$ cFLRP

   $\Rightarrow$ **Applicability**: information networks

2. Hub-Location Problem (HLP) $\times$ Location-Routing Problem (LRP) $\rightarrow$ HLRP

   $\Rightarrow$ **Applicability**: virtualization/cloud networks

3. Mixed-Integer Programming Model for the Multi-Stage Hub Location Problem $\rightarrow$ mHLRP

   $\Rightarrow$ **Applicability**: multi-tenant virtualization/cloud networks

4. Robust cFLP (variant of)

   $\Rightarrow$ **Applicability**: cooperative monitoring point placement and dimensioning

Challenging Problems

1. (Reliable) capacitated Facility Location Problem (cFLP) $\times$ Multicommodity Flow Routing (MCF) $\rightarrow$ cFLRP

   $\Rightarrow$ **Applicability**: information networks

2. Hub-Location Problem (HLP) $\times$ Location-Routing Problem (LRP) $\rightarrow$ HLRP

   $\Rightarrow$ **Applicability**: virtualization/cloud networks

3. Mixed-Integer Programming Model for the Multi-Stage Hub Location Problem $\rightarrow$ mHLRP

   $\Rightarrow$ **Applicability**: multi-tenant virtualization/cloud networks

4. Robust cFLP (variant of)

   $\Rightarrow$ **Applicability**: cooperative monitoring point placement and dimensioning

5. Multi-Period Multicommodity Capacitated Network Design and Routing Problem

Challenging Problems

1. (Reliable) capacitated Facility Location Problem (cFLP) $\times$ Multicommodity Flow Routing (MCF) $\rightarrow$ cFLRP
   $\Rightarrow$ **Applicability**: information networks

2. Hub-Location Problem (HLP) $\times$ Location-Routing Problem (LRP) $\rightarrow$ HLRP
   $\Rightarrow$ **Applicability**: virtualization/cloud networks

3. Mixed-Integer Programming Model for the Multi-Stage Hub Location Problem $\rightarrow$ mHLRP
   $\Rightarrow$ **Applicability**: multi-tenant virtualization/cloud networks

4. Robust cFLP (variant of)
   $\Rightarrow$ **Applicability**: cooperative monitoring point placement and dimensioning

5. Multi-Period Multicommodity Capacitated Network Design and Routing Problem
   $\Rightarrow$ **Applicability**: multi-agent network control (towards self-optimization)

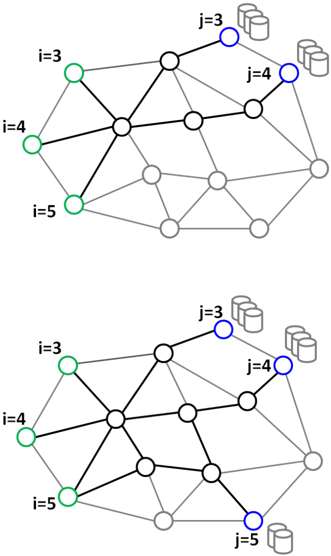
Outline

Facility Location: problems

Objectives to locate facilities

1. Median models:
   Minimize transportation costs between clients and facilities
   - p-median problem: locate p facilities such that sum of distances between vertices and nearest located facility is minimized
   - p-center problem: locate p facilities such that maximum distance is minimized
2. Covering models:
   If facility located within a specified proximity (neighborhood) of demand point/vertex then demand is covered
   - Set covering: minimize number of facilities needed to cover all clients
   - Maximum covering: maximize covered clients with a particular number of facilities
3. **Fixed charge** location models:
   minimize total facility installation/opening and transportation costs
   → Tradeoff between fixed operating and variable delivery cost

Overall Picture



Facility Location: Formulation

Input data and Parameters

- Graph $G = (\mathcal{V}, \mathcal{E})$ where vertex set $\mathcal{V}$ represents
  - Demand originating points $\mathcal{I} \subseteq \mathcal{V}$
  - Set of potential facility locations (sites) $\mathcal{J} \subseteq \mathcal{V}$
- $\forall j \in \mathcal{J}$ of finite capacity b_j
  - Facility opening cost φ_j
  - Assignment cost κ_{ij} (allocation of demand a_i to opened facility j)
  - Distance $d(i, j) = \delta_{ij}$ from demand point i to location j

Facility Location: Formulation

Input data and Parameters

- Graph $G = (\mathcal{V}, \mathcal{E})$ where vertex set $\mathcal{V}$ represents
  - Demand originating points $\mathcal{I} \subseteq \mathcal{V}$
  - Set of potential facility locations (sites) $\mathcal{J} \subseteq \mathcal{V}$
- $\forall j \in \mathcal{J}$ of finite capacity b_j
  - Facility opening cost φ_j
  - Assignment cost κ_{ij} (allocation of demand a_i to opened facility j)
  - Distance $d(i,j) = \delta_{ij}$ from demand point i to location j

Variables

- Binary variable $y_j = 1$ if facility of capacity b_j opened at location j (0 otherwise)
- Real variable $x_{ij} \geq 0$: fraction of demand a_i satisfied by facility (opened at location) j

Facility Location: Formulation

Input data and Parameters

- Graph $G = (\mathcal{V}, \mathcal{E})$ where vertex set $\mathcal{V}$ represents
  - Demand originating points $\mathcal{I} \subseteq \mathcal{V}$
  - Set of potential facility locations (sites) $\mathcal{J} \subseteq \mathcal{V}$
- $\forall j \in \mathcal{J}$ of finite capacity b_j
  - Facility opening cost φ_j
  - Assignment cost κ_{ij} (allocation of demand a_i to opened facility j)
  - Distance $d(i,j) = \delta_{ij}$ from demand point i to location j

Variables

- Binary variable $y_j = 1$ if facility of capacity b_j opened at location j (0 otherwise)
- Real variable $x_{ij} \geq 0$: fraction of demand a_i satisfied by facility (opened at location) j

Task

Select a subset of potential locations where to install/open a facility and assign every client i with known demand a_i to single or to (sub)set of open facilities without exceeding their capacity b_j (capacitated)

Facility Location Problem: MILP Formulation

Find i) set of locations to install/open facilities (location) and ii) assignment of demands to open facilities (allocation) that minimize

- Opening/installation cost of selected facilities: $\sum_{j \in \mathcal{J}} \varphi_j y_j$
- Customer demand supplying cost at each facility: $\sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \kappa_{ij} x_{ij}$
- Connection cost of each demand to subset of selected facilities: $\sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \delta_{ij} x_{ij}$

$$\min \sum_{j \in \mathcal{J}} \varphi_j y_j + \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \kappa_{ij} x_{ij} + \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \delta_{ij} x_{ij} \quad (1)$$

subject to:

$$\sum_{j \in \mathcal{J}} x_{ij} = 1 \qquad i \in \mathcal{I} \quad (2)$$

$$x_{ij} \leq y_j \qquad i \in \mathcal{I}, j \in \mathcal{J} \quad (3)$$

$$\sum_{i \in \mathcal{I}} a_i x_{ij} \leq b_j y_j \qquad j \in \mathcal{J} \quad (4)$$

$$\sum_{i \in \mathcal{I}} a_i \leq \sum_{j \in J} b_j y_j \quad (5)$$

$$x_{ij} \in [0,1] (\text{or} x_{ij} \in \{0,1\}) \qquad i \in \mathcal{I}, j \in \mathcal{J} \quad (6)$$

$$y_j \in \{0,1\} \qquad j \in \mathcal{J} \quad (7)$$

Model Properties (1)

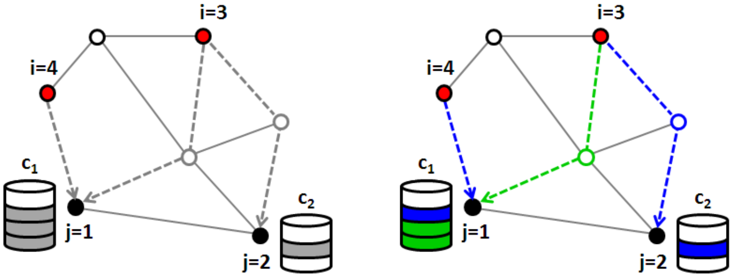
Properties

1. **Hard-capacitated**: only one facility may be installed at each location $j \in \mathcal{J}$ with finite capacity b_j
2. **Multi-source**: each demand a_i may be served by multiple sources (facilities $j \in \mathcal{J}$)
   single-source: each client demand served by a single facility
3. **Multi-product**: each opened facility j offers multiple (k) product types
   product: data object - product type: data object class

Model Properties (1)

Properties

1. **Hard-capacitated**: only one facility may be installed at each location $j \in \mathcal{J}$ with finite capacity b_j
2. **Multi-source**: each demand a_i may be served by multiple sources (facilities $j \in \mathcal{J}$)
   single-source: each client demand served by a single facility
3. **Multi-product**: each opened facility j offers multiple (k) product types
   product: data object - product type: data object class



multi-product Facility Location: MILP Formulation

$$\min \sum_{j\in\mathcal{J}} \varphi_j y_j + \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}} \kappa_{ijk} x_{ijk} + \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}} \delta_{ij} x_{ijk} \tag{8}$$

subject to:

$$\sum_{j\in\mathcal{J}} x_{ijk} = 1 \qquad i\in\mathcal{I}, k\in\mathcal{K} \tag{9}$$

$$z_{jk} \le y_j \qquad j\in\mathcal{J}, k\in\mathcal{K} \tag{10}$$

$$x_{ijk} \le z_{jk} \qquad i\in\mathcal{I}, j\in\mathcal{J}, k\in\mathcal{K} \tag{11}$$

$$\sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{K}} a_{ik} x_{ijk} \le b_j y_j \qquad j\in\mathcal{J} \tag{12}$$

$$\sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{K}} a_{ik} \le \sum_{j\in J} b_j y_j \tag{13}$$

$$x_{ijk} \in [0,1] (\text{or} x_{ijk} \in \{0,1\}) \qquad i\in\mathcal{I}, j\in\mathcal{J}, k\in\mathcal{K} \tag{14}$$

$$y_j \in \{0,1\} \qquad j\in\mathcal{J} \tag{15}$$

$$z_{jk} \in \{0,1\} \qquad j\in\mathcal{J}, k\in\mathcal{K} \tag{16}$$

Model Properties (2)

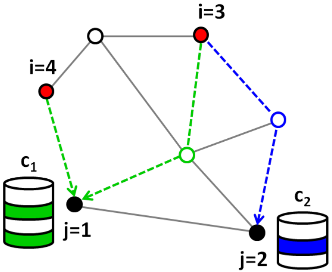
Properties

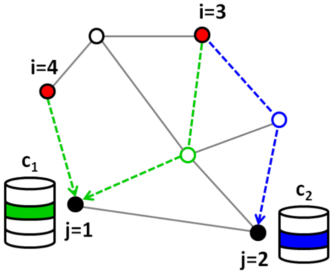
4. **Symmetric transportation cost**: optimal solution to client-to-server problem $\equiv$ optimal solution to server-to-client problem
5. **Shared-capacity**: installed capacity shared among product types hosted by each facility (no dedicated capacity per-product type)
6. **Digital goods**: **single copy** of each object hosted at installed facilities even if assigned to multiple customer demands a_i

Model Properties (2)

Properties

4. **Symmetric transportation cost**: optimal solution to client-to-server problem ≡ optimal solution to server-to-client problem
5. **Shared-capacity**: installed capacity shared among product types hosted by each facility (no dedicated capacity per-product type)
6. **Digital goods**: **single copy** of each object hosted at installed facilities even if assigned to multiple customer demands a_i





multi-digital products Facility Location: Formulation

$$\min \sum_{j\in\mathcal{V}} \varphi_j y_j + \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}} \kappa_{ijk} x_{ijk} + \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}} \delta_{ij} x_{ijk} \tag{17}$$

subject to:

$$\sum_{j\in\mathcal{J}} x_{ijk} = 1 \qquad i\in\mathcal{I}, k\in\mathcal{K} \tag{18}$$

$$z_{jk} \leq y_j \qquad j\in\mathcal{J}, k\in\mathcal{K} \tag{19}$$

$$x_{ijk} \leq z_{jk} \qquad i\in\mathcal{I}, j\in\mathcal{J}, k\in\mathcal{K} \tag{20}$$

$$\sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{K}} a_{ik} \frac{x_{ijk}}{\sum_{\ell\in\mathcal{L}} x_{\ell jk}} \leq b_j y_j \qquad j\in\mathcal{J} \tag{21}$$

$$\sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{K}} a_{ik} \frac{x_{ijk}}{\sum_{\ell\in\mathcal{L}} x_{\ell jk}} \leq \sum_{j\in J} b_j y_j \tag{22}$$

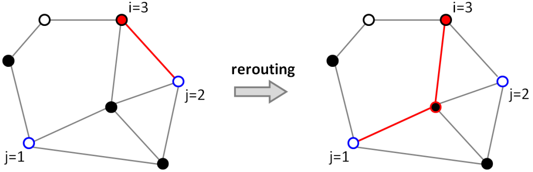
$$x_{ijk} \in [0,1] \qquad i\in\mathcal{I}, j\in\mathcal{J}, k\in\mathcal{K} \tag{23}$$

$$y_j \in \{0,1\} \qquad j\in\mathcal{J} \tag{24}$$

$$z_{jk} \in \{0,1\} \qquad j\in\mathcal{J}, k\in\mathcal{K} \tag{25}$$

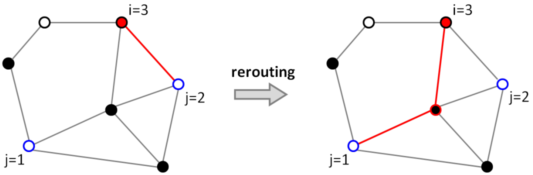
Facility Location-Routing Problem (1)

- When routing topology determined endogenously, more effective to change routing decisions instead of locating additional facilities (or increase capacity on installed facilities) $\rightarrow$ Coupled location and routing decisions



Facility Location-Routing Problem (1)

- When routing topology determined endogenously, more effective to change routing decisions instead of locating additional facilities (or increase capacity on installed facilities) → Coupled location and routing decisions



Main idea

- **Combination** of multi-source multi-product capacitated facility location (MSMP-cFLP) **for digital goods** with flow routing: MSMP-cFLRP
- Modeled and solved **independently** → Modeled and solved **simultaneously**

Facility Location-Routing Problem (2)

Facility location × Flow routing → Facility Location-Routing

- **Conventional cFLP**: models cost of allocating demand a_i originated by a given client i independently of other demands $a_j, \forall j \in \mathcal{I}, i \neq j$

  → Facility location aggregates demands

- **Location-Routing Problem (LRP)**: combines cFLP with routing decisions **removes allocation independence property**

  → Strongly interrelated location and routing decisions

  - Multiple demands may or not be served simultaneously by sharing (some) edges along (partially) common routing path
  - Allocation (transportation, routing) cost not limited to graph distance

Facility Location-Routing Problem (3)

Methods

- Sequential: minimize location and allocation cost (cFLP) then routing cost (min-cost multicommodity flow problem)
- Simultaneous: minimize location, allocation and routing cost (MSMP-cFLRP)

Tradeoff: solution quality vs. computational complexity

Facility Location-Routing Problem (3)

Methods

- Sequential: minimize location and allocation cost (cFLP) then routing cost (min-cost multicommodity flow problem)
- Simultaneous: minimize location, allocation and routing cost (MSMP-cFLRP)

Tradeoff: solution quality vs. computational complexity

Computational complexity dependence

MFP	Allocation	
	Single-sourcing var. x_{ij} in {0,1}	Multi-sourcing var. x_{ij} in [0,1]
Unsplittable: flow variables $f_{uv,ij}$ in {0,1}	+++	++
Splittable: flow variables $f_{uv,ij} \geq 0$	++	+

Input: Data and Parameters

Data

- Finite graph $G = (\mathcal{V}, \mathcal{E})$ with edge set $\mathcal{E}$ and vertex set $\mathcal{V}$
  - Set of demand originating points $\mathcal{I} \subseteq \mathcal{V}, |\mathcal{I}| = I$
  - Set of potential facility locations $\mathcal{J} \subseteq \mathcal{V}, |\mathcal{J}| = J$
- Set $\mathcal{K}$ ($|\mathcal{K}| = K$): family of products that can be hosted by each facility located at $j \in \mathcal{J}$
- Demand set $\mathcal{A}$
  - a_{ik}: size of requested product of type $k \in \mathcal{K}$ initiated by demand point $i \in \mathcal{I} \subseteq \mathcal{V}$
  - Total demand over all product types $k \in \mathcal{K}$: $A = \sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} a_{ik}$

Input: Data and Parameters

Data

- Finite graph $G = (\mathcal{V}, \mathcal{E})$ with edge set $\mathcal{E}$ and vertex set $\mathcal{V}$
  - Set of demand originating points $\mathcal{I} \subseteq \mathcal{V}, |\mathcal{I}| = I$
  - Set of potential facility locations $\mathcal{J} \subseteq \mathcal{V}, |\mathcal{J}| = J$
- Set $\mathcal{K}$ ($|\mathcal{K}| = K$): family of products that can be hosted by each facility located at $j \in \mathcal{J}$
- Demand set $\mathcal{A}$
  - a_{ik}: size of requested product of type $k \in \mathcal{K}$ initiated by demand point $i \in \mathcal{I} \subseteq \mathcal{V}$
  - Total demand over all product types $k \in \mathcal{K}$: $A = \sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} a_{ik}$

Parameters

- b_j: capacity of facility opened at location $j \in \mathcal{J}$ (storage capacity)
- q_{uv}: nominal capacity of arc (u, v) from node u to v

Variables and Costs

Variables

- Real variable x_{ijk}: fraction of demand a_{ik} requested by customer demand node i for product type k satisfied/served by facility j (opened/installed at $u \in \mathcal{V}$)
- Binary variable $y_j = 1$ if facility j of capacity b_j opened/installed at node $u \in \mathcal{V}$ (0 otherwise)
- Binary variable $z_{jk} = 1$ if product type k provided at (opened) facility j (0 otherwise)
- Continuous flow variable $f_{uv,ijk}$: amount of traffic flowing on arc (u, v) in supply of customer demand i for product k assigned to opened facility j

Variables and Costs

Variables

- Real variable x_{ijk}: fraction of demand a_{ik} requested by customer demand node i for product type k satisfied/served by facility j (opened/installed at $u \in \mathcal{V}$)
- Binary variable $y_j = 1$ if facility j of capacity b_j opened/installed at node $u \in \mathcal{V}$ (0 otherwise)
- Binary variable $z_{jk} = 1$ if product type k provided at (opened) facility j (0 otherwise)
- Continuous flow variable $f_{uv,ijk}$: amount of traffic flowing on arc (u, v) in supply of customer demand i for product k assigned to opened facility j

Costs

- φ_j: cost of opening/installing a facility at site j
  $\rightarrow$ Facility location cost: $\sum_{j \in \mathcal{J}} \varphi_j y_j$
- κ_{ijk}: cost of assigning to facility opened at site j the fraction of demand a_{ik} issued by customer demand point i for product k
  $\rightarrow$ Demand allocation cost: $\sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \sum_{k \in \mathcal{K}} \kappa_{ijk} x_{ijk}$
- τ_{uv}: cost of routing one unit of traffic along arc (u, v)
  $\rightarrow$ Traffic routing cost: $\sum_{(u,v) \in \mathcal{E}} \tau_{uv} \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \sum_{k \in \mathcal{K}} f_{uv,ijk}$

MIP Formulation

$$\min \sum_{j\in\mathcal{J}} \varphi_j y_j + \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}} \kappa_{ijk} x_{ijk} + \sum_{(u,v)\in\mathcal{E}} \tau_{uv} \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}} f_{uv,ijk} \quad (26)$$

subject to:

$$\sum_{j\in\mathcal{J}} x_{ijk} = 1 \qquad i\in\mathcal{I}, k\in\mathcal{K}, a_{ik}>0 \quad (27)$$

$$z_{jk} \le y_j \qquad j\in\mathcal{J}, k\in\mathcal{K} \quad (28)$$

$$x_{ijk} \le z_{jk} \qquad i\in\mathcal{I}, j\in\mathcal{J}, k\in\mathcal{K} \quad (29)$$

$$\sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{K}} a_{ik} \frac{x_{ijk}}{\sum_{\ell\in\mathcal{L}} x_{\ell jk}} \le b_j y_j \qquad j\in\mathcal{J} \quad (30)$$

$$\sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{K}} a_{ik} \frac{x_{ijk}}{\sum_{\ell\in\mathcal{L}} x_{\ell jk}} \le \sum_{j\in J} b_j y_j \quad (31)$$

$$f_{uv,ijk} \le a_{ik} x_{ijk} \qquad (u,v)\in\mathcal{E}, i\in\mathcal{I}, j\in\mathcal{J}, k\in\mathcal{K} \quad (32)$$

$$\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}} f_{uv,ijk} \le q_{uv} \qquad (u,v)\in\mathcal{E} \quad (33)$$

$$a_{ik} x_{iik} + \sum_{v\in\mathcal{V}:(i,v)\in\mathcal{E}} \sum_{j\in\mathcal{J}} f_{iv,ijk} = a_{ik} \qquad i\in\mathcal{I}, k\in\mathcal{K}, i\neq j, a_{ik}>0 \quad (34)$$

$$\sum_{v:(v,u)\in\mathcal{E}} \sum_{j\in\mathcal{J}} f_{vu,ijk} = \sum_{v:(u,v)\in\mathcal{E}} \sum_{j\in\mathcal{J}} f_{uv,ijk} + a_{ik} x_{iuk} \qquad i\in\mathcal{I}, u\in\mathcal{V}, k\in\mathcal{K}, u\neq i \quad (35)$$

$$x_{ijk} \in [0,1] \qquad i\in\mathcal{I}, j\in\mathcal{J}, k\in\mathcal{K} \quad (36)$$

$$y_j \in \{0,1\} \qquad j\in\mathcal{J} \quad (37)$$

$$z_{jk} \in \{0,1\} \qquad j\in\mathcal{J}, k\in\mathcal{K} \quad (38)$$

$$f_{uv,ijk} \ge 0 \qquad (u,v)\in\mathcal{E}, i\in\mathcal{I}, j\in\mathcal{J}, k\in\mathcal{K} \quad (39)$$

MSMP-cFLRP Constraints (1)

- **Demand satisfaction constraints**: demand a_{ik} for product type k issued by each customer i shall be satisfied:

$$\sum_{j \in \mathcal{J}} x_{ijk} = 1, i \in \mathcal{I}, k \in \mathcal{K}, a_{ik} > 0 \tag{40}$$

- **Product availability**: product type k available on facility j only if j opened Forbids assigning products to closed facilities:

$$z_{jk} \leq y_j, j \in \mathcal{J}, k \in \mathcal{K} \tag{41}$$

- Demand fraction x_{ijk} satisfiable by facility j only if product k available at j Forbids delivery from facility j of product type k to demand node i if product type k unavailable at facility j

$$x_{ijk} \leq z_{jk}, i \in \mathcal{I}, j \in \mathcal{J}, k \in \mathcal{K} \tag{42}$$

MSMP-cFLRP Constraints (2)

Constraints linking MSMP-cFLP and Flow routing problem:

- **Individual flow constraints** on arc (u, v): traffic flow associated to customer i demand for product type k (a_{ik}) directed to facility j along arc (u, v)

$$f_{uv,ijk} \leq \min(q_{uv}, a_{ik}x_{ijk}), (u, v) \in \mathcal{E}, i \in \mathcal{I}, j \in \mathcal{J}, k \in \mathcal{K} \quad (43)$$

- **Aggregated flow constraints** on arc (u, v): load (sum of traffic flows) on individual arcs $(u, v) \in \mathcal{E}$ does not exceed their nominal capacity q_{uv}

$$\sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \sum_{k \in \mathcal{K}} f_{uv,ijk} \leq q_{uv}, (u, v) \in \mathcal{E} \quad (44)$$

- **Flow conservation constraints**:

$$a_{ik}x_{iik} + \sum_{v:(i,v) \in \mathcal{E}} \sum_{j \in \mathcal{J}} f_{iv,ijk} = a_{ik}, i \in \mathcal{I}, k \in \mathcal{K}, i \neq j, a_{ik} > 0 \quad (45)$$

$$\sum_{v:(v,u) \in \mathcal{E}} \sum_{j \in \mathcal{J}} f_{vu,ijk} = \sum_{v:(u,v) \in \mathcal{E}} \sum_{j \in \mathcal{J}} f_{uv,ijk} + x_{iuk}a_{ik}, i \in \mathcal{I}, u \in \mathcal{V}, k \in \mathcal{K}, u \neq i \quad (46)$$

MSMP-cFLRP Constraints (3)

Facility capacity constraints:

- For physical goods (canonical cFLP):

$$\sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} a_{ik} x_{ijk} \leq b_j y_j, \forall j \in \mathcal{J} \qquad (47)$$

MSMP-cFLRP Constraints (3)

Facility capacity constraints:

- For physical goods (canonical cFLP):

$$\sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} a_{ik} x_{ijk} \leq b_j y_j, \forall j \in \mathcal{J} \tag{47}$$

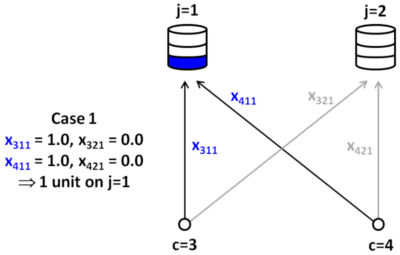
- For **digital goods**:
  - Sum of fractions x_{ijk} assigned to opened facility $j \in \mathcal{J}$ does not exceed its max. capacity b_j
  - Set of d identical demands (same product type k of size s) assigned to j consumes s units of facility capacity at j instead of $d.s$ units

$$\sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} a_{ik} \frac{x_{ijk}}{\sum_{\ell \in \mathcal{L}} x_{\ell jk}} \leq b_j y_j, \forall j \in \mathcal{J} \tag{48}$$

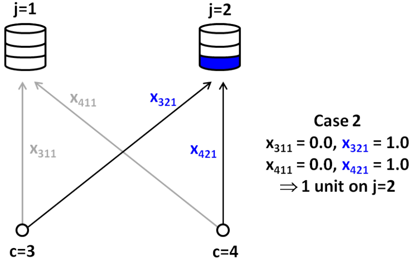
where, $\mathcal{L}(\subseteq \mathcal{I}) \triangleq$ set of identical demands assigned to the same facility j (this set is unknown prior to assignment)

Example (1)

- $\sum_{\ell \in \mathcal{L}} x_{\ell 11} = \sum_{\ell \in \mathcal{L}} x_{311} + x_{411} = 1$
- $\sum_{\ell \in \mathcal{L}} x_{\ell 21} = \sum_{\ell \in \mathcal{L}} x_{321} + x_{421} = 0$



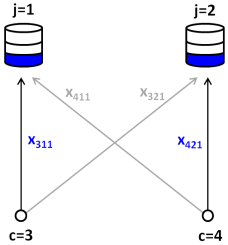
- $\sum_{\ell \in \mathcal{L}} x_{\ell 11} = \sum_{\ell \in \mathcal{L}} x_{311} + x_{411} = 0$
- $\sum_{\ell \in \mathcal{L}} x_{\ell 21} = \sum_{\ell \in \mathcal{L}} x_{321} + x_{421} = 1$



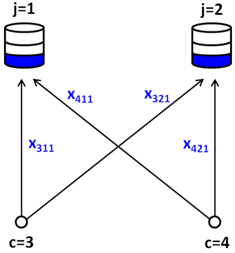
Example (2)

- $\sum_{\ell \in \mathcal{L}} x_{\ell 11} = \sum_{\ell \in \mathcal{L}} x_{311} + x_{411} = 1$
- $\sum_{\ell \in \mathcal{L}} x_{\ell 21} = \sum_{\ell \in \mathcal{L}} x_{321} + x_{421} = 1$

Case 3
$x_{311} = 1$
$x_{321} = 0$
$x_{411} = 0$
$x_{421} = 1$
$\Rightarrow$ 1 unit on j=1 and j=2



- $\sum_{\ell \in \mathcal{L}} x_{\ell 11} = \sum_{\ell \in \mathcal{L}} x_{311} + x_{411} = 1$
- $\sum_{\ell \in \mathcal{L}} x_{\ell 21} = \sum_{\ell \in \mathcal{L}} x_{321} + x_{421} = 1$



Case 4
$x_{311} = 0 < a < 1$
$x_{321} = 0 < 1 - a < 1$
$x_{411} = 0 < 1 - b < 1$
$x_{421} = 0 < b < 1$
$\Rightarrow$ 1 unit on j=1 and j=2

Fractional Constraints (1)

- **Physical goods model**: facility capacity constraints $\sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{K}} a_{ik}x_{ijk} \leq b_j y_j$

Fractional Constraints (1)

- **Physical goods model**: facility capacity constraints $\sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} a_{ik} x_{ijk} \leq b_j y_j$
- **Digital goods model**: capacity sharing between digital objects available on opened facilities leads to fractional term in facility capacity constraints ($\mathcal{L} \subseteq \mathcal{I}$)

$$\sum_{i^* \in \mathcal{I}} \sum_{k \in \mathcal{K}} a_{i^* k} \frac{x_{i^* jk}}{x_{i^* jk} + \sum_{\ell \in \mathcal{L} \setminus \{i^*\}} x_{\ell jk}} \leq b_j y_j \tag{49}$$

Fractional Constraints (1)

- **Physical goods model**: facility capacity constraints $\sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} a_{ik} x_{ijk} \leq b_j y_j$
- **Digital goods model**: capacity sharing between digital objects available on opened facilities leads to fractional term in facility capacity constraints ($\mathcal{L} \subseteq \mathcal{I}$)

$$\sum_{i^* \in \mathcal{I}} \sum_{k \in \mathcal{K}} a_{i^* k} \frac{x_{i^* jk}}{x_{i^* jk} + \sum_{\ell \in \mathcal{L} \setminus \{i^*\}} x_{\ell jk}} \leq b_j y_j \tag{49}$$

  - To linearize these constraints: first define a new variable ξ_{jk} such that

$$\xi_{jk} = \frac{1}{x_{i^* jk} + \sum_{\ell \in \mathcal{L} \setminus \{i^*\}} x_{\ell jk}} \tag{50}$$

  - Condition equivalent to

$$\xi_{jk} \left(x_{i^* jk} + \sum_{\ell \in \mathcal{L} \setminus \{i^*\}} x_{\ell jk} \right) = \sum_{i^* \in \mathcal{L}} \xi_{jk} x_{i^* jk} = 1 \tag{51}$$

  - In terms of ξ_{jk}, facility capacity constraints can then be rewritten as ($i^* \to i$)

$$\sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} a_{ik} \xi_{jk} x_{ijk} \leq b_j y_j \tag{52}$$

$$\sum_{i \in \mathcal{L}} \xi_{jk} x_{ijk} = 1 \tag{53}$$

Fractional Constraints (2)

- Theorem: polynomial mixed term $z = x.y$ ($x \triangleq$ binary variable, $y \triangleq$ continuous variable such that $L \leq y \leq U$) can be represented by linear inequalities:
  1) $Lx \leq z \leq Ux$
  2) $y - U(1-x) \leq z \leq y - L(1-x)$

Fractional Constraints (2)

- Theorem: polynomial mixed term $z = x.y$ ($x \triangleq$ binary variable, $y \triangleq$ continuous variable such that $L \leq y \leq U$) can be represented by linear inequalities:
  1) $Lx \leq z \leq Ux$
  2) $y - U(1-x) \leq z \leq y - L(1-x)$

$\Rightarrow$ Introduce auxiliary variable $\zeta_{ijk} = \xi_{jk} x_{ijk}$, where $L(=0) \leq \xi_{jk} \leq U(=1)$, to obtain:

$$\sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} a_{ik} \zeta_{ijk} \leq b_j y_j \tag{54}$$

$$\sum_{i \in \mathcal{L}} \zeta_{ijk} = 1 \tag{55}$$

$$0 \leq \zeta_{ijk} \leq x_{ijk} \tag{56}$$

$$\xi_{jk} - (1 - x_{ijk}) \leq \zeta_{ijk} \leq \xi_{jk} \tag{57}$$

Fractional Constraints (2)

- Theorem: polynomial mixed term $z = x.y$ ($x \triangleq$ binary variable, $y \triangleq$ continuous variable such that $L \leq y \leq U$) can be represented by linear inequalities:
  1) $Lx \leq z \leq Ux$
  2) $y - U(1-x) \leq z \leq y - L(1-x)$

$\Rightarrow$ Introduce auxiliary variable $\zeta_{ijk} = \xi_{jk} x_{ijk}$, where $L(=0) \leq \xi_{jk} \leq U(=1)$, to obtain:

$$\sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} a_{ik} \zeta_{ijk} \leq b_j y_j \tag{54}$$

$$\sum_{i \in \mathcal{L}} \zeta_{ijk} = 1 \tag{55}$$

$$0 \leq \zeta_{ijk} \leq x_{ijk} \tag{56}$$

$$\xi_{jk} - (1 - x_{ijk}) \leq \zeta_{ijk} \leq \xi_{jk} \tag{57}$$

Linearization

- Increases complexity: addition of $(I+1).J.K$ auxiliary variables ζ_{ijk} and ξ_{jk} together with $(4.I+1).J.K$ associated constraints
- Works for small-size problems but gap between IP and LP relaxation may become huge for larger problems
- Set $\mathcal{L}$ a priori unknown

Approximation (1)

Facility capacity constraints: $\sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} a_{ik} \frac{x_{ijk}}{\sum_{\ell \in \mathcal{L}} x_{\ell jk}} \leq b_j y_j, \forall j \in \mathcal{J}$

- Explicit dependence on product index k in LHS prevents per-product formulation
- Capacity sharing among K product types leads to more complex structure than superposition of K independent constraints

Approximation (1)

Facility capacity constraints: $\sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{K}} a_{ik}\frac{x_{ijk}}{\sum_{\ell\in\mathcal{L}} x_{\ell jk}} \leq b_j y_j, \forall j \in \mathcal{J}$

- Explicit dependence on product index k in LHS prevents per-product formulation
- Capacity sharing among K product types leads to more complex structure than superposition of K independent constraints

Approximation

- Start from facility capacity constraints formulated as for single-product model ($K = 1$):

$$\sum_{i\in\mathcal{I}} a_i \frac{x_{ij}}{\sum_{\ell\in\mathcal{L}} x_{\ell j}} \leq b_j y_j, j \in \mathcal{J}$$

- Move denominator out of LHS: $\sum_{i\in\mathcal{I}} a_i x_{ij} \leq b_j \sum_{i\in\mathcal{L}} y_j x_{ij}, j \in \mathcal{J}$
- Assume inequality verified for each k independently (dedicated capacity per-product type):

$$\sum_{i\in\mathcal{I}} a_{ik} x_{ijk} \leq b_{jk} \sum_{i\in\mathcal{L}} y_j x_{ijk}, j \in \mathcal{J}, k \in \mathcal{K}$$

- Re-introduce summation over k (in both members):

$$\sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{K}} a_{ik} x_{ijk} \leq \sum_{k\in\mathcal{K}}(b_{jk} \sum_{i\in\mathcal{L}} y_j x_{ijk}), j \in \mathcal{J}$$

Approximation (2)

- Transformation removes fractional term (LHS) but introduces sum over individual product capacity (b_{jk})

  $\Rightarrow$ Question: Gain from this transformation ?

Approximation (2)

- Transformation removes fractional term (LHS) but introduces sum over individual product capacity (b_{jk})

  $\Rightarrow$ Question: Gain from this transformation ?

- **Assumption**: product types homogeneously distributed among installed facilities $\rightarrow b_j = Kb_{jk}$ (remove dependence on per-product capacity distribution)

  $\Rightarrow$ Inequalities for facility capacity constraints (80) when $\mathcal{L} \rightarrow \mathcal{I}$: identical demands assigned to same facility j

$$\sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} a_{ik} x_{ijk} \leq \frac{1}{K} b_j y_j \sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} x_{ijk}, \forall j \in \mathcal{J} \quad (58)$$

  $\Rightarrow$ Inequalities for facility capacity constraints (80) when $|\mathcal{L}| \rightarrow 1$: each product type-size pair assigned to single demand

$$\sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} a_{ik} x_{ijk} \leq \frac{1}{K} b_j y_j \sum_{k \in \mathcal{K}} x_{*jk}, \forall j \in \mathcal{J} \quad (59)$$

Approximation (3)

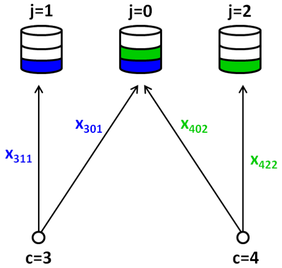
- **Scenario**: Set of disjoint demands wrt product type k of same size s: pairs $(k_1, s), (k_2, s), \dots, (k_K, s)$
  With $K = I$ pairs (one per demand point i): total capacity required $= K.s$
- If $b_j = s$ and facility installation cost low enough to steer local assignment
  Then demands initiated locally should be assigned locally
  $\Rightarrow$ Routing cost should be zero

Approximation (3)

- **Scenario**: Set of disjoint demands wrt product type k of same size s: pairs $(k_1, s), (k_2, s), \ldots, (k_K, s)$
  With $K = I$ pairs (one per demand point i): total capacity required $= K.s$
- If $b_j = s$ and facility installation cost low enough to steer local assignment
  Then demands initiated locally should be assigned locally
  $\Rightarrow$ Routing cost should be zero
- Not verified because per-facility capacity b_j divided by factor K
  $\Rightarrow$ Capacity required on at least one installed facility multiplied by factor $K (= I)$

Approximation (3)

- **Scenario**: Set of disjoint demands wrt product type k of same size s: pairs $(k_1, s), (k_2, s), \ldots, (k_K, s)$
  With $K = I$ pairs (one per demand point i): total capacity required $= K.s$
- If $b_j = s$ and facility installation cost low enough to steer local assignment
  Then demands initiated locally should be assigned locally
  $\Rightarrow$ Routing cost should be zero
- Not verified because per-facility capacity b_j divided by factor K
  $\Rightarrow$ Capacity required on at least one installed facility multiplied by factor $K (= I)$



Case:
x_{311} = a $\rightarrow$ 1
x_{301} = 1 – a $\rightarrow$ 0 (routing cost $\rightarrow$ 0)
x_{422} = b $\rightarrow$ 1
x_{402} = 1 – b $\rightarrow$ 0 (routing cost $\rightarrow$ 0)
$\Rightarrow$ **1 unit on j=1 and 1 unit on j=2**
+ 1 unit on j = 3 (almost unused)

With K products of same size s: over-dimensioning up to factor K

Additional Constraints

Consider simplified objective:

$$\min \sum_{j \in \mathcal{J}} \varphi_j y_j + \sum_{(u,v) \in \mathcal{E}} \tau_{uv} \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \sum_{k \in \mathcal{K}} f_{uv,ijk} \tag{60}$$

Additional Constraints

Consider simplified objective:

$$\min \sum_{j \in \mathcal{J}} \varphi_j y_j + \sum_{(u,v) \in \mathcal{E}} \tau_{uv} \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \sum_{k \in \mathcal{K}} f_{uv,ijk} \tag{60}$$

with additional constraints:

- Aggregated capacity constraints $\sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} a_{ik} \leq \frac{1}{K} \sum_j b_j y_j \sum_i \sum_k x_{ijk}$
- Individual fractions remain within $[0, 1]$, i.e., $0 \leq x_{ijk} \leq 1$
- At least one facility shall be opened $\sum_{j \in J} y_j \geq 1$
  Particular case ($b_j = b, \forall j$): divide total demand size by per-facility capacity b_j such that min.number of facilities $\leq \sum_{j \in J} y_j$
- All product types covered by installed facilities $\sum_{j \in J} \sum_{k \in \mathcal{K}} z_{jk} \geq K$

MSMP-cFLRP: MIP Formulation

$$\min \sum_{j\in\mathcal{J}} \varphi_j y_j + \sum_{(u,v)\in\mathcal{E}} \tau_{uv} \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \sum_{k\in\mathcal{K}} f_{uv,ijk} \tag{61}$$

subject to:

$$\sum_{j\in\mathcal{J}} x_{ijk} = 1 \qquad i\in\mathcal{I}, k\in\mathcal{K} \tag{62}$$

$$z_{jk} \le y_j \qquad j\in\mathcal{J}, k\in\mathcal{K} \tag{63}$$

$$x_{ijk} \le z_{jk} \qquad i\in\mathcal{I}, j\in\mathcal{J}, k\in\mathcal{K} \tag{64}$$

$$\sum_{i\in\mathcal{I}} \sum_{k\in\mathcal{K}} a_{ik} x_{ijk} \le \frac{1}{K} b_j y_j \sum_{i\in\mathcal{I}} \sum_{k\in\mathcal{K}} x_{ijk} \qquad j\in\mathcal{J} \tag{65}$$

$$\sum_{i\in\mathcal{I}} \sum_{k\in\mathcal{K}} a_{ik} \le \frac{1}{K} \sum_{j\in\mathcal{J}} b_j y_j \sum_{i\in\mathcal{I}} \sum_{k\in\mathcal{K}} x_{ijk} \tag{66}$$

$$f_{uv,ijk} \le a_{ik} x_{ijk} \qquad (u,v)\in\mathcal{E}, i\in\mathcal{I}, j\in\mathcal{J}, k\in\mathcal{K} \tag{67}$$

$$\sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \sum_{k\in\mathcal{K}} f_{uv,ijk} \le q_{uv} \qquad (u,v)\in\mathcal{E} \tag{68}$$

$$a_{ik} x_{iik} + \sum_{v\in\mathcal{V}:(i,v)\in\mathcal{E}} \sum_{j\in\mathcal{J}} f_{iv,ijk} = a_{ik} \qquad i\in\mathcal{I}, k\in\mathcal{K}, i\ne j, a_{ik} > 0 \tag{69}$$

$$\sum_{v:(v,u)\in\mathcal{E}} \sum_{j\in\mathcal{J}} f_{vu,ijk} = \sum_{v:(u,v)\in\mathcal{E}} \sum_{j\in\mathcal{J}} f_{uv,ijk} + a_{ik} x_{iuk} \qquad i\in\mathcal{I}, u\in\mathcal{V}, k\in\mathcal{K}, u\ne i \tag{70}$$

$$x_{ijk} \in [0,1] \qquad i\in\mathcal{I}, j\in\mathcal{J}, k\in\mathcal{K} \tag{71}$$

$$y_j \in \{0,1\} \qquad j\in\mathcal{J} \tag{72}$$

$$z_{jk} \in \{0,1\} \qquad j\in\mathcal{J}, k\in\mathcal{K} \tag{73}$$

$$f_{uv,ijk} \ge 0 \qquad (u,v)\in\mathcal{E}, i\in\mathcal{I}, j\in\mathcal{J}, k\in\mathcal{K} \tag{74}$$

Performance benchmark

Goals

- Computational performance evaluation (computational time and solution quality) using `CPLEX 12.6.3`
- Target computational time upper bound of 900s (average roll-out time)

Performance benchmark

Goals

- Computational performance evaluation (computational time and solution quality) using `CPLEX 12.6.3`
- Target computational time upper bound of 900s (average roll-out time)

Method

- Generate set of 12 instances with O(1000) demands (at least O(100) demands per node)
- Network topology of 25 nodes and 90 arcs
- Tuning facility capacity and associated costs

Performance benchmark

Goals

- Computational performance evaluation (computational time and solution quality) using `CPLEX 12.6.3`
- Target computational time upper bound of 900s (average roll-out time)

Method

- Generate set of 12 instances with O(1000) demands (at least O(100) demands per node)
- Network topology of 25 nodes and 90 arcs
- Tuning facility capacity and associated costs

Execution

- Concurrent (Dual simplex and Barrier algorithm) to solve root relaxation (`rootalg = 6`)
- Concurrent (Dual simplex and Barrier algorithm) to solve other MIP subproblems after initial relaxation (`nodealg = 6`)
- Balance feasibility and optimality (`mipemphasis = 1`)

Performance benchmark: results

Scenario	Root Sol. time	Root Proc. time (s)	Total Proc. time (s)	Final Gap (%)
sc-0k75-0k75	84	346	346	0.00
sc-1k-1k	80	340	340	0.00
sc-1k2-1k2	86	346	346	0.00
sc-1k5-1k5	124	383	383	0.00
sc-1k8-1k8	132	392	392	0.00
sc-2k-2k	209	754	754	0.00
sc-2k25-2k25	589	1329	2099	0.00
sc-3k-3k	140	2375	3893	0.00
sc-3k6-3k6	1950	3808	4978	0.00
sc-4k5-4k5	2011	4935	4935	0.00
sc-6k-6k	3514	7310	9782	0.00
sc-9k-9k	5307	9332	9332	0.00
Avg	1101	2461	3048	0.00
Stdev	1660	3019	3744	0.00

Scenario	Root Sol. time	Root Proc. time (s)	Total Proc. time (s)	Final Gap (%)
sc-0k75-0k75	79	336	336	0.00
sc-1k-2k	84	342	342	0.00
sc-1k2-2k	85	343	343	0.00
sc-1k5-2k	128	386	386	0.00
sc-1k8-2k	130	386	386	0.00
sc-2k-2k	207	733	733	0.00
sc-2k25-2k	572	1350	1911	0.00
sc-3k-2k	1390	2407	3581	0.00
sc-3k6-2k	2091	3253	3253	0.00
sc-4k5-2k	1493	2691	2691	0.00
sc-6k-2k	1034	2055	2055	0.00
sc-9k-2k	1047	1823	1823	0.00
Avg	648	1265	1399	0.00
Stdev	688	1060	1215	0.00

Evaluation instances: topologies and demands

- Topologies (SNDLib database)

Topology	Nodes	Links	Min,Max,Avg Degree	Diameter
abilene	12	15	1;4;2.50	3
atlanta	15	22	2;4;2.93	3
france	25	45	2;10;3.60	8
geant	22	36	2;8;3.27	5
germany50	50	88	2;5;3.52	9
india35	35	80	2;9;4.57	7
newyork	16	49	2;11;6.12	2
norway	27	51	2;6;3.78	7

- Links capacity and cost from SNDlib database

Evaluation instances: topologies and demands

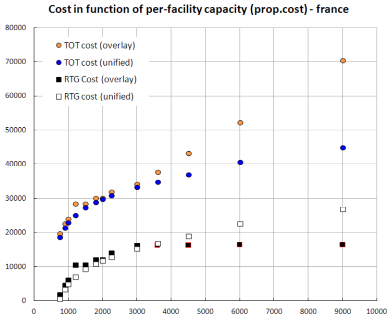
- Topologies (SNDLib database)

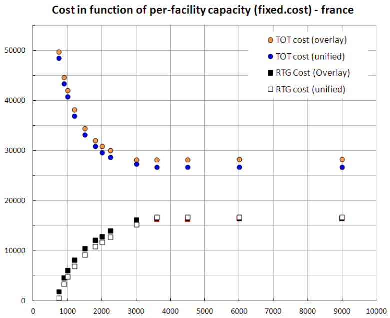
Topology	Nodes	Links	Min,Max,Avg Degree	Diameter
abilene	12	15	1;4;2.50	3
atlanta	15	22	2;4;2.93	3
france	25	45	2;10;3.60	8
geant	22	36	2;8;3.27	5
germany50	50	88	2;5;3.52	9
india35	35	80	2;9;4.57	7
newyork	16	49	2;11;6.12	2
norway	27	51	2;6;3.78	7

- Links capacity and cost from SNDlib database
- Demands
  - Produce set of ten problem instances with 3000 demands
  - Demands generated using following distributions:
    - **Demand size**: Pareto distribution commonly used to model file size $f(x) = \frac{\alpha x_m^\alpha}{x^{\alpha+1}},\ x \geq x_m$
    - **Demand frequence**: Generalized Zipf-Mandelbrot distribution (frequency of event occurrence inversely proportional to its rank)

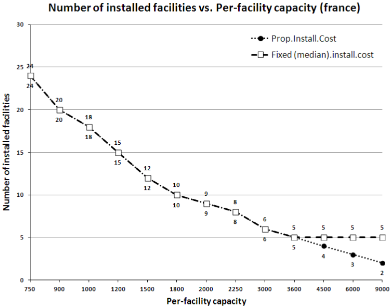
Results: Unified vs. Overlay

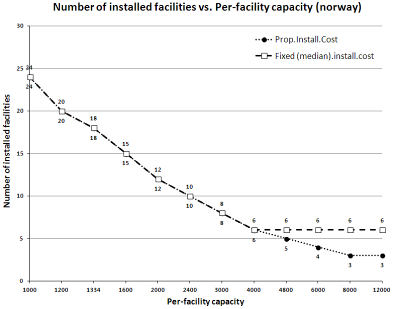
- Overlay (sequential): minimize location and allocation cost (cFLP) then routing cost (MMCF)
- Unified (simultaneous): minimize location, allocation and routing cost (MSMP-cFLRP)



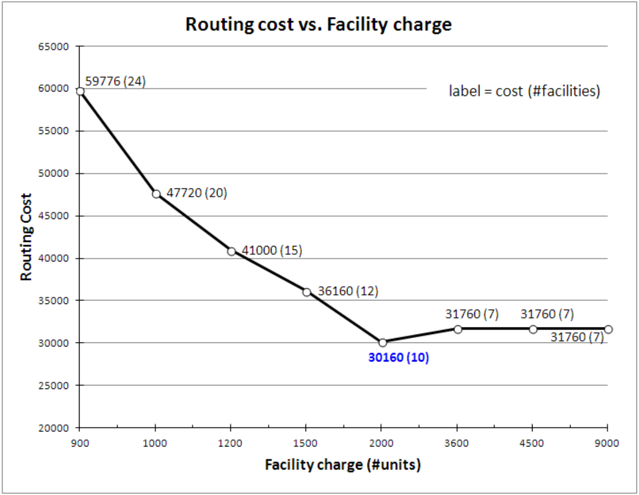


Results: Number of Facilities vs. (Per-)Facility Capacity

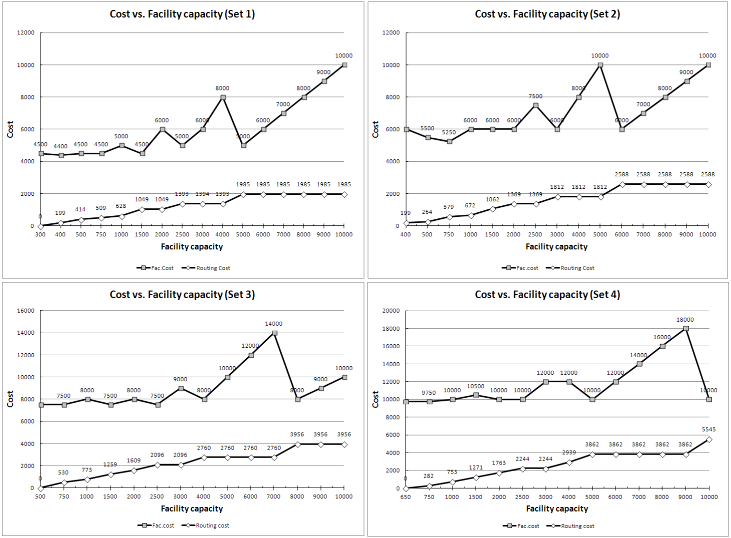




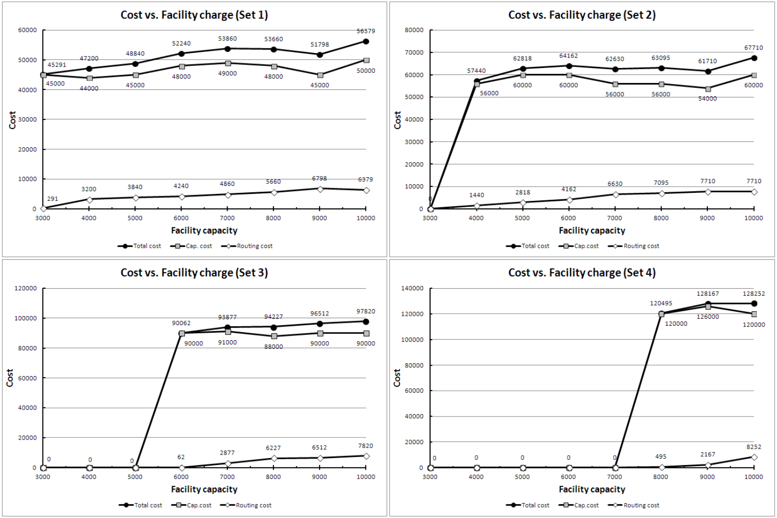
Results: Routing Cost vs. Facility Charge



Deeper look (1): Digital goods model (atlanta)

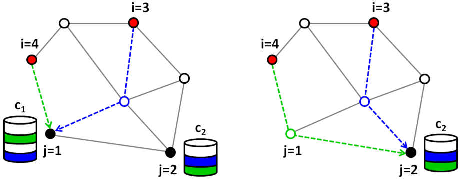


Deeper look (2): Physical goods model (atlanta)



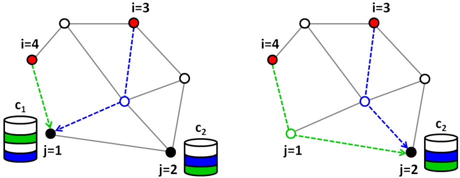
Reliable Facility Location

- **Facility protection**: when choosing a location for a facility, another facility is selected which will serve as its backup when the primary facility fails
  $\Rightarrow$ Demands assigned to same primary facility have same backup facility

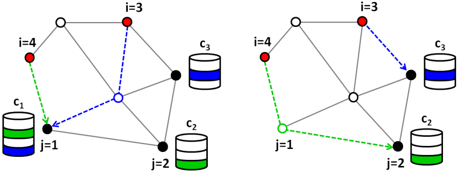


Reliable Facility Location

- **Facility protection**: when choosing a location for a facility, another facility is selected which will serve as its backup when the primary facility fails
  ⇒ Demands assigned to same primary facility have same backup facility

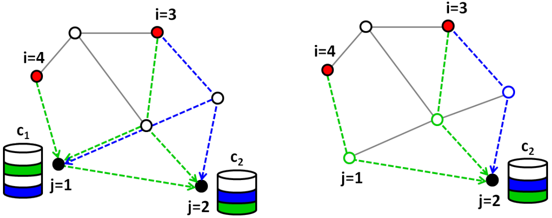


- **Demand protection**: when choosing an allocation for a demand, another facility is assigned which will serve as its backup when the primary facility fails
  ⇒ Demands assigned to same primary facility may have different backup facilities



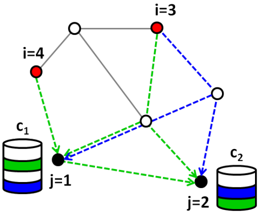
Reliable Facility Location ((c)RFLP) vs. MSMP-cFLRP

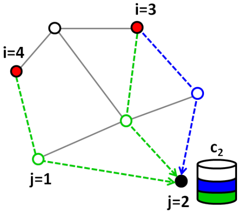
- Demands protection: RFLP (Snyder2005) and capacitated RFLP (Yu2015)



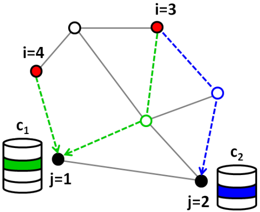
Reliable Facility Location ((c)RFLP) vs. MSMP-cFLRP

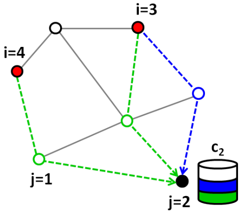
- Demands protection: RFLP (Snyder2005) and capacitated RFLP (Yu2015)





- **Demands rerouting**: MSMP-cFLRP





Demand protection: capacitated Reliable Fixed Charge Location Problem (cRFLP)

- Reliability based on **levels assignments** strategy: r ($r = 0, \ldots, J-1$) level at which a facility serves a given customer demand
  - r=0: primary assignment
  - r=1: first backup
  - and so on

- Reliability based on **levels assignments** strategy: r ($r = 0, \ldots, J-1$) level at which a facility serves a given customer demand
  - r=0: primary assignment
  - r=1: first backup
  - and so on
- If customer i demand level$-r$ assigned facility failed then level$-(r+1)$ assigned facility serves this demand as backup

- Reliability based on **levels assignments** strategy: r ($r = 0, \ldots, J-1$) level at which a facility serves a given customer demand
  - r=0: primary assignment
  - r=1: first backup
  - and so on
- If customer i demand level$-r$ assigned facility failed then level$-(r+1)$ assigned facility serves this demand as backup
- Objective function:

$$\sum_{j \in \mathcal{J}} \varphi_j y_j + \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \sum_{k \in \mathcal{K}} \sum_{r=0}^{J-1} d_{ij} a_{ijk} x_{ijkr} q^r (1-q) \tag{75}$$

- Reliability based on **levels assignments** strategy: r ($r = 0, \ldots, J-1$) level at which a facility serves a given customer demand
  - r=0: primary assignment
  - r=1: first backup
  - and so on
- If customer i demand level$-r$ assigned facility failed then level$-(r+1)$ assigned facility serves this demand as backup
- Objective function:

$$\sum_{j \in \mathcal{J}} \varphi_j y_j + \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \sum_{k \in \mathcal{K}} \sum_{r=0}^{J-1} d_{ij} a_{ijk} x_{ijkr} q^r (1-q) \tag{75}$$

  - First term: total fixed installation cost
  - Second term: expected transport cost where facility j serves customer i demand if
    - its lower-level assigned facilities all disrupted: occurrence probability q^r
    - and facility j still available: occurrence probability $1-q$

Reliable MSMP-cFLRP: MIP Formulation

$$\min \sum_{j\in\mathcal{J}} \varphi_j y_j + \sum_{(u,v)\in\mathcal{E}} \tau_{uv} \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \sum_{k\in\mathcal{K}} f_{uv,ijk} \tag{76}$$

subject to:

$$\sum_{j\in\mathcal{J}} x_{ijk} = 1 \qquad i\in\mathcal{I}, k\in\mathcal{K} \tag{77}$$

$$z_{jk} \le y_j(1-q_j) \qquad j\in\mathcal{J}, k\in\mathcal{K} \tag{78}$$

$$x_{ijk} \le z_{jk} \qquad i\in\mathcal{I}, j\in\mathcal{J}, k\in\mathcal{K} \tag{79}$$

$$\sum_{i\in\mathcal{I}} \sum_{k\in\mathcal{K}} a_{ik} x_{ijk} \le \frac{1}{K} b_j y_j (1-q_j) \sum_{i\in\mathcal{I}} \sum_{k\in\mathcal{K}} x_{ijk} \qquad j\in\mathcal{J} \tag{80}$$

$$\sum_{i\in\mathcal{I}} \sum_{k\in\mathcal{K}} a_{ik} \le \frac{1}{K} \sum_{j\in\mathcal{J}} b_j y_j (1-q_j) \sum_{i\in\mathcal{I}} \sum_{k\in\mathcal{K}} x_{ijk} \tag{81}$$

$$f_{uv,ijk} \le a_{ik} x_{ijk} \qquad (u,v)\in\mathcal{E}, i\in\mathcal{I}, j\in\mathcal{J}, k\in\mathcal{K} \tag{82}$$

$$\sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \sum_{k\in\mathcal{K}} f_{uv,ijk} \le q_{uv} \qquad (u,v)\in\mathcal{E} \tag{83}$$

$$a_{ik} x_{iik} + \sum_{v\in\mathcal{V}:(i,v)\in\mathcal{E}} \sum_{j\in\mathcal{J}} f_{iv,ijk} = a_{ik} \qquad i\in\mathcal{I}, k\in\mathcal{K}, i\ne j, a_{ik} > 0 \tag{84}$$

$$\sum_{v:(v,u)\in\mathcal{E}} \sum_{j\in\mathcal{J}} f_{vu,ijk} = \sum_{v:(u,v)\in\mathcal{E}} \sum_{j\in\mathcal{J}} f_{uv,ijk} + a_{ik} x_{iuk} \qquad i\in\mathcal{I}, u\in\mathcal{V}, k\in\mathcal{K}, u\ne i \tag{85}$$

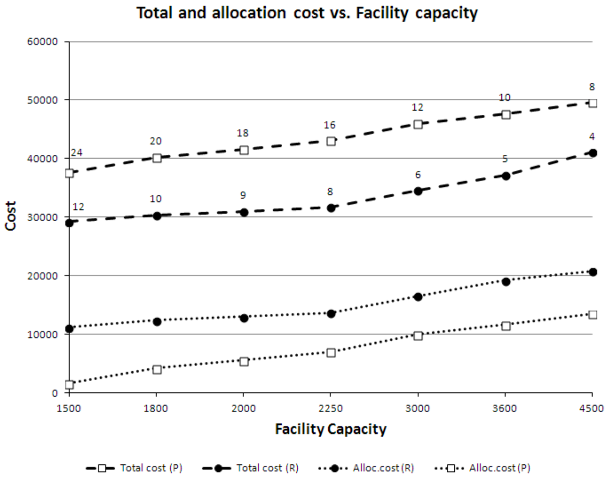
$$x_{ijk} \in [0,1] \qquad i\in\mathcal{I}, j\in\mathcal{J}, k\in\mathcal{K} \tag{86}$$

$$y_j \in \{0,1\} \qquad j\in\mathcal{J} \tag{87}$$

$$z_{jk} \in \{0,1\} \qquad j\in\mathcal{J}, k\in\mathcal{K} \tag{88}$$

$$f_{uv,ijk} \ge 0 \qquad (u,v)\in\mathcal{E}, i\in\mathcal{I}, j\in\mathcal{J}, k\in\mathcal{K} \tag{89}$$

Results: Demand Protection (cRFLP) vs. Rerouting (MSMP-cFLRP)



Main observations (france)

- As facility capacity increases, total cost (R) of re-routing strategy remains lower than total cost (P) of protection strategy (two levels of protection)
  - Higher allocation cost required by cRFLP compared to MSMP-cFLRP because of smaller number of installed facilities
  - Higher routing cost required by MSMP-cFLRP because of load-dependent routing cost instead of graph distance cost

Main observations (france)

- As facility capacity increases, total cost (R) of re-routing strategy remains lower than total cost (P) of protection strategy (two levels of protection)
  - Higher allocation cost required by cRFLP compared to MSMP-cFLRP because of smaller number of installed facilities
  - Higher routing cost required by MSMP-cFLRP because of load-dependent routing cost instead of graph distance cost
- Highest gain (36%) obtained when tradeoff between spatial distribution of facility capacity (over 8 locations) and routing cost to access them reaches its optimal value

Main observations (france)

- As facility capacity increases, total cost (R) of re-routing strategy remains lower than total cost (P) of protection strategy (two levels of protection)
  - Higher allocation cost required by cRFLP compared to MSMP-cFLRP because of smaller number of installed facilities
  - Higher routing cost required by MSMP-cFLRP because of load-dependent routing cost instead of graph distance cost
- Highest gain (36%) obtained when tradeoff between spatial distribution of facility capacity (over 8 locations) and routing cost to access them reaches its optimal value
- Implication: routing metric would require accounting from facility load distribution and data availability

Summary

Summary

- Propose a mixed-integer formulation for combined multi-source multi-product capacitated facility location-flow routing problem (MSMP-cFLFRP)
- Our formulation accounts for specifics of digital object storage and supply
  Note: known formulations translate multi-product problem as single-commodity problem solved separately for each product
- Approximation of fractional constraints enables to solve to optimality small- to medium-size instances with an order of thousands of demands
- Exploitation in demand assignment re-routing scheme (comparison to demand protection scheme)

Future Research Work

Method/Computational Level

- Improve computation method to avoid excessive computation time on (very) large network instances with order of 10k demands

Future Research Work

Method/Computational Level

- Improve computation method to avoid excessive computation time on (very) large network instances with order of 10k demands

Formulation/Modeling Level

- Quadratic assignment (instead of linear assignment): $x_{ijk} \to x_{ijk}^2$
- Multi-period formulation (account for demand dynamics)

Outline

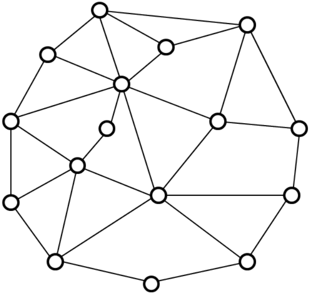
Hub Location Problem (HLP)

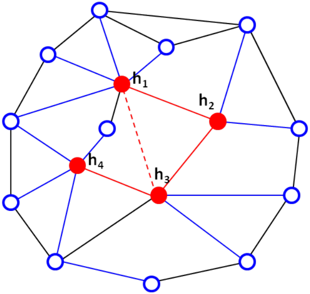
- Undirected graph G with node set V with flow between every pair $(u, v) \in \mathcal{V}$ of nodes
- Subset of central nodes acting as transshipment nodes (hubs); other (terminal or non-hub) nodes connected with an arc (spoke) starlike with one of the hubs
- Flows (u, v) travel directly if both nodes are hubs $(u, v \in \mathcal{H})$ or if one node is a hub and both are connected through a spoke
  Otherwise flow travels via at least another hub h

Introduction: Hub Location Routing Problem (HLRP)

Hub Location Problem (HLP)

- Undirected graph G with node set V with flow between every pair $(u, v) \in \mathcal{V}$ of nodes
- Subset of central nodes acting as transshipment nodes (hubs); other (terminal or non-hub) nodes connected with an arc (spoke) starlike with one of the hubs
- Flows (u, v) travel directly if both nodes are hubs $(u, v \in \mathcal{H})$ or if one node is a hub and both are connected through a spoke
  Otherwise flow travels via at least another hub h





Introduction: Hub Location Routing Problem (HLRP)

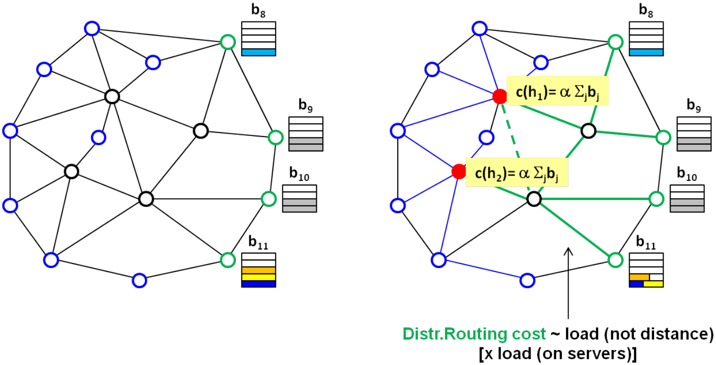
Hub functions

- Connect demand points $i \in \mathcal{I}$
- Demand a_i de/multiplexing (first level)
- Logical composition and/or aggregation of physical of resources from facilities (second level) of finite capacity b_j

Introduction: Hub Location Routing Problem (HLRP)

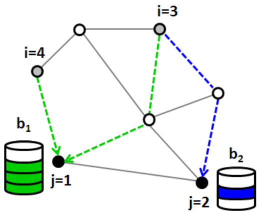
Hub functions

- Connect demand points $i \in \mathcal{I}$
- Demand a_i de/multiplexing (first level)
- Logical composition and/or aggregation of physical of resources from facilities (second level) of finite capacity b_j



Model (1)

- **Objective**: quantitatively assess the tradeoffs between resource abstraction, (al)location, and routing
- **Model**: Combines HLP for demands allocation and cFLP (together with flow routing) for their distribution to multiple facilities
  - Hubs equipped with resource abstraction and aggregation functionality, may split incoming demands over multiple facilities
  - Individual demands d_i are assigned to single hub h offering logical capacity c_h
  - Single hub h may segment demands depending on capacity distribution and consumption at each facility
- **Example**: processing of incoming client demands at a given hub h (red circle) in comparison to cFLP model with single-assignment



Model (2)

- **Single hub-level**: no inter-hub flows but instead hub-to-facility flows
- **Hard location**
  - a single facility of finite capacity may be located at each site
  - a single hub may be located at each site; a given site may either host a facility or a hub (but not both)
- **Two-level**
  - First level: client demands assigned to a single hub
  - Second level: each hub connected to subset of sites where facilities are installed
- **Resource abstraction**: **logical capacity** associated to hubs
  - Minimum $\equiv$ capacity of single facility
  - Maximum (theory) $\equiv$ sum of capacities of all installed facilities
  - In practice: equal distribution of facility capacity between a pair of hubs (minimum level of reliability)
- **Hybrid assignments**: client demands are allocated to a single hub (single-source/-assignment) which can then fraction these demands among multiple facilities (multi-source/-assignment) located at different sites

Model (3)

- **Routing cost**
  - Between demand points and hubs: follow standard graph (hop-count) metric
  - Between hubs and facilities follow minimum cost multi-commodity flow problem: dynamic (re-)allocation of demands to different facilities depending on available capacity on servers they host
- **Combined problem**: facility location (and dimensioning their capacity for customer allocation purposes) but also routing of set of flows corresponding to demands originated by individual customers to set of facilities via single hub h
- Comparisons at two levels depending on i) metric selected and ii) installation of hubs (or not)

Formulation: Data and Parameters

Given finite directed graph $G = (\mathcal{V}, \mathcal{E})$

- $\mathcal{I} \subseteq \mathcal{V}$, ($|\mathcal{I}| = I$): set of client demand points/nodes
- $\mathcal{J} \subseteq \mathcal{V}$, ($|\mathcal{J}| = J$): set of potential locations (or sites) where to host a facility of finite capacity b_j
- $\mathcal{H} \subseteq \mathcal{V}$, ($|\mathcal{H}| = H$): set of potential locations candidate for hosting a hub

  Note: a given location can either host a hub or a facility but not both, $\mathcal{H} \bigcap \mathcal{J} = \emptyset$

Formulation: Data and Parameters

Given finite directed graph $G = (\mathcal{V}, \mathcal{E})$

- $\mathcal{I} \subseteq \mathcal{V}$, $(|\mathcal{I}| = I)$: set of client demand points/nodes
- $\mathcal{J} \subseteq \mathcal{V}$, $(|\mathcal{J}| = J)$: set of potential locations (or sites) where to host a facility of finite capacity b_j
- $\mathcal{H} \subseteq \mathcal{V}$, $(|\mathcal{H}| = H)$: set of potential locations candidate for hosting a hub

  Note: a given location can either host a hub or a facility but not both, $\mathcal{H} \bigcap \mathcal{J} = \emptyset$

Data and Parameters

- Nominal capacity $\kappa_{(u,v)}$ of each arc $(u, v) \in \mathcal{E}$ directed from node u to v
- Demand set $\mathcal{A} = \{a_i\}$ where a_i = size of demand initiated by demand point $i \in \mathcal{I} \subseteq \mathcal{V}$
- Total demand $A = \sum_{i \in \mathcal{I}} a_i$

  Note: in comparison to canonical flow routing problems, demand described by source initiating point but obviously not its destination

Formulation: Variables

Variables

- Real variable $x_{hj} \geq 0$: aggregated amount of traffic which has to be transferred from the hub h to the facility located at site $j \in J$
- Binary variable $y_j = 1$ if facility with capacity b_j opened/installed at location $j \in \mathcal{J}$ and 0 otherwise
- Binary variable $z_{ih} = 1$ if customer demand point i assigned to hub $h \in \mathcal{H}$ and 0 otherwise

  Note: when $i = h$, variable z_{ih} represents installation $(= 1)$ or not $(= 0)$ of a hub at location $h \in \mathcal{H}$
- Real variable $f_{h(u,v)j} \geq 0$: amount of (aggregated) traffic flowing along arc $(u, v) \in \mathcal{E}$ from hub h to facility j

Formulation: Costs and Objective

Costs

- Hub installation cost η_h of installing a hub at location $h \in \mathcal{H}$
- Facility installation cost φ_j of installing a facility at location $j \in \mathcal{J}$
- Routing cost comprises
  1. Cost of routing traffic associated to demand d_i originated by demand point i to hub installed at location $h \in \mathcal{H}$
     Set proportionally to graph distance $d(i, h)$ from i to h, i.e., $\delta_{ih} a_i$
  2. Cost $\tau_{(u,v)}$ of routing one unit of aggregated traffic from hub installed at location $h \in \mathcal{H}$ to facility located at site $j \in \mathcal{J}$
- Solution cost = sum of i) hub location cost, ii) facility location cost, and iii) routing cost of customers demands to a subset of installed facilities via a single installed hub

Objective

Combined problem consists in minimizing sum of all costs while satisfying demand requirements and facility capacity constraints

Formulation: MIP Formulation

$$\min \sum_{h\in\mathcal{V}} \eta_h z_{hh} + \sum_{j\in\mathcal{V}} \varphi_j y_j + \sum_{i\in\mathcal{V}}\sum_{h\in\mathcal{V}} \delta_{ih} a_i z_{ih} + \sum_{(u,v)\in\mathcal{E}}\sum_{h\in\mathcal{V}}\sum_{j\in\mathcal{V}} \tau_{(u,v)} f_{h(u,v)j} \quad (90)$$

subject to:

$$\sum_{h\in\mathcal{V}} z_{ih} = 1 \qquad i\in\mathcal{V} \quad (91)$$

$$z_{ih} \leq z_{hh} \qquad i\in\mathcal{V}, h\in\mathcal{V} \quad (92)$$

$$y_h + z_{hh} \leq 1 \qquad h\in\mathcal{V} \quad (93)$$

$$\sum_{i\in\mathcal{V}} a_i z_{ih} = \sum_{j\in\mathcal{V}} x_{hj} \qquad h\in\mathcal{V} \quad (94)$$

$$\sum_{i\in\mathcal{V}} a_i z_{ih} \leq (\alpha \sum_{j\in\mathcal{V}} b_j y_j) z_{hh} \qquad h\in\mathcal{V} \quad (95)$$

$$\sum_{h\in\mathcal{V}} x_{hj} \leq b_j y_j \qquad j\in\mathcal{V} \quad (96)$$

$$f_{h(u,v)j} \leq x_{hj} \qquad h\in\mathcal{V}, (u,v)\in\mathcal{E}, j\in\mathcal{V} \quad (97)$$

$$\sum_{h\in\mathcal{V}}\sum_{j\in\mathcal{V}} f_{h(u,v)j} \leq \kappa_{(u,v)} \qquad (u,v)\in\mathcal{E} \quad (98)$$

$$\sum_{j\in\mathcal{V}}\sum_{v:(v,u)\in\mathcal{E}} f_{h(v,u)j} = x_{hu} + \sum_{j\in\mathcal{V}}\sum_{v:(u,v)\in\mathcal{A}} f_{h(u,v)j} \qquad h\in\mathcal{V}, u\in\mathcal{V}, h\neq u \quad (99)$$

$$\sum_{j\in\mathcal{V}} x_{hj} = \sum_{j\in\mathcal{V}}\sum_{v:(h,v)\in\mathcal{E}} f_{h(h,v)j} \qquad h\in\mathcal{V} \quad (100)$$

$$f_{h(u,v)j} \geq 0 \qquad h\in\mathcal{V}, (u,v)\in\mathcal{A}, j\in\mathcal{V} \quad (101)$$

$$x_{hj} \geq 0 \qquad h\in\mathcal{V}, j\in\mathcal{V} \quad (102)$$

$$y_j \in \{0,1\} \qquad j\in\mathcal{V} \quad (103)$$

$$z_{ih} \in \{0,1\} \qquad i\in\mathcal{V}, h\in\mathcal{V} \quad (104)$$

Formulation: Constraints (1)

- **Demand satisfaction** constraints $\sum_{h \in \mathcal{V}} z_{ih} = 1$ (91) together with $z_{ih} \in \{0, 1\}$ (104): every demand a_i can be satisfied by reaching a single hub (single-source assignment)

  In turn, every demand a_i can be satisfied by set of installed facilities (provided demand point i connected to single hub h)
- Inequalities $z_{ih} \leq z_{hh}$ (92) for each pair (i, h): no demand a_i assigned to node location other than one where a hub h is located (prevents direct allocation of demands to installed facilities)
- Every location may either host a facility or a hub (but not both): $y_h + z_{hh} \leq 1$ (93)

  Note: each location may remain free from any hub or facility

Formulation: Constraints (2)

- **Conservation constraints** $\sum_{i \in \mathcal{V}} a_i z_{ih} = \sum_{j \in \mathcal{V}} x_{hj}$ (94): sum of fractions assigned from demand points to each hub h = sum (over all facilities j) of aggregated amount from that hub h

  At each hub h, sum of incoming traffic = sum of outgoing traffic (following hub transformation)

- **Regulation constraints** $\sum_{i \in \mathcal{V}} a_i z_{ih} \leq (\alpha \sum_{j \in \mathcal{V}} b_j y_j) z_{hh}$ (95) regulate incoming demands such that each hub h attracts fraction α of demands

  Fraction α set such that sum of processed demands $<$ hub logical capacity $c_h = \alpha \sum_{j \in \mathcal{V}} b_j y_j$

- **Facility capacity** constraints $\sum_{h \in \mathcal{V}} x_{hj} \leq b_j y_j$ (96): sum of fractions x_{hj} reaching every facility j doesn't exceed its capacity b_j

- **Hub-shipping** constraints $x_{hj} \leq min(c_h, b_j) y_j$: regulate amount of aggregated traffic transferable from hub h to facility j

  Logical capacity of each hub h, c_h (at least) $\sim$ capacity of single facility $\Rightarrow x_{hj} \leq b_j y_j$

Formulation: Constraints (3)

Set of constraints linking flow routing to hub-facility location problem:

- **Individual flow** constraints $f_{h(u,v)j} \leq x_{hj}$ (97): traffic flow $f_{h(u,v)j}$ along each arc $(u,v) \in \mathcal{A}$ from hub h to facility j delimited by fraction x_{hj} allocated to facility j
- **Mutual capacity** constraints $\sum_{h \in \mathcal{V}} \sum_{j \in \mathcal{V}} f_{h(u,v)j} \leq \kappa_{(u,v)}$ (98): load (sum of traffic flows) on individual arcs $(u,v) \in \mathcal{E}$ does not exceed nominal capacity $\kappa_{(u,v)}$
- **Flow conservation** constraints $\sum_{j \in \mathcal{V}} \sum_{v:(v,u) \in \mathcal{E}} f_{h(v,u)j} = x_{hu} + \sum_{j \in \mathcal{V}} \sum_{v:(u,v) \in \mathcal{A}} f_{h(u,v)j}$ (99): outgoing traffic flowing from hub h to facility j and entering node u must be equal to fraction served by facility j plus outgoing traffic flow leaving that node towards j
- **Flow conservation** constraints $\sum_{j \in \mathcal{V}} x_{hj} = \sum_{j \in \mathcal{V}} \sum_{v:(h,v) \in \mathcal{E}} f_{h(h,v)j}$ (100): sum over j of fractions x_{hj} transferred by hub h equals to sum of flows $f_{h(h,v)j}$ leaving that hub (a given location j may either host a hub or a facility (93))

Formulation: Constraints (4)

Handling of the RHS of the nonlinear constraints (93):

- Theorem: binary product $c = a.b$, where a, b are binary variables, can be linearized by substituting binary variable c with linear inequalities: 1) $c \leq a$; 2) $c \leq b$; 3) $c \geq a + b - 1$
- As both y_j and z_{hh} are binary: introduce auxiliary variables $\zeta_{hj} = y_j z_{hh}$ $\rightarrow$ Set of constraints:

$$\sum_{i \in \mathcal{V}} a_i z_{ih} \leq \alpha \sum_{j \in \mathcal{V}} b_j \zeta_{hj} \quad (105)$$

$$\zeta_{hj} \leq y_j \quad (106)$$

$$\zeta_{hj} \leq z_{hh} \quad (107)$$

$$\zeta_{hj} \geq y_j + z_{hh} - 1 \quad (108)$$

Linearization procedure

- Increases number of constraints and binary variables up to additive factor of V^2
- Does not significantly increase model complexity since substitutions independent of flow variables $f_{h(u,v)j}$ which dominate formulation complexity ($V^2.E$)

Numerical Experiments (1)

- **HLRP formulation** (collective resource abstraction) and LRP variant solved with CPLEX 12.6.3 (computation time limit of 43200s)
- **LRP formulation**: individual resource abstraction
  - Combines cFLP with minimum cost multi-commodity flow routing problem
  - Allocates demands to facilities without involving intermediate hubs but assuming each facility individually capable to assign local resources to incoming demands
- Executions performed on a dedicated server equipped with 8 x Intel Xeon quad-core processors and 512GB of DDR3 RAM
- Topologies (source: SNDlib library)

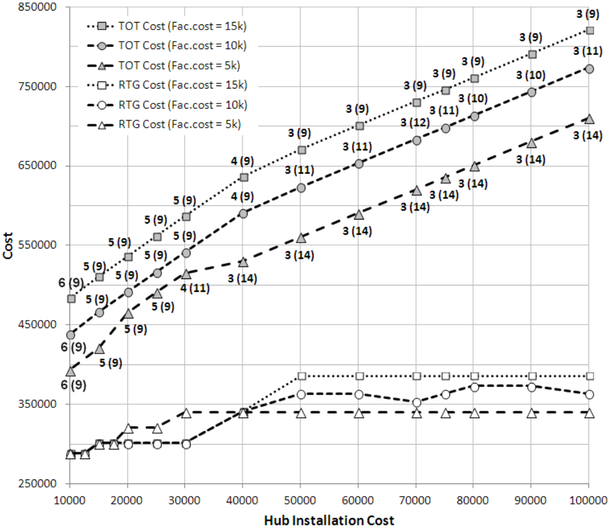
Topology	Nodes	Arcs	Degree			Diameter
			Min.	Max.	Avg	
austria	24	110	2	11	4.58	4
france	25	90	2	10	3.60	5
norway	27	102	2	6	3.78	7
india35	35	160	2	9	4.57	7
giul39	39	344	6	16	8.82	6

Numerical Experiments (2)

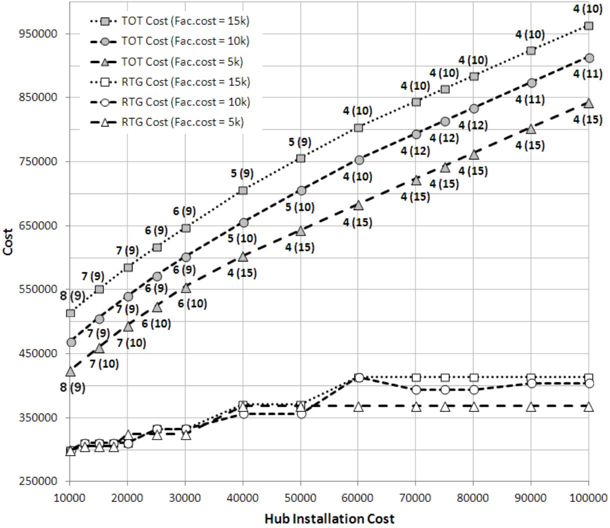
- For each topology, matrix of distances $d(i,j) = \delta_{ij}$ computed from corresponding graph
- SNDlib doesn't provide cost of locating hubs (η_h) and facilities (φ_j)
  - Facility location cost set independently of its physical location in value range $5000, 10000, 15000$
  - Hub installation cost set proportionally to that cost following step increasing factor from 1 to 10
- Capacity distributed over set of (potential) facilities is non-blocking: sum of all demands over all originating points does not exceed total facility capacity $\sum_{j \in \mathcal{J}} b_j$
- Total required capacity homogeneously distributed among installed facilities $b_j = b, \forall j \in \mathcal{J}$
- Demand set $\mathcal{A}$ comprises order of $10.|\mathcal{V}|$ tuples (demand point, demand size) drawn from a truncated Pareto distribution $P(\beta)$ with support $[10, 1000]$ and shape parameter $\beta = 1.4$

  Note: using other patterns, such as step functions (where each step corresponds to a given demand size), results obtained do not show any significant variation

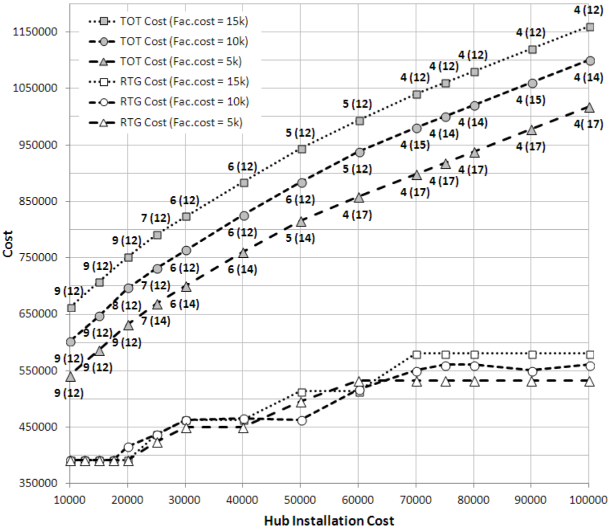
Numerical Experiments: Results (1a)



Numerical Experiments: Results (1b)



Numerical Experiments: Results (1c)



Numerical Experiments: Results (2)

Main observations

- When number of hubs decreases:
  - Number of installed facilities may increase by up to 50% (compared to min.number of facilities required to serve all demands)
  - Value reached when max.number of installed hubs decreases by factor 2
- When facility cost = 15000, number of installed facilities remains almost steady even when number of installed hubs decreases by 50%
- For all topologies
  - Routing cost increases before reaching its maximum when number of installed hubs crosses pivotal value of half max.number of hubs (compared to value obtained with min.installation cost)
  - Max.value $\sim$ number of installed facilities

$\Rightarrow$ Decreasing number of reachable hubs tends to increase number of required facilities at detriment of increasing routing cost

Numerical Experiments: Results (3)

Main observations

- Enforce both number of hubs and facilities to their lowest value by setting them to their minimal value obtained from the previous executions
- Strategy is not beneficial, in particular, when hub installation cost sits in the lower range
- Reason: higher routing cost counter-balances gain in installation cost obtained when number of hubs and facilities take their minimal value

⇒ Decreasing both number of hubs and facilities comes at detriment of higher routing cost

Numerical Experiments: Results (3)

Main observations

- Enforce both number of hubs and facilities to their lowest value by setting them to their minimal value obtained from the previous executions
- Strategy is not beneficial, in particular, when hub installation cost sits in the lower range
- Reason: higher routing cost counter-balances gain in installation cost obtained when number of hubs and facilities take their minimal value

⇒ Decreasing both number of hubs and facilities comes at detriment of higher routing cost

Comparison with LRP model

- For smaller topologies: significant drop in routing cost (following the MMCF strategy) from 27% for *norway* to 38% for *france* with gain in total cost up to 15%
- For larger topologies (e.g., *india35* and *giul39*): routing cost can decrease by factor 2 although total cost itself increases by 30% (distribution of functionality per server)

Performance results

Topology	Nodes	Arcs	Variables		Constraints	Comp. Time (s)
			Continuous	Binary		
france	25	90	56875	1275	117115	376
austria	24	110	63936	1176	131006	411
norway	27	102	75087	1485	154083	2112
india35	35	160	197225	2485	400945	12414
giul39	39	344	524745	3081	1057673	43200

- Main limitation: large number of continuous variables and constraints
- Such dimensions lead to challenging problems which require to process very large number of continuous variables indirectly linked to binary variables by conservation and capacity constraints
- Note: enforcing single-assignments between hubs and facilities (binary variables x_{hj}) would render this relation even tighter following (97)

Perspectives

- Apply Benders decomposition method to design an algorithm whereby
  - binary variables y_j and z_{ih} kept in master problem (MP)
  - continuous variables x_{hj} and $f_{h(u,v)j}$ projected out and used only in subproblems
- Resulting MP ((y_j, z_{ih}) space): single continuous variable and subset of inequalities not involving x_{hj} and $f_{h(u,v)j}$
- Method requires to solve iteratively master and subproblems several times $\Rightarrow$ suitable when decomposed problem much easier than the original one (master reduced to variant of cFLP and subproblems to variant of flow routing)
- Extend model to more complex routing costs defined by increasing convex functions ($\sim$ arc load $\ell_{(u,v)} = \sum_{h,j \in \mathcal{V}} f_{h(u,v)j}$)

Outline

Foreword

Relationships

- (Robust) optimization $\rightarrow$ machine learning
  - (Robust) Model extraction: discover relationships between $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ and y: $y = F(\mathbf{x})$
  - (Robust) Prediction: produce function $F(\mathbf{x})$ such that $\hat{y} = F(\mathbf{x})$ minimizes loss $L(y, \hat{y})$, e.g., $\sqrt{\mathbb{E}[y - \hat{y}]^2}$
    Then knowing F, use new input $\mathbf{x}^*$ to predict $\hat{y}^* = F(\mathbf{x}^*)$
- **Machine learning $\rightarrow$ (robust) optimization**: achieve more than computation task(s) automation tool (e.g., parameter selection)

Foreword

Relationships

- (Robust) optimization $\rightarrow$ machine learning
  - (Robust) Model extraction: discover relationships between $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ and y: $y = F(\mathbf{x})$
  - (Robust) Prediction: produce function $F(\mathbf{x})$ such that $\hat{y} = F(\mathbf{x})$ minimizes loss $L(y, \hat{y})$, e.g., $\sqrt{\mathbb{E}[y - \hat{y}]^2}$
    Then knowing F, use new input $\mathbf{x}^*$ to predict $\hat{y}^* = F(\mathbf{x}^*)$
- **Machine learning $\rightarrow$ (robust) optimization**: achieve more than computation task(s) automation tool (e.g., parameter selection)

Goal

Automate construction of uncertainty sets: one of the two central problems in robust optimization

- **Construction**: combine model-driven (incumbent approach) with data-driven methods
- **Automation**: procedure combining feature extraction from data, stat.hypothesis tests and selection of model which best explains data

$\Rightarrow$ Exploit machine learning techniques

What do we mean by Uncertainty ?

Aleatory uncertainty

- Compared to epistemic uncertainty: physical variability present in the system (endogenous) or its environment (exogenous)
- Properties of aleatory uncertainty
  - **Intrinsic**: variable is random; different value each time it is observed
    $\rightarrow$ additional experiments (observations, data) can only be used to **better characterize variability**
  - **Irreducible**: not strictly due to a lack of knowledge, cannot be reduced
    $\rightarrow$ taking more measurements will not reduce uncertainty in the value of the variable

What do we mean by Uncertainty ?

Aleatory uncertainty

- Compared to epistemic uncertainty: physical variability present in the system (endogenous) or its environment (exogenous)
- Properties of aleatory uncertainty
  - **Intrinsic**: variable is random; different value each time it is observed
    $\rightarrow$ additional experiments (observations, data) can only be used to **better characterize variability**
  - **Irreducible**: not strictly due to a lack of knowledge, cannot be reduced
    $\rightarrow$ taking more measurements will not reduce uncertainty in the value of the variable

Implications

- **Modeling**: (typically) probabilistic framework
- **Examples**: demands variability, (certain) topology failures, etc.

Problem space

Goal (both stochastic and robust optimization): find a solution that will perform well under any possible realization of random parameters, i.e., find solutions which remain valid even if input data changes

Problem space

Goal (both stochastic and robust optimization): find a solution that will perform well under any possible realization of random parameters, i.e., find solutions which remain valid even if input data changes

Stochastic optimization: probabilistic description of uncertainty

- *Random parameters* governed by prob. distributions known by the decision maker, and the objective is to find a solution that minimizes the *expected cost*
- Applied when seeking solutions that perform well in the long run on average, with poor performance at some times balanced by good performance at others
- Decisions are evaluated ex-post, i.e., after uncertainty has been resolved, and costs have been realized (solution quality known at realization time)

Problem space

Goal (both stochastic and robust optimization): find a solution that will perform well under any possible realization of random parameters, i.e., find solutions which remain valid even if input data changes

Stochastic optimization: probabilistic description of uncertainty

- *Random parameters* governed by prob. distributions known by the decision maker, and the objective is to find a solution that minimizes the *expected cost*
- Applied when seeking solutions that perform well in the long run on average, with poor performance at some times balanced by good performance at others
- Decisions are evaluated ex-post, i.e., after uncertainty has been resolved, and costs have been realized (solution quality known at realization time)

Robust optimization: deterministic description of uncertainty

- Prob.distribution of uncertainty not known, and *uncertain parameters* are specified either by discrete scenarios, continuous ranges, sets, etc.
- Time independence (more precisely, solution quality known at computation time)

Robust Optimization

Objective

Find solutions to uncertain problems that remain feasible for all scenarios involving uncertainty (in parameters or even variables) such as to protect/immunize against infeasibility

Robust Optimization

Objective

Find solutions to uncertain problems that remain feasible for all scenarios involving uncertainty (in parameters or even variables) such as to protect/immunize against infeasibility

Properties

- Probability distribution characterizing uncertainty not known
- Time independence: solution quality known at computation time
- *Uncertain parameters* specified by discrete scenarios, continuous ranges, sets, etc.
  $\rightarrow$ representation models for **uncertainty sets**
  - Hypercube uncertainty set (Soyster, 1973)
  - Polytopic uncertainty: ellipsoidal uncertainty set (Ben-Tal-Nemirovski, 1999)
  - Cardinality constrained uncertainty (Bertsimas-Sim, 2004$-\Gamma$-robustness)
  - Data-driven/distributional approaches (Bertsimas, 2006) to build models yielding less conservative uncertainty sets

Data-Driven Robust Optimization (1)

- Uncertain constraints $F(\tilde{a}, \mathbf{x}) \leq 0$
  $\tilde{a}$: uncertainty parameter modeled as random variable whose distribution P^* is unknown (except for some pre-assumed structural features)
- Robust constraints modeled by choosing uncertainty set $\mathcal{U}$ such that $F(\mathbf{a}, \mathbf{x}) \leq 0, \forall \mathbf{a} \in \mathcal{U}$
- Given $\epsilon > 0$, constructed sets $\mathcal{U}_\epsilon$ implies probabilistic guarantee for P^* at level ϵ: for any x^*

$$\textbf{if } F(\mathbf{a}, \mathbf{x}^*) \leq 0, \forall \mathbf{a} \in \mathcal{U}_\epsilon \quad (109)$$

$$\textbf{then } P^*(F(\tilde{a}, \mathbf{x}^*) \leq 0)) \geq 1 - \epsilon \quad (110)$$

  Condition ensuring that feasible solution to robust constraint also feasible with probability $1 - \epsilon$ wrt P^* even if P^* is not known exactly

Data-Driven Robust Optimization (2)

- Assumptions
  - Data set $\{\hat{a}_1, \ldots, \hat{a}_n\}$ drawn i.i.d. according to P^*
  - Structural features of P^* known a priori

Data-Driven Robust Optimization (2)

- Assumptions
  - Data set $\{\hat{a}_1, \ldots, \hat{a}_n\}$ drawn i.i.d. according to P^*
  - Structural features of P^* known a priori
- **Main concept**
  - Pair different **a priori assumptions** and **stat.hypothesis tests** to obtain distinct data-driven uncertainty sets
    - Each with its own geometric and computational properties
    - Capturing features of P^*, e.g., skewness, heavy-tails and correlations
  - Use confidence region of hypothesis test to quantify learning about P^* from data

Data-Driven Robust Optimization (2)

- Assumptions
  - Data set $\{\hat{a}_1, \ldots, \hat{a}_n\}$ drawn i.i.d. according to P^*
  - Structural features of P^* known a priori
- **Main concept**
  - Pair different **a priori assumptions** and **stat.hypothesis tests** to obtain distinct data-driven uncertainty sets
    - Each with its own geometric and computational properties
    - Capturing features of P^*, e.g., skewness, heavy-tails and correlations
  - Use confidence region of hypothesis test to quantify learning about P^* from data
- Using this (general) technique one may consider (Bertsimas, 2013):
  - P^* with finite discrete support (known)
  - P^* with possible continuous support and
    - Components of $\tilde{a}$ are independent
    - Data drawn from marginal distributions of P^* separately (data sampled asynchronously)
    - Data are sampled from joint distribution

Learning traffic uncertainty from data

Machine learning methods

- **Data-driven methods**: inference tasks (density estimation), stat.hypothesis tests, structure and feature extraction from data samples
- **Model-driven methods**: produce and select an hypothesis (approx.function) which best explains the data

Learning traffic uncertainty from data

Machine learning methods

- **Data-driven methods**: inference tasks (density estimation), stat.hypothesis tests, structure and feature extraction from data samples
- **Model-driven methods**: produce and select an hypothesis (approx.function) which best explains the data

Robust optimization problems

- Traffic demand variability → Traffic fluctuations
  Example: network design, capacity (re-)dimensioning, routing decision and action planning under uncertainty (robustified MCF, MMCF, MCND)
- Topology failures → Traffic fluctuations
  Example: re-routing decisions and protection capacity dimensioning
- Resource demands variability (distributed file servers/caches)
  Example: server (re)location decisions
- Quality of service (congestion): bandwidth - delay
  Example: robust multi-objective optimization

Input data: from (distributed) monitoring task

Traffic uncertainty

- **Goal**: find solutions that remain feasible for all scenarios in uncertainty set $\mathcal{U}$ to protect/immunize against infeasibility
- Model: uncertainty in traffic flows (model parameters):

$$\varphi_{ij}^{st} \rightarrow \bar{\varphi}_{ij}^{st} + \xi^{st}\hat{\varphi}_{ij}^{st}, \ \forall s, t \in W \tag{111}$$

Traffic uncertainty

- **Goal**: find solutions that remain feasible for all scenarios in uncertainty set $\mathcal{U}$ to protect/immunize against infeasibility
- Model: uncertainty in traffic flows (model parameters):

$$\varphi_{ij}^{st} \to \bar{\varphi}_{ij}^{st} + \xi^{st}\hat{\varphi}_{ij}^{st}, \ \forall s, t \in W \tag{111}$$

  - $\bar{\varphi}_{ij}^{st} \in \mathbb{R}^+$: nominal value of the traffic flow
  - $\hat{\varphi}_{ij}^{st}$: deviation or perturbation of $\bar{\varphi}_{ij}^{st}$
  - ξ_{ij}^{st}: random variable with unknown distribution (except some structural properties)
  - $W \subseteq V$: subset of nodes defining (s, t) pairs of traffic flows subject to uncertainty

Traffic uncertainty

- **Goal**: find solutions that remain feasible for all scenarios in uncertainty set $\mathcal{U}$ to protect/immunize against infeasibility
- Model: uncertainty in traffic flows (model parameters):

$$\varphi_{ij}^{st} \rightarrow \bar{\varphi}_{ij}^{st} + \xi^{st}\hat{\varphi}_{ij}^{st}, \ \forall s,t \in W \tag{111}$$

  - $\bar{\varphi}_{ij}^{st} \in \mathbb{R}^{+}$: nominal value of the traffic flow
  - $\hat{\varphi}_{ij}^{st}$: deviation or perturbation of $\bar{\varphi}_{ij}^{st}$
  - ξ_{ij}^{st}: random variable with unknown distribution (except some structural properties)
  - $W \subseteq V$: subset of nodes defining (s,t) pairs of traffic flows subject to uncertainty
- **Method**:
  - Build uncertainty set $\mathcal{U}(\xi^{st})$ with $\xi^{st} \in \mathcal{Z}$ such that constraints rewritten by grouping deterministic and uncertain part:

$$\sum_{s,t \in V} \bar{\varphi}_{ij}^{st} y_{ij}^{st} + \sum_{s,t \in W} \xi^{st}\hat{\varphi}_{ij}^{st} y_{ij}^{st} \leq C_{ij}x_{ij}, \forall (i,j) \in A \tag{112}$$

Traffic uncertainty

- **Goal**: find solutions that remain feasible for all scenarios in uncertainty set $\mathcal{U}$ to protect/immunize against infeasibility
- Model: uncertainty in traffic flows (model parameters):

$$\varphi_{ij}^{st} \rightarrow \bar{\varphi}_{ij}^{st} + \xi^{st}\hat{\varphi}_{ij}^{st}, \ \forall s,t \in W \tag{111}$$

  - $\bar{\varphi}_{ij}^{st} \in \mathbb{R}^+$: nominal value of the traffic flow
  - $\hat{\varphi}_{ij}^{st}$: deviation or perturbation of $\bar{\varphi}_{ij}^{st}$
  - ξ_{ij}^{st}: random variable with unknown distribution (except some structural properties)
  - $W \subseteq V$: subset of nodes defining (s,t) pairs of traffic flows subject to uncertainty
- **Method**:
  - Build uncertainty set $\mathcal{U}(\xi^{st})$ with $\xi^{st} \in \mathcal{Z}$ such that constraints rewritten by grouping deterministic and uncertain part:

$$\sum_{s,t\in V} \bar{\varphi}_{ij}^{st} y_{ij}^{st} + \sum_{s,t\in W} \xi^{st}\hat{\varphi}_{ij}^{st} y_{ij}^{st} \leq C_{ij}x_{ij}, \forall (i,j) \in A \tag{112}$$

  - Find solutions that remain feasible for any $\xi \in \mathcal{Z}$

$$\sum_{s,t\in V} \bar{\varphi}_{ij}^{st} y_{ij}^{st} + \max_{\xi^{st}\in\mathcal{Z}} \Big(\sum_{s,t\in W} \xi^{st}\hat{\varphi}_{ij}^{st} y_{ij}^{st} \Big) \leq C_{ij}x_{ij}, \forall (i,j) \in A \tag{113}$$

Perturbation sets

Formulation of robust counterpart depends on construction and selection of **perturbation set** $\mathcal{Z}$:

Base sets

- **Box**: $\mathcal{Z}_\infty = \{\xi \in \mathbb{R} \mid \|\xi\|_\infty \leq \Psi\} = \{\xi \in \mathbb{R} \mid |\xi^{st}| \leq \Psi, \forall (s,t) \in W\}$
- **Polyhedral**: $\mathcal{Z}_1 = \{\xi \in \mathbb{R} \mid \|\xi\|_1 \leq \Gamma\} = \{\xi \in \mathbb{R} \mid \sum_{s,t \in W} |\xi^{st}| \leq \Gamma\}$
- **Ellipsoid**: $\mathcal{Z}_2 = \{\xi \in \mathbb{R} \mid \|\xi\|_2 \leq \Omega\} = \{\xi \in \mathbb{R} \mid \sum_{s,t \in W} (\xi^{st})^2 \leq \Omega^2\}$

  where, Ψ, Γ, Ω are the adjustable parameter controlling the magnitude of the perturbation set

Perturbation sets

Formulation of robust counterpart depends on construction and selection of **perturbation set** $\mathcal{Z}$:

Base sets

- **Box**: $\mathcal{Z}_\infty = \{\xi \in \mathbb{R} \mid \|\xi\|_\infty \leq \Psi\} = \{\xi \in \mathbb{R} \mid |\xi^{st}| \leq \Psi, \forall (s,t) \in W\}$
- **Polyhedral**: $\mathcal{Z}_1 = \{\xi \in \mathbb{R} \mid \|\xi\|_1 \leq \Gamma\} = \{\xi \in \mathbb{R} \mid \sum_{s,t \in W} |\xi^{st}| \leq \Gamma\}$
- **Ellipsoid**: $\mathcal{Z}_2 = \{\xi \in \mathbb{R} \mid \|\xi\|_2 \leq \Omega\} = \{\xi \in \mathbb{R} \mid \sum_{s,t \in W} (\xi^{st})^2 \leq \Omega^2\}$

  where, Ψ, Γ, Ω are the adjustable parameter controlling the magnitude of the perturbation set

Combinations (intersection between box, polyhedral, ellipsoidal sets)

- **Box + Polyhedral**: $\mathcal{Z}_{\infty\cap 1} = \{\xi \in \mathbb{R} | \, \|\xi\|_\infty \leq \Psi, \|\xi\|_1 \leq \Gamma\}$
- **Box + Ellipsoidal**: $\mathcal{Z}_{\infty\cap 2} = \{\xi \in \mathbb{R} | \, \|\xi\|_\infty \leq \Psi, \|\xi\|_2 \leq \Omega\}$
- **Box + Polyhedral + Ellipsoidal**:
  $\mathcal{Z}_{\infty\cap 1\cap 2} = \{\xi \in \mathbb{R} | \, \|\xi\|_\infty \leq \Psi, \|\xi\|_1 \leq \Gamma, \|\xi\|_2 \leq \Omega\}$

Perturbation sets: incremental construction

Tradeoff against **Computational complexity**

Original Problem	LP	MILP	QCQP	SOCP
Polyhedral Set	LP	MILP	MINLP	MINLP
Ellipsoidal Set	SOCP	MISOCP	SDP	SDP

Perturbation sets: incremental construction

Tradeoff against **Computational complexity**

Original Problem	LP	MILP	QCQP	SOCP
Polyhedral Set	LP	MILP	MINLP	MINLP
Ellipsoidal Set	SOCP	MISOCP	SDP	SDP

Procedure

`If` $Z \leftarrow \mathcal{Z}_\infty = \{\xi \in \mathbb{R} \mid \|\xi\|_\infty \leq \Psi\}$ too conservative

`Then` $Z \leftarrow \mathcal{Z}_{\infty\cap 1} = \{\xi \in \mathbb{R} | \, \|\xi\|_\infty \leq \Psi, \|\xi\|_1 \leq \Gamma\}$

- `If` Z too liberal
  `Then` $Z \leftarrow \mathcal{Z}_{\infty\cap 1\cap 2} = \{\xi \in \mathbb{R} | \, \|\xi\|_\infty \leq \Psi, \|\xi\|_1 \leq \Gamma, \|\xi\|_2 \leq \Omega\}$

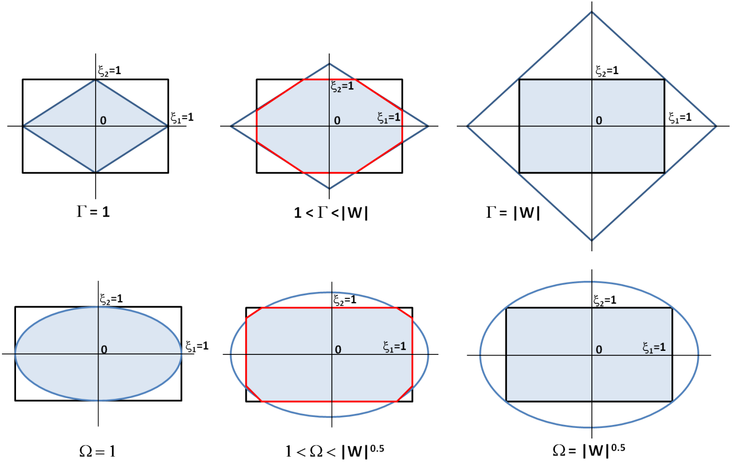
`Else If` $Z \leftarrow \mathcal{Z}_1 = \{\xi \in \mathbb{R} | \, \|\xi\|_1 \leq \Gamma\}$ too conservative

`Then` $Z \leftarrow \mathcal{Z}_{1\cap\infty}$ or $\mathcal{Z}_{1\cap 2}$

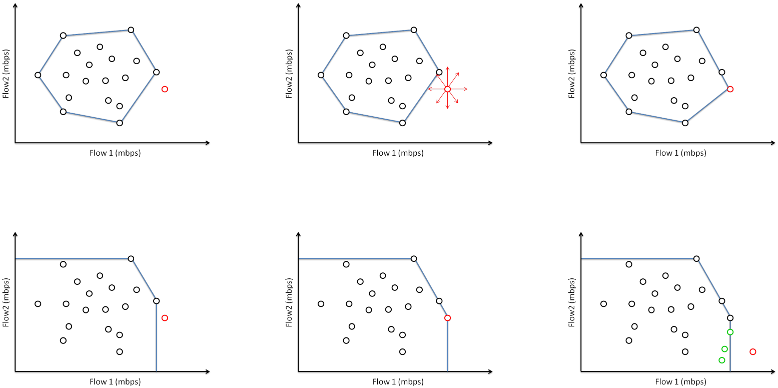
`Else If` $Z \leftarrow \mathcal{Z}_2 = \{\xi \in \mathbb{R} | \, \|\xi\|_2 \leq \Omega\}$ too conservative

`Then` $Z \leftarrow \mathcal{Z}_{2\cap\infty}$ or $\mathcal{Z}_{2\cap 1}$

Geometric interpretation ($\Psi = 1$)



Example: with two traffic flows



Uncertainty set: safety margin model

- **Model assumptions**:
  - Characteristic value $\bar{\varphi}_{ij}^{st}$ for each traffic flow determined a priori from past data (prediction), e.g., (expected) mean
  - Deviation delimited by safety margin in $\left[-\hat{\varphi}_{ij}^{st}, \hat{\varphi}_{ij}^{st}\right]$, with $\hat{\varphi}_{ij}^{st}$ = max.deviation
    Note: such choice may be too conservative (consider instead, e.g., standard deviation or mean deviation)

Uncertainty set: safety margin model

- **Model assumptions**:
  - Characteristic value $\bar{\varphi}_{ij}^{st}$ for each traffic flow determined a priori from past data (prediction), e.g., (expected) mean
  - Deviation delimited by safety margin in $\left[-\hat{\varphi}_{ij}^{st}, \hat{\varphi}_{ij}^{st}\right]$, with $\hat{\varphi}_{ij}^{st}$ = max.deviation
    Note: such choice may be too conservative (consider instead, e.g., standard deviation or mean deviation)
- Then following Eq.111 $\varphi_{ij}^{st} \rightarrow \bar{\varphi}_{ij}^{st} + \xi^{st}\hat{\varphi}_{ij}^{st}, \ \forall s,t \in W$:
  - **Uncertainty set**: $\mathcal{U} = \{\varphi_{ij}^{st} \mid \bar{\varphi}_{ij}^{st} - \xi^{st}\hat{\varphi}_{ij}^{st} \leq \varphi_{ij}^{st} \leq \bar{\varphi}_{ij}^{st} + \xi^{st}\hat{\varphi}_{ij}^{st}, \xi \in \mathcal{Z}_{\infty,1}\}$
  - **Perturbation set**:
    - $\mathcal{Z}_\infty = \{\xi \in \mathbb{R}^{|W|} \mid |\xi^{st}| \leq 1, \forall s,t \in W\}$
    - $\mathcal{Z}_1 = \{\xi \in \mathbb{R}^{|W|} \mid \sum_{s,t\in W} |\xi^{st}| \leq |W|\}$
    - $\mathcal{Z}_{\infty,1} = \{\xi \in \mathbb{R}^{|W|} \mid \sum_{s,t\in W} |\xi^{st}| \leq |W|, |\xi^{st}| \leq 1, \forall s,t \in W\}$
  - **Constraints** reformulated as:

$$\sum_{s,t\in V} \bar{\varphi}_{ij}^{st} y_{ij}^{st} + \sum_{s,t\in W} w_{ij}^{st} + |W| z_{ij} \leq C_{ij} x_{ij}, \ \ (i,j) \in A \tag{114}$$

$$z_{ij} + w_{ij}^{st} \geq \hat{\varphi}_{ij}^{st} y_{ij}^{st} \ \ s,t \in W, (i,j) \in A \tag{115}$$

$$w_{ij}^{st} \geq 0, z_{ij} \geq 0 \ \ s,t \in W, (i,j) \in A \tag{116}$$

Uncertainty set: perturbation model

- **Model assumptions**:
  - Nominal value $\bar{\varphi}_{ij}^{st}$ for each traffic flow
  - Deviation $\hat{\varphi}_{ij}^{st}$ modeled as **bounded perturbation** around that value $\epsilon^{st}\bar{\varphi}_{ij}^{st}, |\epsilon^{st}| \in [0,1]$:
    $\hat{\varphi}_{ij}^{st} \rightarrow \epsilon^{st}\bar{\varphi}_{ij}^{st}$
  - **Goal**: construct less conservative sets

Uncertainty set: perturbation model

- **Model assumptions**:
  - Nominal value $\bar{\varphi}_{ij}^{st}$ for each traffic flow
  - Deviation $\hat{\varphi}_{ij}^{st}$ modeled as **bounded perturbation** around that value $\epsilon^{st}\bar{\varphi}_{ij}^{st}, |\epsilon^{st}| \in [0,1]$:
    $\hat{\varphi}_{ij}^{st} \rightarrow \epsilon^{st}\bar{\varphi}_{ij}^{st}$
  - **Goal**: construct less conservative sets
- Then following $\varphi_{ij}^{st} \rightarrow \bar{\varphi}_{ij}^{st} + \xi^{st}\hat{\varphi}_{ij}^{st}, \ \forall s,t \in W$ (Eq.111):
  - **Uncertainty set**:
    $\mathcal{U} = \{\varphi_{ij}^{st} \mid \bar{\varphi}_{ij}^{st} - \xi^{st}\epsilon^{st}\bar{\varphi}_{ij}^{st} \leq \varphi_{ij}^{st} \leq \bar{\varphi}_{ij}^{st} + \xi^{st}\epsilon^{st}\bar{\varphi}_{ij}^{st}, \xi^{st} \in \mathcal{Z}_{\infty,1}\}$
  - **Perturbation set**:
    - $\mathcal{Z}_\infty = \{\xi \in \mathbb{R}^{|W|} \mid |\xi^{st}| \leq 1, \forall s,t \in W\}$
    - $\mathcal{Z}_1 = \{\xi \in \mathbb{R}^{|W|} \mid \sum_{s,t\in W} |\xi^{st}| \leq |W|\}$
    - $\mathcal{Z}_{\infty,1} = \{\xi \in \mathbb{R}^{|W|} \mid \sum_{s,t\in W} |\xi^{st}| \leq |W|, |\xi^{st}| \leq 1, \forall s,t \in W\}$
  - **Constraints** reformulated as:

$$\sum_{s,t\in V} \bar{\varphi}_{ij}^{st} y_{ij}^{st} + \sum_{s,t\in W} w_{ij}^{st} + |W| z_{ij} \leq C_{ij} x_{ij} \quad (i,j) \in A \qquad (117)$$

$$z_{ij} + w_{ij}^{st} \geq \epsilon^{st}\bar{\varphi}_{ij}^{st}|y_{ij}^{st}| = \epsilon^{st}\bar{\varphi}_{ij}^{st} y_{ij}^{st} \quad s,t \in W, (i,j) \in A \qquad (118)$$

$$w_{ij}^{st} \geq 0, z_{ij} \geq 0 \quad s,t \in W, (i,j) \in A \qquad (119)$$

Probabilistic Guarantees

- Questions (when uncertainty set does not cover the whole uncertainty space)
  - Necessary size of uncertainty set to ensure that the degree of constraint violation does not exceed a certain level ?
  - What is the degree (probability) of constraint violation P_v upon solution of robust optimization problem ?

Probabilistic Guarantees

- Questions (when uncertainty set does not cover the whole uncertainty space)
  - Necessary size of uncertainty set to ensure that the degree of constraint violation does not exceed a certain level ?
  - What is the degree (probability) of constraint violation P_v upon solution of robust optimization problem ?

$$P_v = P[\sum_{s,t \in V} \bar{\varphi}_{ij}^{st} y_{ij}^{st} + \sum_{s,t \in W} \xi_{ij}^{st} \hat{\varphi}_{ij}^{st} y_{ij}^{st} \geq C_{ij} x_{ij}] \tag{120}$$

Probabilistic Guarantees

- Questions (when uncertainty set does not cover the whole uncertainty space)
  - Necessary size of uncertainty set to ensure that the degree of constraint violation does not exceed a certain level ?
  - What is the degree (probability) of constraint violation P_v upon solution of robust optimization problem ?

$$P_v = P[\sum_{s,t \in V} \bar{\varphi}_{ij}^{st} y_{ij}^{st} + \sum_{s,t \in W} \xi_{ij}^{st} \hat{\varphi}_{ij}^{st} y_{ij}^{st} \geq C_{ij} x_{ij}] \tag{120}$$

- Methods to evaluate probabilistic guarantees: probabilistic guarantee on constraint satisfaction or upper bound on probability of constraint violation
  1. Derive probability of constraint violation using the uncertainty set information **before** solving the problem (as much as possible)
     $\rightarrow$ a priori probability bound
  2. Derive the probability directly from the robust counterpart optimization solution (sometimes only possible alternative)
     $\rightarrow$ a posteriori probability bound

Use case (1)

- Measurement of spatio-temporal traffic properties
  - Troubleshoot communication networks performance, quality, etc.
  - Information to traffic-driven processes (predictive routing decisions)

Use case (1)

- Measurement of spatio-temporal traffic properties
  - Troubleshoot communication networks performance, quality, etc.
  - Information to traffic-driven processes (predictive routing decisions)
- Active vs. Passive measurement
  - Active: set of dedicated messages (probes) sent along links / paths
  - **Passive**: dedicated devices (monitoring points) placed on node's outgoing interfaces, sampling outgoing traffic, i.e., capture percentage of traffic following a given configuration (sampling rate)
- Passive monitoring ⇒ adequate placement and configuration of monitoring points (traffic uncertainty and dynamics)

Use case (2)

Cooperative monitoring problem

Adequate placement and configuration of passive monitoring points to jointly realize a given task of monitoring time-varying traffic flows along their respective routing path

Use case (2)

Cooperative monitoring problem

Adequate placement and configuration of passive monitoring points to jointly realize a given task of monitoring time-varying traffic flows along their respective routing path

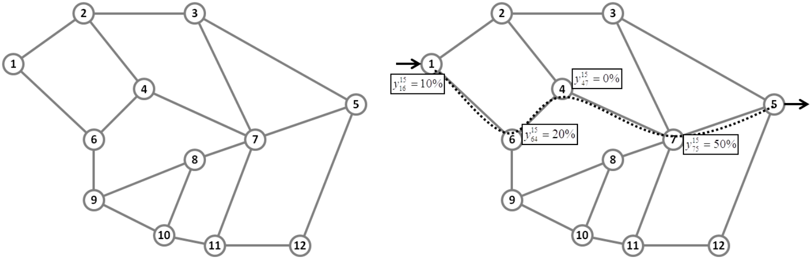
- **Problem**: knowing traffic demands, where to place and how to configure passive monitoring points such that $k\%$ of traffic flowing along each path is monitored ?

Use case (2)

Cooperative monitoring problem

Adequate placement and configuration of passive monitoring points to jointly realize a given task of monitoring time-varying traffic flows along their respective routing path

- **Problem**: knowing traffic demands, where to place and how to configure passive monitoring points such that $k\%$ of traffic flowing along each path is monitored ?
- **Example**: monitoring task $\equiv$ monitor flow f_{ij}^{15} with $k = 80\%$
  Monitoring point installed at head-end of arc (1,6) sampling traffic at 10%
  — arc (6,4) — 20%
  — arc (7,5) — 50%



Optimization Problems

Cost function

- Installation cost $\propto$ spatial distribution of flows
- Configuration cost $\propto$ fraction of traffic sampled at each monitoring point such that along each routing path the total fraction $k \leq 100\%$
- Flow variables $\propto$ strategy adopted for routing of traffic flows, e.g., Minimum cost flow (MCF), Minimum cost multi-commodity flow (MMCF)

Optimization Problems

Cost function

- Installation cost $\propto$ spatial distribution of flows
- Configuration cost $\propto$ fraction of traffic sampled at each monitoring point such that along each routing path the total fraction $k \leq 100\%$
- Flow variables $\propto$ strategy adopted for routing of traffic flows, e.g., Minimum cost flow (MCF), Minimum cost multi-commodity flow (MMCF)

Optimization problems

- **Minimize total monitoring cost**
  such that task of monitoring time-varying traffic flows can be jointly realized
- **Maximize utility of monitoring traffic flows**
  without violating budget constraint imposed on total monitoring cost
- **Cooperation** between monitoring points
  along each routing path, traffic sampled at a given monitoring point NOT sampled again at another point along same path

Prior and Related Work

[Suh2005]

- Problems: i) minimize installation (and operational) cost to achieve given monitoring objective, ii) maximize utility of captured traffic under monitoring budget constraints $\rightarrow$ limit number of devices to be deployed
- Evaluation on small instances (limited to 10-nodes random graphs) and number of flows (synthetic traffic matrices)
- Configurable sampling rate for each flow at each monitoring point but rate adjusted independently along the same routing path

Prior and Related Work

[Suh2005]

- Problems: i) minimize installation (and operational) cost to achieve given monitoring objective, ii) maximize utility of captured traffic under monitoring budget constraints $\rightarrow$ limit number of devices to be deployed
- Evaluation on small instances (limited to 10-nodes random graphs) and number of flows (synthetic traffic matrices)
- Configurable sampling rate for each flow at each monitoring point but rate adjusted independently along the same routing path

[Chaudet2009]

- Monitoring at least k% of total traffic (without necessarily monitoring every path) while minimizing setup cost (device installation) and exploitation cost (sampling ratio assignation to each device)
- When total fraction $k = 100\% \Leftrightarrow$ Minimum Set Cover problem
- Simpler arc-path formulation (although sill non-polynomial)
- No results provided for passive monitoring model with traffic sampling $k < 100\%$

Prior and Related Work

[Suh2005]

- Problems: i) minimize installation (and operational) cost to achieve given monitoring objective, ii) maximize utility of captured traffic under monitoring budget constraints $\rightarrow$ limit number of devices to be deployed
- Evaluation on small instances (limited to 10-nodes random graphs) and number of flows (synthetic traffic matrices)
- Configurable sampling rate for each flow at each monitoring point but rate adjusted independently along the same routing path

[Chaudet2009]

- Monitoring at least k% of total traffic (without necessarily monitoring every path) while minimizing setup cost (device installation) and exploitation cost (sampling ratio assignation to each device)
- When total fraction $k = 100\% \Leftrightarrow$ Minimum Set Cover problem
- Simpler arc-path formulation (although sill non-polynomial)
- No results provided for passive monitoring model with traffic sampling $k < 100\%$

[Cantieni2006]

- Individual nodes apply local decisions in order to minimize their memory usage following a **global sampling strategy** for a specific monitoring goal
- Proposed formulation: multiplies (for each arc) sampling rate with traffic load (aggregate) instead of individual flows

Input data, Parameters and Variables

Input Data

- Network topology modeled by directed graph $G = (V, A)$

  $V \triangleq$ finite set of nodes and $A \triangleq$ finite set of arcs (i, j)
- Demand matrix D: $D(s, t) \triangleq$ total amount of traffic from source s to destination t, $\forall (s, t) \in V, \; s \neq t$

Input data, Parameters and Variables

Input Data

- Network topology modeled by directed graph $G = (V, A)$
  $V \triangleq$ finite set of nodes and $A \triangleq$ finite set of arcs (i, j)
- Demand matrix D: $D(s, t) \triangleq$ total amount of traffic from source s to destination t, $\forall (s, t) \in V, \; s \neq t$

Parameters

- Flow parameters: $\varphi_{ij}^{st} \;\; \forall (i, j) \in A$ (depend on routing strategy, e.g., MCF, MMCF)
- Total fraction of traffic k to be monitored along each path (single path routing)
- When formulation is capacitive: monitoring points associated capacity β_{ij}

Input data, Parameters and Variables

Input Data

- Network topology modeled by directed graph $G = (V, A)$
  $V \triangleq$ finite set of nodes and $A \triangleq$ finite set of arcs (i, j)
- Demand matrix D: $D(s, t) \triangleq$ total amount of traffic from source s to destination t, $\forall (s, t) \in V,\ s \neq t$

Parameters

- Flow parameters: $\varphi_{ij}^{st}\ \ \forall (i, j) \in A$ (depend on routing strategy, e.g., MCF, MMCF)
- Total fraction of traffic k to be monitored along each path (single path routing)
- When formulation is capacitive: monitoring points associated capacity β_{ij}

Variables

- Binary variable $x_{ij} = 1$ if a monitoring point should be installed at head end i along arc (i, j), 0 otherwise
- Continuous variables y_{ij}^{st} = fraction of traffic flow φ_{ij}^{st} sampled on monitoring point installed along arc (i, j)

Cost function and Formulation

Monitoring cost function $\triangleq$ Installation cost + Configuration cost

- **Installation cost**: fixed cost $\mathtt{m}_{ij}$ of installing a monitor at head-end of i arc (i, j)
- **Configuration cost** over all installed monitors: $\sum_{(i,j) \in A} \mathtt{n}_{ij} \ell_{ij}$
    - Fraction of traffic sampled at monitoring point installed along arc (i, j) : y_{ij}^{st}
    - Monitoring load at that point : $\ell_{ij} = \sum_{s,t \in V} \varphi_{ij}^{st} y_{ij}^{st}$
    - Cost per unit of sampled traffic : $\mathtt{n}_{ij}$

Cost function and Formulation

Monitoring cost function $\triangleq$ Installation cost + Configuration cost

- **Installation cost**: fixed cost $\mathtt{m}_{ij}$ of installing a monitor at head-end of i arc (i, j)
- **Configuration cost** over all installed monitors: $\sum_{(i,j)\in A} \mathtt{n}_{ij}\ell_{ij}$
  - Fraction of traffic sampled at monitoring point installed along arc (i, j) : y_{ij}^{st}
  - Monitoring load at that point : $\ell_{ij} = \sum_{s,t\in V} \varphi_{ij}^{st} y_{ij}^{st}$
  - Cost per unit of sampled traffic : $\mathtt{n}_{ij}$

Utility function u

- Utility function $u \propto a - \exp(-b \sum_{(i,j)\in A} y_{ij}^{st})$ where, $a, b \in \mathbb{R}_0^+$
- Non-decreasing (increasing monitoring fraction improves utility) but after reaching a certain threshold, relatively less beneficial to increase monitored fraction of traffic
- For computational purposes, approximate concave continuous function using piecewise-linear continuous fit [Geoffrion1977]

Utility Maximimization Problem

Problem

- **Problem**: Given budget $\mathcal{M}$ for monitoring installation cost and $\mathcal{N}$ for monitoring configuration cost, find localization and configuration of monitoring points which maximizes sum of utilities of monitoring traffic flows without violating budget constraints
- **Objective**: Maximize sum of utilities of monitoring individual traffic flows without violating budget constraints on installation and configuration cost

Utility Maximimization Problem

Problem

- **Problem**: Given budget $\mathcal{M}$ for monitoring installation cost and $\mathcal{N}$ for monitoring configuration cost, find localization and configuration of monitoring points which maximizes sum of utilities of monitoring traffic flows without violating budget constraints
- **Objective**: Maximize sum of utilities of monitoring individual traffic flows without violating budget constraints on installation and configuration cost

$$\max \sum_{s,t \in V} u^{st}\Big(\sum_{(i,j) \in A} y_{ij}^{st}\Big) \quad (121)$$

subject to:

$$\sum_{(i,j) \in A} \mathtt{m}_{ij} x_{ij} \leq \mathcal{M} \quad (122)$$

$$\sum_{(i,j) \in A} \mathtt{n}_{ij} \sum_{s,t \in V} \varphi_{ij}^{st} y_{ij}^{st} \leq \mathcal{N} \quad (123)$$

$$y_{ij}^{st} \leq x_{ij} \qquad (i,j) \in A, s,t \in V \quad (124)$$

$$\sum_{s,t \in V} \varphi_{ij}^{st} y_{ij}^{st} \leq \beta_{ij} x_{ij} \qquad (i,j) \in A \quad (125)$$

$$\sum_{(i,j) \in A} y_{ij}^{st} \geq K_{min} \qquad s,t \in V \quad (126)$$

$$x_{ij} \in \{0,1\} \qquad (i,j) \in A \quad (127)$$

$$y_{ij}^{st} \in [0,1] \qquad (i,j) \in A, s,t \in V \quad (128)$$

Robust Formulation

Motivations

- Capture utility dependency on temporal variability of traffic demands
  - Given a certain monitoring budget, whether the corresponding layout will cope with traffic variability
  - Determine if increasing monitoring budget enables to better cope with traffic variability and to which extend

Robust Formulation

Motivations

- Capture utility dependency on temporal variability of traffic demands
  - Given a certain monitoring budget, whether the corresponding layout will cope with traffic variability
  - Determine if increasing monitoring budget enables to better cope with traffic variability and to which extend

Method

- Assumption: flow variables obtained by resolving the MCF or the single path MMCF problem
- Reformulate utility maximization problem such that uncertainty in traffic demands translates into uncertainty of corresponding flow parameters φ_{ij}^{st} in budget constraints (123) and monitoring capacity constraints (125)
- Scenarios where uncertainty in traffic demands does not lead to spatial flow re-routing

Box+Polyhedral perturbation set $\mathcal{Z}_{\infty,1}$

- Uncertainty in traffic flows (model parameters): $\varphi_{ij}^{st} \rightarrow \bar{\varphi}_{ij}^{st} + \xi^{st}\hat{\varphi}_{ij}^{st}, \ \forall s,t \in W$
- Constraints (123) and (125) rewritten by grouping deterministic and uncertain part:

$$\sum_{(i,j)\in A} \mathtt{n}_{ij} \sum_{s,t\in V} \varphi_{ij}^{st} y_{ij}^{st} \leq \mathcal{N} \rightarrow \sum_{(i,j)\in A} \nu_{ij} \Big\{ \sum_{s,t\in V} \bar{\varphi}_{ij}^{st} y_{ij}^{st} + \max_{\xi^{st}\in\mathcal{U}} \sum_{s,t\in W} \xi^{st}\hat{\varphi}_{ij}^{st} y_{ij}^{st} \Big\} \leq \mathcal{N} \tag{129}$$

$$\sum_{s,t\in V} \bar{\varphi}_{ij}^{st} y_{ij}^{st} + \max_{\xi^{st}\in\mathcal{U}} \sum_{s,t\in W} \xi^{st}\hat{\varphi}_{ij}^{st} y_{ij}^{st} \leq \beta_{ij} x_{ij}, \quad (i,j)\in A \tag{130}$$

where uncertainty set $\mathcal{U}(\xi)$ with $\xi \in \mathcal{Z}$ given by:

$$\mathcal{U} = \Big\{ \varphi_{ij}^{st} = \bar{\varphi}_{ij}^{st} + \xi^{st}\hat{\varphi}_{ij}^{st} \mid \xi^{st} \in \mathcal{Z}_{\infty,1} \Big\} \tag{131}$$

$$\mathcal{Z}_{\infty,1} = \Big\{ \xi \in \mathbb{R}^{|W|} \mid \sum_{s,t\in W} |\xi^{st}| \leq \Gamma, |\xi^{st}| \leq \Psi, \forall s,t \in W \Big\} \tag{132}$$

- Robust counterpart of constraints (123) and (125) equivalently reformulated as:

$$\sum_{(i,j)\in A} \mathtt{n}_{ij} \Big\{ \sum_{s,t\in V} \bar{\varphi}_{ij}^{st} y_{ij}^{st} + \Psi \sum_{s,t\in W} w_{ij}^{st} + \Gamma z_{ij} \Big\} \leq \mathcal{N} \tag{133}$$

$$\sum_{s,t\in V} \bar{\varphi}_{ij}^{st} y_{ij}^{st} + \Psi \sum_{s,t\in W} w_{ij}^{st} + \Gamma z_{ij} \leq \beta_{ij} x_{ij}, \quad (i,j)\in A \tag{134}$$

Box+Ellipsoidal perturbation set $\mathcal{Z}_{\infty,2}$

- Uncertainty in traffic flows (model parameters): $\varphi_{ij}^{st} \rightarrow \bar{\varphi}_{ij}^{st} + \xi^{st}\hat{\varphi}_{ij}^{st},\ \forall s,t \in W$
- Constraints (123) and (125) rewritten by grouping deterministic and uncertain part:

$$\sum_{(i,j)\in A} \mathtt{n}_{ij} \sum_{s,t\in V} \varphi_{ij}^{st} y_{ij}^{st} \leq \mathcal{N} \rightarrow \sum_{(i,j)\in A} \nu_{ij} \Big\{ \sum_{s,t\in V} \bar{\varphi}_{ij}^{st} y_{ij}^{st} + \max_{\xi^{st}\in\mathcal{U}} \sum_{s,t\in W} \xi^{st}\hat{\varphi}_{ij}^{st} y_{ij}^{st} \Big\} \leq \mathcal{N} \quad (137)$$

$$\sum_{s,t\in V} \bar{\varphi}_{ij}^{st} y_{ij}^{st} + \max_{\xi^{st}\in\mathcal{U}} \sum_{s,t\in W} \xi^{st}\hat{\varphi}_{ij}^{st} y_{ij}^{st} \leq \beta_{ij} x_{ij},\ \ (i,j)\in A \quad (138)$$

where uncertainty set $\mathcal{U}(\xi)$ with $\xi \in \mathcal{Z}_{\infty,2}$ given by:

$$\mathcal{U} = \Big\{ \varphi_{ij}^{st} = \bar{\varphi}_{ij}^{st} + \xi^{st}\hat{\varphi}_{ij}^{st} \mid \xi^{st} \in \mathcal{Z}_{\infty,2} \Big\} \quad (139)$$

$$\mathcal{Z}_{\infty,2} = \Big\{ \xi \in \mathbb{R}^{|W|} \mid \sum_{s,t\in W} |\xi^{st}| \leq \Gamma,\ \sum_{s,t\in W} (\xi^{st})^2 \leq \Omega^2 \Big\} \quad (140)$$

- Robust counterpart of constraints (123) and (125) equivalently reformulated as:

$$\sum_{(i,j)\in A} \mathtt{n}_{ij} \Big\{ \sum_{s,t\in V} \bar{\varphi}_{ij}^{st} y_{ij}^{st} + \Psi \sum_{s,t\in W} w_{ij}^{st} + \Omega \sqrt{\sum_{s,t\in W} (\hat{\varphi}_{ij}^{st})^2 (z_{ij}^{st})^2} \Big\} \leq \mathcal{N} \quad (141)$$

$$\sum_{s,t\in V} \bar{\varphi}_{ij}^{st} y_{ij}^{st} + \Psi \sum_{s,t\in W} w_{ij}^{st} + \Omega \sqrt{\sum_{s,t\in W} (\hat{\varphi}_{ij}^{st})^2 (z_{ij}^{st})^2} \leq \beta_{ij} x_{ij},\ \ (i,j)\in A \quad (142)$$

Modeling uncertainty and sets (1)

Gaussian process

$\{X(t) : t \geq 0\} : dX(t) = \sigma dB_H(t) + \mu dt$ with solution $X(t) = \sigma B_H(t) + \mu t$

- Mean function: $E[X(t)] = \mu t$
- Variance function: $V[X(t)] = E[(X(t) - \mu)^2] = \sigma^2 t^{2H}$
- Hurst parameter (index): $H(0 < H < 1)$
  - If $H = 1/2$ (Brownian motion): stationary and independent increments (short-range dependence, autocorrelations decay exponentially)
  - If $H > 1/2$ (Fractional Brownian motion): stationary and positively correlated increments (long-range dependence, autocorrelations decay hyperbolically, self-similarity)
- Definitions
  - Independent increments: for any $0 \leq s_1 < t_1 \leq s_2 < t_2 \leq < \ldots < s_{n-1} \leq t_{n-1} < t_n < \infty, X_{t_i} - X_{s_i}$ are independent random variables
  - Stationary increments: probability distribution of any increment $X(t) - X(s)$ depends only on the length $t - s$ of the time interval (if $\{X(t) - X(s)\}$ independent of s) $\rightarrow$ for any $s < t$, $X(t) - X(s)$ distributionally equivalent to X_{t-s}

Modeling uncertainty and sets (2)

Poisson-Pareto process

- Superposition of independent traffic bursts (H) of variable length
- Bursts lengths follows Pareto distribution with scale parameter δ and shape parameter (decay rate $\alpha = 3 - 2H$) $\rightarrow$ complementary distribution: $P(b > t) = \left(\frac{\delta}{t}\right)^{3-2H}$ if $t \geq \delta$ (1, otherwise)
- Bursts arrival follows Poisson process with rate λ

Gaussian process with $H = \frac{1}{2}$

Traffic parameter: $\varphi_{st} = E[\varphi_{st}] + \xi_{st}(\epsilon_{st} E[\varphi_{st}])$

Polyhedral set: $\mathcal{Z}_1$

- Expected value: $E[\varphi_{st}] \rightarrow \mu_{st}$
- Perturbation (mean abs. deviation): $E[|\varphi_{st} - E[\varphi_{st}]|] = E[|\varphi_{st} - \mu_{st}|] \rightarrow \sqrt{\frac{2}{\pi}}\sigma_{st}$

$\rightarrow \xi_{st} = \frac{\varphi_{st} - E[\varphi_{st}]}{\epsilon_{st} E[\varphi_{st}]} = \sqrt{\frac{\pi}{2}} \left[\frac{\varphi_{st} - \mu_{st}}{\sigma_{st}} \right]$

$\rightarrow \mathcal{Z}_1 = \left\{ \varphi_{st} \in \mathbb{R}^{n \times n} \mid \sum_{(s,t) \in W} \sqrt{\frac{\pi}{2}} \left[\frac{|\varphi_{st} - \mu_{st}|}{\sigma_{st}} \right] \leq \Gamma \right\}$

where, μ_{st} and σ_{st} given by the model (see next slide)

Gaussian process with $H = \frac{1}{2}$

Traffic parameter: $\varphi_{st} = E[\varphi_{st}] + \xi_{st}(\epsilon_{st} E[\varphi_{st}])$

Polyhedral set: $\mathcal{Z}_1$

- Expected value: $E[\varphi_{st}] \rightarrow \mu_{st}$
- Perturbation (mean abs. deviation): $E[|\varphi_{st} - E[\varphi_{st}]|] = E[|\varphi_{st} - \mu_{st}|] \rightarrow \sqrt{\frac{2}{\pi}}\sigma_{st}$

  $\rightarrow \xi_{st} = \frac{\varphi_{st} - E[\varphi_{st}]}{\epsilon_{st} E[\varphi_{st}]} = \sqrt{\frac{\pi}{2}} \left[\frac{\varphi_{st} - \mu_{st}}{\sigma_{st}}\right]$

  $\rightarrow \mathcal{Z}_1 = \left\{\varphi_{st} \in \mathbb{R}^{n \times n} \mid \sum_{(s,t) \in W} \sqrt{\frac{\pi}{2}} \left[\frac{|\varphi_{st} - \mu_{st}|}{\sigma_{st}}\right] \leq \Gamma\right\}$

  where, μ_{st} and σ_{st} given by the model (see next slide)

Ellipsoidal set: $\mathcal{Z}_2$

- Expected value: $E[\varphi_{st}] \rightarrow \mu_{st}$
- Perturbation (standard deviation): $\sqrt{E[(\varphi^{st} - E[\varphi^{st}])^2]} = \sqrt{E[(\varphi_{st} - \bar{\varphi_{st}})^2]} \rightarrow \sigma^{st}$

  $\rightarrow \xi_{st} = \frac{\varphi_{st} - E[\varphi_{st}]}{\epsilon_{st} E[\varphi_{st}]} = \frac{\varphi_{st} - \mu_{st}}{\sigma_{st}} \rightarrow \mathcal{Z}_2 = \left\{\varphi_{st} \in \mathbb{R}^{n \times n} \mid \sum_{(s,t) \in W} \left[\frac{\varphi_{st} - \mu_{st}}{\sigma_{st}}\right]^2 \leq \Omega^2\right\}$

  where, μ_{st} and σ_{st} given by the model (see next slide)

Note: relation standard (L_2-norm) and mean deviation (L_1-norm): $\triangleq \frac{2}{\pi}\sigma \Rightarrow \text{MAD} \leq \text{SD}$

Build corresponding set from data (1)

Procedure (1)

- Parameter estimation: let x_i be i^{th} independent observation of random variable X
  - Sample mean: $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$
  - Sample variance (unbiased): $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$
- If more than one sample then check if they come from the same distribution

  `If` number of samples $= 2$
  `Then` $2-$sample Kolmogorov-Smirnov test (Anderson-Darling test)
  `Else` $r-$sample Kolmogorov-Smirnov test (Anderson-Darling test)
- Hurst parameter test: variance-time plot (more elaborated in frequency domain: Whittle MLE estimator and wavelet-based)
  - Aggregated time series (to level m): $X^{(m)} = \{X_k^{(m)} : k = 1, 2, \cdots\}, m = 1, 2, \cdots$ with $X_k^{(m)} = \frac{1}{m}\sum_{j=km-m+1}^{km} X_j$
  - Estimate variance of $X^{(m)}$: $\hat{\sigma}^2_{(m)} \simeq \sum_k (X_k^{(m)} - \bar{X})$
  - Plot $(log(m), log(\hat{\sigma}^2_{(m)}))$
  - Compute slope $\triangleq 2\bar{H} - 2$ (negatively biased)

Build corresponding set from data (2)

Procedure (2)

- Normality test:
  - `If` $H \to 1/2$
  - `Then` Shapiro-Wilk test to verify if (random) sample comes from specifically a normal distribution
  - `Else` see next slide
- Extract model: (non-linear least squares) curve fitting problem
  - Non-linear regression problem (minimize weighted sum of squared residuals) in stat. referred to as χ^2
  - **Levenberg-Marquardt algorithm**: iterative procedure combining gradient descent and Gauss-Newton algorithm
  - Requires good starting (adjustable) parameter values (μ, σ^2) and choice of damping parameter (influences both descent direction and step-size)
- Goodness of fit test
  - As new samples comes perform (1-sample) KS- or Anderson-Darling test to determine if it can be explained by this model.
  - Adjust the model or build a new model (several aggregates in macro-flows)

Gaussian process with $\frac{1}{2} < H < 1$

Procedure (3)

- Fit data to Pareto distribution (characterizing bursts length): goodness of fit test
  - Estimation of scale parameter δ using the maximum likelihood estimator (MLE) $\equiv$ smallest observation
  - Data transformation: if X follows Pareto distribution with shape parameter α then $Y = \ln(\frac{X}{\hat{\delta}})$ follows exponential distribution with scale parameter α
  - Sum of weighted increments of the form $\bar{u} = \frac{1}{n-1} \frac{\sum_{i=1}^{n-1}(\sum_{j=1}^{i}(X_j - X_{j-1}))}{\sum_{j=1}^{n}(X_j - X_{j-1})}$
  - Test statistic for linear component $Z_1(\bar{u})$ and quadratic component $Z_2(\bar{u})$ such that $Z_0 = Z_1^2 + Z_2^2$
  - Reject null hypothesis if $Z_0 > \chi^2_{2,\alpha}$

Gaussian process with $\frac{1}{2} < H < 1$

Procedure (4)

Case 1: **independent** short-duration bursts Poisson process X (rate λ_1) and long-duration bursts Poisson process Y (rate λ_2) $\rightarrow B = X + Y$ Poisson process (rate $\lambda = \lambda_1 + \lambda_2$)

- Safety margin
  - Per pair (s, t) determine common distribution B (forward recurrence time of Pareto distribution)
  - Set τ such that $\lambda E(B)P(B > \tau)$ captures sufficient large number of (long) bursts to produce LRD
  - Compute $E[B]$ and $Var[B]$ over period $[t, t + \tau]$ (for non-equal non-constant burst rate, less trivial)
  - Derive max.admissible burst size ($\mathrm{b}^{st} = \hat{\varphi}$ as safety margin)
- Expected value
  - Determine if short-burst data follows Poisson distribution: χ^2 goodness of fit test using $\hat{\lambda}_1$ computed from data
  - If λ_1 (model) $\gg 1$ then $\bar{\varphi}$ derived from Gaussian model
  - Else $\bar{\varphi} = E[X] = \lambda_1$

Case 2: **dependent** short-duration bursts Poisson process X (rate λ_1) and long-duration bursts Poisson process Y (rate λ_2) $\rightarrow Z = X + Y|X$ (mixed Poisson model)

Robust formulation: Gaussian model with $\mathcal{Z}_{\infty,1}$

$$\max \sum_{s,t \in V} u^{st}\Big(\sum_{(i,j)\in A} y_{ij}^{st}\Big) \tag{145}$$

subject to:

$$\sum_{(i,j)\in A} \mathtt{m}_{ij} x_{ij} \leq \mathcal{M} \tag{146}$$

$$\sum_{(i,j)\in A} \mathtt{n}_{ij} \Big\{ \sum_{s,t\in V} \mu^{st} y_{ij}^{st} + \Psi \sum_{s,t\in W} w_{ij}^{st} + \Gamma z_{ij} \Big\} \leq \mathcal{N} \tag{147}$$

$$z_{ij} + w_{ij}^{st} \geq \sqrt{\frac{2}{\pi}} \sigma^{st} y_{ij}^{st} \qquad s,t \in W, (i,j) \in A \tag{148}$$

$$y_{ij}^{st} \leq x_{ij}^{st} \qquad (i,j) \in A, s,t \in V \tag{149}$$

$$\sum_{s,t\in V} \mu^{st} y_{ij}^{st} + \Psi \sum_{s,t\in W} w_{ij}^{st} + \Gamma z_{ij} \leq \beta_{ij} x_{ij} \qquad (i,j) \in A \tag{150}$$

$$\sum_{(i,j)\in A} y_{ij}^{st} \geq K_{min} \qquad s,t \in V \tag{151}$$

$$x_{ij}^{st} \in \{0,1\} \qquad (i,j) \in A \tag{152}$$

$$y_{ij}^{st} \in [0,1] \qquad (i,j) \in A, s,t \in V \tag{153}$$

$$w_{ij}^{st} \geq 0 \qquad (i,j) \in A, s,t \in V \tag{154}$$

$$z_{ij} \geq 0 \qquad (i,j) \in A \tag{155}$$

Robust formulation: Gaussian model with $\mathcal{Z}_{\infty,2}$

$$\max \sum_{s,t \in V} u^{st}\Big(\sum_{(i,j)\in A} y_{ij}^{st}\Big) \tag{156}$$

subject to:

$$\sum_{(i,j)\in A} \mathtt{m}_{ij} x_{ij} \leq \mathcal{M} \tag{157}$$

$$\sum_{(i,j)\in A} \mathtt{n}_{ij} \Big\{ \sum_{s,t\in V} \mu^{st} y_{ij}^{st} + \Psi \sum_{s,t\in W} \sigma^{st} |y_{ij}^{st} - z_{ij}^{st}| + \Omega \sqrt{\sum_{s,t\in W} (\sigma^{st})^2 (z_{ij}^{st})^2} \Big\} \leq \mathcal{N} \tag{158}$$

$$y_{ij}^{st} \leq x_{ij}^{st} \qquad (i,j) \in A, s,t \in V \tag{159}$$

$$\sum_{s,t\in V} \mu^{st} y_{ij}^{st} + \Psi \sum_{s,t\in W} \sigma^{st} |y_{ij}^{st} - z_{ij}^{st}| + \Omega \sqrt{\sum_{s,t\in W} (\sigma^{st})^2 (z_{ij}^{st})^2} \leq \beta_{ij} x_{ij} \qquad (i,j) \in A \tag{160}$$

$$\sum_{(i,j)\in A} y_{ij}^{st} \geq K_{min} \qquad s,t \in V \tag{161}$$

$$x_{ij}^{st} \in \{0,1\} \qquad (i,j) \in A \tag{162}$$

$$y_{ij}^{st} \geq 0 \qquad (i,j) \in A, s,t \in V \tag{163}$$

$$z_{ij}^{st} \geq 0 \qquad (i,j) \in A, s,t \in V \tag{164}$$

Robust formulation: Poisson-Pareto model with $\mathcal{Z}_{\infty,1}$

$$\max \sum_{s,t \in V} u^{st}\Big(\sum_{(i,j)\in A} y_{ij}^{st}\Big) \tag{165}$$

subject to:

$$\sum_{(i,j)\in A} \mathtt{m}_{ij} x_{ij} \leq \mathcal{M} \tag{166}$$

$$\sum_{(i,j)\in A} \mathtt{n}_{ij} \Big\{ \sum_{s,t \in V} \lambda^{st} y_{ij}^{st} + \Psi \sum_{s,t \in W} w_{ij}^{st} + \Gamma z_{ij} \Big\} \leq \mathcal{N} \tag{167}$$

$$z_{ij} + w_{ij}^{st} \geq \mathtt{b}^{st} y_{ij}^{st} \qquad s,t \in W, (i,j) \in A \tag{168}$$

$$y_{ij}^{st} \leq x_{ij}^{st} \qquad (i,j) \in A, s,t \in V \tag{169}$$

$$\sum_{s,t \in V} \lambda^{st} y_{ij}^{st} + \Psi \sum_{s,t \in W} w_{ij}^{st} + \Gamma z_{ij} \leq \beta_{ij} x_{ij} \qquad (i,j) \in A \tag{170}$$

$$\sum_{(i,j)\in A} y_{ij}^{st} \geq K_{min} \qquad s,t \in V \tag{171}$$

$$x_{ij}^{st} \in \{0,1\} \qquad (i,j) \in A \tag{172}$$

$$y_{ij}^{st} \in [0,1] \qquad (i,j) \in A, s,t \in V \tag{173}$$

$$w_{ij}^{st} \geq 0 \qquad (i,j) \in A, s,t \in V \tag{174}$$

$$z_{ij} \geq 0 \qquad (i,j) \in A \tag{175}$$

Evaluation instances: Topologies

Topologies (SNDLib database)

Topology	Nodes	Links	Min,Max,Avg Degree	Diameter
atlanta	15	22	2;4;2.93	5
cost266	37	57	2;5;3.08	8
france	25	45	2;10;3.60	5
geant	22	36	2;8;3.27	5
india35	35	80	2;9;4.57	7
newyork	16	49	2;11;6.12	3
nobel-eu	28	41	2;5;2.93	8
norway	27	51	2;6;3.78	7

Evaluation instances: Topologies

Topologies (SNDLib database)

Topology	Nodes	Links	Min,Max,Avg Degree	Diameter
atlanta	15	22	2;4;2.93	5
cost266	37	57	2;5;3.08	8
france	25	45	2;10;3.60	5
geant	22	36	2;8;3.27	5
india35	35	80	2;9;4.57	7
newyork	16	49	2;11;6.12	3
nobel-eu	28	41	2;5;2.93	8
norway	27	51	2;6;3.78	7

- Link capacities and costs provided by SNDlib database
- Traffic demands provided by SNDlib database for these topologies

Execution

Environment

- IBM ILOG OPL modeling language and solved it with CPLEX 12.6
- Execution on dedicated server with 8 Intel Xeon quad-core processors and 512GB DDR3 RAM
- Linux CENTOS 6.5

Execution

Environment

- IBM ILOG OPL modeling language and solved it with CPLEX 12.6
- Execution on dedicated server with 8 Intel Xeon quad-core processors and 512GB DDR3 RAM
- Linux CENTOS 6.5

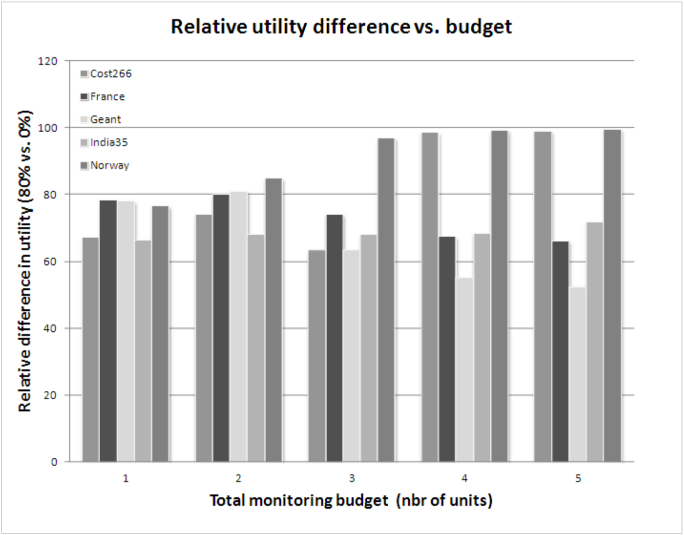
Execution

- Add constraints $\mathcal{M} + \mathcal{N} \leq_{\text{cost}}$ and give total budget $_{\text{cost}}$ as input
- Step-increase of total monitoring cost and determine utility obtained while maximizing total fraction k of monitored traffic
  - Fraction k does not apply equally to each traffic flow (report average monitoring fraction over traffic flows)
  - Each execution runs up to 3600s for each step
- Traffic demands experiencing perturbation from 0% (no perturbation) to 80% with steps of 5%
- utility function (parameter values): $a = 1$ and $b = 6.3$

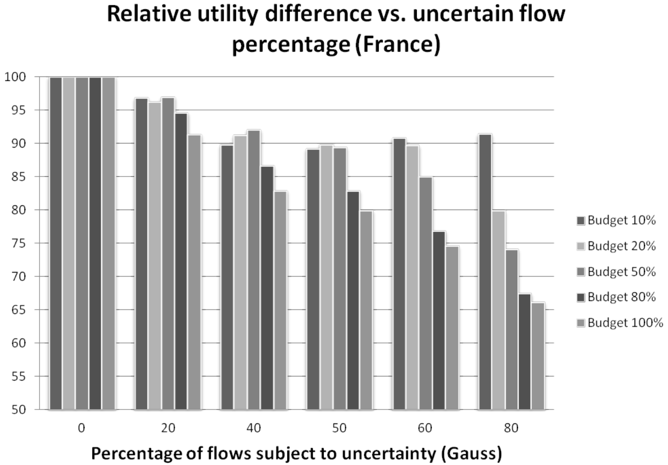
Numerical Results: Gaussian model (H = 1/2)

	0.00			0.20			0.40			0.50			0.60			0.80		
$\mathcal{C}_{\text{tot}}$	avg(k)	N_m	u	avg(k)	N_m	u	avg(k)	N_m	u	avg(k)	N_m	u	avg(k)	N_m	u	avg(k)	N_m	u
France																		
1100	0.13	9	313	0.12	8	304	0.12	6	282	0.12	7	280	0.12	9	285	0.10	8	247
2200	0.24	11	581	0.23	11	559	0.22	9	530	0.22	11	522	0.22	12	521	0.19	13	464
5500	0.55	19	1314	0.53	18	1265	0.51	21	1203	0.49	24	1171	0.47	30	1112	0.41	28	969
8800	0.84	26	1983	0.79	38	1878	0.72	47	1711	0.69	47	1641	0.64	40	1522	0.56	36	1335
11000	0.99	55	2344	0.90	51	2138	0.82	49	1935	0.79	50	1868	0.74	44	1739	0.65	40	1543
India35																		
1000	0.18	16	839	0.16	14	776	0.15	15	702	0.14	14	667	0.13	14	651	0.11	13	558
2000	0.31	24	1483	0.29	23	1370	0.27	25	1263	0.25	24	1205	0.24	22	1166	0.21	21	1004
5000	0.64	42	3022	0.59	39	2796	0.56	42	2624	0.52	38	2470	0.50	40	2370	0.44	40	2064
8000	0.90	50	4212	0.83	55	3889	0.78	58	3656	0.73	53	3427	0.69	50	3259	0.61	47	2876
10000	0.99	86	4636	0.93	71	4375	0.88	69	4117	0.83	63	3894	0.79	57	3712	0.71	54	3336
Norway																		
45000	0.50	14	2769	0.47	14	2613	0.44	12	2459	0.43	12	2391	0.41	10	2294	0.38	10	2133
90000	0.72	25	3987	0.71	24	3949	0.69	22	3804	0.67	23	3723	0.66	21	3667	0.63	20	3490
225000	0.93	60	5131	0.92	58	5112	0.92	55	5081	0.92	53	5078	0.91	55	5048	0.91	51	5018
360000	0.97	87	5377	0.97	86	5360	0.97	82	5376	0.97	83	5372	0.97	83	5348	0.97	79	5349
450000	0.99	99	5491	0.99	99	5479	0.99	96	5494	0.99	97	5473	0.99	95	5469	0.99	93	5466

Results: Safety margin model - Gaussian (1)



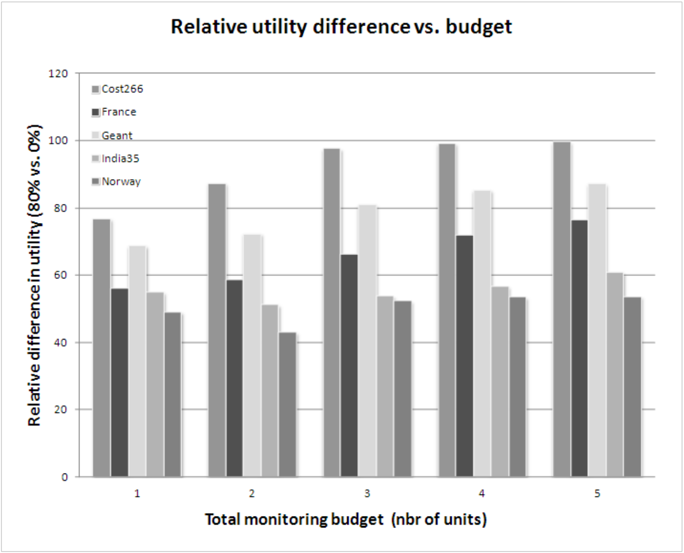
Results: Safety margin model - Gaussian (2)



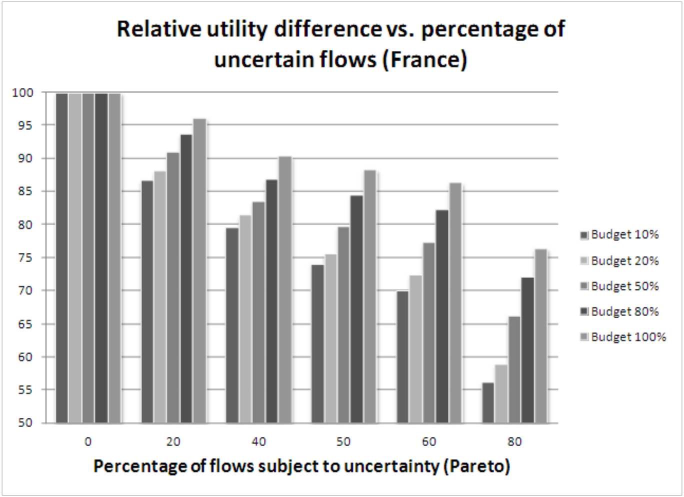
Numerical Results: Poisson-Pareto model (1/2 < H < 1)

	0.00			**0.20**			**0.40**			**0.50**			**0.60**			**0.80**		
$\mathcal{C}_{tot}$	avg(k)	N_m	u	avg(k)	N_m	u	avg(k)	N_m	u	avg(k)	N_m	u	avg(k)	N_m	u	avg(k)	N_m	u
France																		
1100	0.13	9	313	0.12	8	304	0.12	6	282	0.12	7	280	0.12	9	285	0.10	8	247
2200	0.24	11	581	0.23	11	559	0.22	9	530	0.22	11	522	0.22	12	521	0.19	13	464
5500	0.55	19	1314	0.53	18	1265	0.51	21	1203	0.49	24	1171	0.47	30	1112	0.41	28	969
8800	0.84	26	1983	0.79	38	1878	0.72	47	1711	0.69	47	1641	0.64	40	1522	0.56	36	1335
11000	0.99	55	2344	0.90	51	2138	0.82	49	1935	0.79	50	1868	0.74	44	1739	0.65	40	1543
India35																		
1000	0.18	16	839	0.16	14	776	0.15	15	702	0.14	14	667	0.13	14	651	0.11	13	558
2000	0.31	24	1483	0.29	23	1370	0.27	25	1263	0.25	24	1205	0.24	22	1166	0.21	21	1004
5000	0.64	42	3022	0.59	39	2796	0.56	42	2624	0.52	38	2470	0.50	40	2370	0.44	40	2064
8000	0.90	50	4212	0.83	55	3889	0.78	58	3656	0.73	53	3427	0.69	50	3259	0.61	47	2876
10000	0.99	86	4636	0.93	71	4375	0.88	69	4117	0.83	63	3894	0.79	57	3712	0.71	54	3336
Norway																		
45000	0.50	14	2769	0.47	14	2613	0.44	12	2459	0.43	12	2391	0.41	10	2294	0.38	10	2133
90000	0.72	25	3987	0.71	24	3949	0.69	22	3804	0.67	23	3723	0.66	21	3667	0.63	20	3490
225000	0.93	60	5131	0.92	58	5112	0.92	55	5081	0.92	53	5078	0.91	55	5048	0.91	51	5018
360000	0.97	87	5377	0.97	86	5360	0.97	82	5376	0.97	83	5372	0.97	83	5348	0.97	79	5349
450000	0.99	99	5491	0.99	99	5479	0.99	96	5494	0.99	97	5473	0.99	95	5469	0.99	93	5466

Results: Large perturbation model - Pareto (1)



Results: large perturbation model - Pareto (2)



Summary

Summary

- Formulation involves a number of constraints $O(|V|^3)$ $\rightarrow$ Resolution reaches computational limits of MIP solvers CPLEX
  $\rightarrow$ Limit on instances size (in particular for MISOCP)
  $\Rightarrow$ More efficient resolution methods required to cope with combinatorial explosion of monitoring utility maximization problem

Summary

Summary

- Formulation involves a number of constraints $O(|V|^3)$ $\rightarrow$ Resolution reaches computational limits of MIP solvers `CPLEX`
  $\rightarrow$ Limit on instances size (in particular for MISOCP)
  $\Rightarrow$ More efficient resolution methods required to cope with combinatorial explosion of monitoring utility maximization problem

Next steps

- Extend proposed formulation to multiple time period problems (instead of a single period)
- Confront model/results with real data traces (instead of synthetic traces)

Future Research Work

Open research questions

- Combine structural with behavioral properties to automate learning of uncertainty sets
- Predict best fit and combination of uncertainty sets
- Extend set-induced (data-driven) RO to non-i.i.d. data/coefficients (more general data assumptions to construct more representative sets $\Rightarrow$ more difficult to derive a priori probability bounds)

Outline

Conclusion (1)

1. (Reliable) capacitated Facility Location Problem (cFLP) $\times$ Multicommodity Flow Routing (MCF) $\rightarrow$ cFLRP

Conclusion (1)

1. (Reliable) capacitated Facility Location Problem (cFLP) $\times$ Multicommodity Flow Routing (MCF) $\rightarrow$ cFLRP
2. Hub-Location Problem (HLP) $\times$ Location-Routing Problem (LRP) $\rightarrow$ HLRP

Conclusion (1)

1. (Reliable) capacitated Facility Location Problem (cFLP) $\times$ Multicommodity Flow Routing (MCF) $\rightarrow$ cFLRP
2. Hub-Location Problem (HLP) $\times$ Location-Routing Problem (LRP) $\rightarrow$ HLRP
3. Mixed-Integer Programming Model for the Multi-Stage Hub Location Problem $\rightarrow$ mHLRP

Conclusion (1)

1. (Reliable) capacitated Facility Location Problem (cFLP) $\times$ Multicommodity Flow Routing (MCF) $\rightarrow$ cFLRP

2. Hub-Location Problem (HLP) $\times$ Location-Routing Problem (LRP) $\rightarrow$ HLRP

3. Mixed-Integer Programming Model for the Multi-Stage Hub Location Problem $\rightarrow$ mHLRP

4. Robust cFLP (variant of)

Conclusion (1)

1. (Reliable) capacitated Facility Location Problem (cFLP) $\times$ Multicommodity Flow Routing (MCF) $\rightarrow$ cFLRP
2. Hub-Location Problem (HLP) $\times$ Location-Routing Problem (LRP) $\rightarrow$ HLRP
3. Mixed-Integer Programming Model for the Multi-Stage Hub Location Problem $\rightarrow$ mHLRP
4. Robust cFLP (variant of)
5. Multi-Period Multicommodity Capacitated Network Design and Routing Problem

Conclusion (1)

1. (Reliable) capacitated Facility Location Problem (cFLP) $\times$ Multicommodity Flow Routing (MCF) $\rightarrow$ cFLRP
2. Hub-Location Problem (HLP) $\times$ Location-Routing Problem (LRP) $\rightarrow$ HLRP
3. Mixed-Integer Programming Model for the Multi-Stage Hub Location Problem $\rightarrow$ mHLRP
4. Robust cFLP (variant of)
5. Multi-Period Multicommodity Capacitated Network Design and Routing Problem

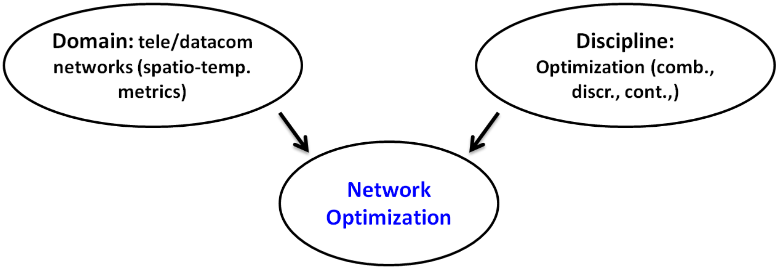
Conclusion (1)

1. (Reliable) capacitated Facility Location Problem (cFLP) × Multicommodity Flow Routing (MCF) → cFLRP
2. Hub-Location Problem (HLP) × Location-Routing Problem (LRP) → HLRP
3. Mixed-Integer Programming Model for the Multi-Stage Hub Location Problem → mHLRP
4. Robust cFLP (variant of)
5. Multi-Period Multicommodity Capacitated Network Design and Routing Problem

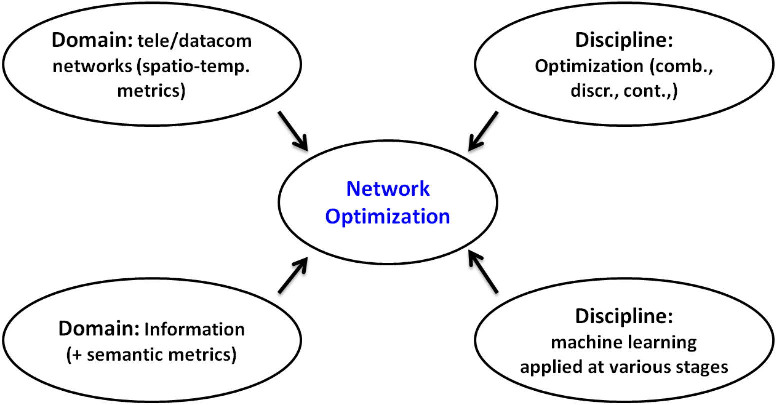
Challenges

- Modeling-level: multi-layer (combined problems/unified operations), multi-period (dynamics), robustification (uncertainty),
- Computational-level: methods/techniques (exact - heuristics - meta-heuristics)

Conclusion (2)



Conclusion (2)



References (1)

- D.Papadimitriou and B.Fortz, Time-dependent combined network design and routing optimization, Proceedings of IEEE International Conference on Communications (ICC), Sydney, Australia, June 10-14, 2014.
- B.Fortz and D.Papadimitriou, Branch-and-cut strategies for a multi-period network design and routing problem, Extended abstract submitted to International Conference on Control, Decision and Information Technologies (CODIT), Metz, France, 3-5 November 2014.
- D.Papadimitriou and B.Fortz, Methods for time-dependent combined network design and routing optimization, Proceedings of IEEE Global Communications (GLOBECOM) Conference, Austin (TX), December 8-12, 2014.
- D.Papadimitriou and B.Fortz, Distributed monitoring problem, Proc. of 7th International Network Optimization Conference (INOC) 2015, Warsaw, Poland, May 18-20, 2015.
- B.Fortz, E.Gorgone, and D.Papadimitriou, Lagrangean relaxation and resolution for the generalized Fixed-charge Capacitated Network Design (FCND) problem, 7th International Network Optimization Conference (INOC) 2015, Warsaw, Poland, May 18-20, 2015.

References (2)

- D.Papadimitriou, B.Fortz, and E.Gorgone, Lagrangian Relaxation for the Time-dependent Combined Network Design and Routing Problem, Proceedings of IEEE International Conference on Communications (ICC) 2015, IEEE International Conference on Communications (ICC) 2015, London, United-Kingdom (U-K), June 8-12, 2015.
- D.Papadimitriou, and B.Fortz, Robust Cooperative Monitoring Problem, Proceedings of 7th International Workshop on Reliable Networks Design and Modeling (RNDM) 2015, 7th International Workshop on Reliable Networks Design and Modeling (RNDM) 2015, Munich, Germany, October 5-7, 2015.
- D.Papadimitriou, et al., Mixed-Integer Optimization for the Combined capacitated Facility Location-Routing Problem, Proceedings of twelfth International Conference on the Design of Reliable Communication Networks 2016, Paris, France, March 15-17, 2016.
- D.Papadimitriou, D.Colle and P.Demeester, Reconciling Resource Abstraction, Allocation and Routing Optimization, Submitted to IEEE Global Communications Conference (IEEE GLOBECOM) 2016, Washington DC, USA, December 4-8, 2016.