On energy-aware and fair routing in telecommunication networks

Edoardo Amaldi

joint work with Antonio Capone, Stefano Coniglio and Luca Gianoli

Dipartimento di Elettronica, Informazione e Bioingegneria
Politecnico di Milano

OR Days, May 2, 2013

Outline

- **Part I**: Energy-aware IP (Internet Protocol) traffic engineering with shortest path routing
- **Part II**: Network routing subject to max-min fair fow allocation

Outline Part I

- Energy-aware IP traffic engineering with shortest path routing
- The problem and related work
- A partial MILP formulation
- A MILP-based heuristic
- Some computational results
- Concluding remarks

Introduction

In 2007 Internet responsible for 5.5% of global energy consumption with estimated annual increment rate of $\approx$ 20-25%

In 2009 the annual comsumption of largest Internet Service Providers $>$ 10 TWh per year

To save energy in IP networks, one can work at different levels:

hardware, protocol, **energy-aware network management**, green energy

Since

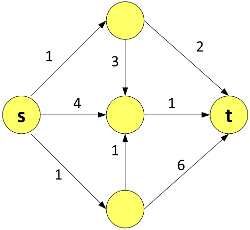
- IP networks are designed to serve the estimated peak traffic demand
- traffic load varies considerably during the day
- power consumption of current network devices is not proportional the traffic load,

aim is to reduce overall energy consumption by putting to sleep links and nodes, according to traffic variations

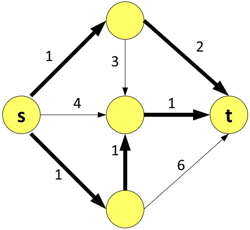
Shortest path routing

Open Shortest Path First (OSPF) is one of the most commonly used intra-domain routing protocol

A **weight** is assigned to each **link**, and each traffic demand d_{st} is routed along the **shortest paths** between the source s and the destination t



$d_{st} = 2$



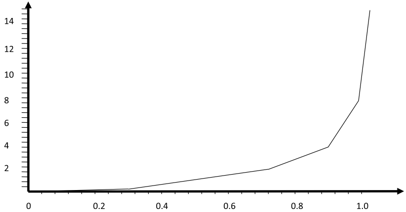
1 unit along each path

Equal cost multi-path (ECMP) rule

Traffic Engineering with shortest path routing

In intra-domanin Traffic Engineering (TE) link weights are adjusted so as to optimize a certain measure of effectiveness

Example: minimize total cost of link utilization [Fortz and Thorup 02]



Other objective functions: residual capacity maximization, load balancing maximization

Related work - energy-aware TE

Shortest paths (OSPF)

- A method based on a variant of the OSPF protocol (neighbouring routers share the shortest path trees) [Cianfrani et al. 10]

  OSPF is modified, weights are fixed

- Lagrangian-based heuristics to find a set of weights that allow to switch-off network links [Lee et al. 12]

  No congestion optimization, switch off only links, worse performance

Flow-based (MPLS)

- Some heuristics to sequentially switch-off network elements [Chiaraviglio et al. 09]
- Simple heuristics based on rate dependent energy profile [Vasic and Kostic 10]

Hybrid (OSPF+MPLS)

- A heuristic based on a MIP formulation that uses previously computed shortest paths [Zhang et al. 10]

  Weights are fixed

Related work - TE with shortest paths

Link weight optimization has been extensively investigated with different objective functions and methods (see e.g. survey [Altin et al. 09])

Local search algorithm (**IGP-WO**) for optimizing link weights so as to minimize the total cost of link utilization, and NP-hardness proof [Fortz and Thorup 02, 04]

IGP-WO available from TOTEM toolbox [Fortz et al.]

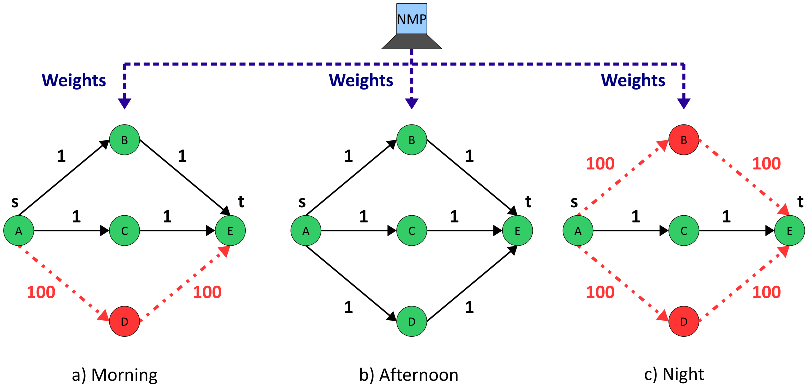
...

"Shortest path routing: modelling, infeasibility and polyhedra" [Call 12]

The approach

Day subdivided into macro periods with "stable" traffic level

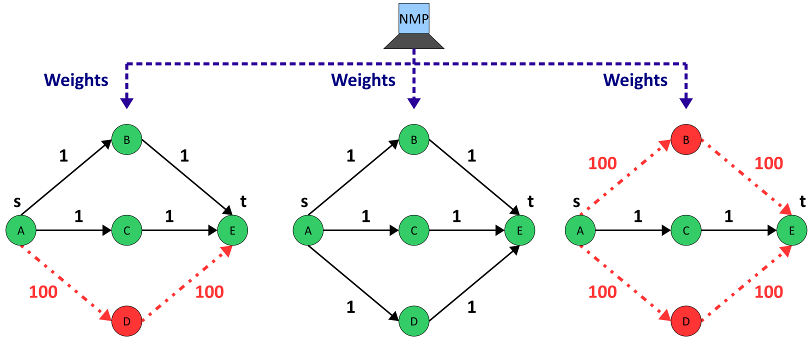
a set of weights is determined for each period



The approach

Links and routers are **switched-off** by excluding them from all shortest paths

by assigning very large link weights

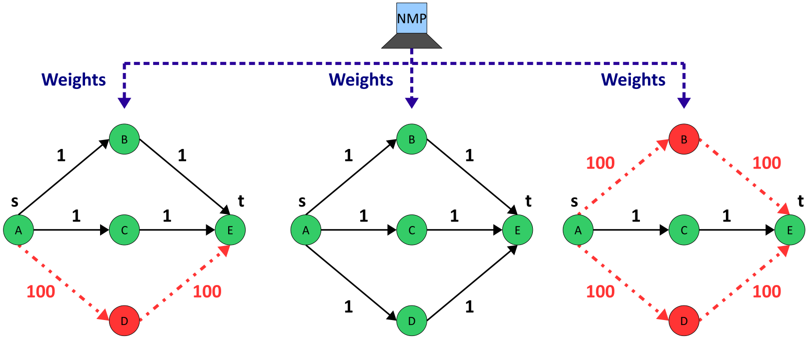


a) Morning b) Afternoon c) Night

The approach

Link weights of active elements must guarantee

the satisfaction of all traffic demands



a) Morning b) Afternoon c) Night

The problem

Given:

- a directed graph $G = (V, A)$ with a capacity assigned to each link $a \in A$
- a traffic matrix D, where d_{st} is the demand from s to t
- a maximum link utilization fraction α, with $0 \leq \alpha \leq 1$

decide which nodes and links to switch off (put in sleeping mode) and determine the link weights so as to minimize

- the total **network energy consumption** (primary objective)
- a measure of the **network congestion** (secondary objective)

while guaranteeing that

- all demads are routed
- the utilization fraction of each link is at most α.

N.B.: network congestion measure = total cost of link utilization

A partial MILP formulation

1. $$x_{ij} = \begin{cases} 1 & \text{if the link } (i,j) \text{ is powered on} \\ 0 & \text{otherwise} \end{cases} \qquad \forall (i,j) \in A$$

2. $$y_i = \begin{cases} 1 & \text{if node } i \text{ is powered on} \\ 0 & \text{otherwise} \end{cases} \qquad \forall i \in V$$

3. $f_{ij}^t \geq 0 \in A$ flow through link (i,j) destined to node t $\forall (i,j)$

4. $w_{ij} \in Z$ weight of link (i,j) $\forall (i,j) \in A$

5. Auxiliary variables needed to define shortest paths

A partial MILP formulation - first part

$$
\begin{array}{llr}
\min & \sum_{(i.j)\in A} p_{ij}x_{ij} + \sum_{k\in V} p_k y_k & \\
s.t. & x_{ij} \le y_i & (i,j)\in A \\
& x_{ij} \le y_j & (i,j)\in A \\
& \sum_{i\in V} f^t_{it} = \sum_{s\in V} d_{st} & t\in V \\
& \sum_{j\in V} f^t_{vj} - \sum_{i\in V} f^t_{iv} = d_{vt} & v,t\in V, t\neq v \\
& \sum_{t\in V} f^t_{ij} \le \alpha c_{ij} x_{ij} & (i,j)\in A \\
& 0 \le z^t_i - f^t_{ij} \le (1-u^t_{ij}) \sum_{v\in V} d_{vt} & t\in V, (i,j)\in A \\
& f^t_{ij} \le u^t_{ij} \sum_{v\in V} d_{vt} & t\in V, (i,j)\in A
\end{array}
$$

Minimize the overall (links + nodes) power consumption

A partial MILP formulation - first part

$$
\begin{array}{llr}
\min & \sum_{(i.j)\in A} p_{ij}x_{ij} + \sum_{k\in V} p_k y_k & \\
s.t. & x_{ij} \leq y_i & (i,j)\in A \\
 & x_{ij} \leq y_j & (i,j)\in A \\
 & \sum_{i\in V} f_{it}^t = \sum_{s\in V} d_{st} & t\in V \\
 & \sum_{j\in V} f_{vj}^t - \sum_{i\in V} f_{iv}^t = d_{vt} & v,t\in V, t\neq v \\
 & \sum_{t\in V} f_{ij}^t \leq \alpha c_{ij} x_{ij} & (i,j)\in A \\
 & 0 \leq z_i^t - f_{ij}^t \leq (1-u_{ij}^t)\sum_{v\in V} d_{vt} & t\in V, (i,j)\in A \\
 & f_{ij}^t \leq u_{ij}^t \sum_{v\in V} d_{vt} & t\in V, (i,j)\in A
\end{array}
$$

Coherence constraints between nodes and incident links

A partial MILP formulation - first part

$$
\begin{array}{lcl}
\min & \sum_{(i.j)\in A} p_{ij}x_{ij} + \sum_{k\in V} p_k y_k & \\
s.t. & x_{ij} \leq y_i & (i,j)\in A \\
 & x_{ij} \leq y_j & (i,j)\in A \\
 & \sum_{i\in V} f_{it}^t = \sum_{s\in V} d_{st} & t\in V \\
 & \sum_{j\in V} f_{vj}^t - \sum_{i\in V} f_{iv}^t = d_{vt} & v,t\in V, t\neq v \\
 & \sum_{t\in V} f_{ij}^t \leq \alpha c_{ij} x_{ij} & (i,j)\in A \\
 & 0 \leq z_i^t - f_{ij}^t \leq (1-u_{ij}^t)\sum_{v\in V} d_{vt} & t\in V, (i,j)\in A \\
 & f_{ij}^t \leq u_{ij}^t \sum_{v\in V} d_{vt} & t\in V, (i,j)\in A
\end{array}
$$

Flow conservation constraints

A partial MILP formulation - first part

$$
\begin{array}{lll}
\min & \sum_{(i,j)\in A} p_{ij}x_{ij} + \sum_{k\in V} p_k y_k & \\
s.t. & x_{ij} \le y_i & (i,j)\in A \\
 & x_{ij} \le y_j & (i,j)\in A \\
 & \sum_{i\in V} f_{it}^t = \sum_{s\in V} d_{st} & t\in V \\
 & \sum_{j\in V} f_{vj}^t - \sum_{i\in V} f_{iv}^t = d_{vt} & v,t\in V, t\neq v \\
 & \sum_{t\in V} f_{ij}^t \le \alpha c_{ij}x_{ij} & (i,j)\in A \\
 & 0 \le z_i^t - f_{ij}^t \le (1-u_{ij}^t)\sum_{v\in V} d_{vt} & t\in V, (i,j)\in A \\
 & f_{ij}^t \le u_{ij}^t \sum_{v\in V} d_{vt} & t\in V, (i,j)\in A
\end{array}
$$

Max utilization constraints

A partial MILP formulation - first part

$$
\begin{array}{llr}
\min & \sum_{(i.j)\in A} p_{ij}x_{ij} + \sum_{k\in V} p_k y_k & \\
s.t. & x_{ij} \leq y_i & (i,j) \in A \\
 & x_{ij} \leq y_j & (i,j) \in A \\
 & \sum_{i\in V} f_{it}^t = \sum_{s\in V} d_{st} & t \in V \\
 & \sum_{j\in V} f_{vj}^t - \sum_{i\in V} f_{iv}^t = d_{vt} & v,t \in V, t \neq v \\
 & \sum_{t\in V} f_{ij}^t \leq \alpha c_{ij} x_{ij} & (i,j) \in A \\
 & 0 \leq z_i^t - f_{ij}^t \leq (1 - u_{ij}^t) \sum_{v\in V} d_{vt} & t \in V, (i,j) \in A \\
 & f_{ij}^t \leq u_{ij}^t \sum_{v\in V} d_{vt} & t \in V, (i,j) \in A
\end{array}
$$

$u_{ij}^t = 1$ if the link (i,j) belongs to a shortest path to t and $u_{ij}^t = 0$ otherwise

z_i^t common value of the flow across all links outgoing from i belonging to shortest paths to t

A partial MILP formulation - second part

$$
\begin{aligned}
0 \le r_j^t + w_{ij} - r_i^t &\le (1 - u_{ij}^t)M && t \in V, (i,j) \in A \\
1 - u_{ij}^t &\le r_j^t + w_{ij} - r_i^t && t \in V, (i,j) \in A \\
u_{ij}^t &\le x_{ij} && (i,j) \in A,\ t \in V \\
w_{ij} &\ge (1 - x_{ij}) w_{max} && (i,j) \in A \\
1 \le w_{ij} &\le w_{max} && (i,j) \in A \\
w_{ij} &\in Z && (i,j) \in A \\
u_{ij}^t &\in \{0,1\} && t \in V, (i,j) \in A \\
x_{ij}, y_k &\in \{0,1\} && (i,j) \in A,\ k \in V \\
f_{ij}^t &\ge 0 && (i,j) \in A,\ t \in V \\
r_i^t,\ z_i^t &\ge 0 && i,t \in V
\end{aligned}
$$

Constraints assuring that

- routing vectors u^t define shortest paths that are consistent with the link weights w_{ij}
- switched off links do not belong to shortest paths and have value w_{max}

A partial MILP formulation - second part

$$
\begin{array}{rl}
0 \leq r_j^t + w_{ij} - r_i^t \leq (1 - u_{ij}^t)M & t \in V, (i,j) \in A \\
1 - u_{ij}^t \leq r_j^t + w_{ij} - r_i^t & t \in V, (i,j) \in A \\
u_{ij}^t \leq x_{ij} & (i,j) \in A,\ t \in V \\
w_{ij} \geq (1 - x_{ij}) w_{max} & (i,j) \in A \\
1 \leq w_{ij} \leq w_{max} & (i,j) \in A \\
w_{ij} \in Z & (i,j) \in A \\
u_{ij}^t \in \{0,1\} & t \in V, (i,j) \in A \\
x_{ij}, y_k \in \{0,1\} & (i,j) \in A,\ k \in V \\
f_{ij}^t \geq 0 & (i,j) \in A,\ t \in V \\
r_i^t,\ z_i^t \geq 0 & i, t \in V
\end{array}
$$

Variables

A partial MILP formulation - second part

$$
\begin{aligned}
0 \le r_j^t + w_{ij} - r_i^t \le (1 - u_{ij}^t)M & \qquad t \in V, (i,j) \in A \\
1 - u_{ij}^t \le r_j^t + w_{ij} - r_i^t & \qquad t \in V, (i,j) \in A \\
u_{ij}^t \le x_{ij} & \qquad (i,j) \in A,\ t \in V \\
w_{ij} \ge (1 - x_{ij}) w_{max} & \qquad (i,j) \in A \\
1 \le w_{ij} \le w_{max} & \qquad (i,j) \in A \\
w_{ij} \in Z & \qquad (i,j) \in A \\
u_{ij}^t \in \{0,1\} & \qquad t \in V, (i,j) \in A \\
x_{ij}, y_k \in \{0,1\} & \qquad (i,j) \in A,\ k \in V \\
f_{ij}^t \ge 0 & \qquad (i,j) \in A,\ t \in V \\
r_i^t,\ z_i^t \ge 0 & \qquad i, t \in V
\end{aligned}
$$

Extension of known formulation (see [Pioro and Medhi 04])

No feasible solution for a small network ($n = 10$, $m = 42$) after 10 hours of cpu time

MILP-based algorithm

Four steps:

1. **Greedy procedure** able to quickly identify a subset of nodes and links that can be switched-off (with $\alpha - 0.1$)

Greedy procedure

Given a network topology, a traffic matrix D, a given α, and a set of "feasible" link weights, we iterate:

1. **Sorting step**: sort network elements (nodes and links) according to a given criterion

   For nodes: Least-Degree (LL), Least-Flow (LF) and Sum-of-Weights (SW)

   For links: Least-Flow (LF) and Weight (W)

2. **Switching-off step**: check whether the first element of the sorted list can be switched off, while respecting maximum utilization constraints (within α)

Nodes before the links, and the 6 combined node-link sorting criteria

Initial weights are obtained with **IGP-WO** (TOTEM toolbox) warm-started with dual variables of a fully-splittable multicommodity flow problem.

MILP-based algorithm

Four steps:

1. **Greedy procedure** able to quickly identify a subset of nodes and links that can be switched-off (with $\alpha - 0.1$)

2. **Switching-off step** based on a **MILP relaxation** with fully splittable routing and no weights (within 3%)

MILP relaxation

$$
\begin{array}{lll}
\min & \sum_{(i.j)\in A} p_{ij}x_{ij} + \sum_{k\in V} p_k y_k & \\
s.t. & x_{ij} \leq y_i & (i,j)\in A \\
 & x_{ij} \leq y_j & (i,j)\in A \\
 & \sum_{i\in V} f_{it}^t = \sum_{s\in V} d_{st} & t \in N \\
 & \sum_{j\in V} f_{vj}^t - \sum_{i\in V} f_{iv}^t = d_{vt} & v,t\in V, t\neq v \\
 & \sum_{t\in V} f_{ij}^t \leq x_{ij}\alpha c_{ij} & (i,j)\in A \\
 & x_{ij}, y_k \in \{0,1\} & (i,j)\in A,\ k\in V \\
 & f_{ij}^t \geq 0 & (i,j)\ \in A,\ t\in V.
\end{array}
$$

Capacitated multicommodity minimum cost flow problem solvable within 3% gap in seconds ($n > 100$ and $m > 300$)

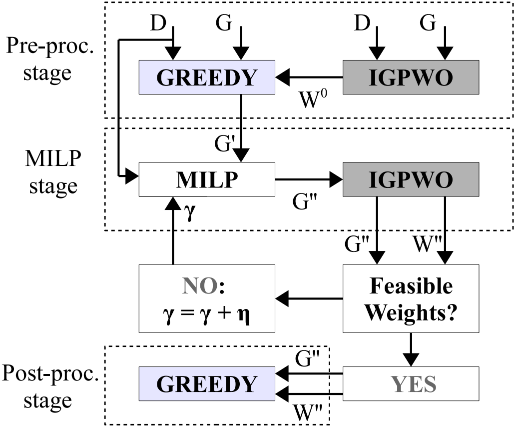
Provides a bound but the resulting reduced network may not admit a feasible set of weights

MILP-based algorithm

Four steps:

1. **Greedy procedure** able to quickly identify a subset of elements that can be switched-off (with $\alpha - 0.1$)
2. **Switching-off step** based on a **MILP relaxation** with fully splittable routing and no weights (within 3%)
3. Existence of **link weights** for reduced network is checked by applying *nb* iterations of **IGP-WO algorithm**. If not found, the traffic matrix is slighty increased and back to Switching-off step.
4. Apply Greedy to the resulting sub-network

MILP-based algorithm



Network topologies

7 real network topologies from the Rocketfuel Project (2 backbone and 5 access networks) and USA28 used in [Lee et al. 12]

Network	Type	Nodes	Links	Edge	Core	%Core
Ebone	Backbone	87	322	31	56	64.4
Exodus	Backbone	79	294	38	41	51.9
Sprint	Access	52	168	52	0	0
AT&T	Access	115	296	115	0	0
Abovenet	Access	19	68	68	0	0
Genuity	Access	42	110	42	0	0
Tiscali	Access	41	174	41	0	0
USA28	Access	28	90	28	0	0

Only edge nodes can be source or destination of traffic

Access networks already used in [Vasic and Kostic 10, Zhang et al. 10, Lee et al. 12]

Network topologies and traffic matrices

Different traffic matrices D for backbone and access networks

Backbone:

- Constant matrices generated with TOTEM
- Poisson matrices generated with TOTEM
- LP matrices generated with linear programming

Access:

- Matrices generated with Gravity model and provided by the authors of [Zhang et al 10]
- Matrices generated with TRUMP and provided by the authors of [Vasic and Kostic 10]
- Matrices generated with the software provided by the authors of [Lee et al. 12]

Computational results - backbone networks

$Inst$	$C-E$	L	$E_c^{tot}\,(W)$	$E_c\,(W)$	gap	N_{off}	L_{off}	$t\,(s)$	$Cong_{\%}$
Ex30	41-38	294	9058.2	4922.2	8.27%	33	176	2189	300%
Ex40	41-38	294	9058.2	5399.2	5.22%	29	158	2024	352%
Ex50	41-38	294	9058.2	6195.0	11.89%	23	120	4983	171%
ExC	41-38	294	9058.2	4704.4	3.67%	34	194	1863	299%
ExP	41-38	294	9058.2	4805.4	3.27%	33	192	1798	253%
Eb30	56-31	322	10126.6	5569.6	0.53%	35	210	2674	302%
Eb40	56-31	322	10126.6	6096.5	3.81%	30	197	2606	279%
Eb50	56-31	322	10126.6	6667.2	5.37%	25	178	3663	213%
EbC	56-31	322	10126.6	5931.0	1.12%	31	198	1868	280%
EbP	56-31	322	10126.6	6010.1	2.47%	31	197	2648	209%

Up to 45% of energy saving and 80% of core nodes switched-off, with moderate increase in congestion

Computing times vary from 30 to 60 minutes

Gap w.r.t. the bound often lower than 5%

Computational results - Zhang et al. access networks

Inst	N	L	$E_c^{tot}\ (W)$	$E_c\ (W)$	L_{off}	$t\ (s)$	$Cong_{\%}$
Spr12	52	168	24972	11950 (48%)	85	578	558%
Spr24	52	168	24972	13339 (53%)	76	584	798%
Spr36	52	168	24972	13795 (55%)	73	1241	428%
AT&T7	115	296	43344	30504 (70%	82	1816	553%
AT&T14	115	296	43344	31026 (72%)	79	1854	415%
AT&T21	115	296	43344	32388 (75%)	70	1990	525%

Up to 50% of energy saving for Sprint network

Lower energy-saving for the AT&T network due to the lower link redundancy and the higher number of leafs/edge nodes

Computing time around 30 minutes in the worst cases

Much larger percentage increase in congestion (up to 8 times) but its absolute value is still reasonable

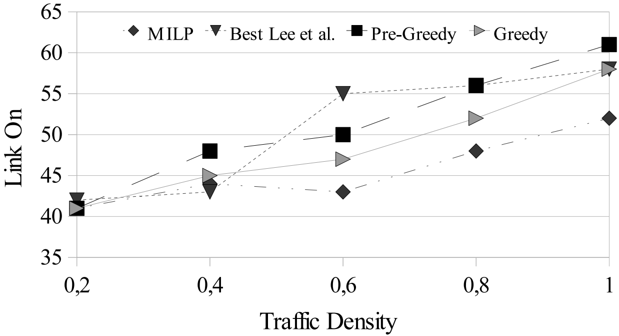
Computational results - Vasic et al. access networks

Inst	*N*	*L*	E_c^{tot} (W)	E_c (W)	N_{off}	L_{off}	t (s)	$Cong_{\%}$
Abov2	19	68	7610	5580 (73%)	0	29	73	219%
Abov6	19	68	7610	5200 (68%)	0	33	42	250%
Abov11	19	68	7610	4910 (64%)	1	35	59	228%
AT&T5	115	296	37970	23270 (61%)	35	135	419	162%
AT&T8	115	296	37970	23060 (60%)	35	138	356	217%
AT&T11	115	296	37970	23130 (61%)	35	137	662	240%
Gen3	42	110	14000	10810 (77%)	4	37	115	191%
Gen6	42	110	14000	10810 (77%)	4	37	251	180%
Gen11	42	110	14000	10670 (76%)	4	39	182	270%
Spr4	52	168	19560	14130 (72%)	4	69	968	175%
Spr8	52	168	19560	13780 (70%)	4	74	298	248%
Spr11	52	168	19560	13430 (68%)	4	79	528	287%
Tis3	41	174	18330	11590 (63%)	2	92	346	184%
Tis7	41	174	18330	11520 (63%)	2	93	270	183%
Tis11	41	174	18330	11660 (64%)	2	91	368	255%

Energy saving around 30% with moderate increase of congestion

Computing time around 15 minutes in the worst cases

Comparison with Lee et al. heuristic



USA28: access network with $n = 28$ and $m = 90$

For each of the five levels of traffic (from 0.2 to 1), average values on ten different randomly generated traffic matrices.

Concluding remarks

- A Greedy Randomized Search Procedure with path relinking provides slightly better solutions than Greedy but computing times are much larger
- The MILP-based heuristic allows to switch off a large number of nodes/links with a reasonable increase in the total cost of link utilization
- Validation on two networks of emulated Linux routers (using Netkit)

  A., Capone, Gianoli, Energy-aware IP traffic engineering with shortest path routing, to appear in Computer Networks

- Future work: extensions to account for traffic uncertainty and link failure

Outline Part II

Network routing subject to max-min fair flow allocation

- Max-min fair (MMF) resource allocation
- Previous work on network optimization with MMF flow allocation
- Network routing subject to MMF flow allocation
- MILP formulation
- Column generation algorithm
- Some computational results
- Concluding remarks

Max-min fair (MMF) resource allocation

Max-min fairness is widely used in a number of areas, including networks

Informally: a resource allocation is MMF if not only the smallest amount allocated is maximized but also the second smallest, the third smallest,...

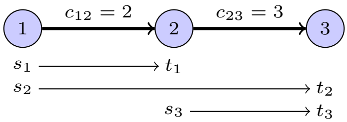
Max-min fair (MMF) resource allocation

Max-min fairness is widely used in a number of areas, including networks

Informally: a resource allocation is MMF if not only the smallest amount allocated is maximized but also the second smallest, the third smallest,...

Network flows: Given graph with arc capacities, m origin-destination pairs $\{(s_k, t_k)\}_{1 \leq k \leq m}$, and predefined routing paths, allocate the flow

e.g. Bertsekas & Gallager 92



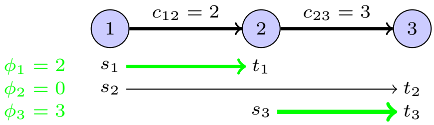
Max-min fair (MMF) resource allocation

Max-min fairness is widely used in a number of areas, including networks

Informally: a resource allocation is MMF if not only the smallest amount allocated is maximized but also the second smallest, the third smallest,...

Network flows: Given graph with arc capacities, m origin-destination pairs $\{(s_k, t_k)\}_{1 \leq k \leq m}$, and predefined routing paths, allocate the flow

e.g. Bertsekas & Gallager 92



Flow allocation $\underline{\phi} = (2, 0, 3)$ with max total throughput $\tau = 5$

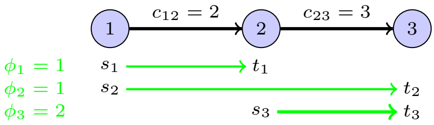
Max-min fair (MMF) resource allocation

Max-min fairness is widely used in a number of areas, including networks

Informally: a resource allocation is MMF if not only the smallest amount allocated is maximized but also the second smallest, the third smallest,...

Network flows: Given graph with arc capacities, m origin-destination pairs $\{(s_k, t_k)\}_{1 \leq k \leq m}$, and predefined routing paths, allocate the flow

e.g. Bertsekas & Gallager 92



Max-min fair flow allocation $\underline{\phi} = (1, 1, 2)$ with total throughput $\tau = 4$

MMF network flow allocation

Let σ be sorting operator permuting ϕ_k, $1 \leq k \leq m$, in nondecreasing order.

Definition: A flow vector $\underline{\phi}$ is **MMF** if,

for any other flow vector $\underline{\phi}'$, $\sigma(\underline{\phi})$ **lexicographically dominates** $\sigma(\underline{\phi}')$,

i.e., $\sigma(\underline{\phi}) = \sigma(\underline{\phi}')$ or $\exists k$ s.t. $\sigma(\underline{\phi})_k > \sigma(\underline{\phi}')_k$ and $\sigma(\underline{\phi})_l = \sigma(\underline{\phi}')_l$, $\forall l < k$.

Previous example: $\sigma(\underline{\phi}) = (1,1,2)$ lex. dominates $\sigma(\underline{\phi}') = (0,2,3)$

Other example: $\sigma(\underline{\phi}) = (1,2,3,4)$ lex. dominates $\sigma(\underline{\phi}') = (1,2,2,5)$

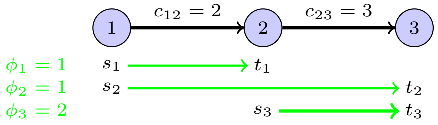
Equivalently: there is no way to increase the flow of any user without decreasing the flow of a user with a smaller or equal flow.

Previous work on MMF network optimization

- When routing paths are given, polynomial-time "progressive filling" algorithm yields a MMF flow (Bertsekas and Gallager 92)

Previous work on MMF network optimization

- When routing paths are given, polynomial-time "progressive filling" algorithm yields a MMF flow (Bertsekas and Gallager 92)



Max-min fair flow allocation $\underline{\phi} = (1, 1, 2)$

Start from $\underline{\phi} = \underline{0}$, simultaneously increase all the flows until one or more arcs are saturated.

Remove such *bottleneck arcs* and all saturating (s, t) pairs, compute the residual capacities, and keep on increasing the remaining flows.

Previous work on MMF network optimization

- When routing paths are given, polynomial-time "progressive filling" algorithm yields a MMF flow (Bertsekas and Gallager 92)
- When routing paths are not known a priori, algorithms to find splittable or unsplittable MMF flow (Nace 02, Tomaszewski 05, Pióro 07,...)
- Single-source variant where paths are selected to maximize fairness:

  - polynomial-time algorithm for fractional flow (Megiddo 74)

  - NP-hardness and approximation algorithms for single path routing (Kleinberg, Rabani, Tardos 01)

So far max-min fairness has been considered as a routing objective

Network routing subject to MMF allocation

Motivation: IP (Internet Protocol) networks

MMF is a reference model for fair capacity allocation when traffic flows are elastic and can adapt their rate based on resource availability.

Due to transport protocols (TCP), the distributed congestion control mechanism leads to an average bandwidth allocation which is well approximated by MMF over the routing paths provided by the IP layer.

In best-effort service the network is expected to provide the best possible service in terms of rate without privileging any traffic flow.

Network routing subject to MMF allocation

The **network operator** aims at optimizing routing based on a classical traffic engineering objective (e.g., throughput), while assuming a MMF bandwidth/flow allocation, which cannot be directly controlled.

Network routing subject to MMF allocation

The **network operator** aims at optimizing routing based on a classical traffic engineering objective (e.g., throughput), while assuming a MMF bandwidth/flow allocation, which cannot be directly controlled.

This network routing problem can be viewed as a **bilevel optimization problem**:

- at the upper level, the network operator selects the routing paths for the origin-destination pairs so as to maximize its utility function,

- at the lower level, the transport protocol allocates the bandwidth/flow to the selected paths according to the MMF paradigm.

Network routing subject to MMF allocation

MMF-constrained Single Path Routing (SPR) problem:

Given

- a directed graph $G = (V, A)$ with a capacity $c_{ij} \geq 0$ for each arc $(i, j) \in A$,
- a set of origin-destination pairs (s_k, t_k), with $s_k, t_k \in V$ and $1 \leq k \leq m$,

select one routing path for each (s_k, t_k) pair so as to

maximize a network utility function (total throughput),

while assuming an MMF flow allocation.

Network routing subject to MMF allocation

MMF-constrained Single Path Routing (SPR) problem:

Given

- a directed graph $G = (V, A)$ with a capacity $c_{ij} \geq 0$ for each arc $(i, j) \in A$,
- a set of origin-destination pairs (s_k, t_k), with $s_k, t_k \in V$ and $1 \leq k \leq m$,

select one routing path for each (s_k, t_k) pair so as to

maximize a network utility function (total throughput),

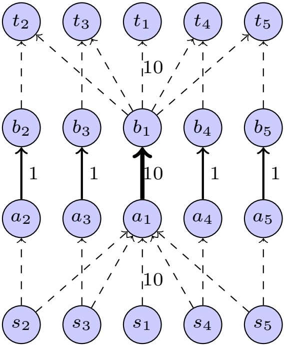
while assuming an MMF flow allocation.

Proposition: NP-hard even when $c_{ij} = 1$ for all $(i, j) \in A$

By reduction from the existence of edge-disjoint paths

Network routing subject to MMF allocation

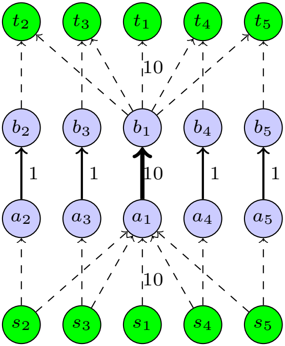
Proposition: The gap (throughput or smallest flow allocation) between optimal solutions of MMF-constrained and MMF-objective SPR can be arbitrarily large.



Capacity of dashed arcs ≥ 5

Network routing subject to MMF allocation

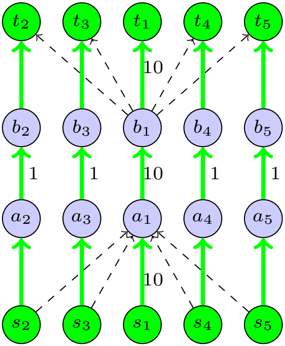
Proposition: The gap (throughput or smallest flow allocation) between optimal solutions of MMF-constrained and MMF-objective SPR can be arbitrarily large.



Capacity of dashed arcs ≥ 5

Network routing subject to MMF allocation

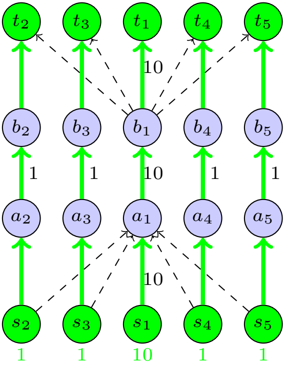
Proposition: The gap (throughput or smallest flow allocation) between optimal solutions of MMF-constrained and MMF-objective SPR can be arbitrarily large.



Capacity of dashed arcs ≥ 5

Network routing subject to MMF allocation

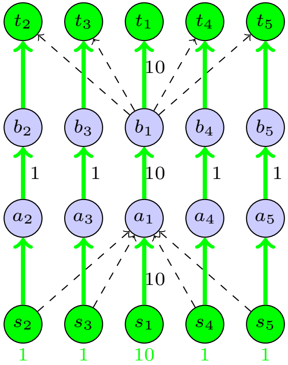
Proposition: The gap (throughput or smallest flow allocation) between optimal solutions of MMF-constrained and MMF-objective SPR can be arbitrarily large.



Capacity of dashed arcs ≥ 5

Network routing subject to MMF allocation

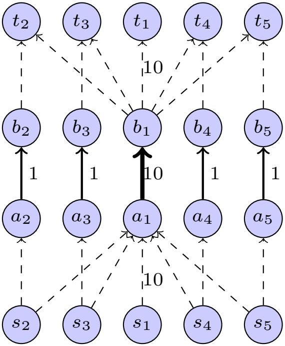
Proposition: The gap (throughput or smallest flow allocation) between optimal solutions of MMF-constrained and MMF-objective SPR can be arbitrarily large.



MMF flow $\sigma(\underline{\phi}) = (1, 1, 1, 1, 10)$
with total throughput $\tau = 14$

Network routing subject to MMF allocation

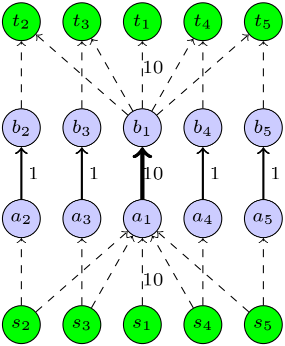
Proposition: The gap (throughput or smallest flow allocation) between optimal solutions of MMF-constrained and MMF-objective SPR can be arbitrarily large.



Capacity of dashed arcs ≥ 5

Network routing subject to MMF allocation

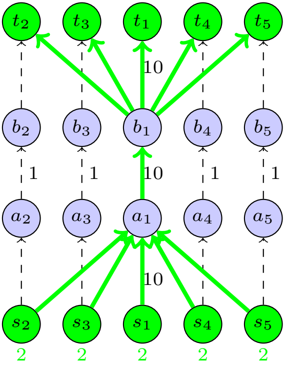
Proposition: The gap (throughput or smallest flow allocation) between optimal solutions of MMF-constrained and MMF-objective SPR can be arbitrarily large.



Capacity of dashed arcs ≥ 5

Network routing subject to MMF allocation

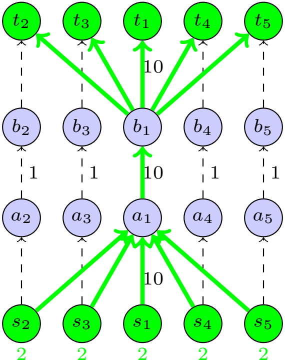
Proposition: The gap (throughput or smallest flow allocation) between optimal solutions of MMF-constrained and MMF-objective SPR can be arbitrarily large.



Capacity of dashed arcs ≥ 5

Network routing subject to MMF allocation

Proposition: The gap (throughput or smallest flow allocation) between optimal solutions of MMF-constrained and MMF-objective SPR can be arbitrarily large.



MMF flow $\sigma(\underline{\phi}) = (2, 2, 2, 2, 2)$
with total throughput $\tau = 10$

Network routing subject to MMF allocation

Proposition: The gap (throughput or smallest flow allocation) between optimal solutions of MMF-constrained and MMF-objective SPR can be arbitrarily large.

$\sigma(\underline{\phi}) = (1, 1, 1, 1, 10)$ is optimal for MMF-constrained SPR

$\sigma(\underline{\phi}) = (2, 2, 2, 2, 2)$ is optimal for MMF-objective SPR

Network routing subject to MMF allocation

Proposition: The gap (throughput or smallest flow allocation) between optimal solutions of MMF-constrained and MMF-objective SPR can be arbitrarily large.

$\sigma(\underline{\phi}) = (1, 1, 1, 1, 10)$ is optimal for MMF-constrained SPR

$\sigma(\underline{\phi}) = (2, 2, 2, 2, 2)$ is optimal for MMF-objective SPR

<u>Extension</u>: $m \gg 5$, $c_{a_1 b_1} = m(1+\delta)$ and capacity $c_{a_i b_i} = 1$ for $2 \leq i \leq m$

$\sigma(\underline{\phi}) = (1, \ldots, 1, m(1+\delta))$ with $\tau = m(1+\delta) + m - 1$

$\sigma(\underline{\phi}) = (1+\delta, \ldots, 1+\delta)$ with $\tau = m(1+\delta)$

Difference in terms of **throughput** $= m - 1$

For every (s_k, t_k) pair, with $2 \leq k \leq m$, **difference** in terms of **flow** $= \delta$

Single-level MILP formulation

Flow variables:

ϕ^{st}: the flow assigned to the (s,t) pair

f_{ij}^{st}: the amount of flow of the (s,t) pair through arc (i,j)

Binary variables: x_{ij}^{st} is 1 if $f_{ij}^{st} > 0$, and 0 otherwise

$$\sum_{(i,j)\in\delta^+(i)} f_{ij}^{st} - \sum_{(j,i)\in\delta^-(i)} f_{ji}^{st} = \begin{cases} \phi^{st} & \text{if } i = s \\ -\phi_{st} & \text{if } i = t \\ 0 & \text{else} \end{cases} \quad \forall i \in V, \forall (s,t) \in K \qquad (1)$$

$$\sum_{(s,t)\in K} f_{ij}^{st} \leq c_{ij} \quad \forall (i,j) \in A \qquad (2)$$

$$f_{ij}^{st} \leq c_{ij} x_{ij}^{st} \quad \forall (i,j) \in A, \forall (s,t) \in K \qquad (3)$$

$$\sum_{(i,j)\in\delta^+(i)} x_{ij}^{st} \leq 1 \quad \forall i \in V, \forall (s,t) \in K \qquad (4)$$

MMF flow-allocation constraints

If routing paths $\{\pi_{st}\}_{(s,t)\in K}$ are known, the progressive filling algorithm implies:

Property: For each (s,t) pair, $\exists$ a **bottleneck arc** (i,j) on path π_{st} such that ϕ^{st} is at least as large as the flow of all other (s,t) pairs using (i,j).

MMF flow-allocation constraints

If routing paths $\{\pi_{st}\}_{(s,t)\in K}$ are known, the progressive filling algorithm implies:

Property: For each (s,t) pair, $\exists$ a **bottleneck arc** (i,j) on path π_{st} such that ϕ^{st} is at least as large as the flow of all other (s,t) pairs using (i,j).

For all $(s,t) \in K$ and $(i,j) \in A$, variable:

y_{ij}^{st} is 1 if (i,j) is a bottleneck arc for the (s,t) pair, and 0 otherwise.

A flow vector $\underline{\phi}$ is MMF iff the following holds:

$$\sum_{(i,j)\in A} y_{ij}^{st} \geq 1 \qquad \forall (s,t) \in K \tag{5}$$

$$\sum_{(o,d)\in K} f_{ij}^{od} \geq c_{ij} y_{ij}^{st} \qquad \forall (i,j) \in A, \forall (s,t) \in K \tag{6}$$

$$f_{ij}^{st} \geq f_{ij}^{od} - c_{ij}(1 - y_{ij}^{st}) \qquad \forall (i,j) \in A, \forall (s,t) \in K, \ \forall (o,d) \neq (i,j) \tag{7}$$

Extended subtour elimination constraints

Subtours may arise:

Consider an (s, t) pair with a routing path. Adding a disjoint subtour defined by $S \subset A$ does not change the value of ϕ^{st}.

Suppose arc $(v, w) \in S$ s.t. $c_{vw} = \min_{(i,j) \in S}\{c_{ij}\}$.

Extended subtour elimination constraints

Subtours may arise:

Consider an (s,t) pair with a routing path. Adding a disjoint subtour defined by $S \subset A$ does not change the value of ϕ^{st}.

Suppose arc $(v,w) \in S$ s.t. $c_{vw} = \min_{(i,j)\in S}\{c_{ij}\}$.

Letting
$$y^{st}_{vw} = 1,\ f^{st}_{ij} = c_{vw}\ \forall (i,j) \in S,$$
$$f^{od}_{ij} = 0\ \forall (i,j) \in S,\ \forall (o,d) \neq (s,t),$$

the MMF constraints are trivially satisfied on the subtour, allowing for a flow on the $s-t$ path that is not MMF.

We add subtour elimination constraints – a variant of Wong's compact (extended) formulation for TSP.

Other valid inequalities

Besides the trivial inequalities

$$y_{ij}^{st} \leq x_{ij}^{st} \quad \forall (i,j) \in A, \forall (s,t) \in K, \tag{8}$$

we add

$$\phi^{st} \geq \frac{\min_{(i,j)\in A}\{c_{ij}\}}{|K|} \quad \forall (s,t) \in K, \tag{9}$$

valid since any MMF flow saturates for each (s,t) pair at least one arc, and the smallest flow value is allocated when an arc is shared by all pairs.

For compactness, the last MMF-constraint is replaced with:

$$u_{ij} \geq f_{ij}^{st} \quad \forall (i,j) \in A, (s,t) \in K \tag{10}$$

$$f_{ij}^{st} \geq u_{ij} - c_{ij}(1 - y_{ij}^{st}) \quad \forall (i,j) \in A, (s,t) \in K \tag{11}$$

involving a new variable $u_{ij} \geq 0$ for each $(i,j) \in A$.

Column generation algorithm

Path formulation:

P^{st} is the set of feasible paths for each commodity $(s,t) \in K$

σ_{ij}^{pst} is 1 if the path $p \in P^{st}$ contains the arc $(i,j) \in A$, and 0 otherwise.

Variables:

- λ_{st}^{p} is 1 if the path p is chosen for the commodity (s,t), and 0 otherwise
- For each commodity $(s,t) \in K$,
  - ϕ^{st} is the amount of flow assigned to it
  - f_{ij}^{st} is its amount on arc $(i,j) \in A$
- u_{ij} is an upper bound on any flow over the arc $(i,j) \in A$
- y_{ij}^{st} is 1 if (i,j) is a bottleneck arc for the (s,t) pair, and 0 otherwise.

$$\max \sum_{(s,t)\in K} \phi^{st} \tag{12}$$

$$\text{s.t.} \sum_{p\in P_{st}} \lambda^p_{st} = 1 \qquad \forall (s,t)\in K \tag{13}$$

$$\sum_{(i,j)\in\delta^+(i)} f^{st}_{ij} - \sum_{(j,i)\in\delta^-(i)} f^{st}_{ji} = \begin{cases} \phi^{st} & \text{if } i = s \\ -\phi^{st} & \text{if } i = t \\ 0 & \text{else} \end{cases} \qquad \forall (s,t)\in K, i\in V \tag{14}$$

$$\sum_{(s,t)\in K} f^{st}_{ij} \leq c_{ij} \qquad \forall (i,j)\in A \tag{15}$$

$$f^{st}_{ij} - c_{ij} \sum_{p\in P^{st}} (\sigma^{pst}_{ij}\lambda^p_{st}) \leq 0 \qquad \forall (i,j)\in A, (s,t)\in K \tag{16}$$

$$\phi^{st} \geq \frac{\min_{(i,j)\in A}\{c_{ij}\}}{|K|} \qquad \forall (s,t)\in K \tag{17}$$

$$\sum_{(i,j)\in A} y^{st}_{ij} \geq 1 \qquad \forall (s,t)\in K \tag{18}$$

$$\sum_{(o,d)\in K} f^{od}_{ij} \geq c_{ij} y^{st}_{ij} \qquad \forall (i,j)\in A, (s,t)\in K \tag{19}$$

$$u_{ij} \geq f^{st}_{ij} \qquad \forall (i,j)\in A, (s,t)\in K \tag{20}$$

$$u_{ij} - f^{st}_{ij} \leq c_{ij}(1 - y^{st}_{ij}) \qquad \forall (i,j)\in A, (s,t)\in K \tag{21}$$

$$\phi^{st} \geq 0, f^{st}_{ij} \geq 0, y^{st}_{ij} \in \{0,1\}, \lambda^p_{st} \in \{0,1\}, u_{ij} \geq 0 \qquad \forall (i,j)\in A, (s,t)\in K. \tag{22}$$

Column generation algorithm

Pricing subproblem for each $(s,t) \in K$:

Given

- w^{st} dual variables associated to $\sum_{p \in P_{st}} \lambda_{st}^p = 1$

- π_{ij}^{st} dual variables associated to $f_{ij}^{st} \leq c_{ij} \sum_{p \in P_{st}} (\sigma_{ij}^{pst} \lambda_{st}^p)$,

find a **longest simple path** from s to t in graph G where the length of each arc (i,j) is equal to $\pi_{ij}^{st} c_{ij}$.

A column is added depending on the sign of $\tilde{c}^* - w^{st}$, where $\tilde{c}^*$ is the maximum path length.

Since very sparse subgraphs, solved as MILP.

Some computational results

Instances:

- Topologies from SND library:
  `nobel-us` ($|V| = 14, |A| = 42$), `polska` ($|V| = 12, |A| = 36$)
  `atlanta` ($|V| = 15, |A| = 44$), `france` ($|V| = 25, |A| = 90$)
- also randomly generated (s, t) pairs
- capacities are uniform or radomly generated (10 Gbs (40%), 5 Gbs (30%), 2 Gbs (20%), 1 Gbs (10%))

Column generation:

- Initial paths are randomly generated with penalities (3 per (s, t) pairs)
- Pricing subproblem solved as MILP
- 1 hour time limit

Intel 2.4 GHz with 16 GB RAM, AMPL + CPLEX 12.3

Network	Inst. Code	$\|K\|$	Complete formulation LB	Complete formulation Int. Gap	Complete formulation Time	Column generation UB	Column generation LB	Gap	Time CG	Time MILP	# Gen. Cols.
atlanta	1-1	12	48.5	0.0	1.4	50.8	48.5	4.8	2.0	0.2	69
atlanta	1-2	20	62.5	0.0	16.2	62.5	59.5	5.0	2.3	27.4	68
atlanta	1-3	30	96.1	2.1	3593.9	98.6	96.5	2.1	6.9	452.1	71
atlanta	1-4	42	-	-	3594.1	76.0	-	-	3.3	3594.0	112
atlanta	1-5	56	-	-	3594.1	87.3	82.1	6.4	3.8	3594.0	96
france	1-2	10	52.5	0.0	58.3	55.0	52.5	4.8	2.2	0.4	100
france	1-3	15	56.5	0.0	482.6	60.1	56.5	6.3	6.2	1.2	199
france	1-4	21	76.5	0.9	3600.0	78.1	76.5	2.0	12.2	94.3	317
france	1-5	28	-	-	3600.0	114.8	113.5	1.1	6.7	33.4	192
france	1-6	36	-	-	3600.0	116.3	116.0	0.2	28.5	184.9	403
nobel-us	1-3	15	63.5	0.0	62.1	63.8	63.5	0.5	2.0	1.5	67
nobel-us	1-4	21	75.0	0.0	2627.0	76.0	75.0	1.3	4.9	47.6	126
nobel-us	1-5	28	104.5	2.4	3600.0	108.0	105.3	2.6	6.5	407.8	220
nobel-us	1-6	36	-	-	3600.0	91.0	90.3	0.8	4.2	2280.1	132
nobel-us	1-7	42	-	-	3600.0	103.9	103.5	0.5	4.5	2414.8	203
polska	1-3	21	90.0	0.0	20.0	90.8	86.7	4.8	0.4	5.3	22
polska	1-4	28	71.4	0.8	3594.1	71.9	68.7	4.6	1.3	5.3	27
polska	1-5	36	81.8	0.8	3593.9	82.4	80.5	2.4	4.2	680.8	69
polska	1-6	42	-	-	3600.0	118.3	111.4	6.2	2.8	2676.4	69
polska	1-7	56	-	-	3600.0	151.6	143.6	5.6	2.3	3594.0	100

Complete formulation: 28 instances out of 80 no feasible solution in 1 hour ($n = 22$ and $m = 72$ gap $10 - 70\%$), CG heuristic only 4

Good quality UB ($2,75\%$ gap w.r.t. to best known solution) found very fast (much smaller than 10 sec. and never exceeds 30 sec.)

Time total = Time CG + MILP

Concluding remarks

- New class of network routing problem where the paths are selected to optimize a network utilization measure subject to MMF flow allocation
- The column generation approach yields promising results
- Ongoing work: energy-aware traffic engineering variant
- Future work: Branch and Price algorithm for medium-size instances and heuristic for larger ones

A., Capone, Coniglio, Gianoli, Network optimization problems subject to max-min fair flow allocation, to appear in IEEE Communications Letters

A., Coniglio, Gianoli, Ileri, On single-path network routing subject to max-min fair flow allocation, to appear in Electronic Notes in Discrete Mathematics