Uncertainty-Aware Optimization for Network Provisioning and Routing

Yingjie Bi

School of Electrical and Computer Engineering
Cornell University

Joint work with Kevin Tang

Mar 21, 2019

Outline

Introduction

Two main goals of traffic engineering:

- Accommodate the demands from customers.
- Minimize the network cost.

Previous approaches to handle traffic fluctuations in traffic engineering:

- Oblivious routing.
- Dynamic routing.

Our Approach

- Combines the ideas from both oblivious routing and dynamic routing.
- Utilizes the different flexibility of network reconfiguration and characteristics of traffic demand in different timescales.
- Applies to wide-area backbone network.

Our Approach

Long-term timescale:

- Typically within a few months.
- Operators can freely reconfigure the network.
- Very limited information on the demand side.

Short-term timescale:

- Approximately from an hour to several minutes.
- Operators cannot change the network topology but can adjust the routing.
- The traffic demand is more predictable.

Transient timescale:

- Seconds or below.
- Operators cannot perform any change to the network.
- Demand information is most complete.

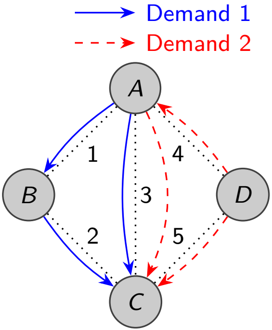
Outline

Notation

- Number of links $L = 5$. Number of demands $N = 2$. Number of available paths for each demand $K^1 = K^2 = 2$.
- The paths of demand i are represented by

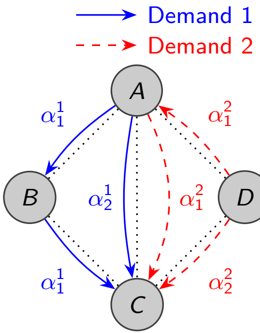
$$R^1 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad R^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Put $R^i_{lk} = 1$ if path k of demand i passes through link l.



Notation

- Link capacity $c = (c_1, \ldots, c_5)^T$.
- The kth entry α^i_k of vector α^i is the split ratio of demand i on path k.



Traffic Model

We adopt the following traffic model (Tune and Roughan, 2013):
The traffic volume of demand i

$$D^i(t) = L^i(t)S^i(t) + \sqrt{\beta^i L^i(t)S^i(t)}W^i(t),$$

where

- $L^i(t)$ is the long-term trend.
- $S^i(t)$ is the seasonal component.
- $W^i(t)$ is the random fluctuations.
- β^i is a constant called the peakedness of the traffic.

Traffic Model

Further assumptions on the traffic model:

- Approximate the seasonal component $S^i(t)$ by Q scenarios, i.e.,
$$S^i(t) = S^i_{q(t)},$$
where $q(t)$ is a periodic and piecewise constant function whose value is an integer between 1 and Q.
- $W^i(t)$ is assumed to be a white Gaussian process with zero mean and unit variance.

Summary

Characteristics of the three timescales in network optimization:

Timescale	Typical Granularity	Known Information	Decision Variables
Long-term	A month	$L^i(t)$	c_l, α_k^i
Short-term	An hour	$L^i(t)$, $S^i(t)$	α_k^i
Transient	A second	$L^i(t)$, $S^i(t)$, $W^i(t)$	None

Outline

Short-Term Optimization

- The actual traffic Y_l on link l

$$Y_l = \sum_{i=1}^{N} \sum_{k=1}^{K^i} R^i_{lk} \alpha^i_k D^i.$$

- The goal is to minimize the probability of link capacity violations, i.e.,

$$\begin{aligned}
\min \quad & \Pr\left(\bigcup_{l=1}^{L} \{ Y_l > c_l \} \right) \\
\text{s. t.} \quad & \sum_{k=1}^{K^i} \alpha^i_k = 1, \quad \forall i = 1, \ldots, N, \\
& \alpha^i_k \geq 0, \quad \forall i = 1, \ldots, N,\ k = 1, \ldots, K^i.
\end{aligned}$$

Long-Term Optimization

- Assume p_l is the cost of link l per unit capacity.
- The goal is to minimize the aggregate cost while the violation probability is not exceeding ϵ, i.e.,

$$\begin{aligned}
\min \quad & \sum_{l=1}^{L} p_l c_l \\
\text{s.t.} \quad & \Pr\left(\bigcup_{l=1}^{L} \{ Y_{l,q} > c_l \} \right) \leq \epsilon, \quad \forall q = 1, \ldots, Q, \\
& \sum_{k=1}^{K^i} \alpha_{k,q}^i = 1, \quad \forall i = 1, \ldots, N,\ q = 1, \ldots, Q, \\
& \alpha_{k,q}^i \geq 0, \quad \forall i = 1, \ldots, N,\ k = 1, \ldots, K^i, \\
& \qquad q = 1, \ldots, Q,
\end{aligned}$$

where $Y_{l,q}$ is the actual traffic on link l at scenario q.

Approximate Solutions

- Both the short-term and long-term optimization problems are difficult to solve.
- For the long-term problem, constraint the violation probability of each link separately by setting

$$\Pr(Y_{l,q} > c_l) \le \epsilon/L, \quad l = 1, \ldots, L.$$

  instead of the original probability constraints.
- $Y_{l,q}$ is subject to Gaussian distribution whose mean $y_{l,q}$ and variance $b_{l,q}$ can be obtained from the split ratio $\alpha^i_{k,q}$ and the traffic model.

Approximate Solutions

- Let $f(x)$ be the survival function of standard normal distribution, i.e.,
$$f(x) = \Pr(X > x),$$
where X is subject to the standard normal distribution.
- Define
$$A = \frac{1}{f^{-1}(\epsilon/L)}.$$
The above constraint
$$\Pr(Y_{l,q} > c_l) \le \epsilon/L$$
is equivalent to
$$\frac{c_l - y_{l,q}}{\sqrt{b_{l,q}}} \ge \frac{1}{A}.$$
- This is a second-order cone constraint which can be handled efficiently by convex optimization techniques.

Performance Guarantees

Let

- sol be the optimal value for the approximate problem and
- opt be the actual optimal cost.

Then the approximation ratio

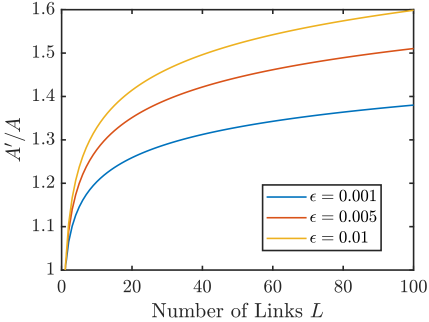
$$\text{sol} / \text{opt} \leq A'/A = O(\sqrt{\log L / \log(1/\epsilon)}),$$

where

$$A = \frac{1}{f^{-1}(\epsilon/L)}, \quad A' = \frac{1}{f^{-1}(\epsilon)}.$$

Performance Guarantees

The actual dependence of the approximation ratio sol / opt on the number of links L and the violation probability ϵ:



Thank You!