A Subgradient Optimization Approach to Inter-Domain Routing in IP/MPLS Networks

A. Tomaszewski[1], M. Pióro[1,2], M. Dzida[1], M. Mycek[1], and M. Zagożdżon[1]

[1] Institute of Telecommunications, Warsaw University of Technology, Poland
[2] Department of Communication Systems, Lund University, Sweden

Abstract. We present a mathematical model for a distributed process of routing optimization that could be run in the control plane of the Internet using existing EGP routing protocols. A more detailed description of the results presented in this paper is given in [1].

1 Introduction

Each domain of the Internet is usually operated as an autonomous system (AS), which, for security, competitiveness, administrative and technical reasons, does not reveal its sensitive internal information concerning topology, link capacities, traffic volumes, and routing parameters. Only a very limited amount of data is made available to other AS's by means of exterior gateway protocols (EGP) such as BGP (see [2] and the discussion there): border routers of neighboring AS's exchange information concerning the capability of reaching a destination, or a group of destinations, identified by a network address.

It is virtually impossible for a single domain to decide on its own on the optimal flow of inter-domain traffic through its routers/gateways using only the information about accessibility of given destinations via particular gateways, and having no knowledge of the overall network topology, routing resources, and end-to-end traffic demand. Even if such information were made available to an AS, its local decisions would still have to be consistent with those in other domains.

It seems that to reach a globally "optimal" traffic routing, it would be required to implement a distributed, network-wide process of routing optimization which should be run in the control plane of the network using the existing EGP protocols. For this, an optimization model is needed that is decomposable into subproblems that could be solved locally in individual domains. Moreover, the results from solving the subproblems exchanged by the domains must be limited to only those pieces of information that directly concern the inter-domain traffic and inter-domain links, and can be distributed using EGP protocols.

Below we present a decomposable mathematical model for the problem of optimizing bandwidth reservation levels on inter-domain links, satisfying the above-mentioned requirements. The decomposition method of Lagrangean relaxation leads to a master problem that uses limited information delivered from individual domains: it concerns only inter-domain traffic and inter-domain links, and is based on detailed intra-domain routing solutions computed locally at AS's. We show how to solve the master problem using a subgradient minimization method and how to recover optimal bandwidth reservation levels.

2   Primal and dual formulation

A network is represented by a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with the set of nodes $\mathcal{V}$ and the set of directed links $\mathcal{E}$ ($\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$). The capacity of link $e \in \mathcal{E}$ is denoted by c_e. Then, for $\mathcal{U} \subseteq \mathcal{V}$ we define the set of links $\delta^+(\mathcal{U})$ outgoing from set $\mathcal{U}$, and the set of links $\delta^-(\mathcal{U})$ incoming to set $\mathcal{U}$ (we write $\delta^\pm(v)$ instead of $\delta^\pm(\{v\})$ when $\mathcal{U} = \{v\}$ is a singleton). $\mathcal{M}$ is the set of network domains. For each domain $m \in \mathcal{M}$, $\mathcal{E}^m$ is the set of links between the nodes in the same domain m (intra-domain links). Further, $\mathcal{E}_\mathcal{O}$ is the set of all inter-domain links, and $\mathcal{E}_\mathcal{I} = \mathcal{E} \setminus \mathcal{E}_\mathcal{O}$ is the set of all intra-domain links. Set $\mathcal{D}$ represents traffic demands between pairs of nodes. The starting and terminating node of demand $d \in \mathcal{D}$ is denoted by $s(d)$ and $t(d)$, respectively, and h_d is the traffic volume of demand $d \in \mathcal{D}$. Also, $\mathcal{D}(s,t)$ denotes the set of all demands from node s to node t. In the sequel, z_d will denote the variable specifying the percentage of volume h_d actually handled in the network, i.e., $z_d h_d$ is the carried traffic of demand d. The set of all demands originating in domain $m \in \mathcal{M}$ is denoted as $\mathcal{D}^m$.

The basic quantity in the considerations of this paper is x_{et}—a variable specifying the amount of aggregated bandwidth (called *flow* in the sequel) reserved on link $e \in \mathcal{E}$ for the traffic destined for (a remote) node $t \in \mathcal{V}$. For each inter-domain link $e \in \mathcal{E}_\mathcal{O}$ we also introduce variables x_{et}^+ and x_{et}^- expressing, respectively, the amount of traffic carried on e and destined for t assumed by domain in which link e originates, and by domain in which link e terminates. We write down one obvious constraint stating that $x_{et}^+ = x_{et}^- = x_{et}, e \in \mathcal{E}_\mathcal{O}, t \in \mathcal{V}$, the flow conservation constraint, and the capacity constraint:

$$\sum_{e \in \delta^+(v) \cap \mathcal{E}_\mathcal{I}} x_{et} + \sum_{e \in \delta^+(v) \cap \mathcal{E}_\mathcal{O}} x_{et}^+ - \sum_{e \in \delta^-(v) \cap \mathcal{E}_\mathcal{I}} x_{et} - \sum_{e \in \delta^-(v) \cap \mathcal{E}_\mathcal{O}} x_{et}^-$$
$$= \sum_{d \in \mathcal{D}(v,t)} z_d h_d, \quad t, v \in \mathcal{V}, v \neq t \tag{1a}$$

$$\sum_{t \in \mathcal{V}} x_{et} \leq c_e, \quad e \in \mathcal{E}. \tag{1b}$$

For each domain $m \in \mathcal{M}$ we introduce flow vectors $\boldsymbol{z}^m = (z_d : d \in \mathcal{D}^m)$, $\boldsymbol{x}^m = (x_{et} : e \in \mathcal{E}^m, t \in \mathcal{V})$, $\boldsymbol{x}^{m+} = (x_{et}^+ : e \in \delta^+(\mathcal{V}^m), t \in \mathcal{V})$, $\boldsymbol{x}^{m-} = (x_{et}^- : e \in \delta^-(\mathcal{V}^m), t \in \mathcal{V})$, and $\boldsymbol{X}^m = (\boldsymbol{x}^m, \boldsymbol{x}^{m+}, \boldsymbol{x}^{m-})$. The set of all pairs of non-negative vectors $(\boldsymbol{z}^m, \boldsymbol{X}^m)$ satisfying constraints (1) will be denoted by $\boldsymbol{\mathcal{Y}}^m$. The routing optimization problem can now be stated as follows.

$$\max \quad F(\boldsymbol{z}) = \sum_{m \in \mathcal{M}} \sum_{d \in \mathcal{D}^m} z_d h_d \tag{2a}$$

$$\text{s.t.} \quad \boldsymbol{Y}^m = (\boldsymbol{z}^m, \boldsymbol{X}^m) \in \boldsymbol{\mathcal{Y}}^m, \qquad m \in \mathcal{M} \tag{2b}$$

$$x_{et}^- = x_{et}^+, \qquad\qquad\quad e \in \mathcal{E}_\mathcal{O}, \ t \in \mathcal{V}. \tag{2c}$$

Now, let us consider the problem dual to (2), obtained by Lagrangean relaxation of constraints (2c). Let $\boldsymbol{\lambda} = (\lambda_{et} : e \in \mathcal{E}_\mathcal{O}, t \in \mathcal{V})$ be a vector of multipliers associated with constraints (2c). Assume that $\mathcal{M} = \{1, 2, \ldots, M\}$ and define

$Y = (Y^1, Y^2, \ldots, Y^M)$ and $\mathcal{Y} = \mathcal{Y}^1 \times \mathcal{Y}^2 \times \ldots \times \mathcal{Y}^M$. As shown in [1], the dual can be written as follows:

$$\min_{\boldsymbol{\lambda}} w(\boldsymbol{\lambda}). \tag{3}$$

where $w(\boldsymbol{\lambda}) = \sum_{m \in \mathcal{M}} w^m(\boldsymbol{\lambda})$ is the dual function defined through subproblems $w^m(\boldsymbol{\lambda}) = \max\{L^m(\boldsymbol{\lambda}; Y^m) : Y^m \in \mathcal{Y}^m\}$. Each $L^m(\boldsymbol{\lambda}; Y^m) = \sum_{d \in \mathcal{D}^m} z_d h_d + \sum_{t \in \mathcal{V}}(\sum_{e \in \delta^-(\mathcal{V}^m)} \lambda_{et} x_{et}^- - \sum_{e \in \delta^+(\mathcal{V}^m)} \lambda_{et} x_{et}^+)$ is a component of the decomposed Lagrangean function $L(\boldsymbol{\lambda}; Y) = \sum_{m \in \mathcal{M}} L^m(\boldsymbol{\lambda}; Y^m)$. Thus, for any given $\boldsymbol{\lambda}$, computation of $w(\boldsymbol{\lambda})$ reduces to solving a set of M mutually independent subproblems where for each domain $m \in \mathcal{M}$ the objective is to find a maximizer $\hat{Y}^m$ of $L^m(\boldsymbol{\lambda}; Y^m)$.

3   Solving the dual and recovering a primal solution

An efficient decomposed approach to the primal problem (2) is applying a subgradient-based procedure resolving the dual problem (3) in variables $\boldsymbol{\lambda}$, and recovering an optimal primal solution in the process (see [4], [5]). An example of such an algorithm (referred to as SPR in the sequel) is as follows.

Step 0: Set $\boldsymbol{\lambda}^0$, γ, and η. Find a maximizer $\hat{Y}_0$ of $L(\boldsymbol{\lambda}^0; Y)$; $Y^* := \hat{Y}_0$. $j := 1$.

Step 1: $\boldsymbol{\lambda}^j := (\boldsymbol{\lambda}^{j-1} - \gamma \nabla w(\boldsymbol{\lambda}^{j-1}))$.

Step 2: Find a maximizer $\hat{Y}_j$ of $L(\boldsymbol{\lambda}^j; Y)$ and modify the current estimate of the optimal primal solution: $Y^* := \eta \hat{Y}_j + (1 - \eta)Y^*$.

Step 3: If stopping criterion satisfied then stop. Else, modify γ and η, put $j := j + 1$ and go to Step 1.

Above, $\nabla w(\hat{\boldsymbol{\lambda}})$ is a subgradient of the dual function at point $\hat{\boldsymbol{\lambda}}$, $\gamma > 0$ is a step-size, and $0 < \eta < 1$ is a weight parameter. We note that subgradients are obtained as a by-product of the distributed computation of the value of $w(\hat{\boldsymbol{\lambda}})$: if $\hat{Y}$ is a maximizer of the Lagrange function $L(\hat{\boldsymbol{\lambda}}; Y)$ for fixed $\hat{\boldsymbol{\lambda}}$, then the corresponding subgradient is $\nabla w(\hat{\boldsymbol{\lambda}}) = (\hat{x}_{et}^- - \hat{x}_{et}^+ : e \in \mathcal{E}_{\mathcal{O}}, t \in \mathcal{V})$ [3]. A stopping criterion is for example no improvement in the value of the dual function $w(\boldsymbol{\lambda}^j)$.

In general, the primal solution Y^* produced by SPR is approximate. This is, however, not a big issue since the input data available at the domains to solve the subproblems is intrinsically not accurate. Consequently, a primal solution computed with accuracy of, say, 5% will be sufficient for the purpose of multi-domain routing optimization.

SPR can be distributed among the domains so that each domain $m \in \mathcal{M}$ is responsible for adjusting the values of $\boldsymbol{\lambda}^{m+} = (\lambda_{et} : e \in \delta^+(\mathcal{V}^m), t \in \mathcal{V})$ for all its outgoing inter-domain links using the partial subgradient $\nabla^m w(\boldsymbol{\lambda}^{m+}) = (x_{et}^- - x_{et}^+ : e \in \delta^+(\mathcal{V}^m), t \in \mathcal{V})$. Observe that for computing $\nabla^m w(\boldsymbol{\lambda}^{m+})$, the domain m has to collect values $x_{et}^-, e \in \delta^+(\mathcal{V}^m), t \in \mathcal{V}$, from its neighboring domains. To make the process consistent, it must be ensured that the same parameters γ and η for adjusting individual components of $\boldsymbol{\lambda}$ and Y^*, for example $\gamma = a/(b + cj)$ and $\eta = 1/j$ (see [5]), are used simultaneously in all the domains.

The information required from each domain $m \in \mathcal{M}$ during the process of resolving the dual problem (3) (called the master problem) is limited to vectors

$\boldsymbol{x}^{m-}$ and $\boldsymbol{x}^{m+}$, and scalar $\sum_{d\in\mathcal{D}^m} z_d h_d$. They concern basically only the inter-domain traffic and do not reveal any confidential data related to a domain.

In fact with a simple SPR (as the one described above), a good approximation of optimal $\boldsymbol{Y}^*$ is hard to reach in a reasonable number of steps. This, however, can be achieved with a more sophisticated approach called conic-bundle method available as freeware in [6].

4   Final remarks

We tested the convergence of SPR on several randomly generated example multi-domain networks with up to 40 nodes, 184 directed links and 1560 directed demands. We used two different SPR algorithms: an advanced deflection algorithm CB taken from [6], and a plain algorithm SGR of the type described in Section 3. Using SGR it was not possible to obtain solutions with 1% accuracy for most of the example networks. It was also not possible to find a 5% quality solution using SGR for the biggest network. Using CB all problems were solved, even for 1% solution accuracy. Still, even though the CB method is more effective than SBR, its application in the distributed environment of the Internet may by difficult. Contrary to CB, implementation of SBR is simple; we hope that a tuned variant of SDR can assure a good tradeoff between implementation feasibility and convergence in a multi-domain environment.

A practical implementation of the presented approach would result in a continuous network-wide process distributed over the domains, responsible for establishing consistent, near-optimal bandwidth reservation levels. A major problem is to provide scalable synchronization of the sub-processes running in individual domains, so that the overall process would converge even in large networks. Specifying such a network-wide distributed optimization process will be a subject of future work.

References

1. A. Tomaszewski, M. Pióro et al., *Towards Distributed Inter-Domain Routing Optimization for IP/MPLS Networks*, Technical Report, Warsaw University of Technology, 2006, http://ztit.tele.pw.edu.pl/TR/NDG/rmd07.pdf.
2. N. Feamster, J. Borkenhagen and J. Rexford, Guidelines for Interdomain Traffic Engineering, *ACM SIGCOM Computer Communications Review*, vol.33, no. 5, pp. 19–30, October 2003.
3. M. Minoux, *Mathematical Programming: Theory and Algorithms*, Wiley, 1986.
4. N.Z. Shor, *Minimization Methods for Nondiff. Functions*, Springer-Verlag, 1985.
5. H.D. Sherali and G. Choi, Recovery of Primal Solutions when Using Subgradient Optimization Methods to Solve Lagrangean Duals of Linear Programs, *Operations Research Letters*, 19(1996), pp. 105–113, 1996.
6. C. Helmberg, *ConicBundle 0.1*, Fakultät für Mathematik, Techn. Univ. Chemnitz, 2005, http://www-user.tu-chemnitz.de/∼helmberg/ ConicBundle/.