NOVEL METHODS FOR TRAFFIC ENGINEERING IN LEGACY IP NETWORKS

Gábor Rétvári, József J. Bíró, Tibor Cinkler
High Speed Networks Laboratory
Department of Telecommunications and Media Informatics
Budapest University of Technology and Economics
H-1117, Magyar Tudósok körútja 2., Budapest, Hungary
E-mail: {retvari, biro, cinkler}@tmit.bme.hu

Abstract

A short introduction, a comprehensive survey of the current state-of-the-art and an overview of some recent progress in the field of OSPF Traffic Engineering (TE) is given. Particularly, we study a multi-stage approach, where the original problem is divided into three subsequent and independent phases. This approach promises with breaking down the intractability of OSPF TE and also to provide interesting further insights. Finally, a simple heuristic is proposed, whose viability is demonstrated by simulation studies.

1 Introduction

To our days, the Internet has quickly evolved into a highly critical communications infrastructure, supporting significant economic, educational and social activities. Simultaneously, the delivery of Internet communications services has become very competitive and end-users are demanding high quality service from their providers. Consequently, performance optimization of large scale public Internet backbones have become an important problem. Network performance requirements are multi-faceted, complex, and often contradictory making the management of Internet traffic quite a bit of a challenge.

Historically, effective management of traffic has been difficult to achieve in public Internet. The reason for this is the limited functional capabilities inherent to the Internet Protocol (IP) suite. Such functionality range from the assessment, measurement and prediction of the volume of incoming traffic, through the mechanisms to manipulate this traffic with a reasonable granularity, all the way to the ability to observe the resultant state the network is driven into. In response to these challenges, the Internet Engineering Task Force (IETF) has recently come up with the concept of Internet *Traffic Engineering* (TE, [1]).

TE is the aspect of Internet network engineering that addresses the performance optimization of operational networks. It encompasses the application of scientific knowledge and engineering principles to achieve specific performance objectives, including reliable and expeditious movement of traffic through the network, efficient and economical utilization of network resources, and planning of network capacity. Motivation for Internet TE is the realization that architectural paradigms and simple capacity expansion are necessary, but not sufficient, to deliver high-quality Internet services under all circumstances.

The Internet exists in order to transfer information from source nodes to destination nodes. Accordingly, one of its most significant functions is the routing of traffic from ingress nodes to egress nodes. Therefore, the single most distinctive function performed by Internet TE is the control and optimization of the routing functionality, to steer traffic through the network in the most effective way.

In the intra-domain IP routing scheme of our days, each router makes independent routing decisions using a synchronized link state database that describes (only) the network topology. The forwarding of IP traffic occurs according to the conventional *shortest path forwarding paradigm*, that is, an IP packet is sent along the shortest path (in terms of some administrative link weight) that is derived based on solely the destination address prefix of the packet and the network topology information. Present IP routing is a valuable engineering artifact of its time, but it is fundamentally flawed, since it does not consider characteristics of the offered traffic and the network capacity constraints. This usually leads to the concentration of user traffic to some highly utilized shortest paths causing congestion and serious service abruption, while, at the same time, certain alternative routes in the network often remain completely underutilized or even vacant.

Traditionally, flaws in IP routing were eliminated by introducing a completely new connection-oriented network layer – equipped with enhanced Traffic Engineering functionality – into the protocol stack, usually MultiProtocol Label Switching (MPLS). Then, the MPLS layer is used to set up a virtual topology overlayed on top of the existing physical infrastructure, that is seen as an optimal one by the IP layer [2]. However, MPLS is slow to be deployed since the global Internet has huge inertia: there is a large installation base of legacy IP routing hardware and software at the service providers, and a great deal of expertise in operating and maintaining it. Thus, the idea to engineer IP networks *without* the need to use MPLS, and staying completely *within* the IP shortest path forwarding paradigm, has come up as an attractive alternative recently [3]. In this positioning paper, we shall explore the current state-of-the-art in Traffic Engineering of legacy IP networks and report on some recent progress we have achieved in the field.

2 OSPF Traffic Engineering

OSPF Traffic Engineering basically means the careful manipulation of OSPF link weights in order to achieve decreased congestion and balanced traffic distribution in IP networks. In this model (see Figure 1), a suitable Traffic Engineer participates in the signaling of the Open Shortest Path First (OSPF, [4]) routing protocol. Based on the routing information gathered from the network, OSPF link weights are computed such that the resultant shortest paths manifest some sophisticated TE goal, and the link weights are distributed back to the routers. Henceforward, the basic operation of OSPF remains the same, but the traffic in the network will follow the paths assigned by the traffic engineer.

Unfortunately, the full-fledged OSPF TE problem is NP hard [3]. To cope with the intractability, a popular approach amongst researchers and practitioners is to partition the task of OSPF TE into multiple phases [5] and, where necessary, invoke viable heuristics. The first phase, the *OSPF TE Path Selection* phase, concerns the expeditious and deliberate assignment of paths between the important source-destination pairs in the network. In the second phase, the so called *Shortest Path Representation* (SPR) phase, the task is to find positive link weights, over which these paths are all shortest paths. In the third phase, the task is to *improve the quality of the link weights* in several different ways. Finally, the emergent link weights are set in the OSPF routers yielding the required routing pattern that was determined by the traffic engineer in the first phase. This "divide-and-conquer" approach promises with an improved OSPF routing table which, while not optimal, still promises with a significant improvement of the utilization and profitability of legacy IP networking hardware and software.

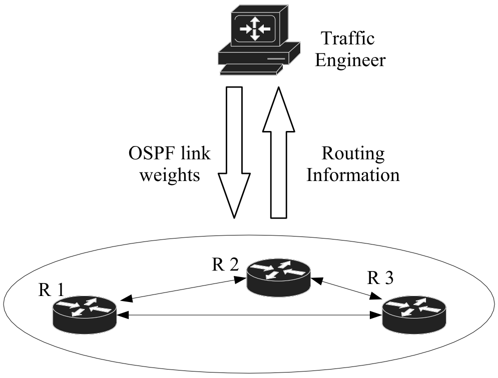


Figure 1: The process model of OSPF Traffic Engineering

2.1 Path Selection

Path selection is the first phase of OSPF TE according to our multi-stage approach. In an OSPF TE deployment, the traffic engineer is given some set of dedicated source-destination pairs and the traffic demands provisioned between them, and the task is to compute a path set that is on the one hand able to satisfy all the demands without overloading any one of the network links, and on the other hand is compatible with the shortest path forwarding paradigm. Path sets fulfilling this latter requirement are called *shortest path representable* paths. In [6] we prove the following important result:

Theorem 1. *The OSPF TE Path Selection problem is NP hard.*

In the same paper we also prove that this first phase is exactly the phase of the OSPF TE problem that hides the real origin of the exponential complexity, and all the remaining phases are polynomially tractable. Therefore, to handle this phase heuristics must be invoked. For a survey and some of our related results see [7].

One might argue that OSPF TE is fundamentally restricted, because it dictates that all traffic must follow the shortest paths in the network. This might limit the variety of path selection strategies compatible with OSPF TE significantly. However, in reality this limitation does not tun out to be too stringent [8]:

Proposition 1. *A set of paths is either shortest path representable by itself, or otherwise, it can be reduced to a shortest path representable one that is on the one hand strictly shorter, and on the other hand provides at least as much bandwidth as the original one.*

Thus, whatever path set one comes up with, either it is immediately shortest path representable, or we can reduce it to a shortest path representable one, and this process usually just improves the quality of the paths.

2.2 Shortest Path Representation

Provided that we have an accurately selected, traffic engineered path set at our disposal, the second phase of the multi-stage OSPF TE model is to compute a set of link weights, which, if set at the OSPF routers, will reproduce the selected paths as shortest paths. This process is called *shortest path representation.* In [9] we give the following characterization of the complexity of this problem:

Theorem 2. *The Shortest Path Representation problem is polynomially tractable.*

In the same paper we also show that solving the Shortest Path Representation problem generally requires solving a linear cost multicommodity flow problem. See for example [10] on how to do this efficiently. However, for real-time applications, where the computational requirements posed by linear programming are too burdensome to tolerate, it is well worth considering approximate algorithms. In [11] we give an iterative algorithm to find the link weights that is based on the successive solving of nothing more than simplistic shortest path problems, for which every commercial IP router contains efficient implementations.

2.3 Improving the Link Weights

As the result of the Shortest Path Representation phase, we obtain a set of "raw" link weights, over which our selected paths (or an improved, shortest path representable path set in the light of Proposition 1) become shortest paths. However, before putting these link weights to use and downloading them to the routers, a post-processing phase is inevitable to make them compatible with OSPF and improve their quality.

First and foremost, OSPF requires that link weights are positive and integer valued, however, our methods so far only guarantee the first requirement, but not the latter. Therefore, in [12] we give a methodology aimed at converting any link weight set to an integral one in polynomial time, without modifying the set of shortest paths produced by the link weights in any regards.

Unfortunately, the process of mapping the selected paths to shortest paths might not be perfectly accurate, which usually turns up in the form of a number of superfluous shortest paths besides the ones selected by the traffic engineer. In [6] we showed that the adverse effects caused by the unintended traffic routed to these superfluous paths might lead to a significant degradation of the efficiency of OSPF routing. Therefore, in [9] we introduced the concept of *minimal shortest path representations*, with the aim of eliminating as many superfluous paths from the set of shortest paths as possible. We summarize our findings as follows:

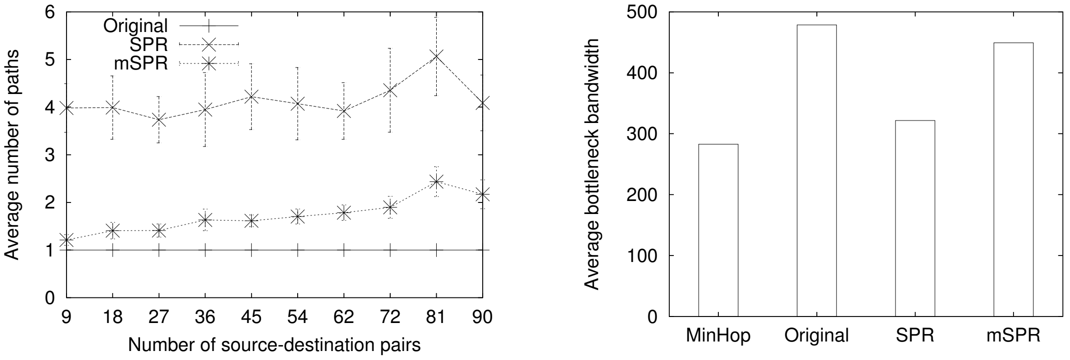
Theorem 3. *A minimal shortest path representation manifests a well-defined theoretical upper bound on the accuracy achievable in mapping paths to shortest paths.*

In [9] we also give an algorithm based on successive linear programming which, given an initial set of paths, first converts it into a shortest path representable one and then computes positive, integer-valued link weights that reproduce the modified path set with the highest accuracy attainable.

2.4 A Viable OSPF TE Solution

In this final section we propose a viable OSPF TE method that follows the multi-stage approach described above. The idea is to use a very efficient path selection heuristic, the shortest-widest-path selection (SWP, [13]) algorithm, which, to each source-destination pair, orders the path with the highest available capacity. Then a minimal shortest path representation is calculated, and the emergent link weights are downloaded to the OSPF routers.

We conducted simulations on a series of 30 random networks of 45 nodes with increasing number of source-destination pairs, to experience how the combination of the SWP algorithm and minimal shortest path representations behave. Consider for instance the average number of paths depicted in Fig. 2. Apparently, traditional methods usually produce more than 4 shortest paths for each source-destination pair, despite of the fact that the traffic engineer only designated one. But,



(a) Average number of paths

(b) Average bottleneck bandwidth

Figure 2: Simulation results for minimum hop-count routing (“MinHop”), the original SWP path set (“Original”), its plain (“SPR”) and minimal (“mSPR”) shortest path representations.

as the figure suggests, minimal representations promise with a serious boost in the precision of this process.

Next, we compared the average bottleneck bandwidth of the paths for minimum hop-count routing (the default OSPF routing scheme), the SWP path set and its “plain” and minimal shortest path representations. While this choice obviously omits interference, the average bottleneck bandwidth is indeed a good measure of the transmission capacity that is made available by the network for the users. The conclusion is that minimal shortest path representations clearly improve the quality of the shortest paths. Our results indicate that the SWP algorithm, combined with shortest path routing, constitutes a really promising traffic engineering platform. Not just that SWP paths can be mapped quite accurately to shortest paths but, in addition, these paths usually provide an abundance of capacity at the same time.

3 Conclusions

IP Traffic Engineering shows really great promise, since it has the potential to boost the performance and profitability of legacy IP networking hardware by several orders of magnitude. And, while doing so, it does not require even the slightest modification of the existing IP network infrastructure, extending its lifetime to a new era of Internet communications. In this paper, we surveyed some recent progress in the field, and proposed a solution that combines the effectiveness of the SWP path selection algorithm with the power of minimal shortest path representations.

References

[1] D. Awduche, A. Chiu, A. Elwalid, I. Widjaja, and X. Xiao. Overview and principles of Internet traffic engineering. RFC 3272, May 2002.

[2] G. Rétvári, J. J. Bíró, T. Cinkler, and T. Henk, “A precomputation scheme for minimum interference routing: the Least-Critical-Path-First algorithm,” in *Proceedings IEEE INFOCOM 2005, Miami, Florida, USA*, vol. 1, pp. 260–268, March 2005.

[3] B. Fortz, J. Rexford, and M. Thorup. Traffic engineering with traditional IP routing protocols. *IEEE Communications Magazine*, 40(10):118–124, Oct 2002.

[4] J. Moy. OSPF Version 2. RFC 2328, April 1998.

[5] M. Pióro, A. Szentesi, J. Harmatos, A. Juttner, P. Gajowniczek, and S. Kozdrowski. On open shortest path first related network optimisation problems. *Performance Evaluation*, 48(1–4):201–223, May 2002.

[6] G. Rétvári, R. Szabó, and J. J. Bíró, "On the representability of arbitrary path sets as shortest paths: Theory, algorithms, and complexity," in *Lecture Notes in Computer Science: Proceedings of the Third International IFIP-TC6 Networking Conference, Athens, Greece*, pp. 1180–1191, May 2004.

[7] G. Rétvári and T. Cinkler, "Practical OSPF traffic engineering," *IEEE Commununications Letters*, vol. 8, pp. 689–691, Nov 2004.

[8] Z. Wang, Y. Wang, and L. Zhang. Internet traffic engineering without full-mesh overlaying. In *Proceedings of INFOCOM 2001*, volume 1, pages 565–571, April 2001.

[9] G. Rétvári, J. J. Bíró, and T. Cinkler, "On Improving the Accuracy of OSPF Traffic Engineering," in *to appear at the Fourth International IFIP-TC6 Networking Conference, Coimbra, Portugal*, May 2006.

[10] R. K. Ahuja snd T. L. Magnanti and J .B. Orlin. *Network Flows: Theory, Algorithms, and Applications*. Prentice-Hall, Englewood Cliffs, NJ, 1993.

[11] G. Rétvári, J. J. Bíró, and T. Cinkler, "A novel Lagrangian-relaxation to the minimum cost multicommodity flow problem and its application to OSPF traffic engineering," in *Proceedings of the The Ninth IEEE Symposium on Computers and Communications (ISCC'2004), Alexandria, EGYPT*, vol. 2, pp. 957–962, June 2004.

[12] G. Rétvári, J. J. Bíró, and T. Cinkler, "On shortest path representation." submitted to IEEE/ACM Transactions on Networking.

[13] Z. Wang and J. Crowcroft. Quality of Service routing for supporting multimedia applications. *IEEE Journal on Selected Areas in Communications*, 14:1228–1234, 1996.