COMPARISON OF INTRA-DOMAIN TRAFFIC ENGINEERING METHODS

R. Suryasaputra*, A. A. Kist*, H. L. Ferra†, R. A. Palmer†, M. J. Dale† and R. J. Harris‡

* School of Electrical and Computer Engineering
RMIT University GPO Box 2476V, Victoria, Australia
Ph: +613 9925 5380 Email: {rsuryasaputra, kist}@ieee.org
† Telstra Research Laboratories, Melbourne, Australia
Email: {herman.ferra, rob.palmer, michael.dale}@team.telstra.com
‡Institute of Information Sciences and Technology
Massey University, Private Bag 11 222 Palmerston North, New Zealand
Email: R.Harris@massey.ac.nz

Keywords: Traffic Engineering, Weight Setting, Neural Networks, Marginal Increase Heuristics

Abstract

This paper presents a comparison study of two Traffic Engineering (TE) methods in an autonomous system (AS). These two methods are novel compared to the existing methods for they are able to provide TE solutions in order of seconds. One could always argue that better solutions can always be found given more time. However, the emphasis here is both methods are made suitable for online TE and it is shown here that they greatly outperform the common IP routing. Herein, we also discuss the consequence of carrying weight changes during the network operation. We believe that weight changes still outweigh the disturbance caused.

1 Introduction

TE has received growing interest by Internet Service Providers (ISPs) because it allows efficient use of the existing network resources. As a result of TE, networks can sustain more traffic, give customers an improved end-to-end network service and help ISPs to meet Service Level Agreement (SLA). This gives us the motivation to compare the two TE methods, namely link weight setting and Multiprotocol Label Switching (MPLS) explicit routing. The two methods compared here are suitable for online TE, hence allows network operators to react in a timely manner to changes in the network.

Intra-domain routing protocols such as Open Shortest Path First (OSPF) or Intermediate System-Intermediate System (IS-IS) compute the routes to the destinations based on link metrics (weights). Links that have a lower weight are treated by the routing protocols as preferable routes. Hence, these links are more prone to congestion. Congestion management, in these cases, can be achieved by employing routing protocol weight setting. Fortz and Thorup [6] suggested that changing a few link weights can re-balance the link loads in the network. In [9], it is shown that a solution of the weight setting problem can be obtained in order of seconds by using an LP formulation.

MPLS allows network administrators to re-direct traffic streams on selected paths via its explicit routing capability. The traffic engineering process is achieved by judiciously selecting these paths such that every traffic stream receives a good service. In the literature, the optimal path selection problem is formulated as a Mixed Integer Linear Programming (MILP) [12]. However, it cannot be guaranteed that the MILP problem can be solved in a "reasonable" time frame.

In [5], the authors devised Marginal Increase Heuristic (MIH) that identifies suitable MPLS paths to carry the traffic demand while conforming to the performance objective. [5] proposes a trained neural network (NN) to generate the initial solution for the MIH heuristic. As a result of combining these techniques, MPLS paths can be re-configured in seconds.

What is the impact of weight setting? Changing weights during a network normal operation are not recommended because of the disturbance caused. The disturbance happens during the convergence period in which routers in the network do not have the same network-wide view. As the consequence, there may be routing loops, packet loss, TCP back-off during this period, which then reduce the network throughput. However, we argue here that it may be worth to optimise the network. In the event of network congestion, the improvement wrought by a weight change, say, may outweigh the short disturbance. In this paper we are trying to quantify the disturbance caused by a single and multiple weight changes.

The contribution of this paper is two folded. Firstly, the performance comparison of two TE methods, namely LP-based Weight setting [9] and NN/MIH [5] for MPLS path selection. It introduces simulation results that demonstrate the performance in terms of packet loss of the two TE methods and compares them to the standard shortest path routing common in IP networks. The experiments are done by using the popular simulator ns-2. A brief summary of both methods is given for completeness. Secondly, the study to investigate the consequences of a weight change event during the convergence period.

2 Optimisation Methods Overview

2.1 LP-based Weight Setting

Using the complementary slackness conditions for the path-flow formulation of the Multi-Commodity Flow (MCF) problem from [3] [11], a set of "modified" link metrics (the original OSPF metrics plus the non-zero link dual prices) can be obtained such that the optimal paths are shortest paths with respect to these modified metrics, for a given traffic matrix. These modified link metrics are then used as the OSPF weights. Earlier results and more details can be found in [9].

By varying the capacity limit by a scaling factor, a limited number of weight changes can be attained. Intuitively, it works as follows: when the capacity scaling factor is very big, the shortest path solution is feasible. As the scaling factor is reduced, the shortest path solution is no longer feasible, because one of the links carries traffic more than its' scaled capacity. Once the problem is solved, the first dual price will appear on that link. Reducing scaling factor further will cause the second link to overload. This link will then have a dual price. This is how one can control a limited number of weight changes [13].

2.2 Neural Networks/Marginal Increase Heuristic (NN/MIH)

An MPLS path selection problem that minimises the maximum link utilisation [12] is formulated and solved by an MILP solver. Intuitively, this has the effect of "spreading" the traffic around the network to avoid congestion. The solutions, which are then used as training data for NN, must conform to all the physical constraints of the network (i.e. the total traffic on any link must be no more than the link capacity) as well as the offered traffic constraints (i.e. capacity allocated to any ODpair must exceed or

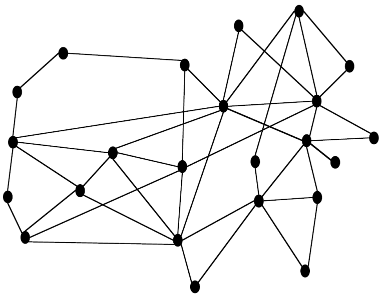


Figure 1: Test network topology

equal the offered traffic for the ODpair).

In operation, the traffic demands are first presented to the “trained” NN, which finds a candidate route configuration for the given demand scenario. The NNs choice of solution may not satisfy all the demands or may violate some other constraints that must be enforced.

The role of the MIH is to correct any flaws due to unsatisfied demand or other broken constraints. Given the actual traffic demands, the trained NN provides a solution, based on its training data, which is the best configuration of paths for the existing traffic conditions. This solution is then used as input to an MIH to generate the final solution. The guiding principle behind the MIH is to select an ODpair and a corresponding path (starting at the origin and ending at the destination), and allocate capacity to the pair and path adding one unit at a time until either all ODpairs have their capacity requirements satisfied or there is no path capable of transporting the extra unit of capacity. Each time a unit of capacity is allocated the link utilisation of some links in the network increase. If the goal is to minimise the maximum link utilisation, then each additional allocation can either leave the objective value unchanged or increase the maximum link utilisation. Which of these two outcomes happens is dependent on the choice of ODpair and path we allocate capacity to. Thus at each iteration the MIH selects the pair and path which will have the least effect on the maximum link utilisation. For details see [5].

3 Experimental Setup

A series of simulations were carried out to benchmark the performance of default OSPF routing and optimised routing based on LP weight setting and the NN/MIH MPLS method. We compare the performance of the three routing schemes based on their packet loss rate as a function of total demand.

For the simulations we have selected a network which consists of 23 nodes, connected by 86 uni-directional links as depicted in Figure 1. The link capacities range from 512 Mbps to 16,384 Mbps. The default OSPF metric for links faster than 100 Mbps is unity (without any changes to the reference bandwidth). The link latency is chosen to be 5 ms. We assume that the traffic represents aggregated flows from every origin to every destination in the network. Thus there are 506 individual flows which need to be accommodated, where each flow represents the offered traffic for one ODpair. Each flow is modelled as a constant bit rate source sending 500 bytes packets from the origin to the destination. The interval rate is determined by the ODpair demand which is randomly generated using a Gaussian distribution with a computed mean and standard deviation. The simulation length is 3 seconds including 0.1 second of a warm-up period. The statistic collection starts after the warm-up period elapsed.

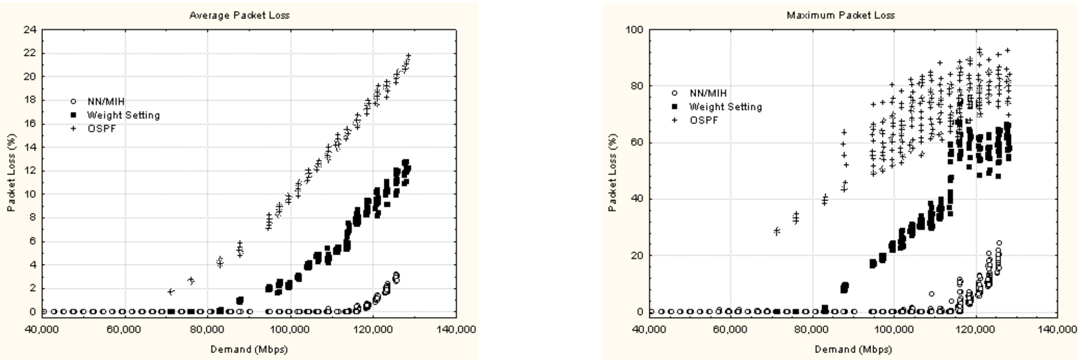
The simulation is done across a number of different traffic matrix instances. There are 374 instances. There are 506 traffic elements in each of the instances. Each of these elements is generated based on the node’s activity as a sender and a receiver. To ensure some randomness, traffic elements are drawn from a distribution as described in [5]. A node with high capacity links originated from it will generate more traffic. Conversely, a node with high capacity links terminated at it will sink more traffic. In practice, the traffic matrix is obtained using historical data, direct traffic measurements, estimation techniques [15] or combination of these techniques.

The inputs for these optimisation techniques are the following information: the source and the destination node of the ODpairs and its associated demand, the network topology and the link capacities. This information is then used to formulate the LP weight setting problem. An LP solver, CLPEX, gives the dual solution, which is then used to modify OSPF metrics. This process is repeated for every 374 traffic matrix (instances) and each will have a different set of weight changes.

To obtain training data for the NN, the path selection problem is solved using MILP solver. The solution is a series of MPLS paths for each corresponding ODpair for a given instance. The same process is repeated over a number of instances to give sufficient training data. Once the NN training is completed, this NN can generate an initial solution for the MIH. For 374 given instances, each of these initial solutions can be attained by NN. Each of them are then refined by the MIH and each will have a set of MPLS paths to route the ODpairs.

4 Results and Discussion

In these experiments, the parameters of interest are the average packet loss and maximum packet loss across all 506 ODpairs. The average packet loss is computed by averaging the loss figure for all 506 ODpairs. The maximum packet loss is taken as the maximum figure across 506 ODpairs. These figures are plotted as a function of the total demand in Mbps.



(a) Average ODpair loss

(b) Maximum ODpair loss

Figure 2: Simulation results with varying traffic demands

Figure 2(a) and 2(b) depict the average ODpair loss and the maximum ODpair loss across different traffic loading. Both TE methods extend the break away point (a point where packet loss starts to occur) compared to OSPF routing significantly. Simulation with OSPF default metric starts to experience packet loss when the total demand approaches 70 Gbps (see Figure 2(b)). The break away points for weight setting and NN/MIH are 88 Gbps and 109 Gbps. These indicate that the network can sustain at least about 26% and 56% more traffic.

There are artifacts in simulations where traffic is routed based on the link metrics (ie. OSPF and weight setting OSPF). This is shown by the wide spread in the maximum packet loss (see Figure 2(b)). The simulator chooses a random outgoing link whenever there are multiple equal length paths from a node to a destination. Hence, one cannot guarantee that traffic is not routed on an already-congested link.

Both TE methods are able to extend the non-loss region. However, as the load increases further, the advantage seems to be slowly diminishing. At this point, the network is approaching its saturation point, where there is not much room left to reroute the traffic. When there is no more room to reroute the traffic, none of TE method would work.

In networks running a shortest path based protocol, a weight setting technique can be readily deployed without the need to upgrade the existing network infrastructure. However, a fine-grain traffic flow control cannot be attained. Weight setting affects flows at the aggregate level.

On the other hand, MPLS gives network administrators full control of routing traffic flows at all levels of granularity. When many equal cost paths exist in a network, it appears that MPLS explicit routing is more suitable for TE purposes. It allows aggregate traffic, or even a single traffic stream, to be routed on a specific path. Although some automated processes can help in configuring MPLS paths, there is a fair amount of configuration required when networks are setup the first time.

5 Discussion on Disturbances caused by Weight Changes

In networks running a link state protocol, such as OSPF or ISIS, weight changes are considered bad because of the disturbances caused to the network. From the network point of view, extra information has to be flooded to disseminate new network information. In OSPF and IS-IS, the information is conveyed in an Link State Advertisement (LSA) and Link State Packet (LSP) respectively. From the end customers' point of view, the disturbances may show in a brief reduction of download speed or an increased delay.

Convergence time: is defined as the time required for all the routers in the network to have the same forwarding information after a certain event in the network. The convergence time due to a weight change can be estimated by adopting that of a link failure [8]. The contributing factors are as follow

1. Initial delay to generate an LSP to inform other routers in the network. The LSP flooding rate may need to be limited in a large network due to an excessive number of LSPs (LSPThrottle).
2. LSP flooding throughout the network, which comprises the LSP queuing delay and transmission delay.
3. Delay between arrival of an LSP and the start of SPF computation in a router (SPFDelay). The period is used to collect few LSPs together and just do the SPF computation once.
4. Shortest Path Forwarding (SPF) computation to compute the shortest route to the destination based on the updated topology (SPFTime).
5. Updating Forwarding Table with the result of SPF computation.

We argue that the disturbance during the convergence process due to a weight change is much less than that of a link failure. The duration of forwarding information unavailability because of a weight change is much shorter than due to a link failure. In the event of link failure, the forwarding information to some prefixes is lost once the link is down. Although a weight change requires the routers a similar process, the forwarding information is not lost during this process. The next subsection addresses the dynamic factors of LSP flooding time. We discuss the lower limit of the LSP flooding time in a single or multiple weight changes event.

We use ns-2 [1] patched with IS-IS routing module [14] to carry out the simulation work. A partially meshed long-type network that consists of 13 nodes and 36 uni-directional links is used as a test network. All link transmission time is set to be 5 ms. The link capacity is set to 10 Mbps. Assuming there is no prioritisation for LSPs forwarding, LSPs will experience a delay due to transmission through network and queuing delay in routers. To observe the queuing delay of an LSP in the ns-2 test network, burst traffic was required because packets are queued during the burst period and released in the idle period. A burst traffic demand was randomly generated from a uniform distribution, and each flow of the traffic demand was used as a mean arrival rate of poisson traffic source to simulate burst traffic.

Queuing Delay: Theoretically, the average delay of a packet on a queue is proportional to the arrival rate of traffic. Therefore, the queuing delay of the LSPs is expected to increase as the arrival rate of traffic increases. However, due to the characteristic of the burst traffic, it does not guarantee that a linear relationship between the LSP queuing delay and the arrival rate always exists. Although the arrival rate is high, if an LSP flooding happens during the idle period, the LSP will experience less delay. Our observation suggested that the queuing time tend to increase with some fluctuations with the increasing traffic arrival rate. However, the magnitude of the queuing delay is not significant at all. This agrees with the measurement study on Sprint backbone network that 50% of packets do not experience queuing delay [10].

Transmission Delay: Another interesting observation is the lower bound of transmission time of the LSPs. The total transmission time of an LSP increases as it travels more links. When a weight change happens in a link, LSPs are flooded through the whole network to synchronise new information. Let's consider two different cases where weight change happens in the centre and in the edge of a network. The maximum number of links that an LSP has to travel throughout the network in the former case is likely to be less than that of links in the latter case. In other words, the flooding takes less time when a weight change happens in the centre of network. The lower bound estimation assumes that none of LSPs are lost. This seems to be a reasonable assumption given the current IP network performance that virtually has 0% packet loss [4].

The flooding process uses hop-by-hop transmission. A router will send an LSP that it just received to its adjacent neighbours except to the neighbour where the LSP came from. This helps to ensure that all routers in the network maintain the same network view, regardless the network status (e.g. routing loops, partial blackout).

There is a common belief that multiple weight changes disturb network more than single weight change because multiple weight change events take a longer time to flood. It may be true in the sense of the queuing delay (although more LSps only cause an insignificant increase in the queuing delay). However, it is not true in the sense of the transmission time. The completion time of flooding all LSPs caused by the multiple weight changes is the same as the completion time of flooding an LSP which travels the most number of links. It implies that multiple weight changes do not necessarily increase the transmission time. Therefore, the impact in terms of transmission time due to multiple weight changes might be the same (or even less) than a single weight change in an extreme scenario (imagine when multiple weight changes happen in the centre of network, and a single weight change does in the edge of network).

Interaction with Inter-Domain Problem: It is possible that the exit point of the AS outgoing (inter-AS) traffic changes when a link weight is altered. This happens when the "hot-potato" routing, which selects the nearest exit point based on IGP metric, is used to route the inter-AS traffic. This must be taken into consideration because according to measurement study in Sprint backbone network [2], 70% of traffic can be potentially affected by hot-potato routing. Recent study, such as [7], considers the joint optimisation of intra-domain and inter-domain routing. However, this may only be applicable to Tier-1 ISPs which have many connectivity to the Internet backbone. Hot potato routing may not be relevant at all for lower level ISPs.

6 Conclusion

We have presented a comparison study of LP-based weight setting and NN/MIH MPLS path selection method. Using MPLS framework, one can achieve better traffic management control at the expense of considerable amount of network configuration work. Furthermore, the routers have to be MPLS capable. On contrast, the weight setting can be readily deployed. The amount of network configuration is minimal with the drawback of less control can be attained. The significance of these two TE methods presented here are their ability to provide TE solution in real-time. Both methods provide a significant improvement in the amount of traffic that can be carried by the network over the standard OSPF routing for a given packet loss figure.

Acknowledgements

The authors wish to thank Australian Telecommunication Co-operative Research Centre (ATcrc) for generous financial support for the on-going work. Robert wishes to thank Dr. Leith Campbell, ATcrc CEO, for his invaluable guidance and to Mr. Suyong Eum for useful discussion.

References

[1] "ns-2 Simulator." [Online]. Available: http://www.isi.edu/ns/nsnam

[2] S. Agarwal, A. Nucci, and S. Bhattacharyya, "Controlling Hot Potatoes in Intradomain Traffic Engineering," Sprint, Technical Report RR04-ATL-070677, July 2004.

[3] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, *Network Flows: Theory, Algorithms and Applications*, 1st ed. Prentice Hall, 1993.

[4] AT&T, "AT&T Global IP Network," 2003. [Online]. Available: http://ipnetwork.bgtmo.ip.att.net/pws/index.html

[5] M. J. Dale, H. L. Ferra, and R. A. Palmer, "Fast MPLS Network Optimisation using Machine Learning," in *TENCON 2005. IEEE International Region 10 Conference*, Melbourne, 2005.

[6] B. Fortz and M. Thorup, "Optimizing OSPF/IS-IS Weights in a Changing World," *Selected Areas in Communications, IEEE Journal on*, vol. 20, no. 4, pp. 756–767, 2002.

[7] K.-H. Ho, N. Wang, P. Trimintzios, and G. Pavlou, "Multi-objective Egress Router Selection Policies for Inter-domain Traffic with Bandwidth Guarantees," in *NETWORKING 2004, Networking Technologies, Services, and Protocols; Performance of Computer and Communication Networks; Mobile and Wireless Communications: Third International IFIP-TC6 Networking Conference.* Springer-Verlag GmbH, 2004, pp. 271–283.

[8] G. Iannaccone, C. N. Chuah, R. Mortier, S. Bhattacharyya, and C. Diot, "Analysis of Link Failures in an IP Backbone," in *Second ACM SIGCOMM Workshop on Internet measurment*, 2002.

[9] J. Murphy, R. Harris, and R. Nelson, "Traffic Engineering Using OSPF Weights and Splitting Ratios," in *Interworking 2002*, ser. IFIP Conference Proceedings, C. McDonald, Ed., vol. 247. Perth, Australia: Kluwer, 2003, pp. 277 – 287.

[10] K. Papagiannaki, S. Moon, C. Fraleigh, P. Thiran, F. Tobagi, and C. Diot, "Analysis of measured single-hop delay from an operational backbone network," in *INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 2, 2002, pp. 535–544.

[11] M. Pióro and D. Medhi, *Routing, Flow, and Capacity Design in Communicaiton and Computer networks*, 1st ed. Morgan Kaufmann Publishers, 2004.

[12] F. Poppe, S. V. den Bosch, P. de La Vallèe-Poussin, H. V. Hove, H. D. Neve, and G. H. Petit, "Choosing the Objectives for Traffic Engineering in IP Backbone Networks Based on Quality-of-Service Requirements," in *QofIS '00: Proceedings of the First COST 263 International Workshop on Quality of Future Internet Services.* London, UK: Springer-Verlag, 2000, pp. 129–140.

[13] R. Suryasaputra, J. Murphy, R. Harris, and W. Lloyd-Smith, "Dousing Hot Spots in OSPF Networks," in *ATNAC*, Sydney, 2004.

[14] K. Svensson, “Implementation of a simulation platform for IS-IS and simulation of Telia’s IP network,” Master’s thesis, Chalmers University of Technology, 2001.

[15] Y. Zhang, M. Roughan, N. Duffield, and A. Greenberg, “Fast accurate computation of large-scale IP traffic matrices from link loads,” *In ACM SIGMETRICS*, June 2003.