Adaptive Multi-topology IGP Based Traffic Engineering with Near-Optimal Network Performance

Ning Wang, Kin-Hon Ho, and George Pavlou

Center for Communication Systems Research, University of Surrey
Guildford, United Kingdom GU2 XH
{N.Wang,K.Ho,G.Pavlou}@surrey.ac.uk

Abstract. In this paper we present an intelligent multi-topology IGP (MT-IGP) based intra-domain traffic engineering (TE) scheme that is able to handle unexpected traffic fluctuations with near-optimal network performance. First of all, the network is dimensioned through offline link weight optimization using Multi-Topology IGPs for achieving maximum path diversity across multiple routing topologies. Based on this optimized MT-IGP configuration, an adaptive traffic engineering algorithm performs dynamic traffic splitting adjustment for balancing the load across multiple routing topologies in reaction to the monitored traffic dynamics. Such an approach is able to efficiently minimize the occurrence of network congestion without the necessity of frequently changing IGP link weights that may cause transient forwarding loops and routing instability. Our experiments based on real network topologies and traffic matrices show that our approach has a high chance of achieving near-optimal network performance with only a small number of routing topologies.

1 Introduction

Intra-domain Traffic Engineering (TE) based on IGPs such as OSPF and IS-IS has recently been receiving numerous attentions in the Internet research community [1-5]. In order to achieve near-optimal or even optimal network performance, it is suggested that both IGP link weights and traffic splitting ratio need to be optimized simultaneously [2, 3, 5] based on the traffic matrix (TM) and the network topology as input. However, this is only applicable to offline TE where knowledge of the estimated TM is assumed *a priori*. Unfortunately, this assumption is usually not valid in real operational networks given frequent presence of traffic dynamics such as unexpected traffic spikes that are difficult to anticipate [6]. As a result, the absence of accurate traffic matrix estimation may lead the offline TE approaches to perform poorly. The most straightforward approach for handling this is to reassign IGP link weights dynamically in reaction to the monitored dynamics. However, re-assigning link weights on the fly may cause transient forwarding loops during the convergence phase, which often leads to service disruptions and traffic instability.

In this paper, we propose *AMPLE* (Adaptive Multi-toPoLogy traffic Engineering), a novel IGP TE approach that is capable of adaptively handling traffic dynamics in operational IP networks. Instead of re-assigning IGP link weights in response to

A. Das et al. (Eds.): NETWORKING 2008, LNCS 4982, pp. 654–666, 2008.

traffic fluctuations, we adopt multi-topology IGPs (MT-IGPs) such as MT-OSPF [7] and M-ISIS [8] as the underlying routing platform to enable path diversity, based on which adaptive traffic splitting across multiple routing topologies is performed for dynamic load balancing. *AMPLE* consists of two distinct phases to achieve our TE objectives. First, the offline phase (e.g., at a weekly or monthly timescale) focuses on the static dimensioning of the underlying network, with MT-IGP link weights computed for maximizing intra-domain path diversity across multiple routing topologies. Since the objective is to obtain diverse IGP paths between each source/destination pair, the computation of MT-IGP link weights is actually agnostic to any traffic matrix. Once the optimized link weights have been deployed in the network, an adaptive TE algorithm performs traffic splitting ratio adjustment for load balancing across diverse IGP paths in multiple routing topologies, according to the up-to-date monitored traffic conditions. This adaptive TE aims to efficiently handle traffic dynamics at short time-scale such as hourly or even in minutes. Given the fact that traffic dynamics are common in operational IP networks, our proposed approach provides a promising and practical solution that allows network operators to efficiently cope with these dynamics that normally cannot be anticipated in advance.

The contributions of our work can be summarized as follows. First of all, *AMPLE* does not require frequent and on-demand re-assignment of IGP link weights, thus minimizing the undesired transient loops and traffic instability. Second, the optimization of the MT-IGP link weights does not rely on the availability of traffic matrix *a priori*, which plagues existing IGP TE solutions due to inaccuracy of traffic matrix estimations. Finally, our experiments based on real network topologies and traffic matrices have shown that *AMPLE* has a very high chance of achieving *near-optimal* performance with only a small number of routing topologies.

2 Related Work

Based on the Equal Cost Multi Path (ECMP) capability in IGP routing, Fortz and Thourp [1] first proposed a local search heuristic to determine a set of IGP link weights that is able to achieve 50%-110% higher service capability in comparison to conventional IGP configurations. Some variations of this TE algorithm were also designed to obtain a set of link weights that is robust to traffic demand uncertainty [11] and link failures [12]. Sridharan et al. [3] revealed that near-optimal TE solutions can be achieved by carefully assigning traffic to some selected next-hops over ECMP paths at each eligible router. By modifying the next-hop entry in the forwarding table of a limited number of routers, uneven traffic splitting based on individual routing prefixes can be emulated. Another approach proposed in [4] is to enable arbitrary traffic splitting at the network edge only. The key idea of this approach is to optimally partition traffic demand into multiple sub-sets at ingress routers and route each of them within dedicated IP routing topologies. As far as resilience is concerned, the authors of [10, 13] proposed to use multiple MT-IGP routing topologies for fast IP recovery while balancing the load in the network after failures. However, none of the proposals above have considered dynamic TE based on the IGP. Most recently, D. Xu et al proposed *PEFT*, a new IGP that is able to achieve *optimal* TE performance by applying traffic splitting with exponential penalties [5]. However, this approach requires changes to the existing intra-domain routing protocols and it is not able to handle traffic dynamics. A more comprehensive TE survey is available in [9].

3 *AMPLE* Overview

As we have already mentioned, *AMPLE* encompasses two distinct tasks, namely (1) offline network dimensioning through link weight optimization for achieving maximum intra-domain path diversity across multiple MT-IGP routing topologies; and (2) adaptive traffic splitting ratio adjustment across these routing topologies for achieving dynamic load balancing in case of unexpected traffic dynamics.

Regarding the first task, our link weight setting scheme is agnostic to traffic matrices, meaning that the input to the MT-IGP link optimization only includes the physical network topology. Fig. 1 illustrates how MT-IGP link weights can be assigned to provide intra-domain path diversity across three routing topologies between a single source/destination pair. The path shown in solid lines in each topology represents the shortest IGP path from the source node *I* to the destination node *E*. It can be seen that disjoined IGP paths are achieved through this link weight setting. Of course, the problem of maximizing IGP path diversity becomes much more complicated if the MT-IGP link weights are computed for *all* source-destination pairs, for instance, in a Point-of-Presence (PoP) topology where every node may send/receive traffic. Once the optimized MT-IGP link weights have been configured in the network, dynamic traffic control through traffic splitting at source nodes can be performed through the provisioned diverse IGP paths.

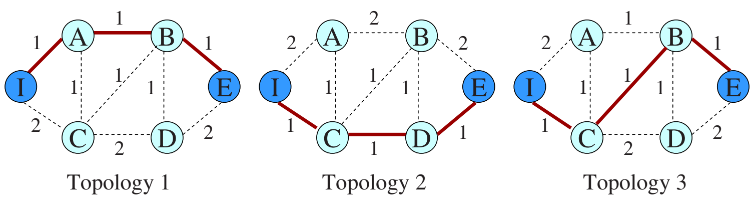

Fig. 1. An example of MT-IGP link weight setting for path diversity

Arbitrary traffic splitting is usually required for achieving optimal TE performance [2, 3, 5]. However, in plain OSPF/IS-IS networks, this prerequisite is not possible since these protocols only allow *equal* traffic splitting onto multiple equal cost paths. Another concern is that changing link weights may cause transient forwarding loops and traffic stability problems. Therefore, this should not be done frequently in reaction to traffic dynamics. By taking these issues into account, a solution that enables unequal traffic splitting while avoiding changing link weights frequently would be attractive. In *AMPLE*, the MT-IGP configuration produced in the offline phase provides an opportunity to use multiple diverse IGP paths for carrying traffic with arbitrary splitting across multiple routing topologies. More specifically, each source node can adjust the splitting ratio of its local traffic (through remarking the Multi-Topology ID field of the IP packets) according to the monitored traffic and network conditions in order to achieve sustainable optimized network performance (e.g. minimize the Maximum Link Utilization, MLU). It is worth mentioning that the computing of new traffic splitting ratios at each source PoP node is performed by a central traffic engineering manager. This TE manager has the knowledge about the entire network topology and periodically gathers the overall network status such as the current utilization of each link and traffic matrices based on which the new traffic splitting ratio is computed and thereafter enforced at individual source nodes.

4 Proposed Algorithms

4.1 Offline Link Weight Optimization

The network model for MT-IGP link weight optimization is described as follows. The network topology is represented as a directed graph $G=<V, E>$, where V and E denote the set of PoP nodes and inter-PoP links respectively. Each link $l \in E$ is associated with bandwidth capacity C_l. In an MT-IGP based paradigm with routing topology set R, each link is also assigned with $|R|$ distinct link weights (denoted by $w_l(r), r \in R$) where $|R|$ is the number of MT-IGP topologies to be configured. In MT-IGP based routing, an IGP path between each pair of nodes (u, v) in routing topology r, denoted by $P_{u,v}(r)$, is the shortest path according to the link weight configuration $W(r)$ for that routing topology. Our definition of path diversity across multiple routing topologies is as follows. For each source-destination pair (u, v) we denote *Degree of Involvement* (*DoI*) for each link l as the number of routing topologies that include l in their shortest IGP paths between the node pair, formally:

$$DoI_l^{u,v} = \sum_{r \in R} x_l^{u,v}(r) \tag{1}$$

where $x_l^{u,v}(r)$ indicates whether link l constitutes the shortest IGP path between u and v in routing topology r:

$$x_l^{u,v}(r) = \begin{cases} 1 & \text{if } l \in P_{u,v}(r) \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

Our ultimate objective is to minimize the chance that a single link is shared by *all* routing topologies between each source-destination pair. The objective is to avoid introducing critical links with potential congestion where the associated source-destination pairs cannot avoid using it no matter which routing topology is used. Towards this end, we define the *Full Degree of Involvement* (*FDoI*), which indicates whether a critical link l is included in the IGP paths between source-destination pair (u, v) in *all* routing topologies:

$$FDoI_l^{u,v} = \begin{cases} 1 & \text{if } DoI_l^{u,v} = |R| \\ 0 & \text{Otherwise} \end{cases} \tag{3}$$

In summary, MT-IGP link weight optimization problem is formally described as follows. To calculate $|R|$ sets of positive link weights $W(r)=\{w_l(r)\}: w_l(r)>0, r \in R$ in order to minimize:

$$\sum_{u,v \in V} \sum_{l \in E} FDoI_l^{u,v} \tag{4}$$

We designed and implemented a Genetic Algorithm (GA) based scheme to compute the MT-IGP link weights for the problem formulated above. The cost function (fitness) is designed as:

$$\frac{\lambda}{\sum_{u,v\in V}\sum_{l\in E} FDoI_l^{u,v}} \tag{5}$$

where λ is a constant value. In our GA based approach each chromosome C is represented by a link weight vector for $|R|$ routing topologies: $C=\{W(r)|\, r\in R\}$. The total number of chromosomes in each generation is set to 100. According to the basic principle of Genetic Algorithms, chromosomes with better fitness value have higher probability of being inherited in the next generation. To achieve this, we first rank all the chromosomes in descending order according to their fitness, i.e., the chromosomes with high fitness are placed on the top of the ranking list. Thereafter, we partition this list into two disjoined sets, with the top 50 chromosomes belonging to the upper class (*UC*) and the bottom 50 chromosomes to the lower class (*LC*). During the crossover procedure, we select one parent chromosome from *UC* and the other parent from *LC* in generation i for creating the child C^{i+1} in generation $i+1$. Specifically, we use a crossover probability threshold $K_C \in [0,0.5)$ to decide the genes of which parent to be inherited into the child chromosome in the next generation. We also introduce a mutation probability threshold K_M to randomly replace some old genes with new ones. The optimized MT-IGP link weights are pre-configured in the network as the input for adaptive traffic engineering which will be detailed in the next section.

4.2 Adaptive Traffic Control

In this section, we present an efficient algorithm for adaptive adjustment of traffic splitting ratio at individual PoP source nodes. In a periodic fashion at a relatively short-time interval (e.g., hourly), the central TE manager needs to perform the following three operations:

1. Measure the incoming traffic volume and the network load for the current interval.
2. Compute new traffic splitting ratios for all PoP nodes based on the measured traffic demand and the network load for dynamic load balancing.
3. Instruct individual PoP nodes to enforce the new traffic splitting ratio over their locally originated traffic.

We start by defining the following parameters:

- $t(u,v)$ – traffic between PoP node u and v.
- $\phi_{u,v}(r)$ – traffic splitting ratio of $t(u,v)$ at u on routing topology r, $0.0\leq \phi_{u,v}(r)\leq 1.0$.

The algorithm consists of the following steps. We define an iteration counter k which is set to zero initially.

Step-1: Identify the most utilized link l_{max} in the network.
Step-2: For the set of traffic flows that are routed through l_{max} in *at least one but not all* the routing topologies (i.e. $\{t(u,v)|\, 0 < DoI_{l_{max}}^{u,v} < |R|\}\ \forall u,v\in V$), consider each at a time

and compute its new traffic splitting ratio among the routing topologies until the first feasible one is identified. A feasible traffic flow means that, with the new splitting ratios, the utilization of l_{max} can be reduced without introducing new hot spots with utilization higher than the original value.

Step-3: If such a feasible traffic flow is found, accept the corresponding new splitting ratio adjustment. Increment the counter k by one and go to Step-1 if the maximum K iterations have not been reached (i.e. $k \leq K$). If no feasible traffic flow exists or $k = K$, the algorithm stops and thereafter the TE manager instructs individual PoPs with the currently computed traffic splitting ratios.

The parameter K controls the algorithm to repeat at most K iterations in order to avoid long running time. In Step-2, the task is to examine the feasibility of reducing the load of the most utilized link by decreasing the splitting ratios of a traffic flow assigned to the routing topologies that use this link, and shift a proportion of the relevant traffic to alternative paths with lower utilization in other topologies. More specifically, the adjustment works as follows. First of all, a deviation of traffic splitting ratio, denoted by δ where $0.0<\delta\leq1.0$, is taken out for trial. For the aggregate traffic flow $t(u,v)$ under consideration, let R^+ be the set of routing topologies in which the IGP paths from u to v traverse l_{max}. The main idea is to decrease the sum of traffic splitting ratios on all the routing topologies in R^+ by δ and at the same time to increase the sum of the ratios on other topologies that do not use l_{max} by δ (We denote this set of topologies by R^- where $R^-=R\backslash R^+$). Specifically, for all the topologies in R^+, which share a common link with the same (maximum) utilization, their traffic splitting ratios are evenly decreased. Hence, the new traffic splitting ratio for each routing topology in R^+ becomes:

$$\phi_{u,v}(r)' = \phi_{u,v}(r) - \delta / |R^+| \qquad \forall r \in R^+$$

On the other hand, let μ_r be the bottleneck link utilization of the IGP path in routing topology $r \in R^-$. The traffic splitting ratio of each routing topology in R^- increases in an inverse proportion to its current bottleneck link utilization, i.e.

$$\phi_{u,v}(r)' = \phi_{u,v}(r) + \left(\frac{1-\mu_r}{\sum_{r \in R^-} 1-\mu_r} \times \delta\right) \quad \forall r \in R^-$$

The lower (higher) the bottleneck link utilization, the higher (lower) the traffic splitting ratio will be increased.

An important issue to be considered is the value setting for δ. If not appropriately set, it may lead to either slow convergence or overshoot of the traffic splitting ratio, both of which are undesirable. On one hand, too large value of δ may miss the chance to obtain desirable splitting ratios due to the large gap between each trial. On the other hand, too small (i.e. too conservative) value of δ may cause the algorithm to perform many iterations before the most appropriate value of δ is found, thus causing slow convergence to the equilibrium. Taking these considerations into account, we apply an algorithm to increase δ exponentially starting from a sufficiently small value. If this adjustment is able to continuously reduce the utilization of l_{max} without introducing negative new splitting ratios on R^+, the value of δ will be increased exponentially for the next trial until no further improvement on the utilization can be

Notation: $U(l)$ is the utilization of link l

Require: A set of MT-IGP topologies R, constants K and Ω

$glb\_improve = TRUE$, $k = 0$
while ($glb\_improve$ & $k < K$) **do**
    $l_{max} \leftarrow$ the most utilized link in the network
    Let T' be the set of traffic flows routed over l_{max} in at least one but not all of the routing topologies
    $t(u,v) \leftarrow$ the first traffic flow in T'
    $feasible\_fnd = FALSE$
    while ($!feasible\_fnd$ & not all flows in T' are examined) **do**
        $R^+ \leftarrow$ the set of routing topologies that uses l_{max} for $t(u,v)$
        $R^- \leftarrow R \setminus R^+$, $\omega = 0$, $\mu_{max} = U(l_{max})$
        $best\_dlt = 0$, $loc\_improve = TRUE$
        while ($\omega \le \Omega$ & $cont$) **do**
            $\delta = \frac{1}{2^{\Omega-\omega}}$
            $\phi_{u,v}(r)' = \phi_{u,v}(r) - \delta / |R^+|) \quad \forall r \in R^+$
            $\mu_r \leftarrow$ the bottleneck link utilization of the path for $t(u,v)$ in topology $r \in R^-$
            $\phi_{u,v}(r)' = \phi_{u,v}(r) + (\frac{1-\mu_r}{\sum_{r \in R^-} 1-\mu_r} \times \delta) \quad \forall r \in R^-$
            $l'_{max} \leftarrow$ the most utilized link among those traversed by $t(u,v)$ in all the routing topologies if $\phi_{u,v}(r)'$ is to be implemented
            if ($U(l'_{max}) < \mu_{max}$ & $\phi_{u,v}(r)' \ge 0 \ \forall r \in R^+$) **then**
                $\mu_{max} = U(l'_{max})$, $best\_dlt = \delta$, $\omega = \omega + 1$
            else
                $cont = FALSE$
            end if
        end while
        if $\mu_{max} < U(l_{max})$ **then**
            accept the adjusted splitting ratios based on $best\_dlt$
            $feasible\_fnd = TRUE$
            $k = k + 1$
        else
            $t(u,v) \leftarrow$ next traffic flow in T'
        end if
    end while
    $l^*_{max} \leftarrow$ the current most utilized link in the network
    if $U(l^*_{max}) \ge U(l_{max})$ **then**
        $glb\_improved = FALSE$
    end if
end while

Fig. 2. Pseudo code - Adaptive traffic splitting ratio adjustment algorithm

made or the value of δ reaches 1.0 (i.e. the maximum traffic splitting ratio that can be applied). The exponential increment of δ works as follows.

$$\delta = \frac{1}{2^{\Omega-\omega}}$$

where Ω is a constant that can be set by the network operator, and ω is the iteration counter. The pseudo code for the algorithm is shown in Fig. 2.

Finally, we discuss some implementation issues on the adaptive traffic control functionality. First of all, traffic measurement plays an important role which provides necessary information on the network status to the central TE manger. In *AMPLE*, we adopt the hop-by-hop based monitoring mechanism that is similar to the proposal of [16]. The basic idea is that, each PoP node within the network is responsible for monitoring the performance (e.g., bandwidth utilization) of the associated outgoing inter-PoP links. In a synchronized fashion, they periodically report the link conditions to the central TE manger. Compared to other end-to-end and distributed traffic monitoring mechanisms, this approach has distinct advantages such as the easiness of identifying the most utilized link in the network.

5 Performance Evaluation

5.1 Experiment Setup

In order to evaluate the performance of *AMPLE*, we use the real topologies and traffic matrices from the GEANT [14] and Abilene [15] networks. The GEANT network topology contains 23 PoP nodes and 74 links, most of which are OC48 (2.5Gbps) and OC192 (10Gbps), but it also has a few link with low capacity of 155Mbps. The traffic matrices have been derived every 15 minutes for several months. We present results based on a 7-day long traffic matrices dataset obtained from the TOTEM Project [17]. The Abilene network topology contains 12 nodes and 30 links, most of which are OC192, but the link between Indianapolis and Atlanta has 2.5Gbps capacity.

5.2 Path Diversity Performance

In this section we present our simulation results for the offline MT-IGP link weight optimization. The performance metric we use to evaluate path diversity is the proportion of source-destination pairs that can successfully avoid any critical link with *FDoI* (i.e., shared by all routing topologies). Fig. 3(a) shows the path diversity performance (i.e. the proportion of source/destination pairs that can fully avoid critical links) in the GEANT topology with: (1) optimized MT-IGP link weight setting for maximizing intra-domain path diversity (*Optimized*), and (2) random link weight setting in all routing topologies (*Random*). From the figure we can see that the optimized link weight setting substantially outperforms the random solution in terms of path diversity. More specifically, our algorithm is able to guarantee 100% avoidance of critical links shared by all topologies with only three routing topologies. In this case, in an event of network congestion, the associated sources are always able to remark their local traffic to enforce alternative IGP path selection to bypass the congested link. Fig. 3(b) shows the path length distribution performance of individual schemes, including the actual link weight setting (*Actual*), proportional to inverse bandwidth capacity (*InvCap*) and our proposed GA-based scheme (*Optimized*), all with three routing topologies. We can see that our proposed algorithm leads to some longer paths due to the efforts for maximizing path diversity across multiple routing topologies, which accounts for some increment in the overall network cost.

Fig. 4 shows the corresponding performance in the Abilene network. Again, we can see that larger number of routing topologies lead to higher path diversity. It should also be noted that the overall path diversity performance is not as good as that of the GEANT topology. This is mainly due to the fact that one of the PoP nodes in Atlanta has node degree of one, thus it is not possible to provide path diversity for it. If we ignore this node, the corresponding path diversity performance can reach 100% with four routing topologies. The lower path diversity of Abilene is also due to its lower mean node degree (2.5) compared to that of GEANT (3.2).

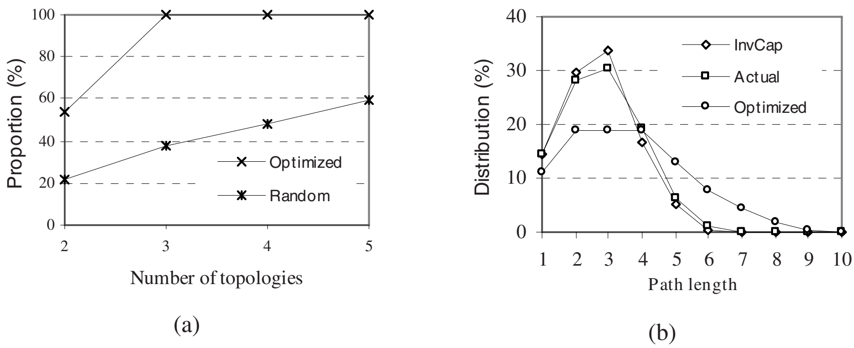

Fig. 3. MT-IGP link weight setting performance (GEANT)

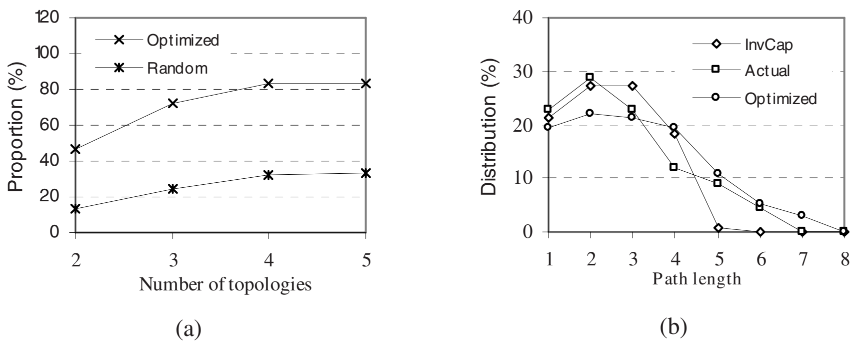

Fig. 4. MT-IGP link weight setting performance (Abilene)

5.3 Adaptive Traffic Engineering Evaluation

We compare the following approaches in our adaptive TE evaluation:

- ***Actual*:** The actual link weight setting in the current operational networks.
- ***InvCap***: Setting link weights proportional to inverse capacity.
- ***Multi-TM*:** We use the TOTEM toolbox [17] to compute a set of link weights for multiple traffic matrices. The objective is to make the IGP TE robust to traffic demand uncertainty. Specifically, the link weights are computed at the beginning

of each day based on the sampled traffic matrices (one per hour) on the same day of the previous week.

- ***AMPLE-n*:** Our proposed adaptive TE algorithm that runs on top of *n* optimized MT-IGP routing topologies.
- ***Optimal*:** As the baseline for our comparisons, we use the GLPK in the TOTEM toolbox to compute optimal MLU for the given topologies and traffic matrices.

We first present in Fig. 5 the MLU achieved by *Actual* and *Optimal* for all the TMs of the GEANT network during a week. As shown in the figure, the performance gap between *Actual* and *Optimal* is very large, which reveals that network resources are far from being utilized at the maximum efficiency. In order to minimize or avoid if possible potential congestion caused by unexpected traffic spikes, it is desirable to maintain the network utilization as close to *Optimal* as possible.

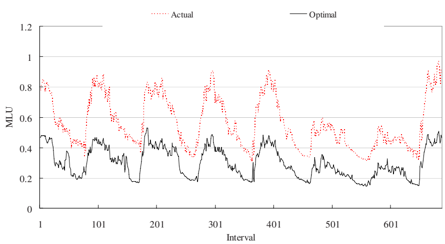

Fig. 5. MLU in GEANT

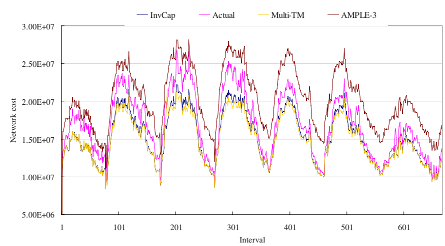

Fig. 6. Network cost in GEANT

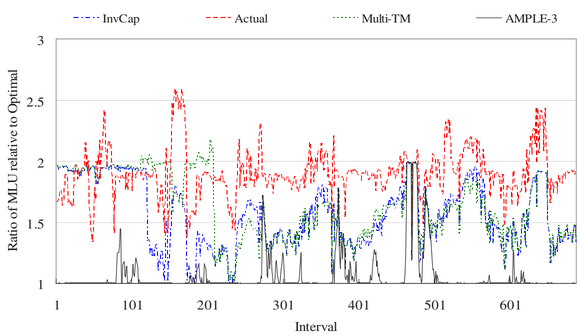

Fig. 7. Ratio of MLU relative to *Optimal* in GEANT

Fig. 7 plots the ratio of MLU relative to *Optimal* based on the same traffic matrices used in Fig. 5. The ratio is calculated as the MLU of a specific method divided by that of *Optimal*. For illustration purposes, we only included *AMPLE* with three topologies (*AMPLE-3*). Nevertheless, Table 1 shows additional statistics on both GEANT and Abilene networks (not shown in figures due to space limit) with the number of topologies varying from 2 to 4. Specific MLU performance metrics are defined:

- **Average maximum link utilization (AMU) –** the average value of the MLU across all the traffic matrices during the seven-day period;
- **Highest maximum link utilization (HMU) –** the highest value of the MLU across all the traffic matrices during the period.
- **Proportion to near-optimal performance (PNO) –** the percentage over all the TMs in which *AMPLE* can achieve near-optimal performance. We define here the meaning of near-optimal to be the MLU that is within 3% gap to the optimality.

A common observation from the figure is that *based on the optimized MT-IGP link weights, AMPLE can achieve near-optimal MLU for most of the traffic matrices.* On average, the MLU of *Actual* is nearly twice that of *Optimal*. This can be explained by the fact that the actual setting of link weights in the GEANT network mainly takes delay into account [18]. Although *InvCap* and *Multi-TM* perform better than *Actual*, their gaps from *Optimal* are still significant with the relative ratio being around 1.5. We further analyze the relevant statistics in Table 1. In the GEANT network, the *Actual* link weight approach produces AMU that is 86% higher than that of the optimal value. Using *InvCap* achieves AMU that is 52% higher than the optimal one, whereas with *AMPLE* the value varies between 0.1% and 43% depending on the number of routing topologies that are used. The larger the number of routing topologies, the closer to the optimal performance can be achieved. For the PNO metric in Table 1, if *AMPLE* is based on two routing topologies, the value is only 13.1% but it still outperforms significantly the *Actual* and *InvCap* approaches. We can now start to see the practical usefulness of our approach in improving network utilizations: When the number of routing topologies increases to three, the PNO boosts up to 78.3%. With 99.6% of all the traffic matrices, *AMPLE* achieves near-optimal performance with four routing topologies. These results reveal that, for the GEANT network, *AMPLE* has very high chance of achieving near-optimal TE performance under any scenario of traffic matrix with four routing topologies. Our experiments based on the Abilene network also show similar results.

We also observed that the *Multi-TM* approach does not achieve good performance in minimizing the MLU according to Fig. 7. There are two reasons for this. First of all, the ultimate objective of *Multi-TM* is to minimize the network cost represented by a piece-wise linear function [1] rather than specifically minimizing MLU. Second, even if multiple traffic matrices with different pattern characteristics are considered in the link weight optimization, unexpected traffic spikes may still introduce poor TE performance. This is especially the case in the Abilene scenario (see HMU in table 1).

Minimizing network cost is another important TE objective. To evaluate this, we adopt the commonly used piece-wise linear function [1] to indicate the actual network cost. By using this cost function, the two objectives of minimizing network bandwidth consumption and load balancing are taken into account simultaneously. Fig. 6 shows the corresponding performance in the GEANT network whose pattern of traffic dynamics pattern is quite regular on the daily basis. In overall, *Multi-TM* is the best performer since it optimizes the network cost as the primary objective. Although *AMPLE* has higher network cost due to the trade-off to path diversity, the increase is small and acceptable. On the other hand, the network cost performance in Abilene is similar to that in GEANT, but due to the space limit it is not shown in the paper.

Table 1. Probability of achieving near-optimality

Optimization Method	GEANT (%)			Abilene (%)		
	AMU	HMU	PNO	AMU	HMU	PNO
Optimal	30.05	52.82	-	12.2	33.42	-
InvCap	45.72	94.41	1.6	19.58	62	0.15
Actual	55.74	96.91	0	19.59	63.24	1.19
Multi-TM	48.56	104.15	0.44	53.2	230	0.15
AMPLE-2	42.9	94.61	13.08	18.61	60.96	64.14
AMPLE-3	31.95	60.36	78.34	12.36	33.44	88.69
AMPLE-4	30.08	52.88	99.56	12.4	49.6	97.77

6 Summary

In this paper we presented *AMPLE*, a novel TE approach that enables dynamic load balancing in operational IP networks. Instead of frequently changing IGP link weights, we use multi-topology IGP routing protocols that allow adaptively splitting traffic across multiple routing topologies. Offline link weight optimization is performed in order to enable path diversity, followed by the adaptive control of traffic splitting across individual routing topologies according to the monitored traffic dynamics. Our simulation experiments based on the GEANT and Abilene networks and the respective traffic matrices have shown that *AMPLE* has high chance of achieving near-optimal network performance with only a small number of topologies.

Acknowledgement

This work was partially funded by the EU FP6 AGAVE Project (IST-027609)

References

1. Fortz, Thorup, M.: Internet Traffic Engineering by Optimizing OSPF Weights. In: Proc. IEEE INFOCOM (2000)
2. Wang, Y., et al.: Internet Traffic Engineering without Full Mesh Overlaying. In: Proc. IEEE INFOCOM (2001)
3. Sridharan, A., et al.: Achieving Near-Optimal Traffic Engineering Solutions for Current OSPF/ISIS Networks. IEEE/ACM Transactions on Networking 13(2) (April 2005)
4. Wang, J., et al.: Edge-based Traffic Engineering for OSPF Networks. Computer Networks 48(4) (July 2005)
5. Xu, D., et al.: Link-state Routing with Hop-by-hop Forwarding Can Achieve Optimal Traffic Engineering. In: Proc. IEEE INFOCOM (2008)
6. Wang, H., et al.: COPE: Traffic Engineering in Dynamic Networks. In: Proc. ACM SIGCOMM (2006)
7. Psenak, P., et al.: Multi-Topology (MT) Routing in OSPF., RFC 4915 (2007)
8. Przygienda, T., et al.: M-ISIS: Multi Topology (MT) Routing in IS-IS. RFC 5120 (2008)

9. Wang, N., et al.: An Overview of Routing Optimization for Internet Traffic Engineering. IEEE Communications Surveys and Tutorials (to appear, 2008)
10. Kvalbein, A., et al.: Post-Failure Routing Performance with Multiple Routing Configurations. In: Proc. IEEE INFOCOM (2007)
11. Zhang, C., et al.: On Optimal Routing with Multiple Traffic Matrices. In: Proc. IEEE INFOCOM (2005)
12. Nucci, A., et al.: IGP Link Weight Assignment for Operational Tier-1 Backbones. IEEE/ACM Transactions on Networking 15(4), 789–802 (2007)
13. Menth, M., Martin, R.: Network Resilience through Multi-topology Routing. In: Proc. International Workshop on Design of Reliable Communication Networks (DRCN) (2005)
14. The GEANT topology: http://www.geant.net/upload/pdf/GEANT_Topology_12-2004.pdf
15. The Abilene topology and traffic matrices dataset: `http://www.cs.utexas.edu/~yzhang/research/AbileneTM/`
16. Asgari, H., et al.: Scalable Monitoring Support for Resource Management and Service Assurance. IEEE Network Magazine (November 2004)
17. Balon, S., et al.: Traffic Engineering an Operational Network with the TOTEM Toolbox. IEEE Transaction on Network Service Management 4(1), 51–61 (2007)
18. Uhlig, S., et al.: Providing Public Intradomain Traffic Matrices to the Research Community. ACM SIGCOMM Computer Communication Review 36(1) (January 2006)