Making Outbound Route Selection
Robust to Egress Point Failure

Mina Amin, Kin-Hon Ho, Michael Howarth and George Pavlou

Centre for Communication Systems Research, University of Surrey, UK

{M.Amin,K.Ho,M.Howarth,G.Pavlou}@eim.surrey.ac.uk

Abstract. Offline inter-domain outbound Traffic Engineering (TE) can be formulated as an optimization problem whose objective is to determine primary egress points for traffic exiting a domain. However, when egress point failures happen, congestion may occur if secondary egress points are not carefully determined. In this paper, we formulate a bi-level outbound TE problem in order to make outbound route selection robust to egress point failures. We propose a tabu search heuristic to solve the problem and compare the performance to three alternative approaches. Simulation results demonstrate that the tabu search heuristic achieves the best performance in terms of our optimization objectives and also keeps traffic disruption to a minimum.

1 Introduction

Inter-domain Outbound Traffic Engineering (TE) [1,2] aims to control traffic exiting a domain by assigning the traffic to the best egress points (i.e. routers or links). Since inter-domain links are the most common bottlenecks in the Internet [2], optimizing their resource utilization is a key objective of outbound TE. In the literature, several outbound TE approaches have been proposed [2,3]. These proposals, however, have neglected the detrimental impact of inter-domain EP failure on the achieved TE performance. In fact, the network performance under failure conditions should ideally be optimized by considering failure as part of the outbound TE optimization.

Failure occurs as part of daily network operations [6]. Inter-domain failures are typically caused by: (1) *physical failures* such as inter-domain link fiber cut and equipment failure, or (2) *logical failures* such as router CPU overload, operation systems problem and maintenance. A recent study [4] discovered that logical inter-domain link failures are common events and are usually transient in nature. When a failure happens on an EP, traffic is shifted to another available EP in accordance to the BGP route selection policies. However, if a large amount of traffic is shifted, congestion is likely to occur on these new serving EPs. This problem has not been considered in the existing outbound TE proposals. An intuitive approach to minimize this congestion is to redirect the traffic to another EP by adjusting BGP routing policies in an online manner until the best available EP has been found. Such online trial-and-error approach may cause router misconfiguration, unpredicted traffic disruption and BGP route flooding, leading to route instability. As a result, an outbound TE approach that produces optimal performance under both normal and failure scenarios so as to minimize online and unpredictable route changes is highly desirable.

In this paper, we propose a *multi-level* outbound TE approach that is robust to EP failure, which achieves reasonably good performance under both Normal State (NS) and Failure States (FS). We refer NS to no failure and each FS to a single EP failure. In multi-level outbound TE, the first level is to select Primary EPs (PEP) under NS, in

a similar fashion to previous work [2,3]. Then, the second level is to select the next best EP as the Secondary EP (SEP) when the PEP fails. This approach can also be repeated for successive EPs. For example, a tertiary EP is used as the traffic exit point when both the PEP and SEP fail. However, since *single* link failure is the predominant form of failure in communication networks [6], we therefore consider a bi-level outbound TE formulation. This problem can be formulated as follows:

Given a network topology, destination prefixes and an inter-domain Traffic Matrix (TM), determine for each traffic demand the PEP and the SEP upon PEP failure. The optimization objective is to minimize the maximum EP utilization under Normal State (NS) and the average of maximum EP utilization across all Failure States (FSs).

Previous work [5,6,7] on making TE robust to link failure has focused on the intra-domain problem. Heuristics have been proposed to compute a set of IGP link weights that is robust to any single intra-domain link failure. Our work is similar to this previous work, but the primary difference is that we focus on inter-domain outbound TE. Given that a significant amount of Internet traffic is routed across domains (e.g. the rapidly increasing peer-to-peer traffic) and inter-domain link failures are common and transient, making outbound TE robust to failures is an important problem. To the best of our knowledge, this issue has yet not been investigated.

To solve the bi-level outbound TE problem, we propose a tabu search heuristic and compare its performance to alternative strategies. Experimental results demonstrate that the tabu search heuristic significantly improves the performance under all FSs (about 1%-3% from the FS lower bound) with a small performance degradation under NS (about 2%-8% from the NS lower bound). The tabu search heuristic also minimizes traffic disruption.

This paper has the following structure. Section 2 presents the bi-level outbound TE problem formulation. We detail the proposed tabu search heuristic in Section 3. Section 4 presents three alternative strategies for solving the problem. Then, we present our evaluation methodology and simulation results in Section 5 and 6 respectively. Finally, we conclude the paper in Section 7.

2 Problem Formulation

TABLE 1. NOTATION USED IN THIS PAPER

NOTATION	DESCRIPTION
K	A set of destination prefixes, indexed by k
L	A set of egress points, indexed by l
S	A set of states $S=\{\varnothing \cup (\forall\, l \in L)\}$, indexed by s
I	A set of ingress points, indexed by i
$t(k,i)$	Bandwidth demand of traffic flows destined to destination prefix $k \in K$ at ingress point $i \in I$
$Out(k)$	A set of egress points that have reachability to destination prefix k
c_{inter}^{l}	Capacity of the egress point l
x_{sk}^{l}	A binary variable indicating whether prefix k is assigned to the egress point l in state s
u_{s}^{l}	Utilization on non-failed egress point l in state s. Its value is zero when $s=l$
$U_{max}(s)$	maximum egress point utilization in state s
U_{Ave}^{FS}	Average of maximum egress point utilization across all failure states

2.1 Primary Egress Point Selection Problem Formulation

We make the following assumptions prior to the problem formulation: (1) we focus the TE optimization objective only on inter-domain resources[1]. (2) we apply our work to the single egress selection case and on a general network model where each EP is composed of an egress router attached to a single inter-domain link[2].

In this section, we review the problem formulation of single egress selection described in [2]. This determines the PEPs under NS ($s=\varnothing$), and is hence the first level of our bi-level outbound TE problem. Table 1 shows the notation used in this paper.

Each element of the inter-domain TM, $t(k,i)$, represents the total volume of traffic from ingress point i towards destination prefix k. Due to the increasing use of multi-homing, a prefix usually can be reached through multiple EPs, thereby allowing outbound TE to select the best PEP for the traffic. Given an inter-domain topology, destination prefixes and an inter-domain TM, the task of single egress selection is to determine the best PEP for each destination prefix[3]. The optimization objective we consider is to minimize the maximum EP utilization, which is defined as the highest utilization among all EPs:

$$Minimize\ U_{max}(\varnothing) = Minimize\ \underset{\forall l \in L}{Max}(u_{\varnothing}^l) = Minimize\ \underset{\forall l \in L}{Max}(\frac{\sum_{k \in K}\sum_{i \in I} x_{\varnothing k}^l t(k,i)}{c_{inter}^l}) \tag{1}$$

subject to the following constraints:

$$\forall k \in K: \sum_{l \in Out(k)} x_{\varnothing k}^l = 1 \tag{2}$$

$$\forall l \in L, k \in K: x_{\varnothing k}^l \in \{0,1\} \tag{3}$$

Minimizing objective function (1) ensures that traffic is moved away from congested to less utilized EPs and attempts to achieve load balancing across all EPs. Constraints (2) and (3) ensure that only one PEP is selected for each destination prefix under NS. In [2], an EP capacity constraint is added to the problem formulation. Nevertheless, we believe that our uncapacitated version is adequate since objective function (1) is effectively similar to the EP capacity constraint. The single egress selection problem has been proven to be NP-hard [2] by reducing it to the Generalized Assignment Problem (GAP), which is itself NP-hard.

2.2 Bi-level Egress Point Selection Problem Formulation

Given the inputs for the single egress selection, the goal of bi-level outbound TE is to determine, for each destination prefix, both a PEP under NS and a SEP that will serve the traffic when the PEP has failed (i.e. under FS). The optimization objective of the bi-level outbound TE problem is to minimize both the maximum EP utilization under NS and the average maximum EP utilization across all FSs. Recall that each FS corresponds to a single EP failure. The number of FSs is hence equal to the number of EPs $|L|$. By adding the NS, the total number of states $|S|$ is $|L| + 1$. The computational complexity of the bi-level outbound TE problem is thus an increasing function of the total number of states. To reduce this complexity, one may take the idea in [7] of performing the TE only on a small subset of FSs whose failures have significant impact on network performance. This set of EPs is referred to as *critical* EPs but we

---

1. This assumption is according to the fact that capacity over-provisioning is usually employed by ISPs within their IP backbones [10]
2. In [2], outbound TE is divided into Single and Multiple Egress Selection. Since the objective of this paper is to demonstrate the principle of robust outbound TE, we consider assumption 2. Nevertheless, our idea is also applicable to multiple egress selection and multiple inter-domain links attached to each EP.
3. Assigning a PEP to a destination prefix is equivalent to selecting that PEP for traffic demands that head towards that destination prefix

leave this as future work. The maximum EP utilization under FS s can be calculated in a similar way to (1) as:

$$\forall s \in S : Minimize\ U_{max}(s) = Minimize\ \underset{\forall l \neq s}{Max}(u_s^l) = Minimize\ \underset{\forall l \neq s}{Max}(\frac{\sum_{k \in K}\sum_{i \in I} x_{sk}^l t(k,i)}{c_{inter}^l}) \tag{4}$$

The term $x_{sk}^l t(k,i)$ consists of flows which are assigned to EP l as their PEP and also flows which are assigned to EP l as their SEP. Since our optimization objective is to minimize the maximum EP utilization under both NS and FSs simultaneously, a bi-criteria optimization problem is formed. However, the two optimization objectives conflict with each other and hence we resort to a weighted sum approach to transform them into a single-criterion optimization problem, which is simpler to solve. The optimization objective function is thus:

$$Minimize\ F = (1-w)U_{max}(\varnothing) + wU_{Ave}^{FS}\ ,\ 0 \leq w \leq 1 \tag{5}$$

$$where\ U_{Ave}^{FS} = \underset{\forall s \in S/\{\varnothing\}}{Ave}(U_{max}(s)) = \frac{\sum_{s \in S/\{\varnothing\}} U_{max}(s)}{|S|-1} \tag{6}$$

subject to the following constraints:

$$\forall k \in K, s \in S : \sum_{l \in Out(k)} x_{sk}^l = 1 \tag{7}$$

$$\forall l \in L, k \in K, s \in S : x_{sk}^l \in \{0,1\} \tag{8}$$

$$\forall l \in L, k \in K\quad if\quad x_{\varnothing k}^l = 1\quad then \begin{cases} x_{sk}^l = 1 & \forall s \in S/\{l\} \\ x_{sk}^l = 0 & \forall s = l \end{cases} \tag{9}$$

By varying weight w and re-solving F, one can generate a trade-off curve between the two objectives using the weighting method of multi-objective programming [10]. If we solve the problem with $w=0$, the problem is simply reduced to the PEP selection problem. If $w=1$, the problem then completely ignores the performance under NS. In this paper, we present results for $w=0.5$ (i.e. equal weight to the objectives optimized under NS and FS), which allows us to achieve significant performance improvement for SEP selection with only a small performance degradation for the PEP selection. Constraints (7) and (8) are equivalent to constraints (2) and (3), ensuring that only one EP is selected for each destination prefix as the PEP under NS ($s=\varnothing$) and only one EP is selected for each prefix as the SEP under FSs. Constraint (9) ensures that if prefix k is assigned to EP l in NS, then this prefix remains on l for all the FSs except when the current FS is the failure on l.

It is not surprising that the bi-level outbound TE problem is NP-hard, since it is an extension of the PEP selection problem, which is itself NP-hard. If the number of FSs is zero, the bi-level outbound TE is reduced to the PEP selection problem. As a result, we resort to using a heuristic approach to solve the problem.

For the implementation of the bi-level outbound TE solution, we can assign for each prefix the largest value of BGP *local-pref* for the selected PEP, the second largest value for the selected SEP and smaller values for the rest of the EPs. Whenever a PEP fails, the EP with the next largest *local-pref* (i.e. the SEP) becomes the exit point for the traffic headed towards the destinations. The solution can also be implemented by the proposal in [4] in which an IP tunnel is established to move traffic from the failed PEP to the precomputed SEP for faster failure recovery.

3 Proposed Tabu Search Heuristic

The Tabu Search (TS) methodology [8] guides local search methods to overcome local optimality and attempts to obtain near-optimal solutions for NP-hard optimization problems. Due to space limitations, the reader is referred to [8] for an overview of TS. In general, our proposed TS heuristic first requires initial PEP and SEP selection solutions, and then proceeds to obtain neighbor solutions by using a neighborhood search strategy in order to gradually enhance the quality of the initial solution.

3.1 Non-TE initial solution

We obtain initial PEP and SEP selection solutions by randomly selecting EPs for the destination prefixes while satisfying constraints (7) to (9). These initial solutions can be regarded as non-TE (i.e. non-optimized) solutions. The rationale of using such initial solutions is to demonstrate the effectiveness of the proposed TS heuristic in producing good performance from poorly performing initial solutions.

3.2 Neighborhood Search Strategy

A *move* transforms the current (initial) solution into a neighbor solution. To perform a move, we apply the SUBROUTINE_BESTMOVE heuristic shown in Figure 1, to first identify the best move for each FS and then select the best one among all the FSs.

SUBROUTINE_BESTMOVE:

1. **For** each $s \in S / \{\varnothing\}$

2.   Store the $PEP_{current}$, $current\_cost \leftarrow (1-w)U_{max}(\varnothing) + wU_{max}(s)$ and $j \leftarrow 0$

3.   **For** each $k \in MUEP_s$

4.   temporarily shift k from $MUEP_s$ to $LUEP_s$ to achieve the new solution PEP_{new}
5.   call SUBROUTINE_GREEDY_HEURISTIC for state s and temporarily make changes for current SEP

6. $new\_cost \leftarrow (1-w)U'_{max}(\varnothing) + wU'_{max}(s)$ and $j \leftarrow j+1$

7. $diff(j) \leftarrow current\_cost - new\_cost$ and restore the $PEP_{current}$

8.   find $\underset{j}{Max\, diff}(j)$ and its coresponding PEP_{new}, $PEP_{state\_best} \leftarrow PEP_{new}$ // the best move for each FS

9. **For** each $s \in S / \{\varnothing\}$

10.   temporarily implement the current PEP_{state_best}
11.   call SUBROUTINE_GREEDY_HEURISTIC for all FSs to achieve SEP_{state_best}, implement it temporarily

12.   calculate $F=(1-w)U_{max}(\varnothing) + wU_{Ave}^{FS}$

13. Find Minimum F // to find the best move among all the FSs (PEP_{state_best}, SEP_{state_best})
14. Accept the changes that yield the Minimum F

Figure 1. SUBROUTINE_BESTMOVE

The following steps explain how to identify the best move for each FS:

Step 1. Store the currently assigned PEP for all prefixes in $PEP_{current}$. Calculate the *current_cost*, i.e. the weighted sum of the maximum EP utilization under both NS and the current FS (Figure 1 line 2). List all the prefixes in $PEP_{current}$ assigned to the Most Utilized EP under the current FS (MUEP$_s$)[4]. Consider each prefix at a time in the list and apply steps 2 to 4 until all the destination prefixes in the list have been considered (Figure 1 lines 3 to 7).

Step 2. Shift the prefix's PEP from MUEP$_s$ to the Least Utilized EP (LUEP$_s$)[5] (the

---

4. MUEP$_s$ is the link that has $\underset{\forall l \neq s}{Max\, u_s^l}$

5. LUEP$_s$ is the link that has $\underset{\forall l \neq s}{Min\, u_s^l}$

goal of this move is to attract traffic towards the $LUEP_s$ and potentially to reduce the load on the $MUEP_s$). This results in a new solution for the PEP selection, which is denoted by PEP_{new}.

Step 3. Reassign the SEPs for the destination prefixes that have been assigned to the failed EP by using the SUBROUTINE_GREEDY_HEURISTIC algorithm. The algorithm works as follows: (a) Sort all the destination prefixes on the failed EP by descending volume of traffic. (b) Take the first of these ordered prefixes and select as its SEP the available EP with the minimum utilization. (c) Repeat step (b) for the rest of the destination prefixes in order.

Step 4. Calculate the *new_cost* in the same way as the *current_cost* for the latest solution (Figure 1 line 6). Then calculate the difference between the *current_cost* and *new_cost* (i.e. *diff = current_cost- new_cost*). Restore the $PEP_{current}$.

Step 5. Identify the prefix that produces the largest value of *diff* (i.e. largest difference between the *current_cost* and *new_cost*). Consider the PEP_{new} that corresponds to this prefix as the best move for the current FS. Store this PEP_{new} in PEP_{state_best}.

Step 6.Repeat steps 1 to 5 for each FS and identify their PEP_{state_best} until all the FSs have been considered (Figure 1 lines 1 to 8).

After identifying the best move for each FS, we now identify the best of the best moves for all FSs by the following steps:

Step 1. For the best move for each FS, reassign the SEPs (SEP_{state_best}) for the corresponding PEP_{state_best} by using the SUBROUTINE_GREEDY_HEURISTIC algorithm for all the FSs. (this calls the subroutine *s* times, once for each FS). Calculate objective function (5). Repeat step 1 for the best move of the next FS until all the FSs have been considered (Figure 1 lines 9 to 12).

Step 2. For all the FSs evaluated in step 1, choose the best move (i.e the PEP_{state_best} and its corresponding SEP_{state_best}) that yields the minimum objective value (Figure 1 lines 13-14).

3.3 Tabu List

The tabu list is a memory list that memorizes the most recent moves, operating as a first-in-first-out queue. As suggested in [8], the size of the tabu list depends on the size and characteristics of the problem. Since in our algorithm the attributes of a move are MUEP, LUEP and shifted destination prefixes, the size of the tabu list is determined by the number of destination prefixes. We define the size of the tabu list to be *total number of destination prefixes / |L|*.

3.4 Diversification

The goal of diversification is to prevent the searching procedure from indefinitely exploring a region of the solution space that consists of only poor quality solutions. It is a modification of the neighbourhood searching strategy and is applied when there is no obvious performance improvement after a certain number of iterations. For diversification, a group of highly and lightly utilized EPs are chosen for shifting destination prefixes under a FS. We define the threshold of obvious performance improvement to be 10% of the best visited solution and the number of iterations to be 10% of the maximum iteration mentioned below.

3.5 Stopping Criterion

Many stopping criteria can be developed depending on the nature of the problem. The most common criterion, used in this paper, is to define a maximum number of iterations. However, we do not arbitrary select the number of maximum iterations since the performance of the TS heuristic mainly depends on how many times the PEPs and SEPs are reassigned. We found that setting the maximum iteration number to be 5 times the number of destination prefixes gives us sufficiently good results.

4   Alternative Strategies

Our proposed TS heuristic is only one of several approaches in solving the bi-level outbound TE problem. In this section, we present three alternative approaches. For these approaches, OPTIMAL-AWARE HEURISTIC is used for the PEP selection and the three alterative approaches only differ in their SEP selection. We remark that the OPTIMAL-AWARE HEURISTIC is our best attempt in solving our PEP selection problem, as no algorithm for solving the problem with objective function (1) has been proposed in the literature. The OPTIMAL-AWARE HEURISTIC works as follows:

Step 1: Calculate the mean utilization by dividing the total traffic volume by the total capacity of all EPs. We regard this mean utilization as the theoretical optimal (i.e. the most load balanced) utilization targeted for each EP to achieve. However, this theoretical result is not a valid solution because it allows arbitrary traffic splitting over any EP, violating constraints (7) and (8). Nevertheless, it is used as an "NS lower bound" solution[6] for comparing performance with other strategies.

Step 2: To ensure that each EP does not exceed the theoretical optimal utilization, set the mean utilization as a capacity constraint on each EP.

Step 3: Sort the destination prefixes in descending order according to the amount of traffic they carry and choose one at a time in order.

Step 4: Select the EP with the minimum utilization as the PEP of this destination prefix if it satisfies the capacity constraint, if not proceed to the next prefix. Repeat this step until all the destination prefixes have been considered.

Step 5: If there exist unassigned destination prefixes because of capacity constraint violation, re-run step 4 without considering the capacity constraint.

4.1 Random Reassignment Strategy

In the Random Reassignment (RANDOMR) strategy, when an EP fails, the prefixes on the failed EP are re-assigned to other available but *randomly* chosen EPs. We illustrate an example of the RANDOMR in Figure 2. In this example there are three EPs (A, B and C) with equal capacity (60Mbps) and an ingress point *i*. The input traffic flows and their traffic volume are shown in Table 2. Figure 2(a) shows a solution of the PEP selection, which can be generated by the OPTIMAL-AWARE HEURISTIC. The solution has the best load balancing over all the EPs. Figure 2(b) shows the solution of the SEP selection under EP A failure produced by the RANDOMR. The figure demonstrates that when EP A is assumed to fail, destination prefixes *k1* and *k5* are then randomly assigned to EP B and C respectively as their SEPs. This random assignment, however, causes heavy load on EP B which could easily lead to congestion (e.g.

---

6. In a similar fashion, we define the "FS lower bound" to be the total volume of traffic divided by the capacity of all EPs excluding the failed one.

$u_A^B = \frac{20+20+10}{60} = 0.833$, $u_A^C = \frac{10+10+10}{60} = 0.5$). Therefore, the RANDOMR performs poorly under any FS since during optimization failures are not taken into account. Nevertheless, since only the affected destination prefixes are reassigned, the level of traffic disruption is minimized (i.e. only prefixes *k1* and *k5* are disrupted when EP A fails).

TABLE 2. INPUT TRAFFIC FLOWS

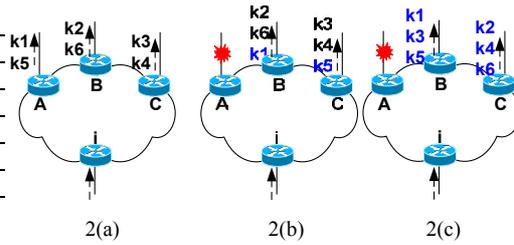| TRAFFIC FLOW	TRAFFIC VOLUME(MBPS)
t(k1,i)	20
t(k2,i)	20
t(k3,i)	10
t(k4,i)	10
t(k5,i)	10
t(k6,i)	10

2(a)   2(b)   2(c)

Figure 2 show the destination prefix assignment according to (a) OPTIMAL-AWARE HEURISTIC , (b) RANDOMR if EP A fails and (c) GLOBALR if EP A fails

4.2 Global Reassignment Strategy

In the Global Reassignment (GLOBALR) strategy, for any EP failure, the OPTIMAL-AWARE HEURISTIC is reapplied to perform PEP selection from scratch. Such network-wide computation can be regarded as the best approach with respect to performance but possible large traffic disruption because the PEPs for most of destination prefixes are likely changed. We use the GLOBALR as a benchmark for comparing the performance to other strategies. Figure 2(c) shows the result of the GLOBALR based on the previous example. As can be seen, when EP A fails, some prefixes are reassigned away from their original EPs. For example, *k2* and *k6* are shifted from EP B to C while *k3* is shifted from EP C to B. Nevertheless, the utilization upon any EP failure is optimal (i.e. $u_A^B = \frac{20+10+10}{60} = 0.666, u_A^C = \frac{20+10+10}{60} = 0.666$).

4.3 Greedy Reassignment Strategy

In the Greedy Reassignment (GREEDYR) strategy, for any EP failure, only the prefixes assigned on the failed EP are re-assigned by a greedy heuristic as follows: the prefix that carries the largest amount of traffic is reassigned to the available EP that has the minimum utilization. This step repeats for the rest of the affected prefixes.

5 Evaluation Methodology

5.1 Network Topology and Inter-domain Traffic Matrices

Our experiment is performed on topologies with 3, 6 and 10 EPs. We note that the 3-EP topology is the smallest scenario where the bi-level outbound TE is applicable. Larger topologies then evaluate performance scalability of the proposed strategies.

We assume the capacity of all the EPs to be OC-12 (622Mbps). For scalability and stability concerns, outbound TE can focus only on a small fraction of Internet destination prefixes, which are responsible for a large fraction of the traffic [1]. In this paper, we consider 30, 60 and 100 such popular destination prefixes for 3, 6 and

10-EP topologies respectively. In fact, each of them may not merely represent an individual prefix but also a group of distinct destination prefixes that have the same set of candidate EPs [12] in order to improve network and TE algorithm scalability. Hence, the number of prefixes we consider could actually represent an even larger value of actual prefixes. Without loss of generality, we assume that each EP has reachability to all the considered destination prefixes.

We generate synthetic traffic matrices for our evaluation. We generate inter-domain traffic from each ingress point towards each of the considered destination prefixes. Previous work has shown that inter-domain traffic is not uniformly distributed [11]. According to [12], the volume of inter-domain traffic demand is top-heavy and it can be approximated by Weibull distribution with the shape parameter equal to 0.2-0.3. We generate the inter-domain TM following this distribution with the shape parameter equal to 0.2. We remark that our TM generation process is just our best attempt to model inter-domain traffic, as no synthetic model for the actual behavior of traffic in real networks can be found in the literature.

5.2 Performance Matrices

The following performance metrics are used to evaluate the proposed strategies. For these metrics, lower values are better than high values.

- *NS maximum EP utilization*: this refers to $U_{max}(\varnothing)$.
- *Average of maximum EP utilization across all FSs*: this refers to U_{Ave}^{FS}.
- *Percentage of the average disrupted traffic volume*: a traffic flow is disrupted if it is shifted to another EP when a failure occurs. We denote the volume of disrupted traffic under FS s by DT_s and the average of disrupted traffic volume by *AveDT*. The percentage of the average disrupted traffic volume (*PerAveDT*) is the ratio of the average disrupted traffic volume to the total traffic volume ($|T|$):

$$PerAveDT = \frac{AveDT}{|T|} \times 100 = \frac{\sum_{s=1}^{|S|} DT_s \Big/ |S|-1}{|T|} \times 100 \quad \text{where } |T| = \sum_{k \in K} \sum_{i \in I} t(k,i) \quad (10)$$

6   Simulation Results

6.1 Evaluation of Normal State Maximum EP Utilization

Figure 3(a) shows the NS maximum EP utilization achieved by different strategies for the 3-EP topology. The x-axis represents the normalized average utilization which is the total traffic volume normalized by the total capacities of all EPs. All the simulation results presented in this paper are the average of 20 trials.

First of all, we can observe from the figure that the performances achieved by RANDOMR, GLOBALR and GREEDYR are identical. This phenomenon is expected since they use the same algorithm (OPTIMAL-AWARE HEURISTIC) for their PEPs selection. The OPTIMAL-AWARE HEURISTIC produces near-optimal performance that is only within 1%-
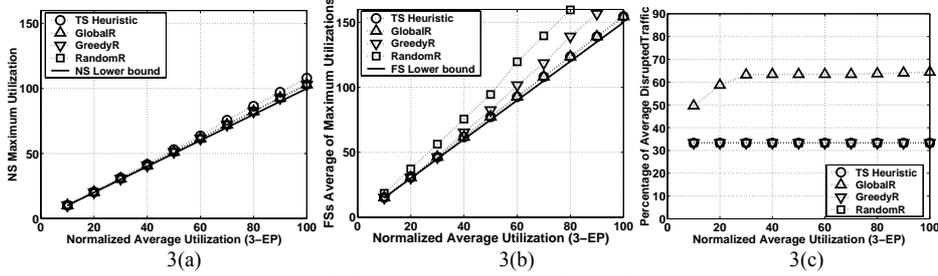
3(a) 3(b) 3(c)

Figure 3. Performance evaluation for 3-EP topology
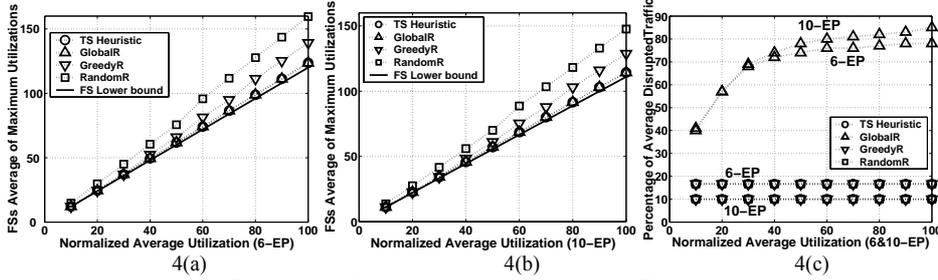

4(a) 4(b) 4(c)

Figure 4. Performance evaluation for 6 and 10-EP topology

3% from the NS lower bound[7]. On the other hand, however, the TS heuristic has slightly higher maximum EP utilization than the others (about 1%-5% compared to **GLOBALR** and 2%-8% compared to the NS lower bound). This can be explained by the reason that the TS heuristic attempts to minimize the maximum EP utilization under both NS and FSs simultaneously, as shown by objective function (5). Since the two objectives do not coincide, there is a performance trade-off between them. Nevertheless, as will be shown next, the TS heuristic significantly improves the performance across FSs at the cost of only a small performance degradation under NS.

6.2 Evaluation of the Average Maximum EP Utilization across all Failure States

Figure 3(b)[8] shows the average of maximum EP utilization across all FSs achieved by different strategies for the 3-EP topology. The figure shows that the TS heuristic and **GLOBALR** have similar results and are within 1%-3% of the FS lower bound. The reason is that both strategies have given attention to optimizing the performance under FSs (i.e. by using objective function (5) for the TS heuristic and EP selection re-computation under each FS for the **GLOBALR**). On the other hand, for the **GREEDYR**, the performance degrades 1%-13% from the **GLOBALR** and the TS heuristic and 2%-16% from the FS lower bound. This performance degradation is expected since the **GREEDYR** only considers minimizing EP utilization under FSs as the second optimization objective. In other words, the performance objectives under NS and FSs are optimized in a lexicographic importance order. As a result, the performance as measured by maximum EP utilization under FSs are not truly optimized: the solution of the PEP may not be a good input for **GREEDYR** to produce the optimal SEP. In addition, the **GREEDYR** performance starts to degrade compared to the TS heuristic and **GLOBALR** when the normalized average utilization (x-axis) exceeds 30%. In fact, with lower normalized average utilization, the EPs comfortably have extra capacity to accommodate the other traffic flows assigned by the **GREEDYR** and keep the utilization balanced.

---

7. Obviously the NS lower bound curve is linear because it is equal to the normalized average EP utilization.
8. For completeness we show performance results for two scenarios, one for below 100% maximum utilization and the other one for over 100% maximum utilization.

However, as the normalized average utilization increases, the residual capacity of EPs reduces and the PEP solution restricts the ability of the GREEDYR to reassign the prefixes of flows from the failed EP. Finally the RANDOMR has dramatic performance degradation, being about 22%-30% worsen than the GLOBALR and the TS heuristic. This performance degradation is primary due to the random SEP selection, which does not optimize any performance objective.

6.3 Evaluation of the Average Disrupted Traffic Volume

Figure 3(c) presents the percentage of the average disrupted traffic volume for the 3-EP topology. The figure shows that the TS heuristic, RANDOMR and GREEDYR have identical and constant performance as the normalized average utilization increases. This is due to the fact that with all these strategies, only the traffic on the failed EP is shifted. With the 3-EP topology, the number of FSs is 3 and this results in minimum (1/3)*100=33% traffic disruption for any single EP failure. However, since the GLOBALR performs network-wide recomputation for any single EP failure, both the affected and unaffected destination prefixes are likely to be reassigned, thereby causing significant traffic disruption in particular when the normalized average utilization is high. The figure shows that the average disrupted traffic volume for the TS heuristic, RANDOMR and GREEDYR are 33%-48% better than the GLOBALR.

6.4 Evaluation of Larger Topologies

We also performed our evaluation on larger topologies with 6 and 10 EPs. We note that the result patterns of NS maximum EP utilization for the 6 and 10-EP topologies are similar to those in Figure 3(a), hence, we proceed our performance analysis in a similar fashion to that in Section 6.1. Figures 4(a) and (b) present the average of maximum EP utilization across all the FSs for 6 and 10-EP topologies respectively. We observe that the higher the total number of EPs the lower the average of maximum EP utilization. As with the 3-EP topology, we can reach a conclusion: the TS heuristic always performs as well as the GLOBALR and better than the others and is very closer to the FS lower bound. This shows that the performance achieved by the TS heuristic scales well for larger topologies.

Figure 4(c) presents the percentage of average disrupted traffic volume for the 6 and 10-EP topologies together. By comparing Figure 3(c) with 4(c), we observe that, as the number of EPs increases, the percentage of the average disrupted traffic volume for the TS heuristic, RANDOMR and GREEDYR decreases. Conversely, as the number of EPs increases, this performance metric for the GLOBALR increases. This is attributed to the fact that, by increasing the number of EPs, the solution spaces for the GLOBALR's prefix reassignment is greatly enlarged. As a result, the likelihood that the prefixes changes from the originally assigned EPs to other EPs increases.

6.5 Overall Performance

In summary, our proposed TS heuristic regarding the average of maximum EP utilization across FSs performs (1) as well as the GLOBALR, (2) almost as well as the FS lower bound, (3) better than the GREEDYR when the normalized average utilization exceeds 30%, and (4) always significantly better than the RANDOMR. The excellent

performance of the TS heuristic under FS is only at the cost of a small performance degradation in the NS maximum EP utilization compared to the other strategies. The TS heuristic also keeps the traffic disruption to a minimum. Hence, overall, it can be regarded as the best among all the strategies.

The GLOBALR performs almost as well as the NS and FS lower bounds. However, it causes very large traffic disruption which leads to frequent BGP configuration changes and route instability. Hence, it is an impractical strategy. The GREEDYR performs as well as the GLOBALR and almost as well as the NS lower bound but it has significant performance degradation in the average of maximum EP utilization across all FSs as the normalized average utilization increases. Finally the RANDOMR is the worst performer in the average of maximum EP utilization across all FSs, which makes it inappropriate for robust TE.

7 Conclusion

In this paper, we have proposed a bi-level outbound TE optimization approach to make outbound TE robust to EP failures. This approach determines for each destination prefix the best PEP and the SEP (the next best EP) upon PEP failure. The optimization objectives are to minimize the maximum EP utilization under NS and the average of maximum EP utilization across all FSs simultaneously. We have proposed a tabu search heuristic to solve the problem and compared its performance to three alternative approaches. Our simulation results show that the tabu search heuristic significantly reduces the average of maximum EP utilization across all the FSs at a cost of only small increases in the NS maximum EP utilization. It also keeps the traffic disruption to a minimum. The other alternative approaches, however, do not achieve all these objectives together. We believe that our work provides insights to network operators on how to make optimal BGP outbound route selection robust to inter-domain EP failures which are common events and transient in nature.

References

1. N. Feamster et al.: Guidelines for Interdomain Traffic Engineering, *ACM CCR*, 2003.
2. B. Bressound et al.: Optimal Configuration for BGP Route Selection, *INFOCOM*, 2003.
3. K. Ho et al.: Multi-objective Egress Router Selection Policies for Inter-domain Traffic with Bandwidth Guarantees, *IFIP Networking*, 2004.
4. O. Bonaventure et al.: Achieving Sub-50 Milliseconds Recovery Upon BGP Peering Link Failures, *ACM CONEXT*, 2005.
5. B. Fortz et al.: Optimizing OSPF/IS-IS Weights in a Changing World, *IEEE JSAC*, 2002.
6. A. Nucci et al.: IGP Link Weight Assignment for Transit Link Failures, *ITC*, 2003.
7. A. Sridharan et al.: Making IGP Routing Robust to Link Failures, *IFIP Networking*, 2005.
8. F. Golver et al. *Tabu Search*. Kluwer Academic Publisher, Norwell MA 1997.
9. T. Telkamp: Traffic Characteristics and Network Planning, *NANOG 2002*.
10. J.L. Cohon. *Multiobjective Programming and Planning*. Academic Press, New York 1978.
11. W. Fang et al.: Inter-AS Traffic Patterns and their Implications, *IEEE GLOBECOM*, 1998.
12. A. Broido et al.: Their Shares: Diversity and Disparity in IP Traffic, *PAM*, 2004.