

Geodiverse Routing with Path Delay and Skew Requirement under Area-Based Challenges

Yufei Cheng^{a,*}, Deep Medhi^b, James P.G. Sterbenz^{a,c,d}

^aInformation and Telecommunication Technology Center, The University of Kansas, Lawrence, KS 66045, USA

^bSchool of Computing and Engineering, University of Missouri-Kansas City, MO 64110, USA

^cSchool of Computing and Communications, Lancaster University, Lancaster LA1 4WA, UK

^dComputing Department, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

Abstract

With the increasing frequency of natural disasters and intentional attacks that challenge communication networks, vulnerability to cascading and regional-correlated challenges is escalating. Given the high complexity and large traffic load of communication networks, these correlated challenges cause substantial damage to reliable network communication. In this work, we extend the GeoDivRP routing protocol to consider delay-skew requirement when using multiple geographically diverse paths for telecommunication networks under area-based challenges. We present a flow-diverse minimum-cost routing multicommodity flow problem. Furthermore, we present a nonlinear delay-skew optimization problem to balance between delay and traffic skew on paths. We investigate the tradeoff between the delay and skew in choosing multiple geodiverse paths. We implement GeoDivRP in *ns-3* to employ the optimized paths given by the two optimization solutions and demonstrate their effectiveness compared to OSPF Equal-Cost Multi-Path routing (ECMP) in terms of overall link utilization. It guarantees the delay-skew constraint provided by the upper layer while satisfies the traffic demand imposed by multiple routing commodities in the telecommunication networks.

Keywords:

network resilience and survivability; physical network topology diversity; path delay and skew optimization; area-based network disasters; multipath geographic routing heuristics; cross-layer routing protocol;

1. Introduction and motivation

Survivability of communication networks under random link and non-correlated failures has been a popular research domain [1, 2]. Recently, the research community has become more concerned about the potential damage caused by large-scale challenges and intentional attacks; efficient mechanisms have been proposed to mitigate their impacts [3, 4, 5, 6]. However, none of these works considers traffic allocation for regional challenges or attacks with a large impact zone, i.e., an earthquake or hurricane can have a challenged radius of up to 500 miles, which can cause failed nodes and links in the vicinity with substantial damage to the normal network communications [7].

It has been observed that a large number of failures in a geographical region can result in catastrophic damage to network communications [6]. When regional challenges or attacks occur, a series of nodes and links in the vicinity can be damaged and removed from the network; these are geographically correlated challenges. Since the challenge effect is frequently long-term [7], a set of backup paths are required for survivable routing. The single location physical challenge scenario has

been analyzed [3, 5, 8], while physical challenges of correlated and simultaneous challenges have been discussed [9]. A random line-cut mechanism has been used to assess the vulnerability to regional-based challenges [4]. Both correlated failures and targeted attacks with simulation results have been modeled [10]. Our previous work has studied different vulnerability area identification mechanisms and routing algorithms to route around the impact zone with a provided threat model [6]; two heuristics were proposed for solving the d -distance separation paths (in which any two nodes on disjoint paths are separated by greater than d distance) problem and demonstrated its effectiveness under regional challenges [11]. However, traffic allocation and delay-skew minimization have not been considered; it is important to understand the mechanism to statistically direct the rerouted traffic onto multiple d -distance separated paths and to better cope with network congestion when large-scale challenges occur.

Multipath routing has been accepted to be advantageous for small networks for the all-commodity traffic scenario. ECMP (Equal-Cost Multipath) is proposed as a multipath routing strategy, which uses equal cost multiple paths for better load-balancing in OSPF (Open Shortest Path First) [12]. Optimization has been done to maximize the flow on each path in an ECMP routing algorithm [13]. Another optimization problem has been formulated by a weighted multipath routing based on ECMP, and its objective function is to minimize the maximum link utilization [14].

*Corresponding author.

Email addresses: yfcheng@ittc.ku.edu (Yufei Cheng),

DMedhi@umkc.edu (Deep Medhi),

jjgs@{ittc.ku.edu, comp.lancs.ac.uk, comp.polyu.edu.hk} (James P.G. Sterbenz)

However, the multipath gain diminishes as the network becomes large [15]. A distributed traffic engineering heuristic, TeXCP, has been proposed and uses four paths for each demand [16]; however, a near-optimal solution may have contributed to the conclusion since at optimality more than one path is rarely needed at any instant as shown in [15] for large networks. An optimization problem has been formulated to model the routing issues in a multi-source-destination multipath routing environment, and it leads to a pseudo-polynomial algorithm based on linear programming in the network with a bounded buffer size and jitter constraint [17, 18]. A multipath flow optimization problem has been formulated with two objectives, total link utilization and bandwidth fairness, and has been solved with a nonlinear programming solver [19]. However, most previous work has focused on multipath routing without challenges. With increasing importance of network resilience under large-scale challenges or attacks, it is imperative to analyze multipath routing efficiency and understand the traffic allocation requirements under these challenges.

Flow-diverse routing mechanism has been proposed to solve the optical network diversity problem. Shared Risk Link Group (SRLG) is a set of links that share a common physical resource, and it has been proposed to address single or multiple physical failures [20]. Minimum-cost diverse SRLG routing is proven NP-complete and an integer linear programming formulation is used to solve the routing problem [21]. Path protection has been proposed to provide two SRLG-disjoint paths using graph transformation techniques [22]. Furthermore, an integer nonlinear programming (INLP) has been proposed to solve the problem of finding two disjoint paths with minimum joint path failure probability in the face of probabilistic physical failures [23]. However, most of the work focused on diverse routing in optical networks with two-diverse path calculation. We extend the flow-diverse routing mechanism into a generic network with three- or more-diverse path calculation considering the geodiversity concept introduced in Section 2.

In this work, we formulate two optimization problems in physical networks under regional challenges to either minimize the traffic cost or the delay-skew of the multiple paths calculated for each node pair. Skew is the difference in time delay across multiple paths. The optimization solution provides better link utilization compared to OSPF with ECMP. With the optimized geodiverse paths from the iWPSP (iterative Way-Point Shortest Path) heuristic [11], our GeoDivRP routing protocol [24, 25, 11, 6] improves the overall link utilization compared to ECMP under large-scale network challenges. Our heuristic does not restrict the maximum path length since it may lead to no usable skew-bounded paths.

We do, however, introduce a trade-off parameter δ to control the path stretch, outage risk, and skews among multiple paths calculated for the same node pair. It balances between short path stretch with high outage risk and the long path stretch with low outage risk. Furthermore, it controls the skew value between multiple geodiverse paths, which is achieved by controlling the d -distance separated paths provided by the GeoDivRP using the iWPSP heuristic. In controlling the delay variation or skew, δ can be either increased or decreased to provide paths

with the required skew value and route around the challenged area.

For various applications, the requirement for path delay or skew is different. For example, data traffic is more sensitive to delay while multimedia traffic is more so to skew. ResTP [26, 27, 28] determines the best combination of delay and skew for a specific application and passes that information down to GeoDivRP through the $[h, t]$ requirement tuple, where h is the desired stretch limit in number of additional hops, and t is the skew target. GeoDivRP calculates the geographically diverse path sets that satisfy the delay-skew requirement using the nonlinear optimization algorithm if permitted. Otherwise, GeoDivRP provides the best path sets returned by the optimization process.

It is rarely feasible to conduct network experiments on a production network, especially at a national scale. Network researchers resort to simulations to study their ideas and proposals. In this paper, we use *ns-3* [29] simulation software to study our protocol. As for traffic optimization, we use the OPENOPT optimization toolkit [30] for solving the two optimization problems and use real-world network topologies from KU TopView [31, 32]. The same physical challenge can cause different damage levels to the network if it occurs at different locations; therefore, we choose the failure regions identified from the previous work [6].

We extend our GeoDivRP routing algorithm to provide d -distance separated paths as well as the optimal traffic allocation information on the multiple paths for all the source-destination node pairs or commodities¹. We formulate a minimum cost routing problem using a linear programming (LP) model and a delay-skew minimization routing problem using a nonlinear programming (NLP) model. The paths for both of the problems are provided by a modified iWPSP routing heuristic explained in Section 2. When the network is under regional challenges, the rerouted traffic has a limited number of backup paths to select from, which raises the potential danger for the network to get congested. The congestion will further cause higher end-to-end delay. We consider the problem of establishing multiple bounded delay-skew geodiverse paths with a given demand matrix when the challenge occurs. We have formulated both of our problems as multicommodity flow problems and solved them.

In the following sections, we introduce our optimization models and problem formulations in Section 2. We present our model implementation details and simulation results in Section 3. Section 4 concludes the paper and suggests future work.

2. GeoDivRP-based multicommodity flow approach

We start our multicommodity optimization discussion with our geodiverse path generation algorithm using our GeoDivRP routing protocol.

2.1. GeoDivRP and geodiverse path generation

We first start with a brief description of GeoDivRP, which fits in the protocol stack as shown in Figure 3 [11]. Con-

¹We use source-destination node pair and commodity interchangeably

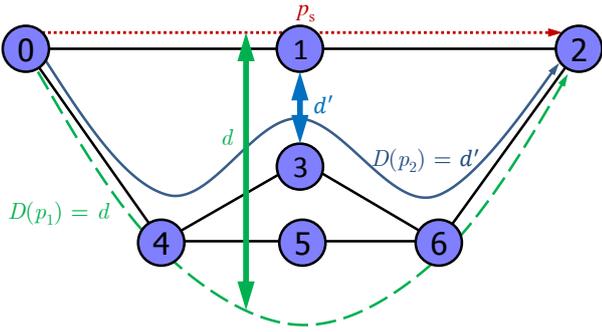


Figure 1: Geographical diversity: distance d

sider *Knobs* \mathbb{K} that are used by higher layers to influence lower layer operation while *dials* \mathbb{D} are the mechanisms for lower layers to provide feedback to higher layers [33]. The application layer passes a service specification and threat model down to our resilient transport layer protocol ResTP (resilient transport protocol) [26, 27, 28]. Upon receiving these parameters, ResTP determines the type of transport service needed (including error control and multipath characteristics) and requests that GeoDivRP calculate geodiverse paths that meet the requirement tuple $(k, d, [h, t])$, where k is the total number of geodiverse paths requested, d is the distance separation criteria, $[h, t]$ are the desired constraints on path stretch h (number of additional hops for diverse paths) and the temporal skew (delay difference) across paths, t . ResTP then establishes a multiflow with error control needed to meet the service spec, including the per-subflow error control (ARQ, hybrid ARQ, FEC, or none) and flow bundle (e.g., 2-of-3 erasure code for real-time critical service or 1+1 redundancy with a hot-standby for delay and loss tolerant service) taking advantage of k d -geodiverse paths $P = p_1 \dots p_k$ provided by GeoDivRP.

All the candidate paths provided to the optimization problems (presented below) are geodiverse. Geographical diversity $D(p_a)$ such that $D \geq d$ is defined as the minimum distance between any node members of vector p_a and that of the shortest path p_s . Consider Figure 1 in which node v_0 is the source and node v_2 is the destination node. The red dotted line shows the shortest path p_s consists of nodes 0–1–2. The green dashed line shows path p_1 and its geodiversity $D(p_1)$ (with respect to p_s) d . The blue solid line shows path p_2 and its geodiversity $D(p_2)$ d' since the minimum distance is d' between node 1 and node 3. There are three paths in total for the commodity (v_0, v_2) in Figure 1. If the requested geodiversity is between d' and d , the returned path set includes p_s and p_1 .

All the geodiverse candidate paths for the multicommodity optimization problem are provided by the modified iWPSP routing heuristic shown in Algorithm 1. The paths returned from iWPSP are all simple paths, with δ controlling the path skew for different geodiverse paths. The skew constraint t is passed down along with the other parameters from ResTP. As shown in Figure 2, when $k = 2$, iWPSP first selects neighbor nodes v_{s_1} and v_{d_1} that are d -distance separated from source node v_s and destination node v_d , respectively (for simplicity in this

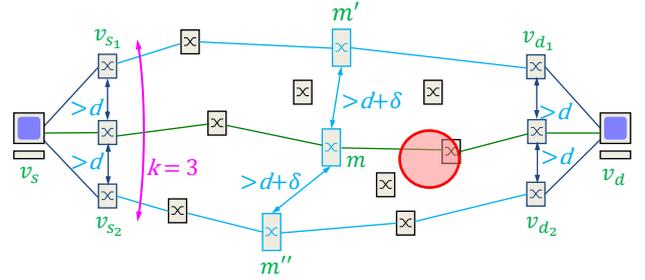


Figure 2: Iterative waypoint shortest path heuristic

presentation we assume that such nodes exist; otherwise the nodes with the greatest distance will be chosen, iterating until nodes d apart are located). Assuming the shortest path connecting v_s and v_d is p_s , iWPSP selects waypoint nodes m' and m'' in the opposite direction that are distance $d + \delta$ apart from the middle node m in the shortest path, where the segment $m'mm''$ intersects the shortest path. Dijkstra's algorithm is performed for the two branches $v_{s_1}m'$ and $v_{d_1}m'$. By connecting the shortest path returned from the two branches, the heuristic obtains the first geodiverse path. The same mechanism repeats for waypoint node m'' for the second geodiverse path. The variable d is a user-chosen parameter based on the threat model, and δ is experimentally chosen for different network topologies to increase the probability of the heuristic to return a d -separated path. The δ parameter is also useful in preventing the links of the two geodiverse paths from interleaving and creating routing loops. By tweaking the value of δ , the heuristic can select a nearby waypoint node if the previous one fails running Dijkstra's algorithm. When the heuristic cannot select paths within the skew bound t , the model increases or decreases δ accordingly. The pseudo code of iWPSP is shown in Algorithm 1.

This heuristic naturally affects the skew for different paths in different commodities with the introduction of δ . By slightly increasing or decreasing δ along each direction of the path calculation, we can indirectly alter the skew value of the returned geodiverse paths. If the returned path set is not bounded by the provided skew requirement, iWPSP uses a different δ value to calculate another set.

2.2. Flow-diverse optimization with minimum cost objective

The problem formulation discussed below is executed in the optimization engine shown in Figure 4. The formulation is based the link-path approach [34, 35] for multicommodity flows. It incorporates the geodiverse candidate paths provided by GeoDivRP discussed above for each commodity in the multicommodity flow formulation.

Formally, a network is represented by a connected directed graph $G(V, E)$, where V is the set of nodes (vertices) and E is the set of links (edges), where each edge allowed to have the maximum flow $u_e, e \in E$ (for example, due to capacity), and where there are W commodities defined by $W^w = (s^w, t^w, h^w)$, where s^w and t^w are the source and destination of commodity w in the graph $G(V, E)$, and h^w is their traffic demand.

Functions:

Calculate k number of geographically d -distance separated skew-bounded paths

Input:

v_{s_k} := source node
 v_{d_k} := destination node
 δ := delta distance when selecting waypoint node
 $[k, d, [h, t]]$:= requirement tuple

begin

shortest path p_s connecting v_s and v_d , with its middle point m ;
choose neighbor node v_{s_k}, v_{d_k} that is at least d distance from $v_{s_{k-1}}, v_{d_{k-1}}$, respectively;

if k is odd then

choose two nodes m_1 and m_2 that are separated by $d + \delta$ on each direction of S , where $m_1 m m_2$ is perpendicular bisector of S ;
 $p_1 = \text{SourceTree}_{v_d v_s} \leftarrow \text{Dijkstra}(v_d, v_s)$;
 $k- = 3$;

else

choose two nodes m_1 and m_2 that are separated by $d/2 + \delta$ on each direction of p_s , where $m_1 m m_2$ is perpendicular bisector of S ;
 $k- = 2$;

end

$p_{m_1 v_{s_1}} = \text{SourceTree}_{v_{s_1} m_1} \leftarrow \text{Dijkstra}(m_1, v_{s_1})$;

$p_{m_2 v_{s_2}} = \text{SourceTree}_{v_{s_2} m_2} \leftarrow \text{Dijkstra}(m_2, v_{s_2})$;

$p_{m_1 v_{d_1}} = \text{SourceTree}_{v_{d_1} m_1} \leftarrow \text{Dijkstra}(m_1, v_{d_1})$;

$p_{m_2 v_{d_2}} = \text{SourceTree}_{v_{d_2} m_2} \leftarrow \text{Dijkstra}(m_2, v_{d_2})$;

while $k > 0$ do

shortest path $p_s =$ newest established path;
choose one node m_k that is separated by distance $d + \delta$ from p_s on the farther direction from the absolute shortest path;

$p_{m_k v_{s_k}} = \text{SourceTree}_{m_k v_{s_k}} \leftarrow \text{Dijkstra}(m_k, v_{s_k})$;

$p_{m_k v_{d_k}} = \text{SourceTree}_{m_k v_{d_k}} \leftarrow \text{Dijkstra}(m_k, v_{d_k})$;

$k- = 1$;

end

if k is odd then

$p_2 = p_{m_1 v_{s_1}} + p_{m_1 v_{d_1}}$;

$p_3 = p_{m_2 v_{s_2}} + p_{m_2 v_{d_2}}$;

...

$p_k = p_{m_{k-1} v_{s_{k-1}}} + p_{m_{k-1} v_{d_{k-1}}}$;

remove path that fails the skew requirement.;

else

$p_1 = p_{m_1 v_{s_1}} + p_{m_1 v_{d_1}}$;

$p_2 = p_{m_2 v_{s_2}} + p_{m_2 v_{d_2}}$;

...

$p_k = p_{m_k v_{s_k}} + p_{m_k v_{d_k}}$;

remove path that fails the skew requirement.;

end

return (p_1, p_2, \dots, p_k)

end

Algorithm 1: Iterative waypoint shortest path heuristic

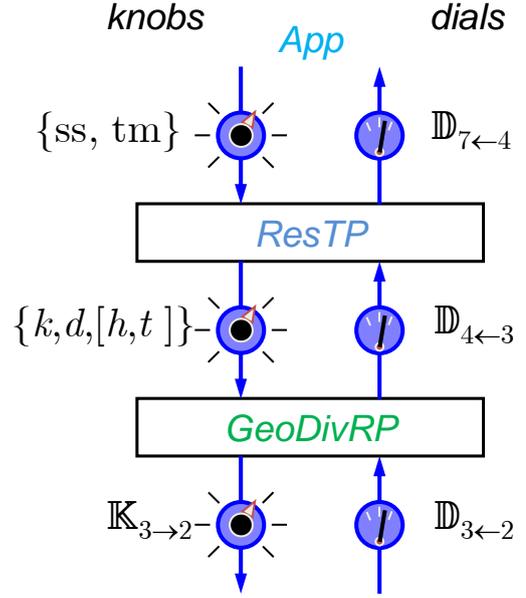


Figure 3: Layered protocol diagram of GeoDivRP and ResTP

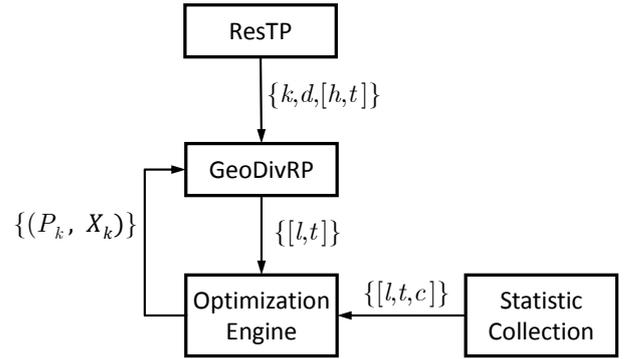


Figure 4: Block diagram of GeoDivRP and optimization engine

A network is represented by a connected directed graph $G(V, E)$, where V is the set of nodes (vertices) and E is the set of links (edges). Each path p for commodity w (that is generated by the procedure described above) has an associated cost $c^w(p)$ that denotes its cost per unit flow.

For each commodity w , let P^w denote the collection of all GeoDivRP-compatible paths from the source node s^w to the destination node t^w . We use variable x_p^w as the flow on path p for commodity w . Link-path indicator variable is defined as $\eta_e(p)$; it is one if link(e) is contained in the path p , and is zero otherwise. We list the important variables used in this optimization model in Table 1.

The GeoDivRP multicommodity linear optimization problem can now be stated as follows:

$$\min \sum_{w \in W} \sum_{p \in P^w} c^w(p) x_p \quad (1)$$

Table 1: Notations used in Problem Formulation

	Description
$G(V, E)$	input graph $ G $ with a set of nodes V
P^w	candidate number of paths to be considered for commodity w
$c^w(p)$	cost per unit flow on path p for commodity w
u_e	flow upper bound on link e
v_s	source node
v_d	destination node
v_{s_k}	neighbor node chosen by source node
v_{d_k}	neighbor node chosen by destination node
k^w	number of geodiverse path requested for commodity w
d	distance separation between each and every node in different disjoint paths
δ	delta distance when selecting waypoint node
h^w	demand for commodity w
$\eta_e(p)$	=1 if link e belongs to path p ($p \in P^w$); 0, otherwise
x_p	traffic flow on path p for commodity w (variable)
y_e	link flow on link e (dependent on x_p)

subject to

$$\sum_{p \in P^w} x_p = h^w, \quad w \in W \quad (2)$$

$$\sum_{w \in W} \sum_{p \in P^w} \eta_e(p) x_p \leq u_e, \quad e \in E \quad (3)$$

$$x_p \geq h^w / k^w, \quad p \in P^w, \quad w \in W. \quad (4)$$

The objective function shown in (1) minimizes the overall cost of flows over different paths for all the commodities. (2) is the flow conservation over all paths $p \in P^w$ of traffic demand h^w for each commodity w . (3) is the link capacity constraint for each link e requiring that the sum of the path flows passing through that link is at most at its capacity upper bound u_e . (4) requires all path flow variables to be greater than or equal to a minimum path flow for traffic diversity, captured by the total traffic demand divided by the minimum number of geodiverse paths k^w to be considered for each commodity w . Note that $k^w \leq \#(P^w)$ and typically, $k^w < \#(P^w)$ (otherwise, the flow will be equally distributed along all the paths for a commodity). Clearly, (4) forces multipath flow, an important requirement for our GeoDivRP approach.

2.3. Flow-diverse Optimization with Delay-Skew Objective

The minimum cost optimization presented above provides the optimum traffic allocation ratio on a required number of diverse paths for each of the commodities while targeting to minimize the overall network cost. However, it does not have a direct control over the path delay or skew; therefore, the above optimization model cannot guarantee on the path delay or skew requirement passed from ResTP. Thus, we propose another formulation that considers both path delay and skew as an weighted objective, which enhances the above optimization model. Furthermore, this enhanced model also allows us to demonstrate the difference between the optimization solutions

focused on the path delay as opposed to the one based on the skew. In other words, it provides a flexible way to manage the weight on either delay or skew depending on the application scenario.

Given the capacity bound u_e on link e , we use the M/M/1 queuing model [36] that states the average packet delay on link e as

$$l_e = \frac{1}{u_e - y_e} \quad (5)$$

where $y_e = \sum_{w \in W} \sum_{p \in P^w} \eta_e(p) x_p$ is the link flow on link e . Then, the average queueing delay l_p^w for path p for commodity w is the sum of the average queueing delay on each link given by

$$l_p^w = \sum_{e \in E} \eta_e(p) l_e \quad (6)$$

Therefore, the average end-to-end delay for a commodity w is given by:

$$l^w = \frac{1}{h^w} \sum_{p \in P^w} \sum_{e \in E} x_p \eta_e(p) l_e \quad (7)$$

Based on the delay for each path for commodity w , we formulate the path skew as:

$$t^w = \sum_{i \in I} |l_{p_s^w} - l_{p_i^w}| \quad (8)$$

where p_s^w is the shortest path for a commodity w , and p_i is the path set I that excludes p_s for that commodity. The overall path skew for all commodities is then given by

$$T = \sum_{w \in W} t^w \quad (9)$$

On the other hand, the total packet delay in the network [36] is given by

$$L = \sum_{e \in E} \frac{y_e}{u_e - y_e} \quad (10)$$

Based on the delay and skew, we formulate the optimization problem as follows:

$$\min \quad [(1 - \gamma)L + \gamma T] \quad (11)$$

subject to

$$\sum_{p \in P^w} x_p = h^w, \quad w \in W \quad (12)$$

$$\sum_{w \in W} \sum_{p \in P^w} \eta_e(p) x_p \leq u_e, \quad e = 1, 2, \dots, E \quad (13)$$

$$x_p \geq h^w / k^w, \quad p \in P^w, \quad w \in W. \quad (14)$$

The objective function in (11) targets minimizing the delay-skew with a tuning parameter γ ($0 \leq \gamma \leq 1$), which controls the weight on either delay or skew in the optimization process. The constraints are the same as the ones used in the minimum cost optimization discussed earlier.

2.4. Complexity analysis

We discuss complexity for both calculating d -distance separated paths as well as both of the optimization problems. iWPSP has a complexity of $2c^2n^2 \log n$, where c is the average number of neighbors for nodes, the complexity for choosing the waypoint node is $O(n)$, where n represents the number of nodes equals $|V|$, and $O(n \log n)$ is for Dijkstra's algorithm to calculate the two shortest paths. Therefore, the worst case scenario is $O(n^2 \log n)$ while the best case scenario is $O(n \log n)$. Most of the physical topologies have an average degree below four [37]. This means that c in our complexity analysis is a small constant. This reduces the best case time complexity of iWPSP to $O(n \log n)$.

The complexity for solving the flow-diverse linear optimization problem is polynomial. Therefore, the complexity of the GeoDivRP routing with minimum cost optimization is dominated by the complexity of geodiverse path calculation. On the other hand, the delay-skew optimization problem is a nonlinear optimization problem that are typically solved using an iterative process, and thus, cannot be directly analyzed from a complexity point of view. We can, however, comment on the cost of running such a problem. In our case, we used the *ralg* solver comes with the *OpenOpt* optimization framework [30]. The total number of variables for the delay-skew optimization problem is the number of commodities plus the number of links for each topology; it is represented as $n\text{Variables} = W + E$. The current implementation of *ralg* solver stores in memory a matrix of size $n\text{Variables}^2$, and each iteration consumes $5 \times n\text{Variables}^2$ multiplication operations. For example, when optimizing a network with 100 commodities and 100 links for a topology, the matrix size is $200 \times 200 = 40000$. Each iteration of the optimization has $5 \times 200^2 = 200000$ multiplication operations. We set the max-iteration of the solver as 1000, which means the worst-case complexity is 0.2×10^9 multiplication operations in total; this is too complex for large real-world networks.

3. Model implementation and simulation results

We use *ns-3* [29] to implement the GeoDivRP routing protocol. The geodiverse paths calculated from the routing protocol are passed to the optimization toolkit using *OpenOpt* optimization framework. After solving the optimization problems, the paths along with their flow allocation information (P_k, X_k) are returned to *ns-3* for network simulation. These optimized paths are used for the network data transmission to guarantee the traffic demand for all the commodities. This mechanism ensures that the paths can achieve the optimum link utilization.

The steps for the routing algorithm to calculate the geodiverse paths is shown as follows:

- Obtain the geodiverse paths using the iWPSP routing heuristic for each node pair that satisfies the skew constraint and d -distance separation criteria.
- Solve multicommodity flow optimization using the linear programming formulation (LP) or nonlinear programming formulation (NLP) for the flow-diverse minimum cost or delay-skew optimization, respectively.
- Use Geodiverse paths with flow allocation returned from the optimization for data transmission in *ns-3* network simulation.

3.1. Simulation results

We now present the multicommodity flow optimization for the flow-diverse minimum cost and delay-skew cases. Then we use these flow allocations to perform *ns-3* simulations over geodiverse paths. Solutions from both of our optimization problems are used for the physical network simulation under challenges.

3.1.1. Flow-diverse minimum cost optimization

In this study, we compare the performance of GeoDivRP-based multicommodity flow approach with OSPF ECMP. We use Level 3 [38] and Sprint [37] physical networks for this study. The capacity for all the links was set to 5 Gb/s, and we use CBR (constant bit rate) traffic, sent from each node to all the others at a data rate varying uniformly from 1 Mb/s to 12 Mb/s as the traffic demand. The varying demand in different networks is to evaluate and demonstrate the maximum traffic demand GeoDivRP can support. We use the challenged area at Kansas City identified in our vulnerability area identification mechanism [6] with a 300 km challenge range.

We record the time for solving the optimization problem in different physical topologies. We further include CORONET [39], Internet2 [40], and TeliaSonera [40] fiber-level networks. We calculate when the challenge occurs around Kansas City with a traffic demand of 10 Mb/s for each commodity. As shown in Table 2, the maximum time for the optimization is about 7 seconds for the Sprint network, while most of the others take less than 1 second. The evaluation is carried on a Linux machine with a 3.16 GHz Core 2 Duo CPU and 4 GB memory.

We further compare GeoDivRP to OSPF with ECMP in terms of the overall link congestion factor. Recall that the link

Table 2: Execution time for optimization algorithm

Network	Number of Nodes	Number of Links	Number of Failed Nodes	Optimization Time (s)
CORONET	39	63	2	0.62
Internet2	16	24	1	0.04
Level 3	63	94	4	2.06
Sprint	77	114	3	6.96
TeliaSonera	18	21	1	0.02

congestion factor is defined as the percentage of the bandwidth that has been used by the network flows. Our minimum cost optimization formulation is not specifically minimizing the link congestion factor; therefore, some links are still using up to 100% link capacity. However, since we specify the capacity upper bound on path flows, GeoDivRP uses the network resources efficiently and does not congest any network link. For OSPF with ECMP, on the other hand, the model always chooses the shortest path, and does not consider the remaining network resources on the network link, which causes congestion by overloading some network links. In the network simulation context, the extra traffic assigned to network links will either be dropped or queued if router buffers are used; traffic loss or delay will occur, respectively.

In Figure 5, we present the link congestion factor for the Level 3 network when the demand is 10 Mb/s for each node pair. GeoDivRP does not overload any link by distributing the traffic load among multiple paths. However, OSPF with ECMP has used some links up to 140%, which means that for each of the overloaded link, the data traffic for 40% of the link capacity will be dropped. Since the link capacity is 5 Gb/s, 2 Gb of traffic is dropped each second on these overloaded links. This causes significant traffic loss to the network communication, and it is especially damaging when the network is under large-scale challenges. The dropped traffic could have been buffered but the end-to-end delay would increase exponentially. Our delay-skew optimization targets at minimizing the path delay and skew and we present its result in the next subsection. The network congestion factor for GeoDivRP is 100% while that for OSPF with ECMP is 140%.

This link congestion analysis has demonstrated that GeoDivRP with flow-diverse minimum-cost optimization can allocate traffic to multiple paths efficiently and avoid overloading any network link; while OSPF with ECMP over-utilizes links which causes the data packets to either be dropped or buffered with increased end-to-end delay.

3.1.2. Delay-skew optimization

For the delay-skew optimization scenario, we set the link capacity at 500 Mb/s, with demand at 10 Mb/s. The total number of commodities is 9. We choose the source nodes in the west coast sending to destination nodes in the east coast. This way the paths calculated represent the highest delay scenario. The geodiverse paths provided by iWSP are the geodiverse

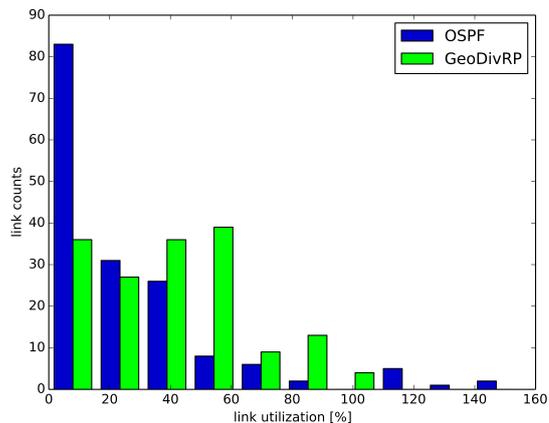


Figure 5: Link utilization in Level 3 network

paths calculated based on the current network topology with area-based challenges. The failure region is the same as that for the minimum-cost optimization: the center of the US around Kansas City. For each commodity, we calculate three geodiverse paths for optimization.

We use a standard nonlinear optimization framework, *OpenOpt* [30], to achieve a local optimum solution, which is an optimization framework using Python and can choose a range of nonlinear programming solvers to solve the nonlinear problem. We use the *ralg* heuristic solver that comes with *OpenOpt*; it is based on the *r*-algorithm with adaptive space dilation [41].

The topologies considered are the structural physical graphs [37] with their properties shown in Table 3. The number of nodes and links are in the same range, and the average node degree for all the topologies are between two and three.

We record the time for solving the optimization problem in different physical topologies for the delay-skew optimization, as shown in Table 4. All the physical topologies show a reasonable optimization time for both the single pair and nine node pair cases. For the nine-node-pairs case, it takes five seconds to solve the problem for Sprint, which is the maximum time among all the topologies since it is the largest one considered. The time for a single traffic pair are all below one second for all the physical topologies. This means that a distributed algorithm for delay-skew optimization is necessary for real-time computation, and we leave the detailed implementation of the

Table 3: Physical topology analysis

Network	Nodes	Links	Degree	Diameter	Radius	Path Length
CORONET	39	63	3.23	9	5	4.08
Internet2	16	24	3.00	6	3	2.63
Level 3	63	94	2.98	14	7	5.68
Sprint	77	114	2.96	16	9	6.47
TeliaSonera	18	21	2.33	7	5	3.58

Table 4: Time for delay-skew optimization algorithm

Network	Number of Nodes	Number of Links	Number of Failed Nodes	Number of Commodities	Single Pair Time (s)	Nine Pair Time (s)
CORONET	39	63	2	9	0.87	4.13
Internet2	16	24	1	9	0.51	3.62
Level 3	63	94	4	9	0.53	8.30
Sprint	77	114	3	9	0.81	5.04
TeliaSonera	18	21	1	9	0.52	3.63

distributed algorithm for future work.

We carry out simulations with varying traffic demand and study the largest demand that GeoDivRP can deliver in a given physical topology. We present the variation of delay and skew when the demand increases for the five physical topologies. We do not include the delay and skew result for OSPF ECMP; the network becomes congested with low demand and the delay becomes too large to present in the same plot with GeoDivRP; γ is set as zero for delay optimization. As shown in Figure 6a, the demand curves for all the topologies begins with a low value around 15 ms and increases slowly when the demand increases. However, when the demand increases beyond the *demand collapse point*, the delay starts increasing exponentially until the optimization cannot provide solutions. For example, if we consider the delay curve for the CORONET network, when the demand increases from 180 Mb/s to 190 Mb/s, the delay increases from 35 ms to over 200 ms, and the network becomes too congested to provide normal service beyond the demand collapse point, which is 190 Mb/s for this case. With the different demand collapse points for the topologies provided to ResTP, better flow allocation decisions can be made and the application can use network resources more efficiently.

In Figure 6b, we present the skew minimization result with physical topologies; γ is set as one to focus on skew optimization. For the low demand case, the skew decreases as the traffic load increases for the demand below 100 Mb/s; each link has low delay and the number of hops for each path in one commodity contributes more to the end-to-end delay. However, when the demand increases past 100 Mb/s, the link delay for the topologies except CORONET begins increasing exponentially. Therefore, the path skew increases exponentially as well.

We continue our simulation with the link congestion analysis with the five physical topologies. The link capacity is set as 500 Mb/s, and the demand is 50 Mb/s; the number of com-

modities is 100. The reason for the demand and number of commodity choice is to have a reasonable amount of traffic going through the network to better demonstrate the effectiveness of GeoDivRP. Before presenting the link utilization result, we present the physical topology of Sprint and Level 3 under the Kansas City area challenge.

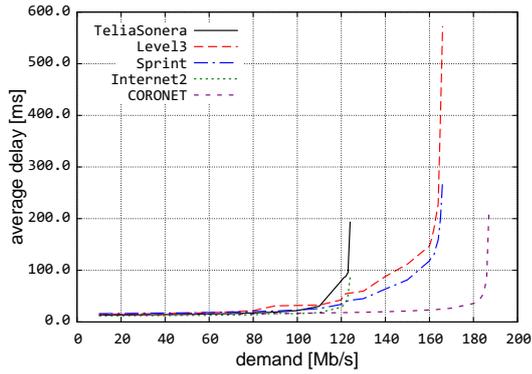
The Sprint physical network contains 77 nodes and 114 links. We use the same challenge scenario as our previous work [11, 6, 24]. The red circle shown in Figure 7a is the challenge area around Kansas City, and the green solid lines are the paths calculated by GeoDivRP.

As shown in Figure 8a, the x -axis presents the link utilization in percentage, and the y -axis shows the number of links with that link utilization level. For example, for 100% link utilization, OSPF with ECMP has five links with this link utilization, while the number for GeoDivRP is six.

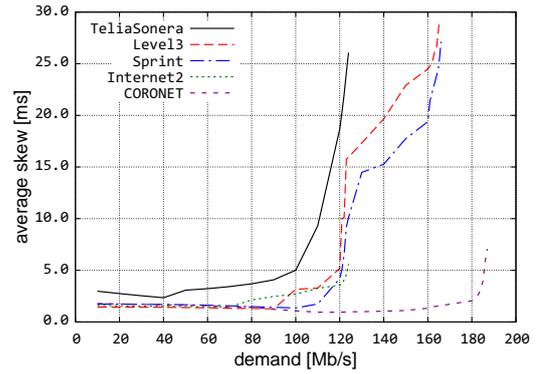
GeoDivRP guarantees that the link utilization for any link is not over 100% and keeps lower link usage whenever possible specified by the objective function from the formulation. On the other hand, OSPF with ECMP simply distributes network traffic among the calculated paths and can easily congest the network when the demand becomes larger. As shown in Figure 8a, OSPF with ECMP congests 6 links; although this is not a large percentage out of the 114 total links, they cause 85% of the commodities and 59% of the paths congested. On the other hand, GeoDivRP guarantees the optimized traffic allocation on all the commodities and presents great performance improvement.

The Level 3 physical network contains 63 nodes and 94 links. The same failure region in Kansas City has been used. As shown in Figure 7b, the challenge causes more damage to the overall connectivity because the Level 3 network lacks some of the nodes and links from Seattle to Chicago.

As shown in Figure 8b, GeoDivRP guarantees the link uti-

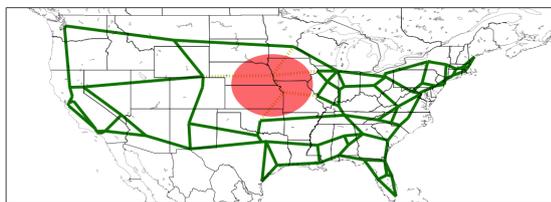


(a) Path delay with varying demand

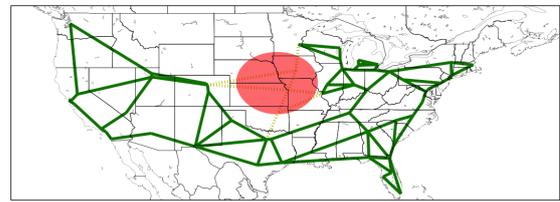


(b) Path skew with varying demand

Figure 6: Delay and skew with varying demand for different topologies

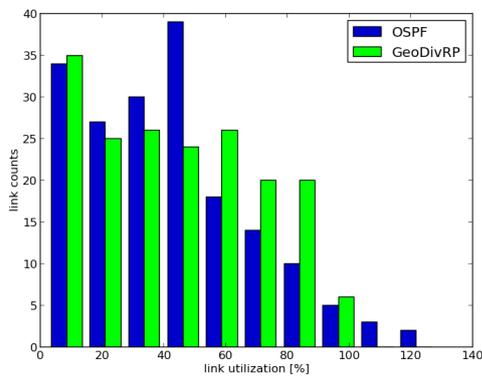


(a) Sprint topology under regional challenges

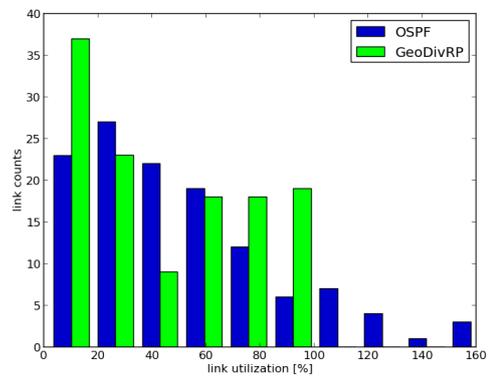


(b) Level 3 topology under regional challenges

Figure 7: Network topology under regional challenges



(a) Sprint network link utilization



(b) Level 3 network link utilization

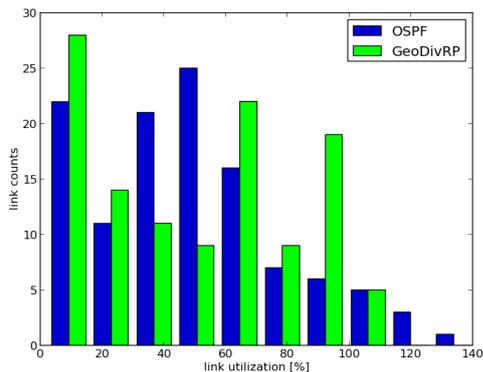
Figure 8: Link utilization of Sprint and Level 3 networks

lization is not over 100%, yet the usage for OSPF with ECMP goes to 160% and therefore greatly congests the network; there are 15 congested links out of 94. Similarly, although not a large percentage, these links cause 91% of commodities and over 71% of the paths congested. On the other hand, GeoDivRP avoids congestion by optimizing the traffic allocation on multiple paths of each commodity.

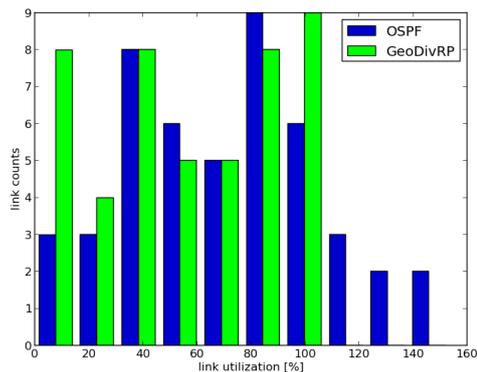
The other three topologies present similar results. Figure 9a presents the link utilization plot for the CORONET network. GeoDivRP statistically distributes the traffic load over the network while OSPF with ECMP congests multiple links. Figure 9b presents the link utilization plot for the Internet2 network. A majority of links are carrying less than 20% of the

traffic for GeoDivRP while OSPF with ECMP congests multiple links. Figure 10 presents the link utilization plot for the TeliaSonera network.

The objective function for the delay-skew optimization formulation is intended to balance the delay and skew in the optimization process through the tuning parameter γ . In Figure 11, we present the average delay and skew change for a single path with the varying γ value using the CORONET network for a single path. The results for the other networks present a similar trend and are not shown. The points on the plot are the γ values ranging from 0 to 1 with 0.1 step increment. The traffic demand and link capacity are 50 Mb/s and 500 Mb/s respectively. As we observe from the figure, when γ increases, the



(a) CORONET network link utilization



(b) Internet2 network link utilization

Figure 9: Link utilization of CORONET and Internet2 network

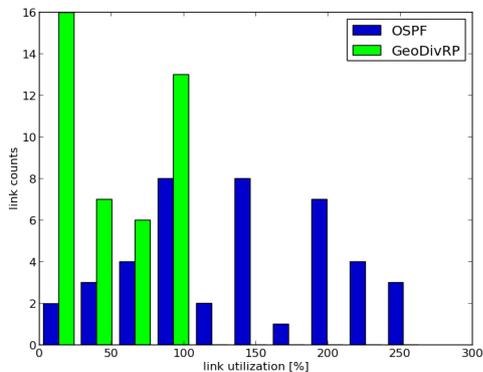


Figure 10: TeliaSonera network link utilization

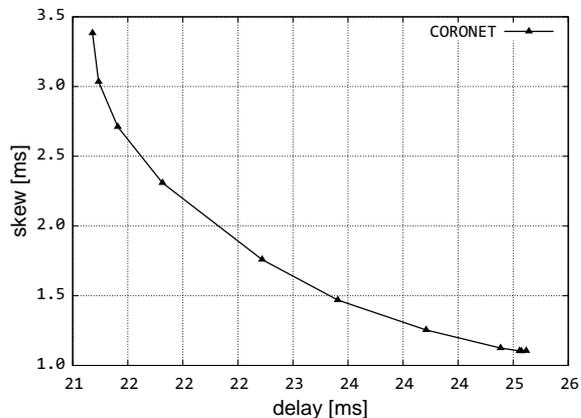


Figure 11: Delay and skew with varying γ for CORONET

average delay for each commodity increases while the average skew decreases. This means that delay and skew work against each other in this optimization process. Based on different application scenarios, we can select different γ for better network communication.

4. Conclusion and future work

We have evaluated the GeoDivRP routing protocol with minimum-cost and the delay-skew requirement. We have generated a linear programming (LP) and a nonlinear programming (NLP) formulation of the problems and successfully solved them. We have incorporated the optimized geodiverse paths in the GeoDivRP and have compared our protocol performance with OSPF ECMP in terms of overall link utilization. Our protocol shows considerably better performance than OSPF with ECMP.

We argue that GeoDivRP performs well in the face of large-scale challenges. First, the iWPSP routing heuristic returns d -distance separated paths with controlled algorithm and time complexity. Second, our previous research [11] presents improved packet delivery ratio (PDR) and delay when compared

to OSPF with ECMP, and this paper presents better link utilization. Finally, the delay-skew requirement guarantees the optimized traffic allocation among different paths and satisfies the delay-skew requirement tuple passed from the upper layer.

For future work, we plan to design a heuristic for calculating the best possible allocation result when the flow pattern is beyond optimization. We can also formulate additional link or additional capacity planning optimization problem. Furthermore, we plan to study different application scenarios and provide the delay-skew combination suggestion for each scenario. The delay-skew optimization becomes complex for large real-world networks, we plan to develop a distributed algorithm to support real-time network communications.

Acknowledgment

This is a significantly extended version and substantial revision of a paper that appeared in the 6th IEEE/IFIP International Workshop on Reliable Networks Design and Modeling (RNDM) 2014 [25]. The authors would like to thank the members of the ResiliNets group for discussions which led to this work. This research was supported in part by US NSF Grants

References

- [1] R. Cohen, K. Erez, D. ben Avraham, and S. Havlin, "Resilience of the Internet to Random Breakdowns," *Phys. Rev. Lett.*, vol. 85, pp. 4626–4628, November 2000.
- [2] D. Magoni, "Tearing down the Internet," *IEEE J.Sel. A. Commun.*, vol. 21, pp. 949–960, September 2006.
- [3] S. Neumayer, G. Zussman, R. Cohen, and E. Modiano, "Assessing the Impact of Geographically Correlated Network Failures," in *Military Communications Conference, 2008. MILCOM 2008. IEEE*, pp. 1–6, 2008.
- [4] S. Neumayer, G. Zussman, R. Cohen, and E. Modiano, "Assessing the Vulnerability of the Fiber Infrastructure to Disasters," *IEEE/ACM Transactions on Networking*, vol. 19, no. 6, pp. 1610–1623, 2011.
- [5] S. Neumayer and E. Modiano, "Network Reliability With Geographically Correlated Failures," in *Proc. of IEEE INFOCOM*, pp. 1–9, March 2010.
- [6] Y. Cheng, J. Li, and J. P. G. Sterbenz, "Path Geo-diversification: Design and Analysis," in *Proceedings of the 5th IEEE/IFIP International Workshop on Reliable Networks Design and Modeling (RNDM)*, (Almaty), September 2013.
- [7] M. J. F. Denise M. B. Masi, Eric E. Smith, "Understanding and Mitigating Catastrophic Disruption and Attack," *Sigma Journal*, pp. 16–22, September 2010.
- [8] W. Wu, B. Moran, J. Manton, and M. Zukerman, "Topology Design of Undersea Cables Considering Survivability Under Major Disasters," in *International Conference on WAINA*, pp. 1154–1159, May 2009.
- [9] P. K. Agarwal, A. Efrat, S. Ganjugunte, D. Hay, S. Sankararaman, and G. Zussman, "Network Vulnerability to Single, Multiple, and Probabilistic Physical Attacks," in *MILITARY COMMUNICATIONS CONFERENCE, 2010 - MILCOM 2010*, pp. 1824–1829, 2010.
- [10] E. K. Çetinkaya, D. Broyles, A. Dandekar, S. Srinivasan, and J. P. G. Sterbenz, "Modelling Communication Network Challenges for Future Internet Resilience, Survivability, and Disruption Tolerance: A Simulation-Based Approach," *Telecommunication Systems*, vol. 52, no. 2, pp. 751–766, 2013.
- [11] Y. Cheng, M. T. Gardner, J. Li, R. May, D. Medhi, and J. P. Sterbenz, "Optimised Heuristics for a Geodiverse Routing Protocol," in *Proceedings of the IEEE 10th International Workshop on the Design of Reliable Communication Networks (DRCN)*, (Ghent, Belgium), pp. 1–9, April 2014.
- [12] C. Hopps, "Analysis of an Equal-Cost Multi-Path Algorithm." RFC 2992 (Informational), Nov. 2000.
- [13] M. Dzida, M. Zagozdzon, M. Zotkiewicz, and M. Pioro, "Flow Optimization in IP Networks with Fast Proactive Recovery," in *Telecommunications Network Strategy and Planning Symposium, 2008. Networks 2008. The 13th International*, vol. Supplement, pp. 1–9, Sept 2008.
- [14] J. Zhang, K. Xi, L. Zhang, and H. Chao, "Optimizing Network Performance Using Weighted Multipath Routing," in *Computer Communications and Networks (ICCCN), 2012 21st International Conference on*, pp. 1–7, July 2012.
- [15] X. Liu, S. Mohanraj, M. Pioro, and D. Medhi, "Multipath Routing from a Traffic Engineering Perspective: How Beneficial Is It?," in *Network Protocols (ICNP), 2014 IEEE 22nd International Conference on*, pp. 143–154, Oct 2014.
- [16] S. Kandula, D. Katabi, B. Davie, and A. Charny, "Walking the Tightrope: Responsive Yet Stable Traffic Engineering," in *Proceedings of the 2005 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, SIGCOMM '05*, (New York, NY, USA), pp. 253–264, ACM, 2005.
- [17] T. Anjali, A. Fortin, and S. Kapoor, "Multi-Commodity Network Flows over Multipaths with Bounded Buffers," in *Communications (ICC), 2010 IEEE International Conference on*, pp. 1–5, May 2010.
- [18] T. Anjali, A. Fortin, G. Calinescu, S. Kapoor, N. Kirubanandan, and S. Tongngam, "Multipath Network Flows: Bounded Buffers and Jitter," in *INFOCOM, 2010 Proceedings IEEE*, pp. 1–7, March 2010.
- [19] T. Ishida, K. Ueda, and T. Yakoh, "Fairness and Utilization in Multipath Network Flow Optimization," in *Industrial Informatics, 2006 IEEE International Conference on*, pp. 1096–1101, Aug 2006.
- [20] D. Papadimitriou, F. Poppe, J. Jones, S. Venkatachalam, S. Dharanikota, R. Jain, R. Hartani, D. Griffith, and Y. Xue, "Inference of Shared Risk Link Groups." Internet Draft, May 2002.
- [21] J. Q. Hu, "Diverse routing in optical mesh networks," *Communications, IEEE Transactions on*, vol. 51, pp. 489–494, March 2003.
- [22] P. Datta and A. K. Somani, "Graph transformation approaches for diverse routing in shared risk resource group (srrg) failures," *Computer Networks*, vol. 52, no. 12, pp. 2381–2394, 2008.
- [23] H.-W. Lee, E. Modiano, and K. Lee, "Diverse Routing in Networks With Probabilistic Failures," *IEEE/ACM Transactions on Networking*, vol. 18, no. 6, pp. 1895–1907, 2010.
- [24] Y. Cheng, M. T. Gardner, J. Li, R. May, D. Medhi, and J. P. Sterbenz, "Analysing GeoPath Diversity and Improving Routing Performance in Optical Networks," *Computer Networks*, vol. 82, pp. 50–67, May 2015.
- [25] Y. Cheng and J. P. Sterbenz, "Geodiverse routing with path jitter requirement under regional challenges," in *Reliable Networks Design and Modeling (RNDM), 2014 6th International Workshop on*, pp. 179–186, Nov 2014.
- [26] J. P. Rohrer, A. Jabbar, E. K. Çetinkaya, and J. P. Sterbenz, "Airborne Telemetry Networks: Challenges and Solutions in the ANTP Suite," in *Proceedings of the IEEE Military Communications Conference (MILCOM)*, (San Jose, CA), pp. 74–79, November 2010.
- [27] J. P. Rohrer, R. Naidu, and J. P. G. Sterbenz, "Multipath at the Transport Layer: An End-to-End Resilience Mechanism," in *Proceedings of the IEEE/IFIP International Workshop on Reliable Networks Design and Modeling (RNDM)*, (St. Petersburg, Russia), pp. 1–7, October 2009.
- [28] T. A. N. Nguyen, J. P. Rohrer, and J. P. G. Sterbenz, "ResTP-A Transport Protocol for FI Resilience," in *Proceedings of the 10th International Conference on Future Internet Technologies (CFI)*, June 2015.
- [29] "The ns-3 Network Simulator." <http://www.nsnam.org>, July 2009.
- [30] "OpenOpt Optimization Framework." <http://openopt.org/> Welcome, October 2007.
- [31] J. P. Rohrer, M. J. F. Alenazi, and J. P. G. Sterbenz, "ResiliNets Topology Map Viewer." <http://www.ittc.ku.edu/resilinetmaps/>, January 2011.
- [32] J. P. Sterbenz, E. K. Çetinkaya, M. A. Hameed, A. Jabbar, Q. Shi, and J. P. Rohrer, "Evaluation of Network Resilience, Survivability, and Disruption Tolerance: Analysis, Topology Generation, Simulation, and Experimentation (invited paper)," *Telecommunication Systems*, vol. 52, no. 2, pp. 705–736, 2013.
- [33] J. P. G. Sterbenz, D. Hutchison, E. K. Çetinkaya, A. Jabbar, J. P. Rohrer, M. Schöller, and P. Smith, "Redundancy, Diversity, and Connectivity to Achieve Multilevel Network Resilience, Survivability, and Disruption Tolerance (invited paper)," *Telecommunication Systems*, vol. 56, no. 1, pp. 17–31, 2014.
- [34] M. Pioro and D. Medhi, *Routing, Flow, and Capacity Design in Communication and Computer Networks*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2004.
- [35] D. Medhi and K. Ramasamy, *Network Routing: Algorithms, Protocols, and Architectures*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2007.
- [36] L. Kleinrock, *Queueing Systems*, vol. II: Computer Applications. Wiley Interscience, 1976. (Published in Russian, 1979. Published in Japanese, 1979.).
- [37] E. K. Çetinkaya, M. J. F. Alenazi, Y. Cheng, A. M. Peck, and J. P. G. Sterbenz, "On the Fitness of Geographic Graph Generators for Modelling Physical Level Topologies," in *Proceedings of the 5th IEEE/IFIP International Workshop on Reliable Networks Design and Modeling (RNDM)*, (Almaty), pp. 38–45, September 2013.
- [38] "Level 3 Network Map." <http://maps.level3.com>.
- [39] "The Next Generation Core Optical Networks (CORONET)." [http://www.darpa.mil/Our_Work/STO/Programs/Dynamic_Multi-Terabit_Core_Optical_Networks_\(CORONET\).aspx](http://www.darpa.mil/Our_Work/STO/Programs/Dynamic_Multi-Terabit_Core_Optical_Networks_(CORONET).aspx).
- [40] "Internet2." <http://www.internet2.edu>.
- [41] N. Shor, N. Zhurbenko, A. Likhovid, and P. Stetsyuk, "Algorithms of Nondifferentiable Optimization: Development and Application," *Cybernetics and Systems Analysis*, vol. 39, no. 4, pp. 537–548, 2003.