Using Multi-Topology Routing to Improve Routing during Geographically Correlated Failures

M. Todd Gardner, Rebecca May, Cory Beard, Deep Medhi
Computer Science & Electrical Engineering Department
University of Missouri–Kansas City
Kansas City, Missouri 64110-2499 USA
Email: {mtgdm6, rm5x8, beardc, dmedhi}@umkc.edu

Abstract—During large geographic events in networks, the routing churn that occurs has been shown to cause significant impacts in routing stabilization following the event. This work addresses some of those problems by proposing a set of algorithms that use Multi-Topology Routing (MTR) to predict large geographic events and switch to virtual topologies that reduce the impact of routing changes that can result in dropped connections until a new link state and shortest path trees can be established. We propose two algorithms to generate virtual topologies, Geographic Coverage MTR (gcMTR) and Geographic Targeted MTR (gtMTR), gcMTR and gtMTR provided methods to create topologies in both a network wide coverage approach and targeted approach that can be used to anticipate a specific event, where knowledge of that event exists. A third algorithm proposed in this work specifies a way to detect a geographic event and select a topology to use.

Using Multi-Topology Routing, we have been able to show that the number of connections that are dropped during a geographic event can be reduced significantly, which reduces the impact to the remaining part of the network.

I. INTRODUCTION

An increasing need has been observed for large networks and the internet to have the ability to quickly recover during large scale geographic events. Natural disasters, major weather events, as well as geo-political events have caused large disruptions in the internet as well as in emergency and mission critical networks. When these events occur several problems can occur. These range from slow convergence of the routing protocols to congestion due to overloaded links caused from rerouting traffic affected by the event.

Large disruptive geographic events cause stability and routing convergence problems in large networks. This is especially problematic in Interior Gateway Protocols (IGP) like OSPF and IS-IS. In both [5] and [1] the authors explore the tradeoff between routing stability and convergence speed in link state protocols. Disruptive outages in networks tend to cause multiple link state advertisements (LSA) to be flooded across the routing area and the routers to perform multiple shortest path calculations and therefore long convergence times. Flapping can occur if link information is propagated through the network too quickly. To prevent flapping, hold down timers are used which slows convergence times [5].

This highlights the tradeoff between convergence speed and routing stability. In addition, major disruptions can cause large numbers of traffic flows to be interrupted or rerouted to other links causing congestion [1].

In this work, we propose to use preemptive reconfiguration through Multi-Topology Routing (MTR) [16], [15] extension of OSPF and IS-IS in order to isolate parts of the network affected by a geographic event. Using this mechanism, the network is less disruptive to existing traffic and the routing convergence time delay due to the basic link-state protocol can be avoided. More specifically, we use Multi-Topology Routing (MTR) to protect against the problems associated with geographically correlated failures. Briefly, in the MTR framework, we create a series of topologies that can be useful to use for potentially disruptive events in the network. We can generated by increasing link weights in specific geographic areas of the network. The general effect is to move the trunk of the shortest path tree (SPT) away from the affected area in that topology. This isolates a region from the rest of the network and limits the impact that disruption in that part of the network can have on the whole network.

To understand why MTR is needed for geographic failures in networks, an understanding of what happens during link or node failures in link-state routing protocols is necessary. First, detection of a link or node state change (like an outage) occurs. Next, routers affected by the change will create LSA's and send them to their neighbors. This information will be flooded across the routing area till all routers have the same link state database. If database changes occurred, each router independently calculates the shortest path tree to all nodes based on the new database.

If multiple outages occur, several things can happen that negatively impact the above restoration process. First, where multiple links are involved, links may exhibit repeated and intermittent failures causing link flapping behavior [1]. This behavior typically causes frequent routing changes and that leads to routing instability. If the outages are significant, multiple LSAs will be flooded across the network causing shortest path first (SPF) calculations to run repeatedly at each node. This is commonly referred to as SPF Throttling and can lead to longer convergence times. A significant impact of route instability is that traffic paths may also become unstable and impacted severely[1]. Long RouterDeadInterval intervals can
Several schemes to improve convergence times and/or improve route stability have been proposed through the years. Modification of OSPF (or IS-IS) timers is the primary way proposed to prevent link or route flapping behavior. Hello timers and Router Dead Interval timers can be extended to dampen flapping behavior. In addition, the interval between successive SPF calculations (SPF hold down timer) can be adjusted to prevent repeated SPF calculations. Clearly, increasing these timers may delay convergence times. These methods are discussed in detail in a number of works including [1] [5] [4].

Improving outage detection times have also been proposed to improve stability and convergence times [5]. This is the most productive if hardware means of detecting outages can be used instead of strictly relying on reducing the Hello timer. False alarms can also cause link flapping and other instability.

Incremental SPF (iSPF) has been proposed as a way to reduce SPF processing at the routers [11]. The idea behind iSPF is to incrementally change the shortest path tree when link or node changes occur near the leaves of the tree. Performance improvements due to iSPF are discussed in both [1] [5]. The algorithms we propose take advantage of iSPF efficiencies.

In [10], one of the interesting observations of the impact of Hurricane Sandy is that the major internet providers moved the majority of the traffic away from New York City during the event. The goal was not to eliminate internet access for the New York City area, but to reduce the possible impact the hurricane was having on traffic that was passing through the New York City area. We propose to use MTR to quickly move traffic from a vulnerable geographic region.

There have been several concepts to modify routing to accommodate disaster planning. In [7], Hanson et al. proposed a method of using routing layers to isolate areas of a network. The idea is similar to our work, except that they do not propose using Multi-Topology Routing to accomplish it. Instead they utilize the ability to segment networks into areas using protocols (like OSPF) that can have multiple areas. Their methods require that routing area be pre-configured for possible geographic events. We felt this was not flexible enough for robust disaster planning. In the next section, we see the same authors went on to propose using MTR to re-route around individual nodes or links. In [12], the authors use a novel hill-climbing algorithm to assign link weights such that the link utilization is maintained at a level so that in the event of a disaster re-routing does not overload links causing congestion. While this is useful for network provisioning, it does not necessarily prevent the routing churn caused by large geographic events.

Several local IP fast re-routing methods have been proposed, these are summarized in [20] and generally include the outage detecting router planning the local alternate path. There are also methods where the packet carries information about the network elements that are to be avoided. These include the not-via methods as described in [18] and others like Localized On-Demand Link State Routing in [17]. These methods are good for a small number of link or node failures, but the mechanisms in these methods either do not support large numbers of correlated node/link failures or the methods do not scale well enough to support large scale outages.

In our MTR-based approach, we create multiple alternate topologies (based the default topology) where link weights are pre-determined in such a way that in the event of a geographic failure, the routing paths created move away from a vulnerable geographic region. In particular, we propose a method to compute the link weights so that the path selection attempts to avoid an affected region. While the MTR-based approach has been used for network resilience [2], [3], [8], our approach considers a geographic vulnerability as the driver for how to determine links weights that can be used in MTR.

The rest of the paper is organized as follows. A description of MTR is in Section II. Following that is our proposal for Geographic MTR. In Sections IV and V, we provide examples and an evaluation of the gcMTR and gtMTR algorithms. Finally, we conclude with a discussion on the complexity of these methods and our conclusions.

II. MULTI-TOPOLOGY ROUTING

Multi-topology Routing (MTR) was originally proposed to enhance traffic engineering in link state protocols like OSPF [16]. The concept was to allow different classes of traffic to use different topologies to cross the network. Traffic that requires higher Quality of Service could use a lower latency path than other traffic. It was soon discovered that MTR could be used for other purposes, like fault tolerant routing.

MTR utilizes the link weights used during shortest path calculation to create a virtual topology, setting link weights for links that will not participate in a particular topology very high compared with other links. In OSPF, the distribution of this additional link weight information is accomplished using the LSA. The Type of Service (TOS) field in the LSA is redefined as the MT-ID field. This provides link weight information for each topology to all of the routers in the network. A default (non-modified weighting) topology is required to be maintained.

In the IP packets to be routed, the Differentiated Services Code Point (DSCP) field is used to note which topology is to be used for that packet. The router can look up the topology ID for that packet, check the shortest path tree (SPT) for that topology, and forward the packet based on that SPT.

MTR can also be used for fault tolerance routing. Concepts for implementing MTR for fault tolerance have been described by several authors including Menth and Martin [13], Cicic [3], and Scheffel, et al. [19], and Kvalbein, Hanson, et al. [9]. Alternate topologies are used in these works primarily to protect against single link or node failure. In [3], Cicic defines a legal topology with the following three criteria. All restricted links are attached to an isolated node and non-isolated node. The links between two isolated nodes are isolated. And the non-isolated nodes have no isolated links. Using an algorithm to create the topologies from this criteria, they are able to minimize the number of topologies.
Scheffel, et al. [19] took an optimization approach to creating topologies. They created a binary integer linear program (BILP) to create topologies that had the goal of being able to mitigate single link failures with a given number of topologies. The objective was to minimize path distance. In both works, once a link outage occurs the detecting node is responsible for finding a topology where the failed link is isolated. The packet can then be rerouted around the failed node or link. We refer to this as local rerouting based on MTR.

III. GEOGRAPHIC MULTI-TOPOLONY ROUTING

With the goal of reducing the impact that a geographic event can have on a network, our approach is different than these approaches (and any other approaches we have seen). The topologies created here are intended to avoid routing (other than local routing) in a specific geographic area. The topology creation algorithm increases link weights around a specific geographic location with a specific radius. This has the affect of pushing the SPT away from the vulnerable area and only allowing leaves of the SPT to exist in the vulnerable area. The weighting algorithm decreases linearly from 100 at the avoided location till the radius boundary and then the weights are set to 1. Other weighting algorithms have been tested (such as exponential decrease). Linear weighting with a boundary at the radius seems to perform well. The advantage of using high (but not infinite) weights is that the SPT will try to avoid the vulnerable area using it only if no other route is available. Multiple topologies are then created with different geographic locations.

Once topologies are created the individual routers have the responsibility of choosing the topology based on the early link state information that arrives at that router. A set of criteria govern the selection of a topology that is not the default topology

A. Topology Creation Algorithm, gcMTR, gtMTR

There are two methods to create multiple topologies proposed here that include the Geographic Coverage Multi-Topology Routing (gcMTR) approach and a Geographic Targeted Multi-Topology Routing approach (gtMTR). In gcMTR, N topologies are created, each that can avoid a geographic area in a network. If these topologies are distributed accordingly, they will provide coverage against most geographic events.

The gtMTR approach uses existing knowledge about specific events to build a topology that targets a specific geographic area. Targets could include natural disasters that have a likely location (like an earthquake fault line or area frequently impacted by hurricanes), geo-politically vulnerable areas, or other anticipated geographic impacts. This enables network planners to anticipate certain geographic impacts. Notations used here are:

- $G(V, E)$: Network Graph with nodes $V$ and edges $E$ (links)
- $N$: number of topologies required gcMTR or gtMTR
- $V_i$: center of vulnerable area $i$
- $R$: radius of any vulnerable area
- $T_i$: MTR Topology $i$
- $D_{ij}$: the distance between $V_i$ for $T_i$ and the closest point on link $j$.
- $W_{ij}$: the Link Weight on link $j$ for $T_i$.

Algorithm 1 (gcMTR) creates a set of topologies that provide coverage across a network to protect against a variety of geographic events. Function $genCoveragePlan$ generates locations to avoid based on a coverage plan that can be configured for a specific network. Algorithm 2 (gtMTR) utilizes knowledge about a particular type of an event to create topologies specifically designed for a certain event. Function $genTargetList$ generates locations to avoid based a custom list of predefined geographic event locations. The function $Norm$ normalizes the maximum weight (at $V_i$) to 100.

Algorithm 1 Create Coverage Topologies - gcMTR

```
for all $i$ in $N$ do
    $V_i$ ← $genCoveragePlan(i)$
    for all $j$ in $K$ links do
        if $D_{ij} < R$ then
            $W_{ij}$ ← $Norm(MaxD_{ij} - D_{ij})$
        else
            $W_{ij}$ ← 1
        end if
    end for
end for
```

Algorithm 2 Create Target Topologies - gtMTR

```
for all $i$ in $N$ do
    $V_i$ ← $genTargetList(i)$
    for all $j$ in $K$ links do
        if $D_{ij} < R$ then
            $W_{ij}$ ← $Norm(MaxD_{ij} - D_{ij})$
        else
            $W_{ij}$ ← 1
        end if
    end for
end for
```

B. Topology Selection and Use

As OSPF (or other IGP) detects link and node failures via hardware methods and Hello protocol, LSAs are generated and flooded across the routing area notifying the other routers of the link and node failures. Each router will use a set of criteria to enter MTR mode and note a Geographically Correlated Event $EV_i$. The proposed criteria is shown here:

1) More than 2 nodes or more than 3 non-adjacent links are out of service simultaneously within radius $R$. These devices form the basis of the geographic event with its center at the geographic mean of the failed nodes or links AND
2) Geographic center of the geographic event is within $R$ distance of any $V_i$

This set of criteria allows for considerable flexibility in topology creation algorithms including overlapping coverage algorithms. If overlapping coverage is used, the possibility of selecting the incorrect topology increases. The criteria also allow for a no MTR mode option that may be desired for certain geographic locations. It is assumed that all routers in the routing area are using the same algorithm (gcMTR or gtMTR). In addition, all routers would use the same criteria and topology selection criteria. Routing loops would be possible only during the time between when the first router has changed topologies to the time when the last router changes topologies.

To achieve fast switching during the event, it would be required that routers maintain SPTs for all topologies that could be used. In order to minimize processing at the routers a couple of approaches could be used. First, the calculation of not-in-use topologies should be prioritized lower than the in-use topology. Second, it may not be required that the not-in-use topologies be calculated as often as the in-use topology. If there were $k$ topologies and we assumed a period of $t$ default topology SPT calculations before not-in-use topologies SPTs are calculated, that would be $k/t$ SPT additional calculations for every normal SPT calculation.

Definition 1. $(EV_k, P_k, T_k)$ is an Geographically Correlated Event meeting MTR mode criteria that is centered at point $P_k$ with topology selected $T_k$.

Algorithm 3 Select Topology

```
if Geographically Correlated Outage Occurs, < EV_k, P_k, 0 > then
    Vdist ← inf
    Tk ← 0
    for all j in N Topologies do
        if D_jk < Vdist then
            Vdist ← D_jk
            Tk ← j
        end if
    end for
    < E_k, P_k, 0 > ← < E_k, P_k, T_k >
end if
```

IV. ILLUSTRATIVE EXAMPLE

We have three topologies to demonstrate the topology generation algorithms (1 and 2). The first is a $5 \times 5$ grid, which is used for illustrative purposes.

A. $5 \times 5$ Grid Network

Fig. 1a shows a $5 \times 5$ grid with default topology (all link weights are one). The link weights are shown in green. The SPT source node is 4. When geographic events occur, we assume several nodes in an area are deleted. This may affect the default SPT in a minimal way as Fig. 1b shows; only routing to nodes 20 and 21 was affected when nodes 10, 11, 15, 16 were deleted. In Fig. 1c, the topology generated when the vulnerability point is $V_t = (22, 57)$ with a radius of $R = 20$. The topology moves the vulnerable point to the leaves of the tree.

However, when nodes 1, 2, 6, 7 are deleted Fig. 1d shows that routing to 10, 11, 12, 15, 16, 17, 20, 21, 22 is affected by the SPT change. Clearly, the location of the vulnerability with respect to the main part of the SPT affects the impact of the outage. Fig. 1e shows the $5 \times 5$ grid with a topology generated with a vulnerability point at $V_t = (40, 24)$ with a radius of $R = 20$.

V. EVALUATION

To evaluate the Geographic Multi-Topology Algorithms, two networks were chosen. The first network is the Cost266 network from the Survivable Network Design Library (SNDlib 1.0) [14]. It is a medium size network example. The second network is the AT&T Layer 1 topology generated by Sterbenz, et al. [21] and available at [6]. It is an example of a large physical layer network with 383 nodes.

Two metrics are computed and analyzed. These include:

1) Average percentage of services disconnected while switching from the default topology to the event topology and while switching from the selected topology to the event topology.

2) Average path length during the event topology, using the default topology on the base network, and using the selected topology on the base network.

For the Cost266 network, SPTs' were generated from all nodes as Dijkstra algorithm sources and the results averaged. The services are defined as the connection from the source to all other nodes. For the AT&T L1 network, SPTs' were generated from a selection of 30 nodes as Dijkstra's algorithm sources that were spread across the network. The 30 largest metropolitan areas were chosen. The SPTs were created from the source to all sites in the AT&T L1 network.

A. Cost266 Network

In the Cost266 network, a series of topologies was created to provide a coverage plan for that network. Fig. 2a shows a grid of 6 locations ($V_i$) and radii ($R_i$) used to create the coverage topologies. The closest ($V_i$) to the event center is the selected topology. Fig. 2b shows the Cost266 network Default Topology (all link weights = 1). For illustrative purposes, the SPT source in the figures is London. Three geographic events were created for evaluation as shown in Table I. These were loosely based on possible disaster scenarios.

Fig. 2c shows the SPT with a London source after event 2. Amsterdam, Dusseldorf, and Brussels and all associated links were removed in this event. Fig. 2d, the SPT with new link weights created by the selected topology using London as a source is shown. The selected topology was 1 with a $V_t$ located at $(2, 52.5)$.

Table II shows the improvements gained by using a selected topology that was similar to the actual event. Generally, it was
Fig. 1. Grid5 Network, SPT Source = 4, SPT links are solid, non-SPT links are dashed. Link weights are shown. (a) Grid5 Default Topology (b) Network with nodes surrounding (22, 57) with radius $R = 20$ deleted (c) Topology created with $V_i = (22, 57)$ with radius $R = 20$ (d) Network with nodes surrounding (40, 24) with radius $R = 20$ deleted (e) Topology created with $V_i = (40, 24)$ with radius $R = 20$

noted that if the event disconnects a major part of the default SPT, the improvements were significantly better than if the event was located in the leaves of the SPT. Event 2 disrupted the default path that crossed northern Europe. The topology that was selected routed connections around this area and improved the number of disconnected services significantly.

Table III shows the average path distance during the event, default topology, and selected topology. The distance was shortest during the default topology having more options than during the event or selected topology.

B. ATT L1 Network

In the ATT L1 network, a series of 8 topologies was created to provide a coverage plan for that network in a manner similar to the coverage plan for the ATT L1 network. Fig. 3a shows a grid of 8 locations ($V_i$) and radii ($R_i$) used to create the coverage topologies. Fig. 3b shows the ATT L1 network Default Topology (all link weights = 1). The SPT source in the figures is New York. Four geographic events were created for evaluation as shown in Table IV. These were loosely based on possible disaster scenarios.

Fig. 3c shows the SPT with a New York source after event 1. Nodes and links in the New Madrid, MO, USA area were removed in this event. Fig. 3d, the SPT created by the selected topology using New York as a source is shown with $V_i = (-92, 33)$, Radius = 3.5. Fig. 3e shows the SPT with a New York source after event 4. Nodes and links in the Chicago, ILL, USA area were removed in this event. Fig. 3f, the SPT created by the selected topology using New York as a source, is shown with $V_i = (-92, 41)$, Radius = 3.5.

In Table V, we see that the percentage of dropped connections actually increased using the selected topology for event 1. This occurred for two reasons. First, the topology selected was not matched well with the area of the event. The topology selected (see Fig. 3a location (-92,33)) was centered significantly south of the center of the actual event. Second, the event itself did not cut either major branch of the SPT that traversed the network from east to west. The topology actually re-routed connections north into the area where the
event occurred. This can be observed on Fig. 3c and Fig. 3d.

Also in Table V, it is noted that when the event did sever a main branch of the SPT, significant improvements were gained by switching to the selected topology during that event. The number of dropped connections changed from 20.5% to 1.2%. The other two events showed significant improvements also. During the evaluations on the larger network, it appears that the selected topology tended to route traffic further around the vulnerability point than occurred during the event. This is confirmed in Table VI, the path lengths were slightly longer during the selected topologies.

<table>
<thead>
<tr>
<th>Event Number</th>
<th>X Location</th>
<th>Y Location</th>
<th>Radius</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-89.5</td>
<td>36.6</td>
<td>3.8</td>
</tr>
<tr>
<td>2</td>
<td>-78.6</td>
<td>34.0</td>
<td>4.0</td>
</tr>
<tr>
<td>3</td>
<td>-119.7</td>
<td>35.6</td>
<td>4.0</td>
</tr>
<tr>
<td>4</td>
<td>-90.0</td>
<td>40.0</td>
<td>3.5</td>
</tr>
</tbody>
</table>

C. Sensitivity to Event Size & Location

One of the areas of interest is how sensitive these methods are to events that are not located directly on a topology from a location and size perspective. The first test to evaluate this was to change the location of Event 4 to locations west, east, and south of the original location by 50% and 100% of the selected topology radius. Fig. 4 shows the average percentage of connections dropped using both the default topology and selected topology when the event center is not aligned with the topology center. The results were as one would expect. The selected topology performed well when the event was centered where the topology is centered. As the event drifts away from the center of the topology, the performance degrades approaching the performance of the default topology.
The second test varied the size of Event 4 from 3% to 186% of the selected topology radius. Fig. 5 shows the average percentage of connections dropped using both the default topology and selected topology when the event size is not aligned with the topology size. It is interesting to note that the performance when the event size was smaller than the topology radius was not drastically different from the matched event size. However, when the event size was significantly larger than the topology radius, the performance quickly degraded. It appears that when the event is smaller than the topology radius, the topology effectively routes around the event. But, when the event is significantly larger than the topology radius, the topology does not route the traffic a suitable distance from the event.

VI. COMPLEXITY OF APPROACH

It is well known that Dijkstra shortest path algorithm has $O(n \log(n) + e)$ complexity, where $n$ is the number of nodes and $e$ is the number of edges (links). With this approach, the shortest path algorithm is executed more often before the event in order to maintain a complete set of SPTs at all nodes when the event occurs. If there are $k$ topologies, $k$ being constant, it will be executed $k(n \log(n) + e)$ times, while the overall complexity remains at $O(n \log(n) + e)$.

Since these topologies are not used until the event occurs, we would recommend that the calculations of the additional SPTs be prioritized lower than the in-use topology (normally default) during SPT calculation. Furthermore, it is possible that the calculation of the SPT of the not-in-use topologies could
even be delayed till the in-use topology has been updated a given number of times \(t\). It would reduce the execution time to \(\frac{1}{t}(n \log(n) + e)\).

VII. CONCLUSION

This work has demonstrated that Multi-Topology Routing (MTR) can be used to improve performance of routing during large geographic events like natural disasters. These improvements are achieved by maintaining additional topologies created by increasing link weights in specific areas of the network. This has the effect of pushing the Shortest Path Tree around this area limiting the amount of connections that are broken when nodes are suddenly out of service in that area.

Multiple topologies can be created to anticipate multiple events in a network. When this is done, a topology is selected when it is apparent that a geographic event has occurred. The traffic is moved to the new SPT quickly and the impact of the event is minimized. Essentially, the amount of network churn is reduced during the geographic event.

Three algorithms were proposed. Two algorithms gcMTR and gTMTR provided methods to create topologies in both a network wide coverage approach and targeted approach that can be used to anticipate a specific event, where knowledge of that event exists. The third algorithm specifies a way to detect a geographic event and select a topology to use.

During evaluation, it was discovered that the number of dropped connections is typically significantly better if a selected topology is used. The exception to this is if the condition exist where (1) the event does not disrupt a significant branch of the default SPT and (2) the selected topology does not fit well with the event from a location perspective. We explored this further by modifying the distance that the event occurs from the center of the topology and by modifying the size of the event compared to the radius of the topology. Although not unexpected, the results emphasize the importance of the algorithm and parameter selection.

Since there is not a large body of research in this area, there are significant opportunities for future research. We plan to implement these algorithms in simulation to assess how they perform in a more lifelike setting. Generating efficient coverage topologies that optimize for size, number, and location of the topologies is needed. Optimizing the frequency of calculating SPTs for not-in-use topologies is also an interesting extension of this work.

ACKNOWLEDGMENT

The authors would like to thank the KU Resilinets Group (J. Sterbenz and Y. Cheng) for assistance with the ATT L1 network and general conversations about this work.

REFERENCES


