Reproducing "Routing on Multiple Optimality Criteria"

Greg Baroni
UC Santa Barbara
gbaroni@ucsb.edu

Erica Buckland
UC Santa Barbara
eric334@ucsb.edu

Matthew Wong
UC Santa Barbara
matthew_wong@ucsb.edu

1 INTRODUCTION

Classic vectoring protocols such as the Border Gateway Protocol (BGP) and 'ignorant' routers have become increasingly scrutinized for being an archaic system. This argument is not without merit, as classical routing protocols are only able to route packages over a single dimension of optimality. *Routing on Multiple Optimality Criteria* by João Sobrinho and Miguel Ferreira at the University of Libosa proposes a new system that takes advantage of multiple pieces of gathered data at each forwarding device in a network.

The concept of an optimal path, as introduced in the paper, is derived from the optimization of individual attributes on each node. This decision on a node level is defined as the 'partial order' which respects and replaces the 'total order', while being isotonic. Traditional protocols use a total order so that each decision for routing is homogeneous in a single dimension on the node level. On the two dimensional level, with width and length defined as bandwidth and propagation delay between routers respectively, moving larger packets across identified 'shortest-widest' paths, and smaller packets through 'widest-shortest' paths is the general optimal solution. However, these are not the only dimensions of optimization explored, and are by no means the 'best' dimensions as they are somewhat arbitrary, which is also what this paper seeks to explore.

This evaluation of packets on a partial order scale and network connections on a multi-dimensional scale allows the system to reduce traffic overall, and increase speed where previously bottleneck hazards were completely unidentified. Sobrinho and Ferreira showed that the vectoring protocols produced from partial orders converge quickly and only return a small set of dominant attributes at each step. Using Complementary Cumulative Distribution Functions (CCDF), they present the probability to have more than one dominant attribute to be high, but the probability of more than seven dominant attributes to be low.

2 RELATED WORK

The simulator for dominant-paths vectoring protocols utilizes Rocketfuel topologies transformed into network data sets as input. Since the exact network topology of ISPs is confidential, the Rocketfuel Project utilizes traceroutes, ISP specific naming conventions, and an alias resolution technique to map the topology of a specific ISP. These topologies are used to simulate the behavior of a supposed real world network when routing with multiple optimality criteria.

Multi-objective pathing problems are a commonly studied space in the networking and mathematics community. *An Empirical Investigation of Some Bicriterion Shortest Path Algorithms* [1] investigates pathing on a two dimensional level, with a primary criteria taking priority on pathing, where the second decides pathing given the primary being homogeneous. As this protocol is a 'primary first' protocol, it is possible that the most efficient solution is not taken in the context of networking, as there could be dual importance of both dimensions. Deciding on 'primary first' is similar to the widths-lengths or lengths-widths dilemma in this paper. As noted, "a partial order on [attributes] is derived, rather than assumed a priori, and does not necessarily coincide with the product order".

Another topic of related work is multipath routing protocols, which can use traditional systems such as BGP for routing, but on multiple routes. *More Flexible Routing Policies While Improving Global Stability* [5] from Wang et al. utilizes BGP for each neighbor rather than on the global scale for path routing. As described in the paper, using a current system such as BGP for multipath routing allows implementation of the advancements on existing hardware. The foundations of this paper is on the algebraic conception of networking, which was idealized in the nineteen fifties, but introduced to internet routing in the early two thousands by papers such as *Algebra and Algorithms for QoS Path Computation and Hop-by-hop Routing in the Internet* [3], which appropriate Dijkstra's algorithm in an environment with strict isotonicity. It was concluded that Dikstra's algorithm could not handle multi-objective pathing problems. Routing with expression based constraints is the, previously exemplified by multiple papers on the network wide level, but on the node connection level performed in *A Programmable System for Performance-Aware Routing* [2], the algorithmic goal of this paper. Hsu et al. utilized this practice to develop a user ranked networking path controlled system to apply policy changes for dynamic path optimization from gathered metrics. Building on these concepts and practices, this paper creates dominant-path vectoring protocols which compute on a partial order of attributes.

3 MULTIPLE OPTIMALITY AND DOMINANCE

The original authors present an algebraic representation of their routing logic, as well as the problem with routing on multiple optimality criteria, which we have summarized. To our knowledge, the assumptions and logical procedures posited by the authors are accurate, and adequately represent the proposed routing problem.

As a summary of their notations, if we consider a set S of attributes that represent some arbitrary performance metrics, we can create an algebraic representation of the optimal path problem. The attribute of a path is obtained through a binary extension operation ($\oplus$). This operation determines the attributes of a path based on the attributes of the links that it passes through. Then, the authors assume that this operation will be associative and commutative with neutral attribute ε. Thus, for a trivial path, with just one node, the attribute is ε. They define $a[uv]$ as the attribute of a link $[uv]$. Then, the attribute of a path P is:

$$a[P] = a[u_o u_1] \oplus \ldots \oplus a[u_{n-1} u_n]$$

Considering a collection of optimality criteria O, with some criteria i in O, there is a binary relation between the set of attributes that is antisymmetric, transitive, and connex. Then, if we have paths a and b, for all paths in the set, if $a \prec_i b$, that is to say that a is i-preferred to b. The null attribute ($\bullet$) is the least preferred of any attribute, in other words, it means there is no valid path available from source to the destination.

In a network, the i-optimal attribute from a source to destination node is given by $a_i(s, d)$. In this case, i is the most preferred attribute considered when choosing among all path attributes between two nodes. When a third attribute is introduced, if $a \preceq i_b$ implies that $c \oplus a \preceq_i c \oplus b$, in the set of attributes S, then isotonicty is satisfied. If this is the case, standard vectoring protocols route based on i-optimal paths on the three attributes a, b, c in S. [4]

For example, we wanted to route on the quickest path possible, assume there is some file with size K. The time required to transmit that file along a path with capacity w (path bandwidth) and propagation delay l (path latency), will be $\frac{K}{w} + l$. The quickest path in the network from said source to destination is the one that minimizes the time required to transmit that arbitrary file.

As another example, consider the scenario in which high-bitrate video is streamed from a source to destination. As with most internet traffic applications, you want to take a shortest path, and lower latency as much as possible. However, assume that this high definition video requires a higher bandwidth. In this case, If we have some desired minimum bandwidth W, we would want to choose the path of a minimum delay, among the paths that have the bandwidth to meet the desired bandwidth W. The authors define this as the *W-wide-shortest order* ($\preceq_{K-W}$). This would be a total order that corresponds to the lowest latency path for paths with a bandwidth greater than or equal to W. If there are no paths that meet that criteria, it is defined as a shortest-widest order, since we would be choosing a path with the lowest latency, but does not meet our desired bandwidth value.

These types of total orders, by definition, do not satisfy isotonicity. To expand this, the authors utilize partial orders (denoted by $\preceq$) on the attributes, is defined by an antisymmetric, transitive, and reflexive binary relation on the current attributes. Reflexive means that attribute a is a partial order of a for all attributes a in S. Transitive is defined as meaning if attribute a is a partial order of b and b is a partial order of c, then a is a partial order of c, in the set S. Anti-symmetric denotes that if attribute $a \preceq b$ and $b \preceq a$, then $a = b$.

Now, rather than being given a network with a set of *optimal* attributes, we are given a network with a set of *dominant* attributes from source to destination. This set of dominant attributes is given by $A^*(s, t)$, and is the set of path attributes from source s to destination t where there is no other path attribute that is more important than any of the attributes in the set, for that pair of source and destination.

In order to route on these attributes, the key idea that the authors utilize is that there needs to be a partial order that satisfies isotonicity within the original total order. The task then, is to produce a set of dominant attributes given a collection of optimality criteria O.

The procedure to obtain these dominant paths from the optimal paths is one that produces partial orders on attributes from the total order, which must satisfy isotonicity, as well as considers all the criteria from the total order. To do so, the authors modify the procedure of computing the aforementioned i-optimal attributes, $a_i(s, t)$ for all criteria i to compute sets of dominant attributes based instead upon the largest possible reduction of the intersection of routing criteria such that the result is also isotonic. One such example is shown in Algorithm 1, from the original paper. [4]

4 EXPERIMENTAL SETUP

4.1 Original Authors' Experimental Setup

The original setup for this experiment is provided by the authors of the paper. They utilize a set of Rocketfuel topologies as a network data set. The simulator considers four sets of attributes, width-length pairs, width-hops pairs, hops-length pairs, and width-hops-length combined.

Algorithm 1: Dominant-paths, non-restarting vectoring protocol. Node u receives set B of attributes from v pertaining to destination t.

$DomTab_u[v,t] := a[uv] \oplus b \mid b \in B$

$Dom_u[t] :=$
$\quad D_{\leq}(\{DomTab_u[v,t] \mid v \text{ an out-neighbor}\})$

if $Dom_u[t]$ has changed, **then**

  for all r an in-neighbor **do**

    send $Dom_u[t]$ to r

Each link in the topology contains an OSPF weight and propagation delay. Because OSPF weights are inverse capacities, the weight assigned to each link to be utilized in the simulator is assigned to the inverse of its OSPF weight. The length of each link is equal to the propagation delay. Each link on a path through the network corresponds to one hop.

The simulator allows advertisements in the network to traverse links in FIFO order, with a random delay from 0.01 to 1 ms. If a non-restarting protocol advertisement utilizes more than 15 hops, that data set is discarded to prevent an unending count.

The authors' primary data set utilized the network topology from AS 1239. Figure 1 on the next page shows the CCDF of the number of dominant attributes for the network. This data set shows which sets of dominant attributes are more common in source to destination routing criteria. As expected, since fewer hops typically means a much shorter path from end to end, path lengths are highly correlated with path hops. This is one factor that we will be modifying in our experimental setup to test the performance and behavior of the system. One interesting observation is that there are more dominant width-length pairs than width-hops pairs. This is largely due to the correlation between path lengths and hops. Since shortest paths usually have the fewest number of hops, width-length pairs are typically more indicative of optimal routing preferences.

The experiment also considers performance of the system after link failures. When restarting, the affected node propagates its updated “unreachable” status to its neighboring upstream nodes. The simulator provides the CCDF of the time it takes for all nodes in the network to refactor their respective attributes with paths that utilize the failed link. The simulator also shows how often the reachability of an end to end connection is disrupted depending on the various attribute sets.

In order to thoroughly test the original setup, it is important to prove baseline similarity, so the artifact was run in default format with two RocketFuel dataset, which are similar in structure. The first dataset was the identical dataset to prove exact results, the second dataset was an identified similar dataset for comparison in similarity of results. Proving baseline comparability, we moved to configure the default format to change different aspects of the protocol. Initially this was performed on RocketFuel datasets, but we moved to change the dataset for outcome variability. Unfortunately there were not similar datasets in the same format, however the datasets themselves can be modified for different outcomes. To generate unique and potentially conclusive results, the datasets can be manipulated to extremes of networking scenarios. As the most utilized attribute pair is widths-hops, the datasets could be altered to have a high width to hop ratio or vice versa. As outputted results change drastically due to extreme conditions this experimental setup is most effective to understand outcomes of changes in the context of this protocol.

4.2 Our Experimental Setup

In order to thoroughly test the original setup, it is important to prove baseline similarity, so the artifact was run in default format with two RocketFuel dataset, which are similar in structure. The first dataset was the identical dataset to prove exact results, the second dataset was an identified similar dataset for comparison in similarity of results. Proving baseline comparability, we moved to configure the default format to change different aspects of the protocol. Initially this was performed on RocketFuel datasets, but we moved to change the dataset for outcome variability.

Unfortunately there were not similar datasets in the same format online, however we were able to easily modify the provided datasets for different outcomes. To generate unique and potentially conclusive results, the datasets can be manipulated to extremes of networking scenarios. As the most utilized attribute pair is widths-hops, the datasets could be altered to have a high width to hop ratio or vice versa. As outputted results change drastically due to extreme conditions this experimental setup is most effective to understand outcomes of changes in the context of this protocol.

5 RESULTS

5.1 Original Authors' Results

The results of their study consisted of the following six graphs:

CCDF of the number of dominant attributes for the product orders for four dominant path vectoring protocols (hops-lengths, width-hops, width-lengths, and width-hops-lengths).

Figure 1 displays the cumulative probability that there will be more than the current number of dominant attributes for a given vectoring protocol at a given number of attributes.

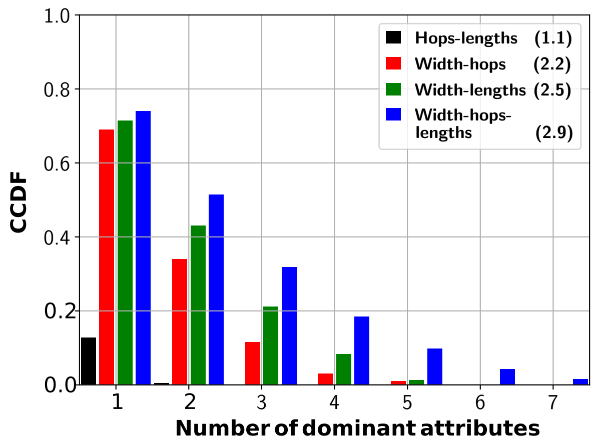

Figure 1: CCDF of the number of dominant attributes in AS1239

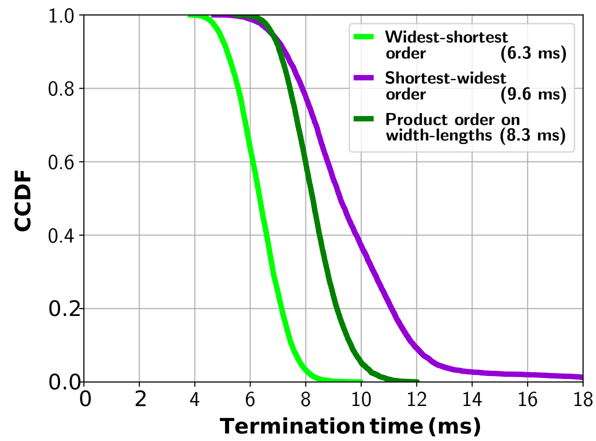

Figure 3: Network-wide announcement of a destination using using standard vectoring protocols

Two sets of CCDFs of the termination time of the non-restarting vectoring protocol during the network wide announcement of a destination and the failure of a link for four dominant paths vectoring protocols (hops-lengths, width-hops, width-lengths, and width-hops-lengths) and two standard vectoring protocols (widest-shortest and shortest-widest). The termination time after a network-wide announcement reflects how long it takes for the protocol to finish.

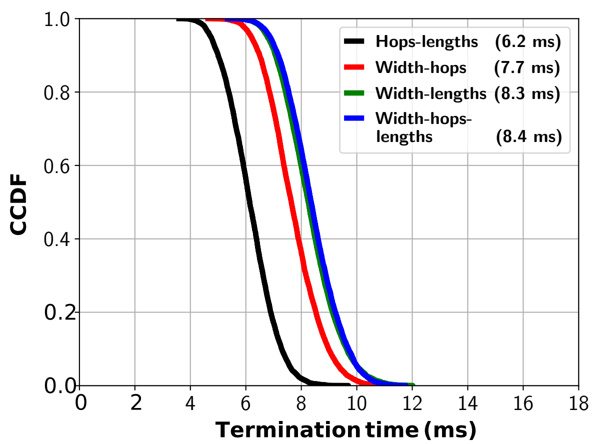

Figure 2: Network-wide announcement of a destination

The termination time after a link failure reflects how long it takes for the protocol to recover from a link failure.

These tests used a non-restarting protocol, meaning that the destination initiates the routing computation only once, by advertising the neutral attribute to all its in-neighbors. While the termination time for the network wide announcement of a destination is equivalent for a non-restarting and restarting protocol, measuring link failure for a restarting protocol requires a different measurement. This is accomplished through the CCDFs of the time to propagate unreachability for a restarting dominant paths vectoring protocol operating on a product order (hops-lengths, width-hops, width-lengths, and width-hops-lengths). The termination time here reflects how long it takes for the restarting protocol to recover from a link failure.The primary difference is the application of the restarting protocol, which creates new computation instances where new attributes replace older versions.

5.2 Our Reproduction Results

Testing the system in default format with the same dataset (AS1239) gave an outcome with the exact results as shown

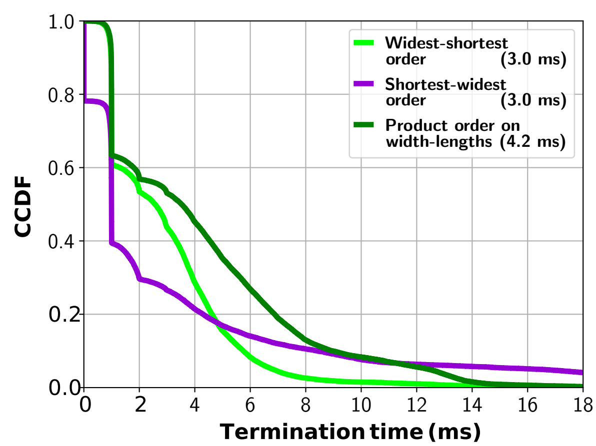

Figure 4: Termination time for the Failure of a link

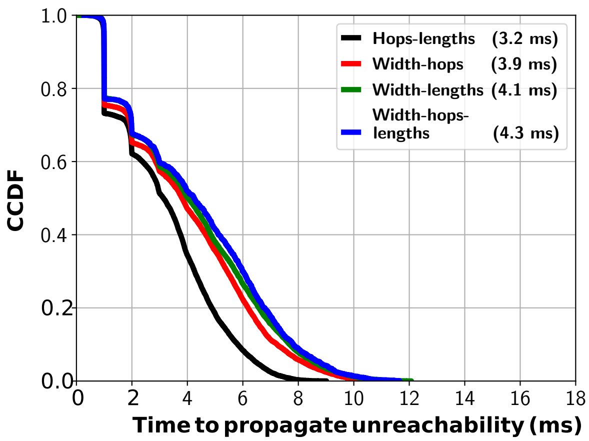

Figure 5: Network-wide reception of an unreachable link

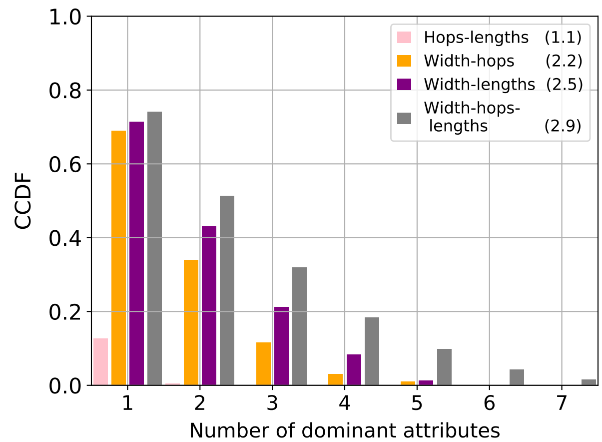

Figure 6: Reproduction of Figure 1

in the original paper. The only difference being that we modified the code to use different colors for the orders to show that our graphs were produced separately. As demonstrated in figure 1 and figure 6, we see that the CCDF is the same for all numbers of dominant attributes. This proves that the outcomes in our tests will be based on a consistent system, and so new tests can be conclusive. In comparison with a similar dataset from Rocketfuel Topology (AS3967), the output in default format from figure 7, we see a slight delta in the CCDF, but insignificant overall. This result demonstrates the protocol's tendency of the cumulative probability of having additional attributes at any given number of attributes is very high at one attribute, and decreases rapidly to almost a probability of zero at seven attributes. The termination times in the AS3967 network also show very similar results as that of the AS1239 network.

Next, we decided to test the properties of the vectoring protocols on networks not provided by the authors of the study. To do this, we created a script to randomly generate network data sets. First we tested the protocols on the AS3967 topology with the weights and lengths randomly generated. The widths were set to a random value between 5 and 55, and the lengths were set to random value between 1 and 18.

Afterwards, we tested the protocols on a completely randomly generated network, which consisted of 50 nodes, each with 2 to 4 links to other random nodes. The widths were set to a random value between 5 and 55, and the lengths were set to a random value between 1 and 24.

For all three sets of AS3967 graphs (figures 7, 8 and 9), network wide announcement of destination: width-lengths and width-lengths-hops are overlapping. Comparing the CCDF of the number of dominant attributes for the product orders graphs from the original AS3967 with the random values

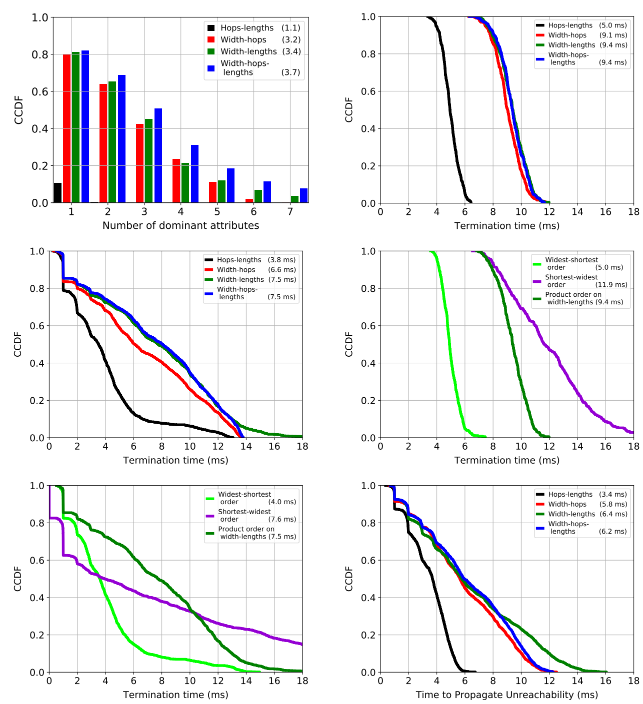

Figure 7: AS3967 CCDF Production

with this given dataset with the completely random topology, there is a notable trend difference, primarily for hops-lengths, but also for other protocols. The cumulative probability of the hops-lengths having more than the current number of attributes increases drastically for the randomized data sets versus the original. The results are similar for the completely random topology (resulting graphs are shown in figure 8) and the randomized original topology, so this value has more

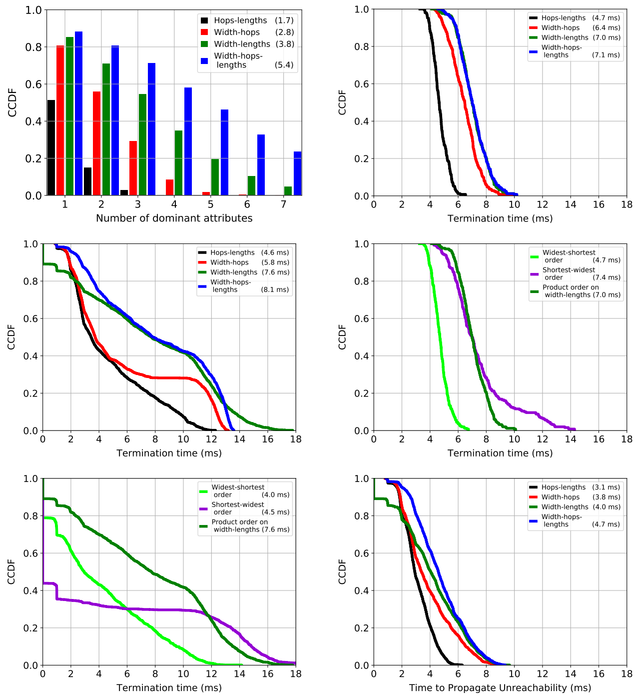

Figure 8: Resulting Graphs for a Completely Random Topology

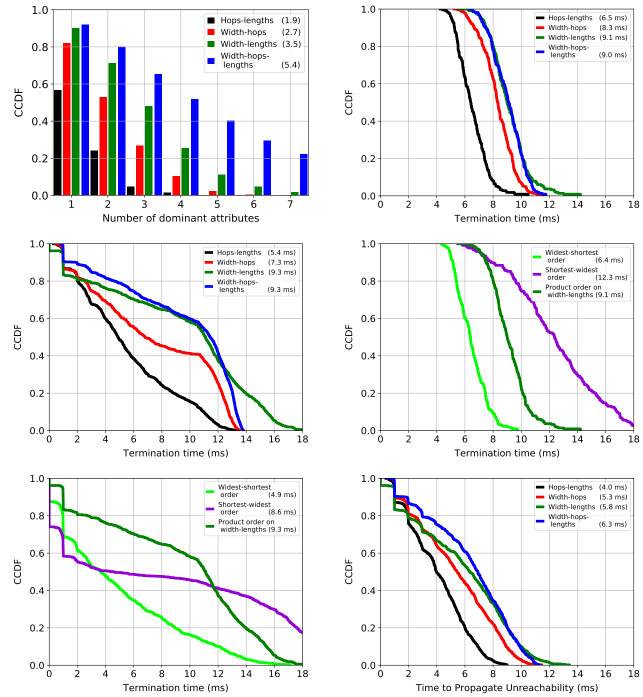

Figure 9: AS3967 with Randomized Width and Length

correlation with variability of hops than it does in variability of topology.

In the original data set, the majority of the lengths were equal to 1, though with some as high as 40. As a result, hops and lengths were very closely correlated as 1 hop would usually translate to 1 unit of length. In the randomly generated data set (resulting graphs shown in figure 9), the lengths were equally distributed between 1 and either 18 or 24, and, as a result, hops and lengths were less directly correlated, which led to higher termination times and a greater number of dominant attributes.

The next experiment we ran was to implement a new routing protocol. Since the authors covered every permutation of a dominant paths vectoring protocol working on widths, hops, and lengths, we implemented a standard protocol that returned the widest of the routes with the least hops. We then compared this widest least-hops protocol to the product order on width hops to see how well the dominant paths protocol version performed. These graphs are shown in figure 10. We found that, compared to the gap between the width lengths protocols, there was a wider gap between the standard and dominant paths vectoring protocols for width hops. While the dominant paths protocol still performs well, this shows that some dominant paths protocols do not scale perfectly, depending on the criteria.

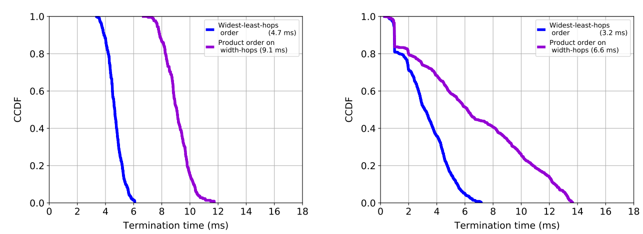

Figure 10: Plots of a newly implemented Widest least-hops order compared to the product order on width-hops

There are multiple given assumptions in the results of the original report and in the testing of our main results. The primary assumption being that there are no given failures in listed links during the network wide announcement. This assumption is reasonable, as link failures are unlikely in a realistic scenario. A secondary assumption is the timing between links, given a random delay from 0.01 to 1 ms. The tertiary assumption being a maximum number of hops of 15.

5.3 System Analysis

In testing, we were successfully able to reproduce the exact results of the original system, and reproduce similar results based on a similar dataset. Through our further testing with randomized data sets and changed parameters, we were able to identify operating trends in the artifact that are valuable for application on the artifact in real world scenarios. Due to the setup of the system and the format of the input network, we were not able to test against other scenarios. Our testing range is limited to RocketFuel topologies and randomized subsets of RocketFuel topologies. This could have prevented us from identifying further problems or further optimizations to be made.

While we were unable to implement new dominant paths vectoring protocols, we did implement the standard path vectoring protocol ‘widest least-hops order’. This allows us to analyze the system’s ability to scale network size for width hops.

In addition, we are utilizing the existing functions of the artifact assuming they are robust in their current state, which gives us no way of knowing if there is failure. Also, the variation of protocols are built within the same analysis structure, which could have unidentifiable issues.

6 FUTURE WORK

Most obviously, it would be best to test the system against a custom, real life dataset, particularly a WAN structure currently active. This should be tested for further analysis of potential utilization of dominant attributes over many nodes. This could be achieved by creating a piece of software to convert a traceroute to the format required by the artifact.

Furthermore, to test the validity of their methods, we could create our own simulation based on the ideas presented in the paper. This would ensure that the success of their simulation is due to the partial orders successfully respecting the total orders they derive from.

In addition, the artifact could be taken and modified to increase verbosity of output, greater consideration of dominant attributes, increased user friendliness in code function explanation and data analysis tools, and modification for allowing default traceroutes for analysis.

The set of protocols chosen for this paper are arbitrary, so any number of protocols could be chosen for analysis. For larger networks, the inclusion of data on the number of neighbor nodes as a protocol could help identify bottlenecks with this system before they even occur. Computation time for a given node could also be an aspect of a protocol. Given that there is almost an infinite set of possibilities for aspects of protocols, an advancement could be made to this paper to include a protocol generator and randomizer to find a possible greater unanticipated solution.

7 CONCLUSION

Routing on Multiple Optimality Criteria can be credited for the application of measuring the capability of multi-dimensional optimality in routing to the current common networking schema. This novel idea spawns discussion on varying protocol usage on an inter-node level. Using Complementary Cumulative Distribution Functions, it is evident that routing on partial orders remains efficient, while allowing for new ways to control the flow of traffic. Through our testing, we have verified that this holds for randomized networks and existing networks with randomized data attributes. For these networks, we found that the rocketfuel datasets tended to have less dominant attributes compared to the randomly generated networks, supporting the claim that partial orders return a small number of dominant attributes on average in a real network. In addition, the fully randomized network has the same probability hierarchy as other tests, but has greater dispersion between cumulative probabilities. This is likely due to the relative placement of connected nodes being naturally inefficient in their placement from randomness. On the other hand, the probability curves remain mostly similar among other protocols. Finally, our implementation of the widest least-hops standard protocol showed that the difference in efficiency between the standard and dominant paths vectoring protocol was not too large to hinder the practicality of the dominant paths protocol. When taken together, our reproduction corroborates the claims of the original paper and highlights how routing on multiple optimality criteria is efficient enough to be practical.

REFERENCES

[1] James Brumbaugh-Smith and Donald Shier. 1989. An Empirical Investigation of Some Bicriterion Shortest Path Algorithms. *European Journal of Operational Research* 43, 2 (1989), 216–224.

[2] Kuo-Feng Hsu, Ryan Beckett, Ang Chen, Jennifer Rexford, Praveen Tammana, and David Walker. 2020. Contra A Programmable System for Performance-Aware Routing. *USENIX NSDI* (2020), 100.

[3] Joao L. Sobrinho. 2002. Algebra and Algorithms for QoS Path Computation and Hop-by-hop Routing in the Internet. *IEEE/ACM Transactions on Networking* 10, 4 (2002).

[4] Joao Luis Sobrinho and Miguel Alves Ferreira. 2020. Routing on Multiple Optimality Criteria. *SIGCOMM ’20* (2020), 214–216.

[5] Yi Wang, Michael Schapira, and Jennifer Rexford. 2009. Neighbor-specific BGP: More Flexible Routing Policies While Improving Global Stability. *Proc. ACM SIGMETRICS* (2009), 217–228.