Evaluation of Algorithms for Multipath Route Selection over the Internet

Fabian Helfert
Faculty of Informatics
Technische Universität München
85748 Garching bei München, Germany
Email: fabian.helfert@tum.de

Heiko Niedermayer
Faculty of Informatics
Technische Universität München
85748 Garching bei München, Germany
Email: niedermayer@net.in.tum.de

Georg Carle
Faculty of Informatics
Technische Universität München
85748 Garching bei München, Germany
Email: carle@net.in.tum.de

***Abstract*—Resilience as well as other objectives like censorship-resistance demand the existence of multiple diverse paths between two hosts on the Internet. On Internet level, this requires the use of an overlay approach. In this paper, we study this problem on the basis of a data set obtained from traces between Internet hosts on Planetlab. We study a variety of path selection algorithms, including adapted versions of Suurballe's algorithm. We find that these can outperform the single-hop relay approach that is more commonly proposed for Internet multipath. Our analysis framework imports traceroute data and it generates graph representations of the overlay and the underlay. We can reduce the granularity of information that the algorithms can access. Topology-aware algorithms like Suurballe profit from better information. We also support a variety of evaluation metrics.**

I. Introduction

Having control over the path your network traffic uses has several advantages. It can allow a client to route the packets around a faulty network segment and therefore improve resilience. Another use case is the enforcement of regional constraints. It can also be used to improve censorship-resistance by routing around the censor. Using separate paths can also improve security. When sending parts of the data over different paths, this reduces the risk to leak the data.

The current Internet infrastructure does not allow the client to specify an alternative path. Routing protocols are responsible to distribute the necessary information and to decide on the best paths for each direction at each router. They are also responsible to route around failures when they occur. However, the process of updating all the routing tables can take time. They also do not solve the other mentioned use cases.

Overlay networks work on top of the Internet infrastructure and can provide alternative paths by rerouting packets over relay server. As a consequence, when considering multipath for Internet-wide communication, we can only split and control the traffic at the subset of machines that we use as relays. This lack of control and also the large size of the Internet makes the situation differ from the optimization of resilience in the networks of large network operators. A variety of issues like the lack of widespread multi-homing question the usefulness of multipath path selection. Can we still benefit from multipath?

Our contribution is as follows. On the basis of a large real-world data set of Internet routes, we explore the potential of multipath routing in this data set. We evaluate different algorithms and questions with respect to multipath routing in this scenario. As a side effect, we also developed a framework that can a) turn traces from the Internet into graph topologies and b) evaluates multipath selection algorithms in these graphs. A limitation is that geolocation data is not yet included.

The paper is structured as follows. First, we present a short vision for Internet-wide multipath. Then, we introduce the framework and the data set. Then, we give an overview of algorithms and metrics. Finally, we present some results.

II. Related Work

A. Overlay Networks

Andersen et al. introduced in [1] a Resilient Overlay Network (RON). Similar to our evaluations, RON uses dedicated nodes to forward packets. These nodes uses a link-state routing protocol to compute different routing tables, based on different metrics: latency, loss rate and throughput. An application can then set metric preferences in the packet header, which a forwarder module at every RON node will use to choose the next hop. The convergence time for RON is lower than for BGP in the internet, because the overlay network is smaller. In contrast to our concept, the client can only specify routing preferences or disallow specific virtual links between RON nodes. The specific overlay path is still chosen by the network and not by the client.

MACHETE was introduced by Raposo et al. in [2]. They use Multipath TCP in combination with an overlay network to improve confidentiality. They send parts of a document over different paths to avoid that the whole document can be restored. Therefore, they chooses relays which have the fewest AS in common. In Section VII-C we evaluate if the AS level is sufficient or if multipath route selection algorithm provide a higher diversity at IP level. A higher diversity would further improve the confidentiality.

Cai et al. intoroduced in [3] Cloud-Routed Overlay Networks (CRONets). By using cloud providers in combination with Multipath TCP, they were able to improve the throughput for 78% of the internet paths with an average improvement factor of 3.27 for 6.600 paths. They also showed improvements in latency and packet loss. Furthermore they showed that the diversity of paths has an influence on the throughput of MPTCP.

In [4] they introduced a modified version of Suurballe's algorithm. We will use this algorithm in our evaluation, as well as the Earliest Divergence Rule presented by Fei et al. [5], [6]. We will explain the algorithms in more detail in Section VI.

B. Alternatives to Overlay Networks

Instead of overlay networks, Cheng et al. proposed with GeoDivRP a new protocol to find geo-diverse paths [7]. They presented in [8] two different route selection heuristics. The Modified Link Weight (MLW) heuristic is, similar to the algorithm in [4], based on Suurballe's algorithm. We differ by using an overlay network and therefore do not require a change in the internet infrastructure. Additionally, they used mostly a single ISP network for evaluations. We use a much larger dataset including multiple ASes and ISPs.

Menth and Martin introduced in [9] multi-topology routing to improve resilience. Similar to the path splicing technique from Motiwala et al. in [10], they run several instances of a routing protocol to provide different paths. Multi-topology routing uses different link costs to ensure that after a link or node failure at least one route still works. The path splicing technique even works for arbitrary failure combinations.

C. Evaluations

Babay et al. identified in [11] three key principals for overlay networks: a resilient network architecture, an overlay software architecture with the ability to deploy new overlay protocols and a flow based processing of packets. Our framework helps to evaluate the first two principals. It can compare different network architectures and allows to develop new concepts for multipath route selection. In [12] Han et al. analyzed the impact of multihoming and overlay networks on path diversity. Both concepts could improve path diversity in their evaluation, but the did not combine both aspects.

In general we noticed that many results are impossible to reproduce and due to different data sets and methods not comparable. Our proposed framework would solve these issues.

III. Vision and Context

The goal of our work is to study and enable Internet-wide multipath communication. The main objective is to improve control over the paths messages take through the Internet. This is in contrast to planned communication within the network of an Internet service provider. As we have no control over the network, an overlay approach has to be taken. Relays are part of an overlay and organize themselve in the overlay. Considering the existance of multiple clouds in at least western countries, the necessary relays could be provided by virtual machines in the data centers of all these cloud data centers.

We assume that the relays perform traceroute operations between them in order to understand the path between them and which routers are shared between which paths. In our analysis in this paper we assume that the relays allow to use all links of a full mesh. However, the future objective would be to mainly utilize short-distance edges that help to control the paths and make them independent of Internet routing. In the analysis, the evaluation will always utilize the full information we have, the algorithms that generate the paths will use a reduced set of information that represents the situation of interest.

IV. Framework

The framework represents the network as a two-layered graph $G = (G_U, G_O)$. The underlay graph G_U represents the network topology as accurate as possible, where each device is represented as a vertex and connections between them as edges. The overlay graph G_O represents the possible paths a packet can travel between clients and relays. An edge represents a possible connection from one vertex to another vertex. As these vertices are also represented in the underlay graph, it is implicit $V_O \subseteq V_U$. Both graphs are directed graphs, to represent the asymmetric characteristic of the internet. We also use only single edges between vertices, because multiple edges can not be discovered with traceroutes. Additionally, the source cannot influence which of these links are used by the packets. Therefore a route which uses one of these links must always be modeled to use all links. Representing them as multiple edges, would therefore not improve the path diversity. Every vertex in the two-layered graph has a name which identifies the vertex uniquely in each layer. Edges between vertices representing the same identity in different layers are called coupling edges. These coupling edges are not stored explicitly in the data model, but can be restored by checking if a vertex with the same name exists in the other layer. Every overlay edge has an annotation, which stores the corresponding path a packet would use in the underlay graph. Fig. 1 shows an example of a two-layered graph with two relays and two clients. The overlay edge annotations are omitted for readability.

The underlay graph can represent the topology at different levels of granularity, like IP or Autonomous System (AS) level. Making the graph more coarse granular, can be done for example by mapping the IP addresses of the vertices to AS numbers and contracting all vertices with the same AS number to a single vertex. These transformations are supported by the framework. Making a graph more fine granular, is often not easily achievable and therefore, the finest granularity is predetermined by the data source. Since we use traceroutes to gather topology information, the IP level is the upper bound.

The algorithms can then be run at these different granularity levels. To evaluate the chosen paths at the same level of granularity, the paths are stored as a list of overlay vertices. With the help of the overlay edge annotations, the underlay paths can then be restored at the desired level. With the same method, it is also possible to evaluate a specific path at different granularity levels.

The framework is written in Python and uses igraph to store the graphs internally. It supports a flexible set of algorithms

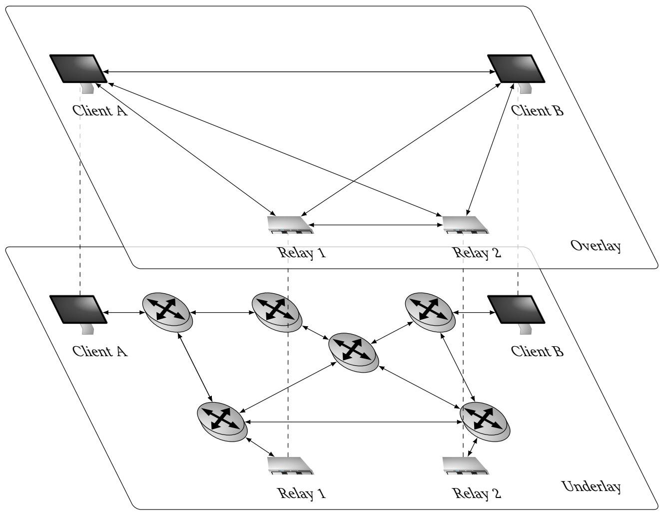

Fig. 1. Example of a two-layered graph.

and metrics. Every algorithm can support different parameters. To speedup evaluations, the framework caches all intermediate results on disk, including generated graphs, selected paths as well as evaluated metrics. A plot module automatically visualizes the metrics with matplotlib.

V. Dataset

Many available traceroute datasets offers traceroutes from one server to thousands of different endpoints. But we also need traceroutes from the relay servers toward the clients, to analyze if the relays offers a diverse path compared to the direct connection. These requirements are well suited for PlanetLab [13]. PlanetLab offers access to over a thousand of university computers, distributed all over the world. The iPlane project performs traceroute measurements between PlanetLab nodes to build a router interface-level atlas of the internet and predict end-to-end performance [14]. As an additional advantage, iPlane already performs IP alias resolution and other paper already used iPlane data for analyzing the impact of IP alias resolution in building traceroute-based internet maps [15]. iPlane provides historical datasets since 2010. Since the number of participating nodes was declining over the time, we have chosen a dataset from 2010-12-23.

This dataset consists of 155470 traceroutes from 208 PlanetLab nodes. This means there are over 700 traceroutes from each node. We then eliminated empty or incomplete traceroutes and resolved IP aliases before building the two-layered graph. The underlay graph consists of 10232 vertices and 23995 edges. We consider every PlanetLab node with at least one incoming and one outgoing edge as relay. The resulting overlay graph consists of 614 overlay vertices, out of them 134 are relays. There are 77062 edges in the overlay graph. The average traceroute has 14.03 hops before reaching the destination.

VI. Algorithms and Metrics

In this paper we evaluate different algorithms for multipath route selection. All algorithms will use the direct connection between source and destination as first path. This represents the case that the current internet infrastructure is used and minimizes the traffic overhead. The algorithms will then choose further alternative paths with the goal to maximize resilience. The number of alternatives paths are configurable in the framework and also depend on the requirements of the application. We will first describe two simple algorithms as reference, and then present a more complex algorithm.

The first algorithm is a random algorithm. The algorithm supports two parameters. The first parameter specifies the number of relays the alternative paths shall use. The second parameter specifies if the first parameter is only an upper bound or if it must be exactly fulfilled. The algorithm then chooses randomly relay servers which fulfill the requirements set by the parameter and connect the source with the destination.

The second algorithm is the Earliest Divergence Rule (EDR) presented in [5], [6]. For this algorithm, only the path from the source to the destination and the paths toward the relays are needed. The first path is the direct connection between source and destination. It chooses then relays whose path have the earliest divergence from the first path. The advantage of this algorithm is that it does not require to exchange topology information. All required information are measurable by the source itself. The original version in [5], [6] only returned alternative paths which have the same earliest divergence point. In case more paths are requested, we modified the algorithm to return also the paths from the second earliest divergence point and so forth until the requirement is met. Since every additional path only improves the later defined metrics, this modification has no negative impact on the performance of the algorithm.

The third algorithm is a modified version of Suurballe's algorithm, which was originally described in [16]. It is an algorithm to find two disjoint paths with minimum total lengths in a nonnegatively-weighted graph. Basically, the algorithm finds the first path using Dijkstra's algorithm, modifies the weight of this path and runs Dijkstra's algorithm a second time.

Engeser uses a similar approach for a two-layered graph [4]. The idea is to give every used path a penalty so that the next iteration of the Shortest-Path Algorithm (SPA) finds a different path in the overlay. In this modification, the overlay edge weight is a composition of two factors. The first factor is the length of the path in the underlay, which is represented by the overlay edge. This shall ensure that paths which are shorter in the underlay are preferred over longer paths. The second factor is the burden in the MLW strategy. Every time an edge in the overlay is used, the burden of every vertex in the underlay path is increased. By this means, the burden in the underlay represents the number of paths which already use

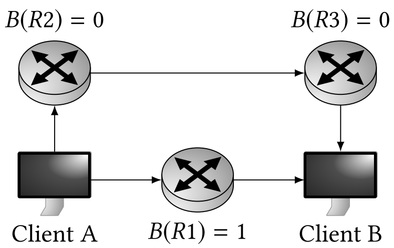

Fig. 2. Influence of θ on the path selection strategy. The direct path from A to B uses the router R1 already.

this vertex. This can be defined as a burden function:

$$B(x) = \begin{cases} \text{annotated burden,} & \text{if } x \text{ is an underlay node} \\ \sum_{n \in X} B(n), & \text{if } x \text{ is a path} \end{cases} \quad (1)$$

The weight function for an overlay edge, which represents the underlay path P, is then defined as:

$$W(P) = |P| + f(B(P)) \quad (2)$$

where the function f is used to balance between the path length and the burden of a path. Previous work has evaluated the functions $f_\theta(x) = \theta * x$ for linear discouragement and $f(x) = 2^x$ for exponential discouragement [4]. We will concentrate on linear discouragement with different values for θ. For example, Figure 2 shows two paths from Client A to B, where $R1$ has already a burden. For $\theta = 1$ both paths are equally preferred. For $\theta < 1$ the shorter path via $R1$ has the lower weight, while for $\theta > 1$ the longer, but diverse path is preferred. In case the path with the lowest weight is new, it is added to the result set. In any cases, the burden of every underlay vertex in the path is increased. Afterwards, the new overlay edge weights are calculated and the SPA can be executed on the overlay graph again. This heuristic process is repeated until k different paths have been found, or 10 consecutive iterations of the algorithm do not find a new path. We will call this strategy topology aware, since it relies strongly on the topology for decision making.

We also evaluated an "ambiance" variant. In this variant, not only the used paths gets a burden, but also all vertices in the neighborhood, inversely proportional to the distance of the path. Vertices in the direct neighborhood get a burden of 0.5, vertices in the second order neighborhood get a burden of 0.25. Since we do not yet have geolocation data in our analysis and since this might also not be reliably available in a real-world application, we hope that this helps to spread the paths a) to different ASes and b) in terms of geolocation.

A. Metrics

1) Terminology: Every packet sent in the network uses a path to travel from the source S to the destination D. This path can then be described by the nodes the packet has traversed:

$$P_{SD} = (S, h1, h2, \ldots, D)$$

When we compare multiple paths from S to D we just denote them as P_a, P_b and so forth. In all of these paths the source and destination are fixed and therefore irrelevant. When evaluating a path under the resilience aspect, it is irrelevant how many times a certain node is used within a path, since the path is down when only one node in the path is down. Therefore we define the node set of a path P, to contain only the unique hops:

$$N(P_{SD}) = \{x \mid x \in P_{SD}\} \setminus \{S, D\}$$

2) Node Diversity: Rohrer et al. defined the path diversity metric in [17]. They argued about whether path diversity is necessary or if diversity in respect to either link or node diversity is sufficient. We share their opinion that link diversity is not sufficient, since link diverse paths, can still have one node in common and therefore have a single point of of failure. However, since the underlay graph is explicitly designed without multiple edges, it is sufficient to evaluate the diversity only in respect to node diversity. We therefore modified the metric to a Node Diversity metric:

$$ND(P_a, P_b) = 1 - \frac{|N(P_a) \cap N(P_b)|}{|N(P_a)|}$$

where $|N(P_a)| \leq |N(P_b)|$ and $N(P_a) \neq \emptyset$.

A node diversity of 1 means both paths are disjoint, a node diversity of 0 means both paths are identical. There was no definition for a direct connection between source and destination since this path has an empty node set. When both node sets are empty, we defined a node diversity of 0, because both paths would be identically on our graph without multiple edges between vertices. If only one of the two node sets is empty, the paths are disjoint, and therefore we define a node diversity of 1.

3) Effective Node Diversity: The authors of [17] extended the path diversity metric to the Effective Path Diversity (EPD) metric. It allows to evaluate the diversity of a set of multiple paths. With the same argumentation as above, we modified the metric to the Effective Node Diversity (END):

$$END(P_a, P_b, \ldots) = 1 - e^{-\lambda k_{sd}}$$

where

$$k_{sd} = \sum_{i=1}^{k} ND_{min}(P_i)$$

and λ is a constant in the range of $(0, \infty)$ which describes the marginal utility of an additional path. $\lambda < 1$ indicates a higher marginal utility for an additional path, while $\lambda > 1$ indicates a lower marginal utility. In our evaluation we use $\lambda = 1$. $ND_{min}(P_i)$ is the minimum node diversity of path i, compared to all previous paths. Due to the symmetric characteristic of the Node Diversity metric, the END is not dependent on the order the paths that are compared. It is important to note, that Node Diversity returns only a value higher than zero, if the path is disjoint in at least one node. So every additional path improves the END, when it differs in at least one hop from the previous paths. This metric will return a value from 0 to 1, while 0 means all paths in the set are identical and 1 would require an infinity number of disjoint paths.

4) Failure Probability: Under the assumption that all vertices fail independent of each other with the same probability p, the framework can calculate the probability that all paths between two vertices fail. With the same argument as for the Node Diversity, we are only interested in the failure probability of the vertices between source and destination. Every vertex v is therefore a random variable which is 1 for vertex failures. These random variables have a Bernoulli distribution with probability p. The probability that a vertex fails in a path, is then a random variable P with a binomial distribution: $P \sim B(l,p)$ where l is the number of distinct vertices in the path. Let F_i be the event that path i fails and W_i the event that path i works. The probability that the path works is then the probability that none vertex fails: $Pr[W_i] = Pr[P = 0] = \binom{l}{0} * p^0(1-p)^{l-0} = (1-p)^l$. It is obviously, that a longer path is more likely to fail than a short path. The probability that at least one of the n paths works can be calculated with the inclusion-exclusion principle:

$$Pr\left[\bigcup_{i=1}^{n} W_i\right] = \sum_{k=1}^{n}\left((-1)^{k-1}\sum_{\substack{I\subseteq\{1,\ldots,n\}\\|I|=k}} Pr\left[\bigcap_{i\in I} W_i\right]\right)$$

Since we are interested in the probability that all paths fails, we can use the complementary of the previous event:

$$Pr\left[\bigcap_{i=1}^{n} F_i\right] = 1 - Pr\left[\bigcup_{i=1}^{n} W_i\right]$$

The probability $Pr\left[\bigcap_{i\in I} W_i\right]$ that multiple paths are working, can then be calculated using a similar binomial distribution as for only one path: $Pr\left[\bigcap_{i\in I} W_i\right] = (1-p)^m$ where m is the number of distinct vertices in all paths.

The benefit of this metric is that it considers the path length and the number of shared nodes in a set of paths. Let us assume we have two set of paths from A to B. The first set M_1 contains 3 completely disjoint paths of length 10, while set M_2 contains 3 completely disjoint paths of length 5. While the END would rate both sets equally diverse, it is obvious that M_2 is preferable over M_1 in respect to resilience. This is expressed by the lower failure probability of set M_2.

The failure probability is also dependent on the granularity of the graph. When evaluated on an IP level graph with resolved IP aliases, it computes the Failure Probability under the assumption that each router will fail independent with probability p. When calculated on an AS level graph, it assumes that each AS will fail with the probability p.

VII. Results

A. Comparing the algorithms

To identify which algorithm provides the most diverse paths, we will request 5 different paths for each of the 15688 source-destination pairs in our dataset. The random strategy using exactly one relay, and the EDR found for only 28 pairs less than 5 paths. This is a result of some incomplete traceroutes in the dataset. The modified Suurballe found with $\theta = 0.5$ for 369 pairs (2.4%) less than 5 paths. A higher value of the parameter θ will find more paths, because it will prefer longer but distinct paths. For $\theta = 10$ only 59 pairs have less than 5 alternative paths.

The paths found by the random strategy have an average length of 25.9 hops and the paths found by the EDR a similar length of 25.6 hops. For the modified Suurballe's algorithm the parameter θ influences the path length. Using $\theta = 0.5$ finds shorter paths with an average length of only 22 hops, while $\theta = 10$ increases it to 26.7 hops, which is similar to the other algorithms. The advantage of the higher value of θ is the increased diversity, which is shown in the histogram of the END metric in Fig. 3. It shows that the most diverse paths are found by the modified Suurballe's algorithm for $\theta = 10$ with an average value of 0.87. With a lower value of θ, the paths are shorter and slightly less diverse with an average value of 0.84. θ allows to choose a tradeoff between diversity and path length. But even the shorter paths of Suurballe with $\theta = 0.5$ are more diverse than the paths found by EDR with an average END of 0.75 and the random algorithm with a value of 0.61.

The ambiance version of Suurballe's algorithm produces for higher values of θ less diverse and longer paths than the normal version. Smaller values of θ produces more diverse paths than the normal version with the same parameter, but they are also longer. We could also archive this result by increasing the θ parameter, since it allows to balance between path length and diversity. Because the ambiance version showed also in other evaluations and metrics slightly lower results, we will omit this algorithm in the rest of our evaluation.

The failure probability metric does not only consider the number of shared vertices like the END metric, but also considers the path length. In Fig. 4 we also added the metric for using only the direct connection without the overlay network. A 1% failure probability of each vertex correlates to the failure of about 100 routers in the underlay network. With this failure probability, the average probability to disconnect the source from the destination without an overlay network is at 13.2%. Using 5 random alternative paths, reduces this risk already to 7.7%, while EDR reduces this risk further to 6.3%. Suurballe minimizes this risk to 5.1% for $\theta = 0.5$ and 4.7% for $\theta = 10$. However, for a higher failure probabilities of each vertex, the path length is getting more important than the path diversity. With a failure probability of 10% for each vertex, the shorter but less diverse paths selected by the lower value of θ are less likely to fail (62.1% vs 65.7%). The random strategy with 70.2% and the direct connection with 76% are more likely to fail.

B. Benefit of multiple hops vs 1-hop

Previous work has shown that the number of shared routers for single- and multi-hop routing are nearly identical, when comparing the alternative paths to the direct connection [18]. However, we will demonstrate that multi-hop routing provides a better failure resistance than single-hop routing, when using a topology-aware algorithm on a real-world dataset. We will

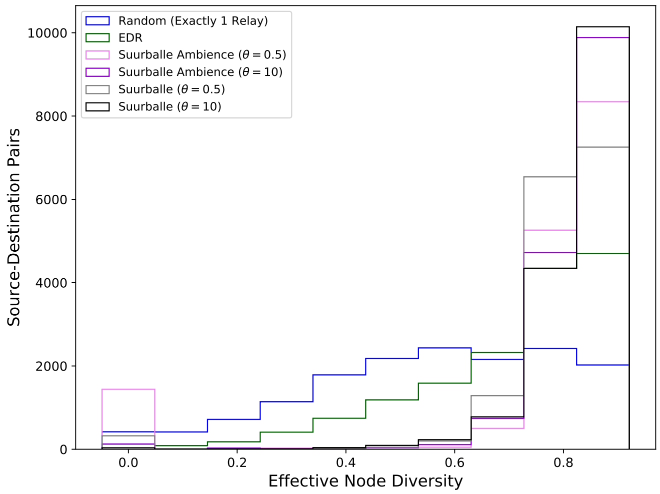

Fig. 3. Effective Node Diversity (END) for different algorithms.

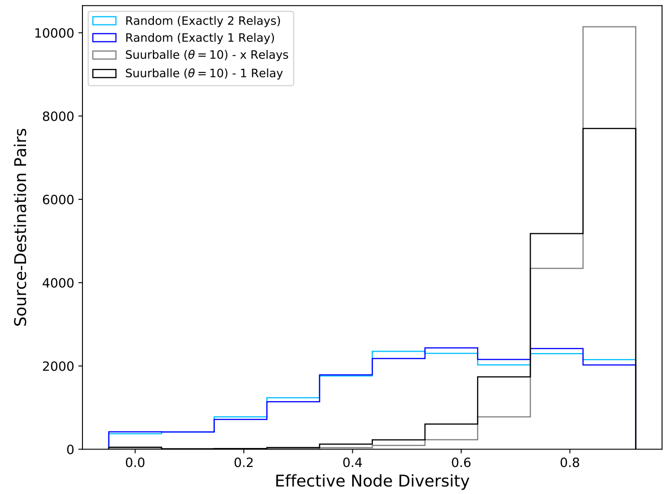

Fig. 5. Effective Node Diversity (END) for single-hop vs multi-hop.

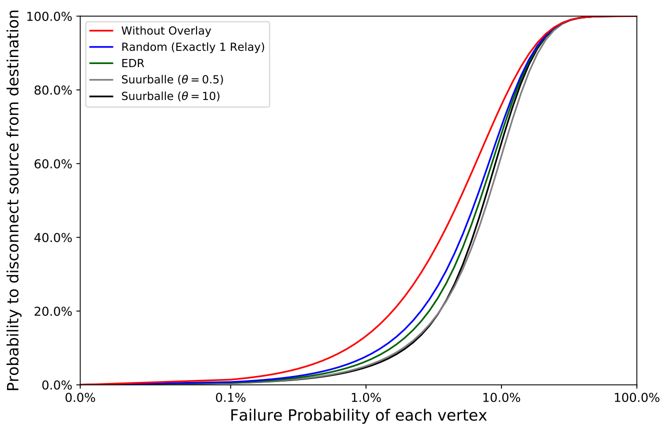

Fig. 4. Failure Probability for different algorithms.

also consider the destination AS in our metric evaluations, which was explicitly excluded in the previous work.

For the evaluations we compare the random algorithm using exactly one or two relays, with the modified Suurballe ($\theta = 10$). The EDR is not defined for multiple relays and therefore omitted. By default, Suurballe uses a flexible number of relays. To limit it to maximal one relay, we will remove all edges between relays in the overlay graph. Since the algorithm will search for paths in the overlay from source to destination, it will then return only paths with one relay.

The END metric in Fig. 5 shows that the random algorithm does not produce more diverse path with multihop routing. Both versions have an average END of about 0.60. However, the average path length increases from 25.9 to 37.5 with two relays. This leads to a higher failure probability of 8% instead of 7.7% even with a relative low failure probability of 1% for each vertex. With a failure probability of 10% this gap increases from 70.2% to 74.3% when using multiple relays. In contrast, the modified Suurballe, benefits from using multiple relays. Suurballe uses 1.57 relays on average for the alternative paths. Using multiple relays, increases the END metric in Fig. 5 from 0.83 to 0.87. Similar to the random algorithm, the path length increases. However, the algorithm takes the path length into consideration. Therefore the average path length increases only from 23.5 hops to 26.7 hops. Since this is a smaller increase than for the random algorithm, the failure probability shows an improvement for using multiple relays. The failure probability decreases from 5.0% to 4.7% when assuming a failure probability of 1% for each vertex and using multiple relays. Only with a failure probability of 10% for each vertex, the probability to disconnect the source from the destination increases from 63.5% to 65.7%.

As a conclusion, topology-aware path selection algorithms can benefit from multiple relays when assuming a low failure probability for each vertex. Additionally, using multiple relays adds a further dimension of flexibility to route selection. This can be used to reroute around censorship.

C. Granularity - from AS to IP

As stated in Section IV, the underlay graph can represent the network topology at different levels of granularity. While a more fine-grained topology like the IP level is more accurate, it also needs a more detailed data source and produces a larger underlay graph. In this Section we want to evaluate if a more coarse-grained topology like AS level is also sufficient for topology-aware path selection strategies. Therefore, we execute the algorithms on the same topology at two different levels: IP and AS. We then evaluate the diversity at IP level, since it is the most accurate representation of the topology in our data.

At AS level, the overlay graph has the same properties. In the underlay, every vertex represents an AS instead of an IP. The resulting graph has only 1464 vertices instead of 10232 and only 4185 edges instead of 23995. The average hop length decreases from 14 to 5.83 hops. We could not map the IP address of 480 hops to an AS number. We let these nodes at

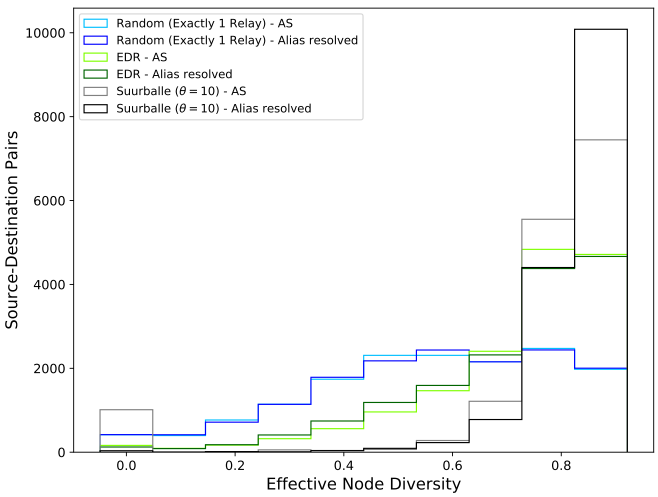

Fig. 6. Effective Node Diversity (END) at different levels of granularity.

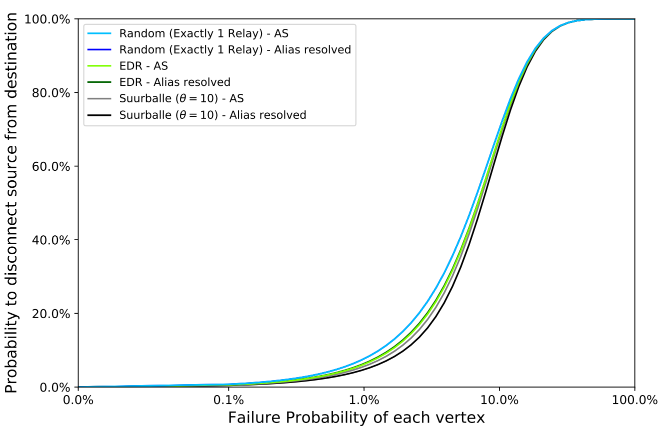

Fig. 7. Failure Probability at different levels of granularity.

IP level to prevent incomplete paths. With an ideal mapping of 100% the effects of the following results would be intensified.

The END in Fig. 6 shows that the random algorithm does not benefit from a more accurate topology representation. The average END value for this strategy is at 0.60. Similar, the EDR does not benefit from using IP level. On the contrary, it is even better at AS level with an average value of 0.76 instead of 0.75. However, a topology-aware path selection algorithm benefits from a more accurate topology. The modified Suurballe improves from 0.80 to 0.87. A similar results shows the failure probability. The lines for the random algorithm and the EDR in Fig 7 overlap.

As a conclusion, the granularity is not important when using a simple algorithm that cannot utilize it. However, a topology-aware path selection algorithm benefits from a more accurate topology. Since traceroutes are often the first choice to build up the topology dataset, we recommend to work at this level and not simplifying it to AS level.

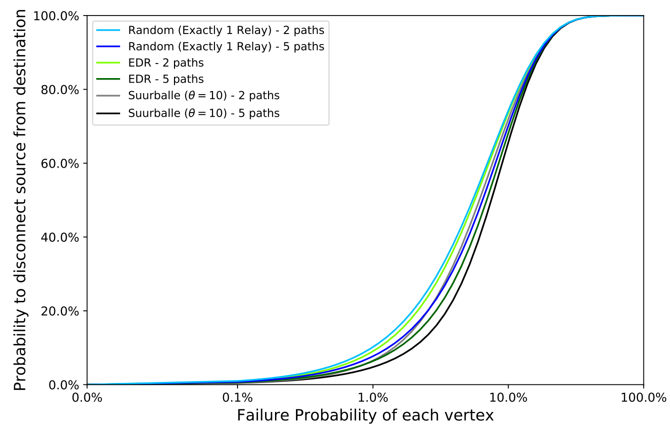

Fig. 8. Failure probability for different number of paths.

D. Number of backup paths

We assumed that increasing the number of alternative paths also increases the resilience and path diversity. To evaluate this theory we configured the algorithms to return two and five paths, including the direct connection.

As the failure probability in Fig. 8 shows, we can improve the resilience by using multiple alternative paths. When using the random algorithm with 5 paths, the failure probability is nearly at the same level as the topology-aware path selection algorithm with 2 paths. Especially with a higher failure probability for each vertex, the random strategy outperforms with 5 paths, the modified Suurballe with only 2 paths. But the EDR algorithm and the modified Suurballe also benefit from more paths. For a failure probability for each vertex of 1%, the probability to disconnect the source from the destination decreases for the random algorithm from 10.2 to 7.7%, the EDR from 9 to 6.4% and the modified Suurballe from 6.5% to 4.7%. A similar result was shown by the END metric, where increasing the number of paths increases the diversity.

As a conclusion, providing multiple alternative path increases the resilience of the overall network for all route selection algorithm. After a failure, all paths can be used in parallel for TCP packets, until a working path has been found.

VIII. Conclusion

We have demonstrated that overlay networks with relay servers can provide alternative paths, which are diverse and could improve resilience. The common approach for multipath where one random relay is used already reduces the failure probability. More advanced strategies can reduce it more. We have also shown that the modified Suurballe's algorithm finds not only more diverse paths than the random algorithm, but also shorter ones. With the parameter theta every application can balance between the length of the paths and the required diversity. This could also be done dynamically and depending on the failure scenario. The ambiance version showed slightly worse results, but might perform better when geolocation would be anticipated in the analysis.

We have further demonstrated that multihop routing shows no benefit when using a random algorithm. Yet a topology-aware algorithm, can use multiple relays to find more diverse paths. Using multiple relays provides an additional degree of flexibility to route around suspicious routers, failures, or anomalies. We illustrated that the granularity of the topology is an important factor when using topology-aware path selection algorithm and that using multiple backup paths can improve the resilience of all route selection algorithms.

In future work, we want to include geographical and latency information in the framework in order to anticipate geographical co-location. Given this information, we could analyze interesting algorithms like GeoDivRP. Furthermore we consider to investigate optimal relay placement and develop new path selection strategies.

References

[1] D. Andersen, H. Balakrishnan, F. Kaashoek, and R. Morris, "Resilient overlay networks," in *Proceedings of the Eighteenth ACM Symposium on Operating Systems Principles*, ser. SOSP '01. New York, NY, USA: ACM, 2001, pp. 131–145. [Online]. Available: http://doi.acm.org/10.1145/502034.502048

[2] D. Raposo, M. L. Pardal, L. Rodrigues, and M. Correia, "Machete: Multi-path communication for security," in *2016 IEEE 15th International Symposium on Network Computing and Applications (NCA)*, Oct 2016, pp. 60–67.

[3] C. X. Cai, F. Le, X. Sun, G. G. Xie, H. Jamjoom, and R. H. Campbell, "Cronets: Cloud-routed overlay networks," in *2016 IEEE 36th International Conference on Distributed Computing Systems (ICDCS)*, June 2016, pp. 67–77.

[4] B. Engeser, "Informed route selection strategies for multipath routing," Master's Thesis, Technische Universität München, 2016.

[5] T. Fei, S. Tao, L. Gao, and R. Guerin, "How to select a good alternate path in large peer-to-peer systems?" in *Proceedings IEEE INFOCOM 2006. 25TH IEEE International Conference on Computer Communications*, April 2006, pp. 1–13.

[6] T. Fei, S. Tao, L. Gao, R. Guerin, and Z. l. Zhang, "Light-weight overlay path selection in a peer-to-peer environment," in *Proceedings IEEE INFOCOM 2006. 25TH IEEE International Conference on Computer Communications*, April 2006, pp. 1–6.

[7] Y. Cheng, M. M. Rahman, S. Gangadhar, M. J. Alenazi, and J. P. Sterbenz, "Cross-layer framework with geodiverse routing in software-defined networking," in *Network and Service Management (CNSM), 2015 11th International Conference on*. IEEE, 2015, pp. 348–353.

[8] Y. Cheng, M. T. Gardner, J. Li, R. May, D. Medhi, and J. P. G. Sterbenz, "Optimised heuristics for a geodiverse routing protocol," in *2014 10th International Conference on the Design of Reliable Communication Networks (DRCN)*, April 2014, pp. 1–9.

[9] M. Menth and R. Martin, "Network resilience through multi-topology routing," in *DRCN 2005). Proceedings.5th International Workshop on Design of Reliable Communication Networks, 2005.*, Oct 2005, pp. 271–277.

[10] M. Motiwala, M. Elmore, N. Feamster, and S. Vempala, "Path splicing," *SIGCOMM Comput. Commun. Rev.*, vol. 38, no. 4, pp. 27–38, Aug. 2008. [Online]. Available: http://doi.acm.org/10.1145/1402946.1402963

[11] A. Babay, C. Danilov, J. Lane, M. Miskin-Amir, D. Obenshain, J. Schultz, J. Stanton, T. Tantillo, and Y. Amir, "Structured overlay networks for a new generation of internet services," in *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*, June 2017, pp. 1771–1779.

[12] J. Han and F. Jahanian, "Impact of path diversity on multi-homed and overlay networks," in *International Conference on Dependable Systems and Networks, 2004*, June 2004, pp. 29–38.

[13] (2017) Planetlab. [Online]. Available: https://www.planet-lab.org/

[14] H. V. Madhyastha, T. Isdal, M. Piatek, C. Dixon, T. Anderson, A. Krishnamurthy, and A. Venkataramani, "iplane: An information plane for distributed services," in *Proceedings of the 7th Symposium on Operating Systems Design and Implementation*, ser. OSDI '06. Berkeley, CA, USA: USENIX Association, 2006, pp. 367–380. [Online]. Available: http://dl.acm.org/citation.cfm?id=1298455.1298490

[15] M. H. Gunes and K. Sarac, "Resolving ip aliases in building traceroute-based internet maps," *IEEE/ACM Transactions on Networking*, vol. 17, no. 6, pp. 1738–1751, Dec 2009.

[16] J. W. Suurballe, "Disjoint paths in a network," *Networks*, vol. 4, no. 2, pp. 125–145, 1974. [Online]. Available: http://dx.doi.org/10.1002/net.3230040204

[17] J. P. Rohrer, R. Naidu, and J. P. G. Sterbenz, "Multipath at the transport layer: An end-to-end resilience mechanism," in *2009 International Conference on Ultra Modern Telecommunications Workshops*, Oct 2009, pp. 1–7.

[18] J. Han, D. Watson, and F. Jahanian, "Topology aware overlay networks," in *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, vol. 4. IEEE, 2005, pp. 2554–2565.