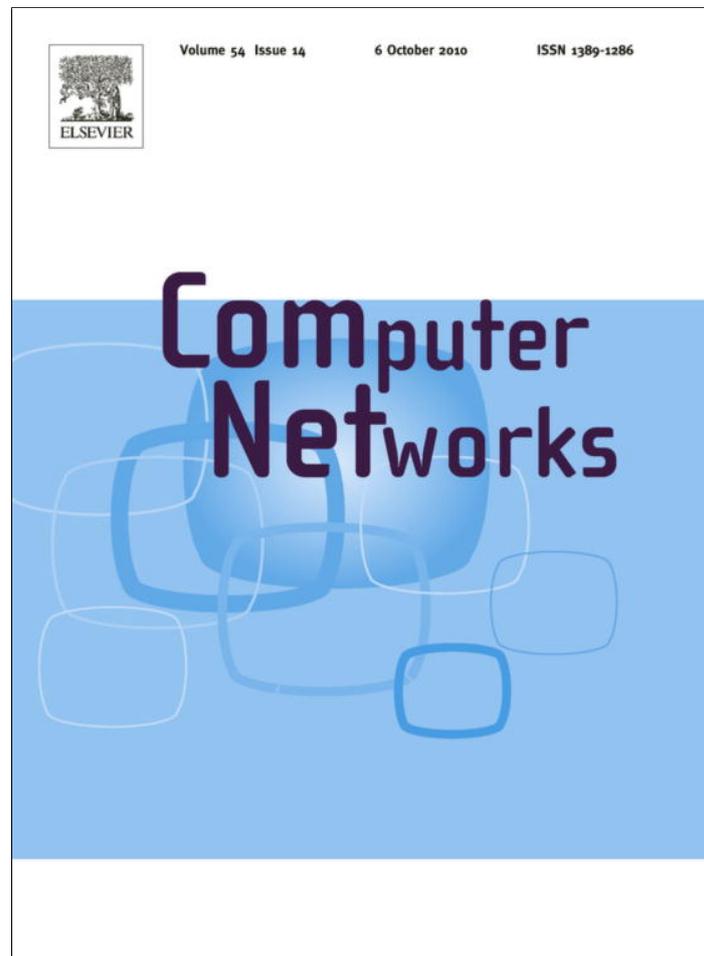


Provided for non-commercial research and education use.  
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at ScienceDirect

## Computer Networks

journal homepage: [www.elsevier.com/locate/comnet](http://www.elsevier.com/locate/comnet)

## Impact of routing parameters on route diversity and path inflation

Wolfgang Mühlbauer<sup>a,\*</sup>, Steve Uhlig<sup>b</sup>, Anja Feldmann<sup>b</sup>, Olaf Maennel<sup>c</sup>,  
Bruno Quoitin<sup>d</sup>, Bingjie Fu<sup>e</sup><sup>a</sup>ETH Zürich, Inst.f.Techn.Informatik u.Kommunik.Netze ETZ G 90, Gloriastrasse 35, 8092 Zürich, Germany<sup>b</sup>TU Berlin, Deutsche Telekom Laboratories, Berlin, Germany<sup>c</sup>Loughborough University, Leicestershire, United Kingdom<sup>d</sup>Université Catholique de Louvain, Computer Science and Engineering Dept., Louvain-la-Neuve, Belgium<sup>e</sup>Alcatel Lucent, The Hague, The Netherlands

## ARTICLE INFO

## Article history:

Received 18 September 2009

Received in revised form 22 January 2010

Accepted 6 April 2010

Available online 11 April 2010

Responsible Editor: J. Zhang

## Keywords:

Inter-domain routing

BGP

Sensitivity analysis

Route diversity

Path inflation

## ABSTRACT

Years after the initial development of the current routing protocols we still lack an understanding of the impact of various parameters on the routes chosen in today's Internet. Network operators are struggling to optimize their routing, but the effectiveness of those efforts is limited.

In this article, we study sensitivity of routing stretch and diversity metrics to factors such as policies, topology, IGP weights, etc. using statistical techniques. We confirm previous findings that routing policies and AS size (in number of routers) are the dominating factors. Surprisingly, we find that intra-domain factors only have marginal impact on global path properties.

Moreover, we study path inflation by comparing against the paths that are shortest in terms of AS-level/router-level hops or geographic distances. Overall, the majority of routes incur reasonable stretch. From the experience with our Internet-scale simulations, we find it hard to globally optimize path selection with respect to the geographic length of the routes, as long as inter-domain routing protocols do not include an explicit notion of geographic distance in the routing information.

© 2010 Elsevier B.V. All rights reserved.

## 1. Introduction

Routing in the Internet is inherently complex. It is controlled by diverse policies, decided *locally* by each Autonomous System (AS), but acting *globally* across the entire system [1]. Furthermore, it depends on protocols for routing between and within individual ASs, on the router-level topology inside Internet domains, and on the peering structure between ASs.

In the past, models have been suggested to verify the correctness of routing inside an autonomous system [2],

to predict the paths between two ASs or routers [3], etc. These models require topology information as input, typically gathered from the routing protocols themselves. The information collected from the current routing system is not only topologically incomplete and biased [4], but it also does not provide detailed enough information to estimate the path diversity available in today's Internet [3]. Even today, there is still no thorough understanding of which and how routing parameters impact the diversity and optimality of the chosen routes.

In this article, we make an important contribution towards understanding how sensitive route stretch and diversity metrics are two factors such as policies, number of routers per AS, IGP weights, location of peerings, iBGP connectivity, etc. To comprehensively explore all factor settings and their consequences on the routes computed by BGP (Border Gateway Protocol), we rely on *full factorial*

\* Corresponding author. Tel.: +49 89 8353 58378.

E-mail addresses: [wolfgang.muehlbauer@tik.ee.ethz.ch](mailto:wolfgang.muehlbauer@tik.ee.ethz.ch) (W. Mühlbauer), [steve@net.t-labs.tu-berlin.de](mailto:steve@net.t-labs.tu-berlin.de) (S. Uhlig), [anja@net.t-labs.tu-berlin.de](mailto:anja@net.t-labs.tu-berlin.de) (A. Feldmann), [olaf@maennel.net](mailto:olaf@maennel.net) (O. Maennel), [bruno.quoitin@uclouvain.be](mailto:bruno.quoitin@uclouvain.be) (B. Quoitin), [bingjie.fu@alcatel-lucent.com](mailto:bingjie.fu@alcatel-lucent.com) (B. Fu).

design. One major contribution of this article is to quantify sensitivity using *analysis of variance* (ANOVA) [5].

Given the inherent limitations of observable routing data, we rely on simulations. Within simulations we can control all parameters, and can compute any desired metric on the simulated routes, since we created the topology and therefore know the routing tables of each individual router. Simulations allow us to compare paths chosen by individual routers with paths that are globally optimal in terms of AS-level, router-level hops, or geographical distance. To the best of our knowledge there is no work on Internet-scale simulations, which is comparable to ours in terms of comprehensiveness, level-of-detail in modeling and the size of the used topologies.

The benefits of our sensitivity analysis are twofold: first, they give insight into the relevant factors, i.e., the ones which need to be accurately reproduced when modeling routing in the Internet – is it really true that state-of-the-art routing models neglect aspects that do have an influence on route computation? Second, both the sensitivity analysis and the simulation framework, we propose in this article, will prove useful for evaluating and comparing new routing protocols and architectures, and will provide hints on how to improve the path selection of BGP.

Surprisingly, we find that the impact of intra-domain parameters, including IGP weights or iBGP connectivity, on *global* route propagation is low compared to other parameters. Note that this does not imply that individual routes chosen by BGP are insensitive to hot-potato routing [6,7]. Consistent with the design of BGP, we find that routing policies and the size of ASs in terms of the number of routers are the dominating factors. Hence, future routing models should focus on these two aspects.

Moreover, we find that the majority of routes incur reasonable stretch in terms of AS-level/router-level hops or geographic distance. Policies slightly increase path length. Overall we did not manage to optimize geographical distance within the space of current configuration alternatives in our simulations. However BGP, and in particular BGP policies, cannot be blamed for the sub-optimality of the current paths. BGP is a policy-routing protocol, for which short distances have not been a design constraint. We will face this limitation of BGP if end-to-end path quality is to become more important in the future.

The structure of this article is as follows. We review the Internet routing system (Section 2) and give an overview of our experimental design (Section 3), before we present implementation details (Section 4). Section 5 presents our comprehensive sensitivity analysis, followed by the study of (sub) optimality of selected routes in Section 6. Finally, we discuss related work (Section 7) and conclude (Section 8).

## 2. Background

When reviewing in this section the current routing system, we focus on the aspects that impact the choice of routes: the structure of the Internet in Section 2.1 and BGP route selection and propagation in Section 2.2.

### 2.1. Structure of the internet

The Internet is divided into a collection of independently administered autonomous systems (ASs). Routing through the Internet is accomplished on a prefix basis and depends on protocols for routing within individual ASs. We review the AS-level and the intra-domain topology as well as the AS-interconnectivity at the router-level.

#### 2.1.1. AS-level graph

The Internet has a tiered structure [8] which reflects AS business relationships, e.g., [9,10,8,11]. A few tier-1 Internet Service Providers (ISP) form the core. A larger number of transit providers buy service from other providers, including tier-1 providers, and provide connectivity to other ASs. Stub ASs get their connectivity from transit or tier-1 providers.

Each business agreement between two ASs corresponds to an AS-level edge and implies some physical connectivity, one or multiple links, between the ASs. It is implemented by imposing routing policies for these links that determine which paths are propagated and selected [9]. Indeed, business relationships lead to constraints on path propagation, e.g., a path learned from a provider is not exported to another provider. While routing policies can be rather specialized [12], it is common to classify AS relationships into three main types: customer–provider (c2p), peer-to-peer (p2p), siblings (s2s) [13].

#### 2.1.2. Intra-domain graph

ASs are not atomic entities, but are composed of many routers [14,15]. Internet Service Providers (ISP) employ an intra-domain routing protocol, such as OSPF or IS-IS, to select paths inside their network [16]: routers exchange link-state information and forward packets along shortest paths, based on the sum of the link weights chosen by the ISP. As such, the internal structure of each network together with its choice of IGP weights impacts the BGP routing decisions.

External BGP sessions are not established between all routers of an AS but only on the so-called border routers. To redistribute these routes to all routers, iBGP sessions are used between all BGP routers on top of the router-level topology. There are multiple ways to realize such iBGP structure, ranging from an iBGP full mesh to route-reflectors or confederations. The latter two are more scalable but reduce the visibility of routes within the AS [17] compared to a full mesh.

#### 2.1.3. Inter-domain connectivity

Peering agreements between ASs usually enforce multiple physical links at multiple different locations in different regions. This has the advantage of minimizing the time a packet stays within an AS as long as hot-potato routing [6,18] is used. Multiple physical links at different locations are also often used to increase reliability, e.g., in the form of customer backup links. Note that these multiple links between two ASs are not reflected in any AS-level topology.

### 2.2. BGP path selection and propagation

In BGP, routers establish BGP sessions over which they receive and propagate routes for destination prefixes. We distinguish between external BGP (eBGP) sessions over inter-domain links and internal BGP (iBGP) sessions between the routers of a given AS. As illustrated in Fig. 1, once a route advertisement coming from a neighbor (called peer) is accepted by the configured input filters, it is placed in the incoming Routing Information Base (RIB-In) for the peer, possibly after some of the route attributes have been modified according to the local routing policies.

The BGP decision process is used to select the best route for each prefix from among the available routes. For this purpose the BGP decision process considers several of the BGP route attributes, amongst others shorter AS-paths are preferred. Later on in the decision process, routes are ranked according to the IGP cost of the intra-domain path towards the next-hop, preferring routes with smaller IGP cost. Therefore, intra-domain routing protocols do influence the path-selection process in BGP. Finally, if there is still more than a single route left, the router breaks ties, for example by selecting the route to the neighbor which has the lowest router-id (typically one of its IP addresses).

Then the selected route is placed into the BGP routing table. Finally, administrators may specify output filters for each peer, which are used to decide which best routes to propagate to a BGP neighbor and whether to change some of their attributes.

### 3. Factorial design

To comprehensively explore all factor settings and their consequences on the routes computed by BGP, we rely on factorial design. A *factorial design* [5] experiment allows to study the effect of each *parameter* (factor) on the evaluation *metrics* (response variables), as well as the effects of interactions between factors on response variables. For each factor there is a set of discrete possible values or *levels*. A *full factorial experiment* explores all combinations of these levels across all factors. We rely on simulations, since within simulations we can control all parameters, and can compute any desired metric on the simulated routes. Analysis of Variance (ANOVA) provides statistic means to describe the impact of individual factors on the metrics. We now provide a short introduction to ANOVA and describe the metrics and factors for our sensitivity study.

### 3.1. Analysis of Variance (ANOVA)

ANOVA is an analysis of the variation present in an experiment. It helps to determine which factors strongly contribute to the metric variations observed across the different level settings. The total *sum of squares*  $SS_{total}$  measures the total, absolute variability across *all* the obtained metric values  $x_i$  in the sample (where  $\bar{x}$  denotes the average metric value):

$$SS_{total} = \sum (x_i - \bar{x})^2 = \frac{\sum x_i^2 - (\sum x_i)^2}{n}$$

The basic idea behind ANOVA is to partition this variability of the whole data set into two components: Variability across different levels of a factor  $SS_{factors}$  and variability for different repetitions of the same level configuration  $SS_{repetitions}$ . This can be expressed as follows:

$$SS_{total} = SS_{factors} + SS_{repetitions}$$

To estimate the impact of a certain factor on the response variable, we can compute the percentage of variability, that a certain factor explains, by dividing  $SS_{factors}$  by  $SS_{total}$  (used in Section 5). Note that ANOVA is not restricted to experiments with two factors. For more details on the computation of the sum of squares, the partitioning of the variation, the regression models, etc., we refer to Jain's book [5].

ANOVA provides means to determine whether the fraction of variance that is explained by a certain factor is *statistically significant*. For this purpose, we will compute in Section 5 the *F-ratios*: the variation due to a factor divided by the variation due to experimental error or noise. The null hypothesis is that this ratio equals 1.0 and is rejected if the *F-ratio* is significantly large enough, i.e. the likelihood that it is equal to 1.0 is smaller than a certain percentage (e.g., 5%). Again, we refer the reader to Jain's book [5].

### 3.2. Metrics

We concentrate on metrics that quantify the results of route computation, path inflation, and route diversity.

**AS path length** (*ASLength*) reflects the number of AS-level hops of the AS path of the route.

**Router-level path length** (*RouterLength*) considers the number of router-level hops of the selected route.

**Geographic path length** (*GeoLength*): Our topology model includes geographic coordinates. Therefore, we

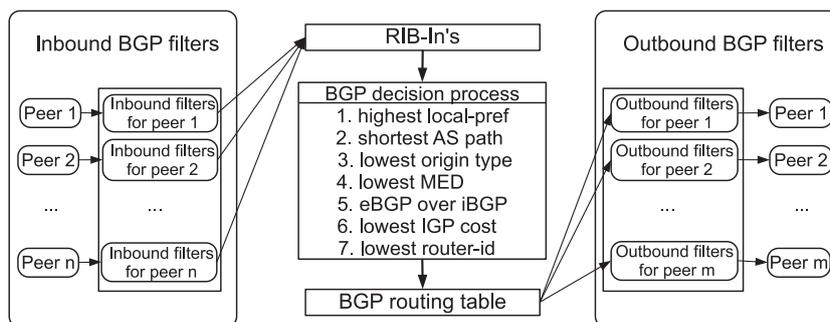


Fig. 1. Operation of a BGP router.

can compute the geographical distance for each path by summing up the distances of each link on the router-level path.

**Number of selected paths** (*NumSelPaths*) counts the total number of distinct AS paths that are selected as best routes by the routers of an AS for a single prefix.

**Number of learned paths** (*NumLearnedPaths*): Same as *NumSelPaths*, but we consider all AS paths learned by an AS.

**Ratio: # Learned paths/# neighbor ASs** (*LearnedNeighborRatio*) measures the average route diversity learned per neighbor AS. We divide *NumLearnedPaths* by the number of neighbor ASs.

**Disjointness** (*Disjointness*): For every AS, we consider all distinct AS paths, learned by its routers and compute their edge disjointness. For this we compute for each pair of AS paths  $p_1$  and  $p_2$ :

$$m = 1 - \frac{\# \text{ AS edges } p_1 \text{ and } p_2 \text{ have in common}}{\# \text{ AS edges in longer path}}.$$

We refer to the average of  $m$  over all these pairs of AS paths as *Disjointness*. A value of 0 implies that all learned AS paths are identical while a value close to 1 indicates high edge disjointness. If an AS learns a single AS path, we set the metric to 1. This metric is similar to the *Novelty* metric used by Motiwala et al. [19] to characterize AS path disjointness.

### 3.3. Factors

We rely both on *Internet-scale* simulations ( $\approx 30,000$  ASs) and on *small-sized* simulations (150 ASs). Each AS can have its own router-level topology. These router-level topologies include a location within a point of presence (PoP) for each router throughout the world. Hereby, we keep the design choices, available to network operators [20], in mind.

**Routing policies** (*Pol*): We identify two levels: no policy and routing policies according to AS relationships. This allows to study the degree of AS path inflation due to the use of routing policies [21,22] and the impact of policies on AS-level route diversity.

**AS router size** (*ASsize*) can vary between ASs. Tier-1 ASs typically have more routers in their backbone part than small transit ASs. Nevertheless, previous work [3] found that only a few routers are necessary to account for route diversity as seen from BGP data. Our level values range from 1 to 5 routers or depend on the tier of the AS.

**Multiple peering links between ASs** (*Peer*) are common in the Internet and increase the diversity of routes. [3] has shown that this factor is necessary to reproduce path diversity observed in the Internet. Our levels for this parameter range from 1 to 5 depending on the AS size.

**IGP weights** (*IGP*) often reflect the delay, the geographic distance or the link capacity or are chosen to minimize network congestion [23]. Moreover, route choices are known to be sensitive to them [7]. Our levels for IGP

weights range from random weights to uniform weights and even geographic distance.

**iBGP topology** (*iBGP*) can limit the visibility of routes inside an AS. We consider iBGP full meshes as well as route-reflector hierarchies that follow the PoP structure as suggested by configuration guidelines [24,25].

**Border routers** (*Border*) are not chosen arbitrarily in practice. Rather, each AS often has a limited number of them in specific locations. Moreover, peerings or customer-provider links are frequently established between close by routers, typically routers at the same location, in order to minimize distances and costs. We use two levels: one which tries to minimize this cost and one which selects border routers randomly.

Note that AS-level connectivity is not a factor in our design for two reasons. First, it is non-trivial to combine other connectivity models such as randomly generated AS graphs with a notion of routing policies. Second, it is widely agreed that the Internet has a tiered structure [8], which is unlikely to change in the foreseeable future. Therefore, we use an AS-level map obtained from [26] as basis for our Internet-scale simulations. For our smaller-sized simulations, we rely on AS graphs that show similar properties as those of the actual Internet.

## 4. Simulation setup and choice of levels

Now we dive into more details about the simulation setup and the generation of the topologies.

### 4.1. Simulation

Our objective is to understand how sensitive the outcome of the routing process is to the different choices of levels for the topology factors. For this purpose, we run both *Internet-scale* simulations ( $\approx 30,000$  ASs) based on collected AS-level maps from Caida [26] and *small-sized* simulations (150 ASs). To study the outcome of global route selection, we do not have to consider the BGP dynamics. Accordingly, we use the C-BGP simulator [27] in order to compute the paths that routers know once the BGP routing has converged [28].

With C-BGP it is possible to run *Internet-scale* simulations with more than 30,000 ASs and with complex inter-domain structures and inter-domain peerings. However, this takes a lot of time even if we restrict the number of prefixes to 100 and originate one prefix per AS for a selected set of ASs including tier-1, tier-2 and stub ASs. We found that such complex simulations may take up to 5 days to complete and may require more than 13 GB memory.<sup>1</sup> For this reason, we decide to run these *Internet-scale* simulations based on AS topologies derived from CAIDA [26] data of April 2009 only for a limited set of parameter settings, amongst others for *defaultSetup* and slight variations of *defaultSetup*, see Section 4.2.4.

Our full factorial design results in 216 different combinations of the levels of our six factors. For each choice of

<sup>1</sup> On AMD Opteron 865 multi-core machine with 32 GB.

factor levels, we generate the topology, execute the simulation, and compute the metrics. After all simulations have been finished and analyzed, we perform the statistical sensitivity analysis. Together with the 10 iterations we run per simulation, a full factorial design would require more than 2000 C-BGP simulations. Obviously, we have to limit the size of the topology in terms of number of ASs for our sensitivity study<sup>2</sup> in Section 5. With regards to the size of the topologies, they have to be large enough to accommodate a tiered structure and allow for complex interconnections between ASs. Yet, they need to be small enough to allow efficient simulation of *all* routing tables and *all* prefixes. We find that a single simulation with 150 ASs (*small-sized*), 150 originated prefixes and some 1000 routers uses roughly 300 MB of memory and finishes within few minutes.

Despite the reduced topology size, the computation of the metrics is time-consuming, lasting roughly 30 min per simulation, due to the computation of all shortest paths with the Floyd-Warshall algorithm<sup>3</sup> in weighted and non-weighted graphs. By parallelizing the processing, it is possible to finish one set of simulations and to compute the metrics within a few days, which is a reasonable time. For our *Internet-scale* simulations, the computation of the metrics and Dijkstra's algorithm for the 100 destination prefixes take up to several hours for a single simulation.

#### 4.2. Topology generation/choice of levels

Our guideline is a comprehensive design of simulations. Our full factorial design includes simulations configured based on configurational practices used in today's Internet (see Section 4.2.4). But we also choose levels to cover extreme cases. For example, as part of our sensitivity analysis, we run simulations where no policies are configured at all. In doing so, we are able to explore the complete space of current configuration alternatives offered by BGP.

The generation of topologies is not done in one step, rather we follow a top-down approach with three sub-tasks: generating an AS-level topology, the intra-domain topology, and the inter-domain interconnectivity.

##### 4.2.1. AS-level topology

For the AS-level topology, we either rely on an AS-level map from [26] (*Internet-scale*) or on a *small-sized* "tiered" topology with 150 ASs – 5 tier-1s, 20 tier-2s, and 125 stub ASs. While the latter does under-represent the number of stub ASs, our goal was to include a reasonable complex set of tier-1 as well as tier-2 ASs and to roughly keep the proportion of transit to stub ASs as observed in the Internet [26]. Spot checks against our *Internet-scale* simulations (see Sections 5.2 and 6) show that using down-scaled topologies does not change the results drastically.

In the *small-sized* topology, all tier-1 ASs are fully interconnected, while tier-2 ASs are randomly connected

to a subset of tier-1 ASs and also peer with a subset of each other. Stub ASs are either single- or dual-homed and connect to either tier-2 or tier-1 ASs. Note that the generation of the AS-level topology is non-deterministic. For example, the probability for a stub AS to be dual-homed is 50% and we randomly determine the number of tier-1 ASs to which a tier-2 AS attaches. In order to account for the non-determinism, we decide to run 10 iterations for each factor combination of our factorial design.

If the factor *Pol* is active, we configure policies according to the tiered architecture. For example, edges between two tier-2 ASs are peer-to-peer (p2p) links, while some edges between a tier-1 and a tier-2 are provider-to-customer (p2c) links and others are p2p links. Routing policies are realized via BGP filters, communities, and local-preference values and implement the no-export and no-valley property of AS relationships [9].

##### 4.2.2. Intra-domain topology

Since intra-domain topologies and link weights of actual ASs are unknown and cannot be inferred easily [14,15], we rely on the functionality offered by IGen [29], a structural topology generator which produces plausible topologies using network design heuristics [20].

The level of the factor *ASsize* determines the number of routers for each AS. We consider four levels, three with a constant numbers of routers inside all ASs, i.e., 1, 2, 5. The fourth level varies the number of routers according to the position of the AS in the AS hierarchy, i.e., 30 routers for tier-1's, 15 for tier-2's and 1 per stub AS.

To be able to consider the geographic distance for the factor *IGP*, each router is placed at a specific geographic location (longitude, latitude).<sup>4</sup> While large ASs are assumed to have routers all over the world, stub ASs are only present in a geographically limited area. Accordingly, the routers of tier-1 ASs and some tier-2 ASs are distributed world-wide, while other tier-2s and stub ASs are mainly restricted to a single continent.

As next step we setup the physical links by first constructing Points of Presence (PoP) and then interconnecting them. Each PoP corresponds to a cluster of geographically-close nodes identified with the help of the K-Medoid algorithm [30]. Within each PoP we distinguish between backbone nodes for connecting to other PoPs, and access nodes. The backbone nodes within the different PoPs are interconnected via a clique. The access nodes are connected to the backbone nodes within the PoP using at least two edges. We need this complex intra-AS topology only if the AS is composed of a significant number of nodes and the level of the factor *iBGP* is "route reflector". If the level of *iBGP* is "route reflector", we configure a set of iBGP sessions which follow the physical topology. Otherwise we choose a full mesh of iBGP sessions.

The three levels of the factor *IGP* are: "uniform" which assigns uniform costs of 1 to all intra-domain links; "random" which assigns random weights of 1 to 100; and "geographic" which assigns weights that correspond

<sup>2</sup> For later discussions on "optimality" we rely on the Internet-scale simulations.

<sup>3</sup> Here we use an Intel Xeon Quad-Core platform with 2.4 GHz processor and 8 GB of memory.

<sup>4</sup> The locations in our simulations do not correspond to actual locations of routers in the Internet. However, the core results of our sensitivity study still hold for the actual Internet.

to the geographical distance between the source and the destination.

#### 4.2.3. Inter-domain connectivity

As next step, we need to determine the router-level connectivity for AS-level edges. The factor *Peer* determines how many router-level peering links are to be used for each AS-level edge. We again choose four levels, three with constant numbers of links per edge, i.e., 1, 2, 5. The fourth level varies the number of links according to the number of routers in the incident ASs. The larger the number of routers, the larger the number of peering links.

The factor *Border* determines on which routers the peerings are placed: either randomly or according to geographic distance. The base assumption of the latter case is that two domains prefer to connect at places which are geographically close to each other. For this purpose, we search among the  $N_i \times N_j$  routers of AS  $i$  and  $j$  and establish the link between the geographically closest ones.

#### 4.2.4. Default simulation

In spite of recent work on topology discovery [14,15] and policy inference [10,9,31], the actual routing policies and the router-level topology of the Internet are largely unknown. Yet, in order to obtain a simulation setup that is as realistic as possible, we consider configuration practices that correspond to common design guidelines in today's Internet and try to consider insights obtained from measurements of ISP topologies [14,15].

For the remainder of this article, we will frequently refer to *defaultSetup*, a choice of factor levels that is as close as possible to actual configurations in the Internet. Accordingly, such a setting uses AS business relationships, has route reflectors inside large ASs and considers geographic locality when determining inter-domain connections and assigning IGP weights. The tier-1 and tier-2 ASs are reproduced with multiple routers, belonging to different route reflector clusters. In addition to *small-sized* simulation setups with 150 ASs, we also run *Internet-scale* simulations for *defaultSetup*, see Sections 5.2 and 6.

## 5. Sensitivity analysis

After identifying our metrics and factors, we now examine the results of the sensitivity analysis.

### 5.1. ANOVA

We perform *small-sized* simulations for all possible combinations of factors and levels as outlined in Section 3. We then rely on the ANOVA technique, see Section 3.1, to identify those factors that strongly contribute to the variations of the metrics observed across the different choices of levels.

While the outcome of simulations with C-BGP is deterministic, a certain non-determinism is involved in the setup of our topologies, see Section 4.2.1. For example, when generating the underlying AS-level topology we only enforce certain properties such as a tiered structure but randomly determine the actual interconnection of ASs.

Likewise, our factors only describe how many routers and what strategy to use for selecting border routers, but not which exact border routers to use for implementing the peering between two ASs. To capture such non-determinism, we run 10 iterations per factor combination. If we keep the values of levels constant, we observe only marginal differences in the results from one simulation to another. Spot-checks with even more than 10 iterations do not change this picture. Hence, we regard 10 repetitions per factor combination as sufficient to determine those factors that significantly impact route propagation. Additionally, confining ourselves to 10 iterations, allows to complete simulations for all our 216 different combinations of factors and levels within a reasonable time. Note that the variability in the results for repetitions of a given simulation setup is captured in the *residual* values.

To explore whether differences in the metrics are caused by the configuration factors or by statistical noise (due to running multiple simulation instances for each factor combination) we partition the variance observed for each metric into its components using *sum of squares*. Moreover, we check for *statistical significance*, relying on *F*-tests, see Section 3.1.

The influence of the six factors on each of the eight metrics is shown in Fig. 2, while Table 1 presents the corresponding numbers and F-values. The percentage values represent the fraction that a certain factor contributes to the total variance of the metric. Hence, a high percentage suggests that a given metric is very sensitive to that factor.

We point out that the values in each row of Table 1 do not always add up to 100%. For example, the fraction of variance for *NumSelPaths* (first row of Table 1) which can be explained by our six factors and the residuals is only 84.1%. The remaining 15.9% are due to *interactions*. Such interactions occur when it is not possible to distinguish the impact of two or more factors. In the above example the value of the metric *NumSelPaths* is sensitive to the combination of two factors: routing policies (*Pol*) and the size of ASs (*ASsize*). However, we find that the impact of such “intertwined” factors is low when compared to the impact of each individual factor.

We also observe that the residual values for the majority of our metrics are rather small (less than 4%), see Table 1. This suggests that the error or statistical noise of our setup is low compared to the variation across different factor configurations. This supports our choice of the factors. The factors indeed impact route computations more than the statistical variations. Nevertheless, for *Disjointness* and *ASLength*, we observe higher residual values with 16.2% and 20.6%, respectively. This is not that surprising as both metrics strongly depend on the AS-level graph which is chosen non-deterministically (see Section 4.2.1).

From Fig. 2 we immediately notice that the majority of our metrics are dominated by two factors: routing policies and the number of routers inside each AS. *NumLearnedPaths* are mostly affected by policies (96.5%). For *NumSelPaths*, 25.2% of the total variation can be attributed to policies and 54.3% to the numbers of routers per AS. This is not too surprising given that policies determine which paths among those learned are preferred and the AS size limits the number of distinct AS paths that can be selected as best.

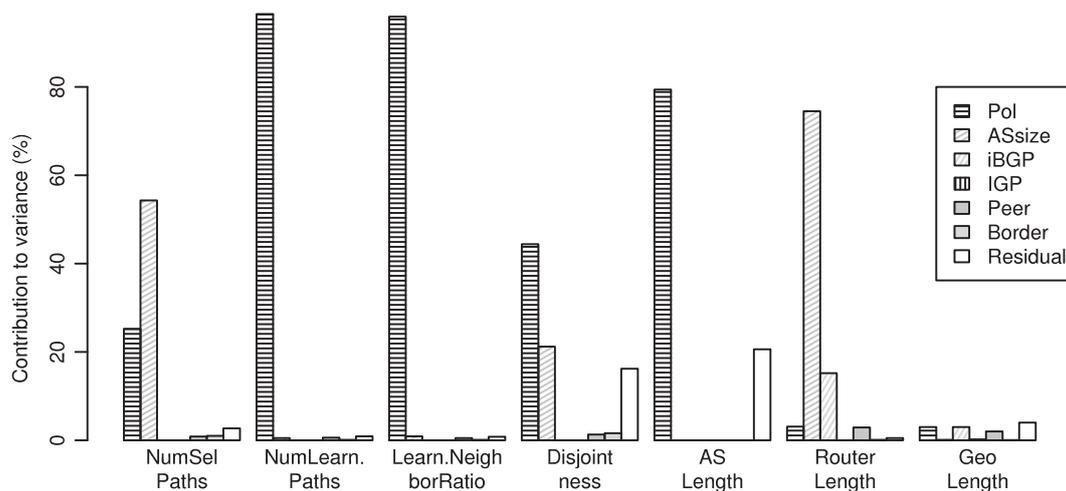


Fig. 2. Contribution of individual factors to total variation (sum of squares).

Table 1

ANOVA – percentage of variation explained by individual factors (“%”) and statistical significance ( $F$ -value, “ $F$ ”) for all factors and metrics.

	<i>Pol</i>		<i>ASsize</i>		<i>iBGP</i>		<i>IGP</i>		<i>Peer</i>		<i>Border</i>		<i>Residual</i>
	%	$F$	%	$F$	%	$F$	%	$F$	%	$F$	%	$F$	%
<i>NumSelPaths</i>	25.2	$10^4$	54.3	$10^4$	0.0	0	0.0	0	0.9	252	1.0	919	2.7
<i>NumLearnedPaths</i>	96.5	$10^5$	0.5	420	0.0	0	0.0	1	0.6	529	0.1	142	0.9
<i>LearnedNeighborRatio</i>	95.9	$10^5$	0.9	885	0.0	0	0.0	0	0.5	446	0.1	611	0.8
<i>Disjointness</i>	44.4	6434	21.2	1025	0.0	1	0.0	1	1.3	65	1.6	228	16.2
<i>ASLength</i>	79.4	9023	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	20.6
<i>RouterLength</i>	3.1	$10^4$	74.5	$10^5$	15.2	$10^4$	0.0	5	2.9	4181	0.1	397	0.5
<i>GeoLength</i>	3.0	1786	0.1	2450	3.00	1867	0.2	54	2.0	329	0.05	$10^4$	4.0

The two metrics that quantify the length of the selected routes in terms of AS-level and router-level hops, display a different behavior. *RouterLength* is dominated by *ASsize* which explains 74.5% of the total variation. The AS-level length of the selected routes is only sensitive to whether policies are used or not. We note that the length of the router-level paths is impacted by the choice of the intra AS topology, i.e., if the topology has a full mesh of iBGP sessions or a route reflector hierarchy (15.2%). Finally, the observed variation for *GeoLength* is not due to a single factor but rather needs to be explained by all factors and in particular their interactions. Overall, the absolute variation for *GeoLength* is low compared to other metrics, which is also supported by our findings of Section 6.3.

In summary, we find that the two most dominant factors for our metrics are routing policies and the number of routers per AS. While some of the other factors are statistically significant, their overall contribution to the variance is nevertheless small. In particular, intra-domain factors, including IGP weights or iBGP connectivity only have weak impact on global path properties.

## 5.2. Impact of individual factors

While the ANOVA technique helps us to understand which factors are responsible for the observed variations, they do not tell us how different parameter settings impact a certain metric. For example, we have not yet quantified

how policies limit the number of paths that are learned by the routers of an AS. Therefore, we now explicitly study the impact of routing policies, intra-domain configuration, and inter-domain connectivity and we investigate potential differences between tier-1, tier-2, and stub domains. The approach we adopt is to rely on a single instance of *defaultSetup* as a reference, vary one factor at a time and then check for differences. Contrary to the preceding section, we now use **Internet-scale** simulations based on AS-level maps obtained from CAIDA [26], see also Section 4.1. As will be shown in the following sections, our findings confirm the results of the general sensitivity analysis in Section 5.1. Table 2 presents the mean values for three of our metrics – *NumLearnedPaths*, *Disjointness* and *RouterLength* – across all ASs and all routes, respectively.

### 5.2.1. Impact of policies

If we repeat the same simulation but do not enforce AS relationships (“no policies”), the number of distinct AS paths learned by the routers of an AS increases from 2.58 to 6.84.

On the one hand, this can be seen as confirmation of the wide-spread belief that policies significantly restrict route diversity. On the other hand, the value of the *Disjointness* metric only slightly increases after removing policies from the simulation. We conclude that BGP does not necessarily propagate alternative, disjoint paths for routing. This can be seen as motivation for proposals such as path splicing [19].

**Table 2**

Impact of policies, intra-domain configuration, and inter-domain connectivity (mean values across all routers). Standard deviation is around 4.5 for *NumLearnedPaths*, around 0.2 for *Disjointness* and around 5.4 for *RouterLength* metrics.

	<i>defaultSetup</i>	No policies	Full mesh	Random IGP weights	1 peering link	2 peering link	Random border router
<i>NumLearnedPaths</i>	2.58	6.84	2.59	2.59	2.56	2.62	2.62
<i>Disjointness</i>	0.81	0.83	0.81	0.81	0.81	0.81	0.81
<i>RouterLength</i>	11.67	7.80	8.72	11.80	12.34	10.65	11.43

For the metrics *RouterLength* and *ASLength* (not shown in Table 2), we observe that ignoring routing policies generally leads to shorter routes. Without policies the average distance in AS-level hops (router-level hops) is 5.67 (7.80) per selected route versus 4.30 (11.67) without routing policies. Although policies lead to longer selected routes, the consequent stretch is not dramatic, confirming previous results [21,22,32].

### 5.2.2. Impact of intra-domain configuration

The ANOVA results indicate that intra-domain design choices, e.g., iBGP full mesh versus route reflector hierarchy (iBGP) or the strategy for assigning IGP weights (IGP) have a comparatively low impact on the results of route propagation and computation. Nevertheless, we find that 70% of all route decisions in *defaultSetup* are reached based on other steps of the BGP decision process than local-preference and shortest AS path, confirming previous findings [6,33]. To investigate, we now compare *defaultSetup* against a simulation where we replace the route reflector hierarchies inside large domains with a full mesh (factor iBGP) and another simulation where we assign IGP weights randomly rather than applying geographic-aware weights (factor IGP), see Table 2.

Varying only the settings of the two factors iBGP and IGP in *defaultSetup* does only induce minor changes in the values of our metrics. The only exception is the metric *RouterLength*. Here we observe that the average number of hops per route decreases from 11.67 to 8.72 when replacing the route reflector hierarchy with a full mesh.

Given these results, we conclude that it is not mandatory for studying inter-domain route propagation to emphasize intra-domain aspects such as assignment of IGP weights or the topological structure of a domain. This relative unimportance of intra-domain aspects is surprising given their importance for the BGP decision process.

### 5.2.3. Impact of inter-domain connectivity

Finally, we quantify the impact of the two factors *Peer* and *Border* on our metrics. In addition to *defaultSetup*, where the number of peering links between two ASs depends on the size of the neighboring ASs, we run a simulation where we realize each AS-level edge with 1 or 5 peerings. We find only minor differences for the metrics *NumLearnedPaths*, *Disjointness* and *RouterLength*, see Table 2. In terms of *RouterLength*, the average number of router hops increases from 11.67 in *defaultSetup* to 12.34 with only one peering link per AS-level edge, but decreases to 10.65 for five peering links: If there are many peering links per AS-level edge, each router can choose between many egress routers. Consequently, the likelihood that a packet

can be forwarded over a direct link to the neighboring AS increases.

Note however that irrespective of how the IGP weight assignment is done, geographic-aware for *defaultSetup* or via random assignment (factor *Border*), the average number of router-level hops only slightly changes.

In summary, path diversity is almost always insensitive to the realization of the AS-level edges (factor *Peer* and *Border*): neither the selection of border routers nor the number of peering links per AS-level edge have a dominating impact.

### 5.2.4. Tier-1 vs. tier-2 vs. stub ASs

We briefly explore if there are any significant differences between tier-1, tier-2, and stub domains. Once more, we rely on the *Internet-scale* simulation of *defaultSetup* as basis and compute the results of our metrics separately for each type of AS.

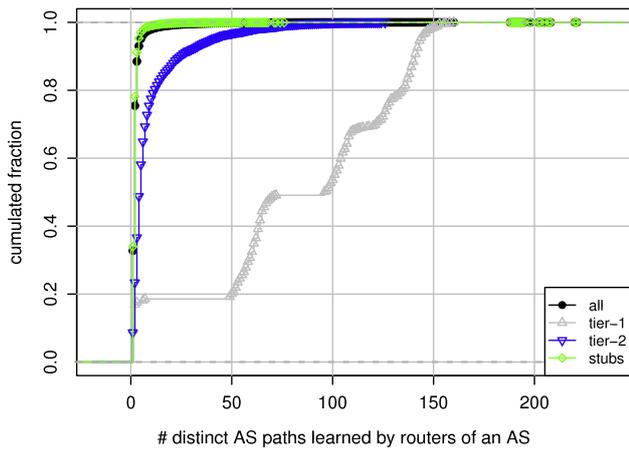
Fig. 3(a) shows the cumulative distributions of the number of distinct paths, learned by the routers of tier-1, tier-2, and stub domains (*NumLearnedPaths*). We observe that path diversity is highest for tier-1 ASs (mean: 81.8), followed by tier-2s (mean: 9.8) and stubs (mean: 2.2). This is mainly due to the AS-level connectivity of today's Internet. While stub ASs<sup>5</sup> are only connected to a small number of upstream ASs, tier-1 and tier-2 domains have peerings with many neighbors.

Fig. 3(b) shows for tier-1, tier-2, and stub domains the distribution of path lengths in terms of router-level hops. In general, tier-1 ASs can use shorter paths towards destinations. While tier-1 domains have paths with an average of 9.3 router hops, stub (tier-2) domains have a mean path length of 11.6 (12.8) hops. Given that we rely on a *Internet-scale* topology from Caida [26], it is not surprising that AS path lengths are comparable to those reported by Huston et al. [32]. Again, tier-1 ASs have shorter paths towards the destination. Overall, the differences observed among the types of ASs are consistent with what we would intuitively expect.

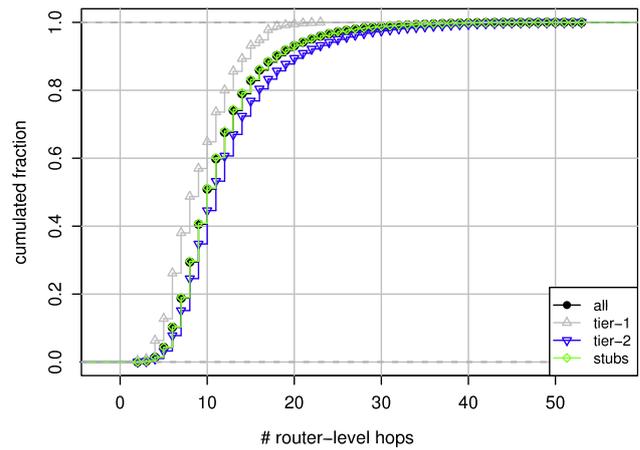
## 6. Path Inflation

After the general sensitivity analysis we now ask: How “good” are the computed routes in terms of their length? The metrics we use are AS-level (*ASLength*) and router-level hops (*RouterLength*), as well as geographical distance (*GeoLength*). These path properties contribute to the delay that a packet sees when routed along a path.

<sup>5</sup> a considerable fraction of them is still only single-homed.



(a) Number of learned paths per AS (*NumLearnedPaths*)



(b) Number of router-level hops (*RouterLength*)

Fig. 3. Differences between tier-1, tier-2, and stub ASs for *defaultSetup*.

The key idea is to compare for a given source and destination against the route that is globally shortest in terms of AS-level hops (Section 6.1), router-level hops (Section 6.2) or geographical distance (Section 6.3). Such shortest routes can be determined by applying Dijkstra’s algorithm. Note that we again rely on *Internet-scale* simulations based on collected AS maps from [26], see also Section 4.1.

### 6.1. AS-level path stretch

The AS-level path stretch is determined by comparing the number of AS hops of the selected AS path with the minimal path length in terms of AS hops that connects the same pair of routers. This is done for more than 3 million pairs of source and destination ASs. While we take the length of the selected AS path directly from the simulation results, we use Dijkstra’s algorithm to obtain the shortest paths from the more than 30,000 source ASs to the 100 destination prefixes in our Internet AS graph [26].

Fig. 4 shows the distribution of the stretch factor, i.e. number of simulated AS hops divided by the optimal AS path length. While the stretch ratio averages 1.3, we observe for almost half of the source/destination pairs a simulated path with same length as the optimal path. There exists only one case with a stretch factor of 8.<sup>6</sup> Overall, the observed AS-level stretch is reasonable. We point out that the observed AS-level path stretch is exclusively due to the use of AS relationship policies, which imply “preference” rules and the “no-valley” property [13].

### 6.2. Router-level path stretch

Contrary to AS-level paths, router-level paths may incur stretch for two reasons: Routing policies and ASs with multiple routers/intra-domain configuration. Both aspects are investigated in this section. The general approach for analyzing the router-level path stretch is similar to Section 6.1. Again, we rely on *defaultSetup* and our AS-level topology from CAIDA [26].

<sup>6</sup> Note that we rely here on the Internet-scale topology from Caida [26]. Such outliers have already been reported in the past.

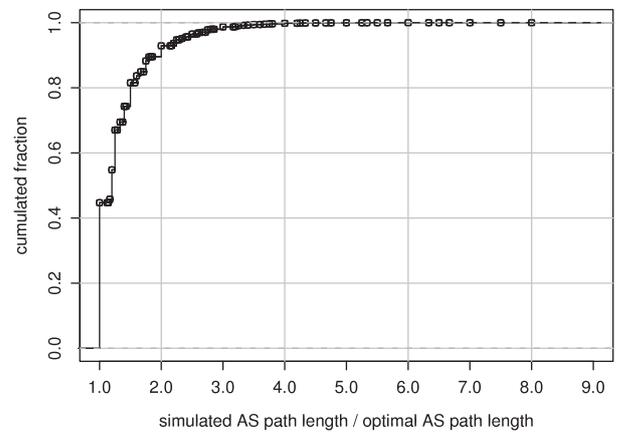


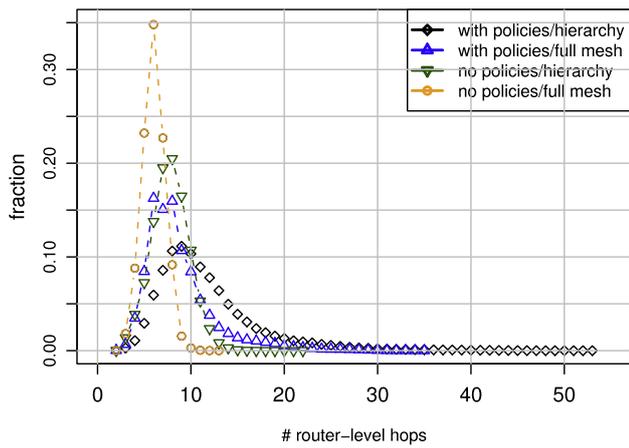
Fig. 4. AS-level path stretch for *defaultSetup*.

Fig. 5(a) shows the distribution of router-level path lengths for 4 different setups. We compare *defaultSetup* against a setup (i) where we ignore routing policies, (ii) where we replace the route reflector hierarchy with a full mesh of iBGP sessions and (iii) where we have both no policies and a full mesh instead of a route reflector hierarchy. Apart from that the *defaultSetup* remains unchanged.

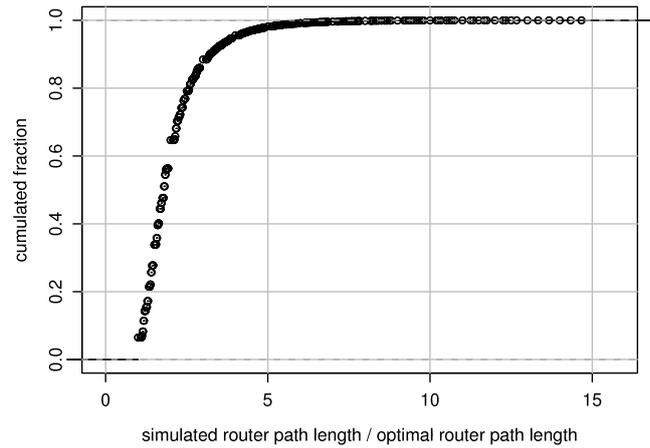
The general observation is that the longest router-level paths occur for the setup with policies, using a route reflector hierarchy, followed by “with policies/full mesh” and “no policies/hierarchy”. If we replace the intra-domain structure by a full mesh and if we remove policies, the length of the routes selected in our simulation decreases from 11.7 to 7.8.

The router-level stretch factor, i.e., number of router-level hops in simulated path divided by the optimal path length is shown in Fig. 5(b). Overall, the majority of the routes in our simulations experience moderate stretch, with an average stretch of 2.1. There are only very few extreme cases with a stretch ratio of more than 2.4 (3rd quantile).

These observations clearly demonstrate that both policies as well as having ASs with multiple routers impact the length of the paths of the best routes. With regards to the intra-domain configuration, we observe that routes incur stretch if ASs consist of multiple routers. In such a



(a) Length of router-level paths



(b) Router-level stretch of *defaultSetup*

Fig. 5. Optimality of router-level paths in terms of router-level hops.

case, packets may have to visit more hops in order to traverse a transit AS domain. Routing policies can cause BGP to prefer longer AS paths over a shorter one, if local-preference values are applied. Hence, we observe a correlation between the number of AS-level and router-level hops of a path in our simulation. Moreover, we find for more than 5% of the routes in *defaultSetup* that local-preference values are used as final criterion to select a best route among the set of learned ones. While this looks like a small percentage, it is actually significant as such routing decisions are then potentially propagated to many neighbors. This underlines the impact of policies on the “optimality” of router-level paths.

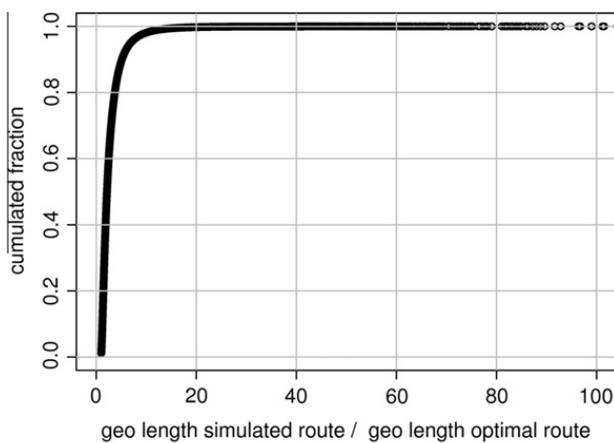
Again, we study the BGP decision process and find that criteria such as “hot-potato routing” are frequently used as decision criteria – in some ASs for more than 60% of the decisions. This underlines the importance of intra-domain routing structures on inter-domain routing and explains why the router-level path stretch is sensitive to the choice of the use of a full mesh or a route reflector hierarchy. Nevertheless, BGP path choice as a whole is mostly insensitive to intra-domain factors such as IGP weights (see Section 5.1).

### 6.3. Geographical path stretch

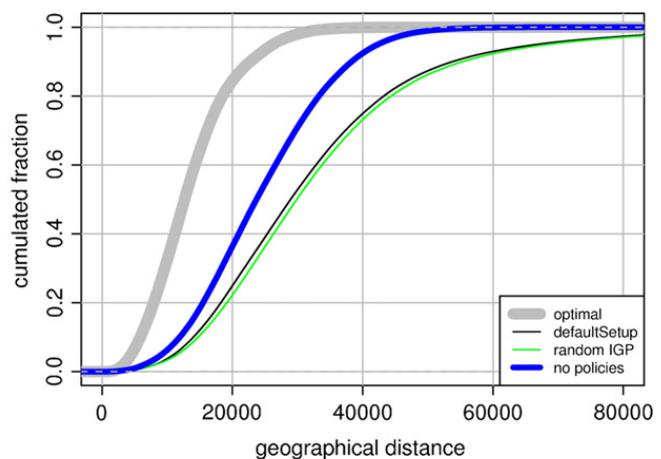
In general, geographic properties have only indirect impact on BGP routing decisions. For example, IGP weights may be chosen to prefer routes with geographically short distances. Given that the geographic distance and the delay that packets experience on the path are correlated, we now study the geographical path stretch.

Again, we rely on *defaultSetup* and run Internet-scale simulations based on this choice of levels. Since we assigned routers to geographic locations in our topologies, We can compute the geographic distance of each route, by adding up the distances of the individual links it traverses to reach its destination. To determine the geographically shortest paths from more than 30,000 source ASs to our 100 prefixes, we apply Dijkstra’s algorithm on a weighted router-level graph, where weights correspond to the distances between adjacent routers.

Fig. 6(a) shows the cumulative distribution of the geographical stretch ratio, i.e., the geographic length of simulated path divided by length of the geographically shortest path, for the more than 3 million pairs of source and destination ASs. The results suggest that most routers incur a stretch



(a) Distribution of geographical stretch



(b) Distribution of geographical lengths

Fig. 6. Optimality of routes in terms of geographical length.

ratio that is comparable to that of the router-level stretch: The average geographic stretch is 2.9 (average router-level stretch: 2.1), while for 25% of the source/destination pairs, we observe a stretch ratio of more than 3.3 (router-level: 2.4). Since BGP route decisions are based on the BGP-level topology and on routing policies that do reflect business goals rather than geographical distances, we can even observe paths with a stretch ratio of 100.

We now check how routing policies as well as different strategies for IGP weight assignment impact the geographical stretch of the selected routes. To find out, we modify the strategy for assigning IGP weights or remove routing policies, but otherwise we keep the simulation setup identical. Fig. 6(b) shows the cumulative distribution of the geographical length of the selected routes. The four curves display the results for *defaultSetup*, which assigns IGP weights to reflect geographical distances, for modified simulations where random IGP weights are applied or no policies are configured, and for the “optimal” lengths.

At first glance, we infer from Fig. 6(b) that the curves for all our simulations are considerably below the curve for the “optimal” lengths. This implies that neither removing policies nor changing the strategy for IGP weight assignment significantly shortens the geographical lengths of the selected paths. The simulation run closest to the “optimal” curve is the configuration where we ignore routing policies. Comparing different IGP assignments, the setup of *defaultSetup*, where IGP weights reflect the geographical distance, performs only slightly better than random IGP weights.

Overall, it seems hard to *globally* optimize route propagation and selection with respect to the geographic length of the selected routes.<sup>7</sup> The reason is that BGP has been designed to support flexible routing policies for implementing business objectives. If the geographical length of paths is to be *globally* optimized, routing policies of individual ASs need to be consistent and need to optimize the same objective: Geographically short routes and *not* routes that agree with existing business agreements.

## 7. Related work

Previous work, e.g., [22,21,34] has mainly focused on the impact of policies on the properties of the selected routes. For example, Gao et al. [34] have studied the extent to which routing policies inflate AS paths. Other work, e.g., [35] investigates how the routing system and its mechanisms affect the end-to-end performance of the selected paths. Finally, there exist several approaches for the measurement [14,15] and generation of router-level topologies, e.g., [36,37]. For their analysis they rely on traceroute measurements and use round-trip time, loss rates, and bandwidth to judge the optimality of paths.

Overall, we do not only confirm the results of [22,21,34,35], but rather we substantiate them. Existing work only relies on paths visible from measured data while it is widely agreed that measurement-based approaches are inherently limited regarding observability, e.g., [38].

<sup>7</sup> Yet, IGP weights or other routing policies are effective means to locally optimize routing for a specific objective within an AS.

In contrast, our completely controlled simulation environment provides a comprehensive picture of path choices for the given setup. Moreover, any sensitivity analysis requires means to vary parameters, something which can only be achieved via simulation in our case. In this respect, we claim that our results are not simply pure speculations as results on path stretch in previous work [22,21,34].

## 8. Conclusion

In this article we study the sensitivity of routing stretch and diversity metrics to factors such as policies, topology, IGP weights, etc. We rely both on Internet-scale and smaller-sized simulations, and use statistical techniques to quantify sensitivity. We confirm previous findings that routing policies and AS size (in number of routers) are the dominating factors. Surprisingly, we find that intra-domain factors only have marginal impact on global path properties.

Our results reveal that it is hard to improve the global properties of route selection by purely tweaking BGP attributes or changing iBGP graphs, etc. This is consistent with inter-domain routing design, that mainly supports the flexible implementation of routing policies, but not the propagation of optimal paths. Improving global properties of Internet paths will require more than tuning BGP.

Our work is an important step towards understanding which and how parameters impact the optimality of inter-domain routing. This is crucial for a wide variety of tasks, e.g., for building scalable and meaningful models of routing, for designing or improving routing protocols, etc. Both our sensitivity analysis and our simulation framework will prove useful for the evaluation and comparison of routing protocols and architectures.

## Acknowledgments

The research results presented herein have received support from Trilogy (<http://www.trilogy-project.eu>), a research project (ICT-216372) partially funded by the European Community under its Seventh Framework Programme. The views expressed here are those of the author (s) only. The European Commission is not liable for any use that may be made of the information in this document. Furthermore, the research results presented herein are partly funded by a grant from Deutsche Telekom Laboratories.

## References

- [1] T. Griffin, F.B. Shepherd, G. Wilfong, The stable paths problem and interdomain routing, *IEEE Trans. Networking (TON)* 10 (2) (2002) 232–243.
- [2] N. Feamster, H. Balakrishnan, Detecting BGP configuration faults with static analysis, in: *Proc. NSDI*, 2005.
- [3] W. Mühlbauer, A. Feldmann, O. Maennel, M. Roughan, S. Uhlig, Building an AS-topology model that captures route diversity, in: *Proc. ACM SIGCOMM*, 2006.
- [4] R. Oliveira, D. Pei, W. Willinger, B. Zhang, L. Zhang, In search of the elusive ground truth: the Internet's AS-level connectivity structure, in: *Proc. ACM SIGMETRICS*, 2008.
- [5] R. Jain, *The Art of Computer Systems Performance Analysis*, Wiley, New York, 1991.

- [6] R. Teixeira, A. Shaikh, T. Griffin, J. Rexford, Dynamics of hot-potato routing in IP networks, in: Proc. ACM SIGMETRICS, 2004.
- [7] R. Teixeira, T. Griffin, G. Voelker, A. Shaikh, Network sensitivity to hot potato disruptions, in: Proc. ACM SIGCOMM, 2004.
- [8] L. Subramanian, S. Agarwal, J. Rexford, R. Katz, Characterizing the Internet hierarchy from multiple vantage points, in: Proc. IEEE INFOCOM, 2002.
- [9] F. Wang, L. Gao, Inferring and characterizing Internet routing policies, in: Proc. ACM IMC, 2003.
- [10] G. Battista, M. Patrignani, M. Pizzonia, Computing the types of the relationships between autonomous systems, in: Proc. IEEE INFOCOM, 2003.
- [11] Z. Mao, L. Qiu, J. Wang, Y. Zhang, On AS-level path inference, in: Proc. ACM SIGMETRICS, 2005.
- [12] M. Caesar, J. Rexford, BGP routing policies in ISP networks, *IEEE Network* 19 (6) (2005) 5–11.
- [13] L. Gao, On inferring autonomous system relationships in the Internet, in: Proc. IEEE Global Internet, 2000.
- [14] N. Spring, R. Mahajan, D. Wetherall, Measuring ISP topologies with rocketfuel, in: Proc. ACM SIGCOMM, 2002.
- [15] R. Sherwood, A. Bender, N. Spring, DisCarte: a disjunctive Internet cartographer, in: Proc. ACM SIGCOMM, 2008.
- [16] R. Mahajan, N. Spring, D. Wetherall, T. Anderson, Inferring link weights using end-to-end measurements, in: Proc. ACM IMW, 2002.
- [17] S. Uhlig, S. Tandel, Quantifying the impact of route-reflection on BGP routes diversity inside a tier-1 network, in: Proc. of IFIP Networking, 2006.
- [18] R. Teixeira, N. Duffield, J. Rexford, M. Roughan, Traffic matrix reloaded: impact of routing changes, in: Proc. of PAM, 2005.
- [19] M. Motiwala, M. Elmore, N. Feamster, S. Vempala, Path splicing, in: Proc. ACM SIGCOMM, 2008.
- [20] L. Li, D. Alderson, W. Willinger, J. Doyle, A first-principles approach to understanding the Internet's router-level topology, in: Proc. ACM SIGCOMM, 2004.
- [21] H. Tangmunarunkit, R. Govindan, S. Shenker, D. Estrin, The impact of routing policy on Internet paths, in: Proc. IEEE INFOCOM, 2001.
- [22] N. Spring, R. Mahajan, T. Anderson, Quantifying the causes of path inflation, in: Proc. ACM SIGCOMM, 2003.
- [23] B. Fortz, M. Thorup, Internet traffic engineering by optimizing OSPF weights, in: Proc. IEEE INFOCOM, 2000.
- [24] R. Zhang, M. Bartell, BGP Design and Implementation, Cisco Press, USA, 2003.
- [25] T. Bates, R. Chandra, E. Chen, BGP route reflection – an alternative to full mesh IBGP, Internet engineering task force, RFC4456, 2006.
- [26] The CAIDA web site. Available from: <<http://www.caida.org/data/>>.
- [27] B. Quoitin, S. Uhlig, Modeling the routing of an autonomous system with C-BGP, *IEEE Network*.
- [28] T. Griffin, G. Wilfong, An analysis of BGP convergence properties, in: Proc. ACM SIGCOMM, 1999.
- [29] B. Quoitin, V.V. den Schrieck, P. Francois, O. Bonaventure, IGen: generation of router-level Internet topologies through network design heuristics, in: Proc. of the 21st International Teletraffic Conference, 2009.
- [30] L. Kaufman, P. Rousseeuw, Finding Groups in Data: An Introduction to Cluster Analysis, Wiley, New York, 2005.
- [31] W. Mühlbauer, S. Uhlig, B. Fu, M. Meulle, O. Maennel, In search for an appropriate granularity to model routing policy, in: Proc. of the ACM SIGCOMM, 2007.
- [32] G. Huston, BGP routing table analysis reports. Available from: <http://bgp.potaroo.net/>.
- [33] M. Buob, M. Meulle, S. Uhlig, Checking for optimal egress points in iBGP routing, in: Proc. of the 6th IEEE International Workshop on the Design of Reliable Communication Networks (DRCN), 2007.
- [34] L. Gao, F. Wang, The extent of AS path inflation by routing policies, in: Proc. of the GLOBECOMM'02, 2002.
- [35] S. Savage, A. Collins, E. Hoffman, J. Snell, T. Anderson, The end-to-end effects of Internet path selection, in: Proc. of the ACM SIGCOMM, 1999.
- [36] A. Medina, A. Lakhina, I. Matta, J. Byers, An approach to universal topology generation, in: Proc. of the IEEE International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems (MASCOTS), 2001.
- [37] S.M.K. Calvert, J. Eagan, A. Namjoshi, J. Stasko, E. Zegura, Extending and enhancing GT-ITM, in: Proc. of the ACM SIGCOMM Workshop on Models, Methods and Tools for Reproducible Network Research, 2003.
- [38] R. Bush, O. Maennel, M. Roughan, S. Uhlig, Internet optometry: assessing the broken glasses in Internet reachability, in: Proc. of the ACM IMC, 2009.



Since November 2009 **Wolfgang Mühlbauer** is a senior researcher in the Communication Systems Group at ETH Zürich, Switzerland. He received a diploma degree from the Technische Universität München, Germany, in 2005, and a Ph.D. degree in Computer Science (Informatik) from Technische Universität Berlin, Germany, in 2009. Between 2006 and 2009 he was a research assistant and teaching assistant in the network architectures groups, first at Technische Universität München and then at Technische Universität Berlin/Deutsche Telekom Laboratories. His research interests include inter-domain routing, alternative routing architectures for the future Internet, network virtualization and design of network testbeds and of distributed systems.



QoS, and network robustness.

**Steve Uhlig** obtained a PhD degree in Applied Sciences in March 2004 from the University of Louvain, Louvain-la-neuve, Belgium. He was a Postdoctoral Fellow of the Belgian National Fund for Scientific Research (F.N.R.S.) between 2004 and 2006. Between 2006 and 2008, he was an assistant professor at Delft University of Technology, Netherlands. He is currently with TU Berlin/Deutsche Telekom Laboratories, Berlin, Germany. His research interests encompass traffic characterization, routing protocols, traffic engineering, inter-domain



member of the scientific boards of Inria and the Swiss center on mobile information and communication systems and a member of the technical advisory board of Endace. She received a M.S. degree in Computer Science from the University of Paderborn, Paderborn, Germany, in 1990 and M.S. and Ph.D. degrees in Computer Science from Carnegie Mellon University in Pittsburgh, USA, in 1991 and 1995, respectively. Her current research interests include understanding the current Internet and its application for the purpose of performance debugging and intrusion prevention as well as new network architectures.

**Anja Feldmann** is a full professor at TU Berlin/Deutsche Telekom Laboratories. From 2000 to 2006 she headed the network architectures group first at Saarland University and then at TU München. Between 1995 and 1999 she was a member of the Networking and Distributed Systems Center at AT&T Labs – Research in Florham Park, New Jersey. She has published more than 50 papers and has served on more than 40 program committees, including as Co-Chair of Sigcomm 2003 and as Co-PC-Chair of Sigcomm 2007. She is a



of IP networks, with a focus on network virtualization, routing, router testing, configuration management and large-scale simulations.

**Olaf Maennel** is currently a lecturer at Loughborough University, UK. From 2008–2009, he was with TU-Berlin/Deutsche Telekom Laboratories, from 2005–2008 he was a senior researcher at the University of Adelaide. He was affiliated with the Centre for Defence Communications and Information Networking (CDCIN) and TRC Mathematical Modelling. Olaf obtained his Ph.D. at the Technical University of Munich in Germany in the networking group of Prof. Anja Feldmann. His current research interests are in the area



**Bruno Quoitin** is a senior researcher in the IP Networking Lab at Université catholique de Louvain (UCLouvain), Belgium. His research interests include amongst others inter-domain routing, inter-domain traffic engineering, software development and embedded systems. He received his Ph.D. in Applied Sciences in 2006 at UCLouvain.



**Bingjie Fu** graduated with a Bachelor of Science (B.Sc.) degree in July 2003 from the Electronic Information Engineering Department of Tsinghua University, Beijing, China. In October 2004, she finished her work as a Master of Philosophy (M.Phil) candidate, and in 2005, she received her M.Phil. Degree from the Electrical Engineering Department of University of Bath, UK. From 2004 to 2008 she worked on her Ph.D. at Delft University of Technology in the Network Architectures and Services (NAS) group. Currently, she is with Alcatel Lucent in The Hague, Netherlands.