# Violation of interdomain routing assumptions

Riad Mazloum, Marc-Olivier Buob, Jordan Auge, Bruno Baynat, Dario Rossi, Timur Friedman

# Violation of Interdomain Routing Assumptions

Riad Mazloum[1], Marc-Olivier Buob[1], Jordan Augé[1], Bruno Baynat[1], Dario Rossi[2], and Timur Friedman[1] *

[1] UPMC Sorbonne Universités
[2] Telecom ParisTech

**Abstract.** We challenge a set of assumptions that are frequently used to model interdomain routing in the Internet by confronting them with routing decisions that are actually taken by ASes, as revealed through publicly available BGP feeds. Our results quantify for the first time the extent to which such assumptions are too simple to model real-world Internet routing policies. This should introduce a note of caution into future work that makes these assumptions and should prompt attempts to find more accurate models.

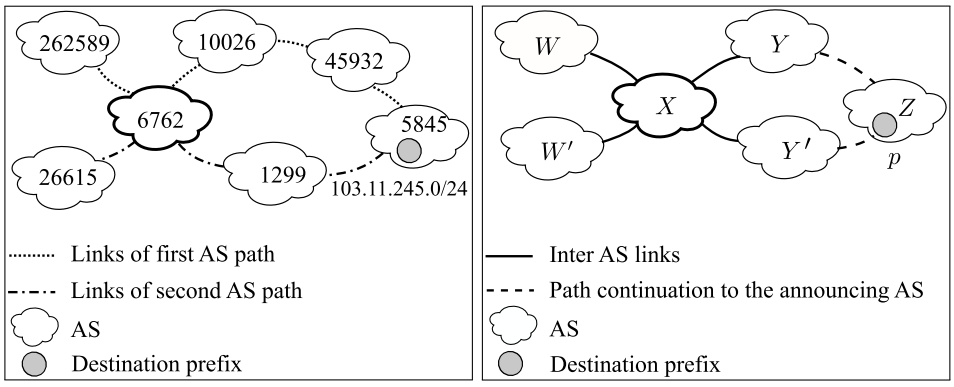
## 1 Introduction

Figure 1a illustrates a case of what is called *multi-exit routing* in the Internet. From BGPmon's [1] publicly-available feed of the BGP interdomain route updates of numerous routers, we know that the autonomous system (AS) in the middle of the figure, AS6762, has two different routes by which to reach the address prefix `103.11.245.0/24`, the advertisement for which is originated by AS5845, on the figure's far right. One route, on top, goes via AS10026 and AS45932, while the other, on the bottom, goes via AS1299. Which of these routes will AS6762 advertise to the ASes that neighbor it on the left, AS262589 and AS26615?

The AS in the middle is Telecom Italia's Sparkle, the world's 9th most important AS as reported by CAIDA's AS Rank service [2]. The top route goes via Pacnet, which is a customer of Sparkle according to CAIDA's AS Relationships database [3]. The bottom route is via TeliaNet, which the database tells us is Sparkle's peer. The standard assumption is that an AS will always route through a paying customer rather than a peer, from which it receives no revenue. And indeed Sparkle advertises the route via its paying customer Pacnet to the top-left neighbor, INTERNEXA. However, the BGP feeds also tell us that Sparkle advertises a different route, the one via its peer TeliaNet, to the bottom-left neighbor, Tim Cellular. It appears that the assumption does not hold.

What is wrong? Could there be an error in the AS Relationships database that we are relying upon? Suppose, for instance, that TeliaNet was in fact a

* Collaboration through the LINCS laboratory. Full institutional affiliation of UPMC Sorbonne Universités authors: Sorbonne Universités, UPMC Univ Paris 06, UMR 7606, LIP6, F-75005, Paris, France.




(a) Routing example (b) General case

Fig. 1: Multi-exit routing example and general case

paying customer of Sparkle, rather than its peer. Then, Sparkle's routing through both Pacnet and TeliaNet would be perfectly coherent with the assumption that Sparkle will prefer to route through its customers.

However, this scenario would violate another common assumption: that an AS with two customers will route through the one that offers a shorter sequence of AS hops to the destination prefix. Since the route via TeliaNet is just two hops, it should be chosen instead of the route via Pacnet, which takes three hops, but this is not the case. If Sparkle were to override this choice, which BGP practices allow, it would be to select Pacnet in place of TeliaNet, and not advertise routes via both of them, as it does.

The scientific community already knows that network operators do not always implement interdomain routing policies in ways that are consistent with the simplifying assumptions that are made for modeling purposes. However, *the degree to which reality defies the assumptions has not previously been quantified.* This paper looks at 4 million routes that we collected from IPv4 BGP feeds, and in particular at 204 thousand instances of multi-exit routing that those feeds reveal. In 33% of the multi-exit cases, the assumption about routing preferentially to customers over peers and to peers over providers is not coherent with the relationships that are described by CAIDA. In fully 57% of the cases, the path length assumption does not hold.

This paper proceeds in Sect. 2 by providing some background for readers who are not familiar with the details of BGP. In this context, we formalize four commonly-held assumptions, and cite examples in the literature where they are made. (The assumptions described above are composites of these four assumptions.) Sect. 3 describes our methodology for confronting the assumptions with the data. Results appear in Sect. 4. The paper wraps up with related work (Sect. 5) and a conclusion pointing to future work (Sect. 6).

Our contributions are to formalize commonly-held assumptions about interdomain routing and AS relationships and propose two methods to identify violations of the models. Also, we provide the first quantification of such violations to be based upon publicly-available data.

## 2 Interdomain Routing and our Set of Assumptions

### 2.1 BGP Background

BGP is the interdomain routing protocol that allows an AS to learn how to route to destinations in other ASes. A BGP route describes the *AS Path*, or sequence of ASes, to be traversed on the way to a *prefix*, which is a set of contiguous IP addresses. The *BGP next hop* is the egress point to use at the IP level in order to follow the route. Routes are exchanged between routers in the same AS through *iBGP* sessions, and between routers in different ASes via *eBGP*.

In the general case, a BGP router learns several routes toward a given destination. It is free to accept just some of them and to modify these. The router then elects one route (the *best route*) by following the *selection steps* of the *BGP decision process* [4], typically modeled as in Table 1. At each step, routes dominated by at least one other route are discarded. When, after one of these steps, there remains just one element in the set, this element is the best route. The router is free to modify a best route before forwarding it to its neighboring

Table 1: Selection steps of the BGP decision process

| | |
|---|---|
| 1. Highest local preference | 5. eBGP over iBGP |
| 2. Shortest AS Path length | 6. Lowest IGP cost |
| 3. Lowest origin type | 7. Tie break rules |
| 4. Lowest multi-exit discriminator | |

routers and it is free to select which of those routers will receive the route.

One modifiable parameter that affects the choice of best routes is the *local preference*. If a router receives two routes $\varrho$ and $\varrho'$ toward the same destination with a higher value of `local_pref` assigned to $\varrho$, then $\varrho$ is preferred to $\varrho'$.

### 2.2 AS Relationships

ASes use BGP to implement their contractual commercial agreements, which are typically modeled by three types of economic relationship.

- *Customer-to-provider (c2p)*: a customer pays a provider for transit service to the rest of the Internet for its traffic and its customers' traffic.
- *Peering (peer)*: a pair of ASes transit traffic between them or their customers to destinations belonging to them or their customers, free of charge.

– *Sibling-to-sibling (s2s)*: a pair of ASes transit traffic for each other and for their respective clients to every destination in the Internet, free of charge.

Gao [5] proposed a way to infer AS relationships based upon observed BGP routes, opening the way to much subsequent work.

### 2.3 A Set of Interdomain Routing Assumptions

This section describes four common assumptions about interdomain routing, citing selected papers that make each assumption.

***(A1) iBGP valid***

The assumption is that any BGP route has the potential to be propagated within an AS to all routers of that AS. In other words, route propagation is *only* governed by routing decisions taken by the different routers in the AS and there are no parts of the AS to which a route cannot be forwarded.

This assumption seems justified since an AS should guarantee this property in order to assure that all of its routers are selecting the best routes [6–10].

***(A2) Policy through eBGP only***

Routing policy is only applied by routers through their participation in interdomain (i.e., eBGP) sessions. This assumption implies that the `local_pref` value is not modified by routers through their iBGP sessions. If a router were to modify the `local_pref` value for some or all of the routes in an iBGP session, this could affect the choices of all routers in the AS to which this route is forwarded.[1]

This assumption is made to simplify the model of route propagation in an AS [6, 8–10].

***(A3) Customer over peer, peer over provider***

The assumption that an AS always prefers to send traffic through a customer over a peer and through a peer over a provider so as to maximize the presumed economic benefits. Sending traffic through a customer means that the customer will pay for it, while sending though a provider means that one has to pay the provider. [6, 8, 10].

An AS will implement this hierarchy by assigning a higher `local_pref` value to routes learned from a customer than to routes learned from a provider.

***(A4) Only one relationship type***

In the literature, each AS interconnection is typically modeled as a single economic relationship [5,8–16]. This assumption rules out, for instance, an AS being the peer of another AS in one part of the world, while being that AS's customer in another location. This is a convenient assumption to make because the main source of data consists of AS paths conveyed on BGP routes. These paths provide only AS-level information, and do not reveal, for instance, in cases where there are several possible egress points through which traffic can pass from one AS to another, which ones are used.

[1] There is a way to influence a routing decision before the local preference step, which is to use a vendor-specific *weight* attribute. It allows a router to prefer routes based upon which router it received them from. For the purposes of (A2), modifying weights through iBGP sessions has the same violation impact as modifying local preferences.

## 3 Methodology

If we had detailed knowledge of the routing decisions made by BGP routers, it would be possible to challenge, and possibly invalidate, the individual assumptions described in the previous section. Unfortunately, this information is unavailable to us. However, the publicly-available BGP feeds do allow us to challenge combinations of assumptions.

The novelty of our approach lies in the way that we use observed instances of multi-exit routing as a means to identify assumption violations. An instance, which we call a *multi-exit*, arises when an AS uses multiple next-hop ASes to reach a given destination prefix. Briefly, we process the feeds to identify multi-exits (Sect. 3.1), and then we examine each one for incoherencies in either the AS path length, the AS relationships, or both (Sect. 3.2). Each incoherency reveals a case in which one or more common assumptions have been violated.

### 3.1 Observing Multi-exits

Not all multi-exits can be observed through BGP feeds, but we can see them when an AS advertises two or more routes to a common destination prefix to its neighboring ASes. Fig. 1b illustrates the general case: an AS $X$ announces to its neighbors $W$ and $W'$ different routes to a destination prefix $p$, each route having a different next-hop AS, $Y$ or $Y'$.

We observe multi-exits as follows. A BGP snapshot at a given instant $t$ is the set of all of the BGP routes being used by the vantage points at that time. The AS Path of a route is a sequence of AS numbers ($AS_1$, ..., $AS_i$, ..., $AS_k$). For each AS $AS_i$ of the AS Path and for each destination prefix $p$ related to this path, we extract the next-hop AS $AS_{i+1}$ used by $AS_i$ to reach the destination $p$. In this way we build the set of *BGP triplets*, $\mathcal{T}_{\text{BGP}} = \{(AS_i, AS_{i+1}, p)\}$. Looking at these triplets, a multi-exit is observed whenever we detect two (or more) triplets of the form ($AS_i$, $AS_{i+1}$, $p$) and ($AS_i$, $AS'_{i+1}$, $p$).

### 3.2 Observing Incoherencies in Multi-exits

We now present simple criteria for detecting, in a multi-exit, two types of incoherency with a set of common assumptions. Each incoherency reveals an instance in which one or more assumptions have been violated. Note that while observed incoherencies allow us to reveal assumption violations, the inverse is not necessarily the case. If an assumption is violated by an AS for which there is no multi-exit in our database, our techniques will not reveal this violation. Furthermore, it is possible, even in a multi-exit, for a violation to not manifest itself as an observable incoherency. Hence, our results provide a lower bound on the number of actual violations present at the time of the BGP snapshot.

**Incoherent AS Path Lengths.** We observe incoherent AS Path lengths as follows. Assume that $X$, in Fig. 1b, through a router $R$ (not shown), announces

to its neighbor $W$ a route $\varrho$ that it has received from $Y$, and simultaneously, through another router $R'$ (not shown), announces to its neighbor $W'$ a route $\varrho'$ that it has received from $Y'$. If any of the first four steps of the BGP decision process (see Table 1) had been decisive, assumptions (A1) and (A2) require that $R$ and $R'$ will have selected the same route. Since each has selected a different route, the decision process will have passed steps 1 and 2, meaning that routes $\varrho$ and $\varrho'$ had the same `local_pref` values and identical AS Path lengths.

Our first criterion is thus to check the AS path length of routes identified in a multi-exit. If an AS announces two routes $\varrho$ and $\varrho'$ toward the same destination, and the AS Path lengths of $\varrho$ and $\varrho'$ differ, we deduce that either (A1) or (A2), or both, have been violated. Since our observations do not allow us to distinguish violations of (A1) from violations of (A2), we state merely that a path length incoherency reveals a violation of the *composite assumption* $(\mathrm{A1} \oplus \mathrm{A2})$.

**Incoherent AS Relationships.** We observe incoherent AS relationships as follows. According to (A3), an AS assigns higher values of `local_pref` to its customers than to its peers, which in turn receive higher values than do the providers. Also, according to (A4), there is only one relationship between two ASes, which means that there is one value of `local_pref` per neighboring AS and that this value further corresponds to the type of the relationship. Further, according to (A1), if an AS $X$ is observed to do multi-exit routing through two different ASes $Y$ and $Y'$ then routes learned from those ASes have the same value of `local_pref`. Finally, according to (A2), $Y$ and $Y'$ *must* have identical types of relationship with $X$ (e.g., they are both customers of $X$).

As a consequence, our second criterion is to examine the relationships between an AS and its next-hop ASes in a multi-exit. This requires the availability of an AS relationship database. We consider *c2p* and *peer* relationships, leaving out the special case of *s2s* without affecting our conclusions. If the relationships differ, then we can infer that at least one of the assumptions in the composite set $(\mathrm{A1} \oplus \mathrm{A2} \oplus \mathrm{A3} \oplus \mathrm{A4})$ is violated.

## 4 Results

### 4.1 Data Sources

Our study is based on two types of data: BGP updates and AS relationships. We parsed IPv4 BGP updates from BGPmon, which gathers data provided by RouteViews[2] and peers to some other BGP routers [1].

We ran our analysis on snapshots taken in August 2012, then January, March, and August 2013. Results presented here are based on a snapshot taken on 24 March 2013 at 10:00:00 GMT. Table 2 lists some snapshot statistics and results. The other snapshots were similar.[3]

[2] http://www.routeviews.org/

[3] All of our data is publicly available at http://top-hat.info/routing-assumptions/.

To increase the likelihood that each route that has been introduced has indeed had a chance to propagate to all of the vantage points, we apply a route stability filter. We consider a route stable if it is the last one received by a BGP router concerning a prefix and it has been received at least 24 hours ago without being withdrawn. The filter causes us to slightly undercount multi-exits, and its effect on the overall results is negligible.

We also remove from AS paths any ASNs reported by CAIDA [3] to belong to Internet exchange points (IXPs). In principle, these do not play a role in the routing policy of the ASes they interconnect.

Table 2: Snapshot statistics and results

| | | |
|---|---|---|
| **routes** | | 3,948,447 |
| **stable routes** | | 3,493,673 |
| **prefixes** | | 459,532 |
| **vantage points** | | 35 routers in 32 ASes |
| **triplets** | | 13,852,998 |
| **unique triplets** | | 8,257,351 |
| **transit ASes** | 6,762 | 100% |
| **transit ASes having multi-exits** | 1,441 | 21% |
| **MEs (multi-exits)** | 204,423 | 100% |
| **MEs with incoherence** | 129,590 | 63% |
| **MEs with incoherent path length only** | 62,051 | 30% |
| **MEs with incoherent relationships only** | 12,229 | 6% |
| **MEs with both incoherencies** | 55,310 | 27% |

There is limited publicly available ground truth for AS relationships. From the projects that aim to infer them, we chose CAIDA's relationship dataset [3] since it is the only one we know to have a fully public methodology. For the 34.6% of their inferences that they were able to validate against either public or privately-obtained ground truth, they report accuracy of 99.6% for c2p relationships and 98.7% for peer relationships [17].

### 4.2 Quantifying Multi-exits

We observed 204,423 multi-exits, each having usually 2, but in some cases as many as 5, next-hop ASes. These constitute 2.7% of the (AS, destination prefix) pairs in our database (the remainder having just one next-hop AS), so by this metric multi-exits might seem to be rare. However, of the 6,762 transit ASes in our dataset, we observed 21% to be performing multi-exit routing. We found multi-exits in the ASes ranked 1 through 38 in CAIDA's AS ranking [2], including in all of the dozen or so ASes that are generally considered to be tier 1. So, multi-exit observations reveal information about ASes that play a central role in Internet routing.

### 4.3 Quantifying Incoherences

Fully 63% of the multi-exits in our dataset show incoherencies. AS Path length incoherencies, implying a violation of composite assumption (A1 $\oplus$ A2), showed up in 57% of multi-exits. AS relationship incoherencies, implying a violation of (A1 $\oplus$ A2 $\oplus$ A3 $\oplus$ A4), appeared in 33% of the multi-exits. There is overlap, with 27% of multi-exits revealing both kinds of incoherency.

### 4.4 Possible Causes for Violations

We speculate on reasons for these assumptions to be violated.
**Traffic engineering.** From our conversations with people familiar with large operators, we believe that the assumptions don't fully capture contemporary traffic engineering practices. An AS might prefer, for example, to send some traffic through a peer rather than a customer, or through a provider rather than a peer, intentionally violating (A3). This could happen when the customer has insufficient bandwidth. It could also arise when a router in a large AS is geographically closer to a peer than to a customer, and the revenue that would be generated by routing via the customer is outweighed by the cost of carrying the traffic internally to the egress point for that customer.
**Complex or hybrid AS relationships.** Previous work [15,17] has highlighted the existence of complex or hybrid relationships, in which, for example, one large AS might be another's peer on one continent and its customer elsewhere. Such relationships violate (A4), and to be implemented (A2) must be violated. The CAIDA AS relationship database [17] is built using an understanding of this sort of relationship, but it provides as output only one relationship per AS pair.
**Misconfigurations.** A router misconfiguration might cause any one of the assumptions to be violated. For example, an incorrect value of `local_pref` could result in an AS inadvertently favoring a provider over a customer, violating (A3).
**Erroneous AS relationships.** An alternative is that assumptions are not violated as often as our results indicate, but rather that CAIDA's database is not indicating the correct AS relationships, despite its high accuracy in cases where it has been validated. However, it would need to be incorrect in a large portion of cases in order to change our overall conclusions.

## 5 Related Work

As we have described in previous sections, many papers in the literature [5–17] have employed various assumptions about interdomain routing. Some of these papers, as well as others, have looked at violations of these assumptions.

Feamster et al. [18] give some examples of violations of (A1) that can appear. Gill et al. [19] queried 100 network operators for their private data, finding that 77% of ASes do not modify the value of local preference, i.e., they are coherent with (A2). The same survey reports that 87% of the concerned ASes are also coherent with (A3). Mühlbauer et al. [9] compared the routes that

actually propagate to vantage points with the routes that ought to propagate, revealing violations of (A3). Giotsas et al. [20] show that relationships between pairs of ASes for IPv4 routes differ in 13% of the cases from those for IPv6 routes. Roughan et al. [21] summarize lessons about modeling ASes based on an extensive study of common assumptions. They observe, notably, that modeling an AS interconnection by a single connection is insufficient. Mühlbauer et al. [8] similarly show the weaknesses in modeling an AS as an atomic entity.

Our work goes further by providing a method for detecting violations of commonly employed assumptions using publicly available data. We supply the first quantification of the extent of observable violations.

Our finding that violations can be observed in a large portion of transit ASes, including all of the biggest ones, does not mean, however, that previous work that made simplifying assumptions should be considered invalid. Most work on AS relationship inference [5, 11–17] makes only assumption (A4). As we have noted, our method does not allow us to specify precisely which of a set of assumptions have been violated, and so we cannot say how often (A4) in particular does not hold. Furthermore, if (A4) is indeed violated, it might not be to a degree that would change previous results.

Our results might pose more serious questions for other work. Javed et al. [10] use the four assumptions to reduce the set of ASes that may be the root cause for an routing event in the network. If the assumptions are violated, the final set might not contain the root cause AS. Buob et al. [6] aim to solve a problem in which the assumptions are respected.

## 6 Conclusion and Future Work

This paper formalized four assumptions about interdomain routing in the Internet that are commonly used in the literature. We employed a data-driven method to challenge these assumptions, making novel use of so-called "multi-exit" scenarios to reveal incoherencies between sets of these assumptions and actual interdomain routing decisions. We observe multi-exits in 21% of transit ASes in a BGP snapshot from March 2013, and find that that in 63% of these multi-exits at least one assumption is violated. Other snapshots showed similar results. Given this, we believe that future work should use these assumptions with caution.

We expect that our technique of using multi-exits to reveal characteristics of interdomain routing behavior can be further developed. Studying how they change over time could, for instance, tell us more about how ASes perform traffic engineering. We also believe that much more can be revealed by combining the BGP data with IP level measurements, which is part of our future work.

**Acknowledgments.** We thank Martin Levy of Hurricane Electric for taking the time to impress upon us the weaknesses in assumptions (A2) and (A3). We also thank the anonymous reviewers and Matthew Luckie for their feedback. The research leading to these results has received funding from the European

Union's Seventh Framework Programme (FP7/2007-2013) under grant agreements no. 287581 – OpenLab, and no. 318627 – mPlane.

## References


1. Yan, H., Oliveira, R., Burnett, K., Matthews, D., Zhang, L., Massey, D.: BGPmon: A real-time, scalable, extensible monitoring system. In: Proc. CATCH. (2009)
2. CAIDA: The CAIDA AS Ranking service, http://as-rank.caida.org/
3. CAIDA: The CAIDA AS Relationships dataset, http://www.caida.org/data/active/as-relationships/
4. Rekhter, Y., Li, T.: A border gateway protocol 4 (BGP-4). RFC 1771, Internet Engineering Task Force (March 1995)
5. Gao, L.: On inferring autonomous system relationships in the Internet. IEEE/ACM Trans. Netw. **9**(6) (December 2001) 733–745
6. Buob, M.O., Meulle, M., Uhlig, S.: Checking for optimal egress points in iBGP routing. In: Proc. DRCN. (2007)
7. Teixeira, R., Shaikh, A., Griffin, T., Voelker, G.M.: Network sensitivity to hot-potato disruptions. In: Proc. SIGCOMM. (2004)
8. Mühlbauer, W., Feldmann, A., Maennel, O., Roughan, M., Uhlig, S.: Building an AS-topology model that captures route diversity. In: Proc. SIGCOMM. (2006)
9. Mühlbauer, W., Uhlig, S., Fu, B., Meulle, M., Maennel, O.: In search for an appropriate granularity to model routing policies. In: Proc. SIGCOMM. (2007)
10. Javed, U., Cunha, I., Choffnes, D., Katz-Bassett, E., Anderson, T., Krishnamurthy, A.: PoiRoot: Investigating the root cause of interdomain path changes. In: Proc. SIGCOMM. (2013)
11. Subramanian, L., Agarwal, S., Rexford, J., Katz, R.: Characterizing the Internet hierarchy from multiple vantage points. In: Proc. Infocom. (2002)
12. Di Battista, G., Patrignani, M., Pizzonia, M.: Computing the types of the relationships between autonomous systems. In: Proc. Infocom. (2003)
13. Xia, J., Gao, L.: On the evaluation of AS relationship inferences. In: Proc. Globecom. (2004)
14. Dimitropoulos, X., Krioukov, D., Huffaker, B., claffy, k., Riley, G.: Inferring AS relationships: Dead end or lively beginning? In: Proc. WEA. (2005)
15. Dimitropoulos, X., Krioukov, D., Fomenkov, M., Huffaker, B., Hyun, Y., claffy, k., Riley, G.: AS relationships: inference and validation. ACM SIGCOMM CCR. **37**(1) (January 2007) 29–40
16. Shavitt, Y., Shir, E., Weinsberg, U.: Near-deterministic inference of AS relationships. In: Proc. ConTEL. (2009)
17. Luckie, M., Huffaker, B., Dhamdhere, A., Giotsas, V., claffy, k.: AS relationships, customer cones, and validation. In: Proc. IMC. (2013)
18. Feamster, N., Balakrishnan, H.: Detecting BGP configuration faults with static analysis. In: Proc. NSDI. (2005)
19. Gill, P., Schapira, M., Goldberg, S.: A survey of interdomain routing policies. ACM SIGCOMM CCR. (2014, to appear)
20. Giotsas, V., Zhou, S.: Detecting and assessing the hybrid IPv4/IPv6 As relationships. In: Proc. SIGCOMM. (2011)
21. Roughan, M., Willinger, W., Maennel, O., Perouli, D., Bush, R.: 10 lessons from 10 years of measuring and modeling the Internet's Autonomous Systems. IEEE JSAC **29**(9) (2011) 1810–1821