Treewidth and Hyperbolicity of the Internet

Fabien de Montgolfier*
LIAFA, Université Paris Diderot
Email: fm@liafa.jussieu.fr

Mauricio Soto*
LIAFA, Université Paris Diderot
Email: mausoto@liafa.jussieu.fr

Laurent Viennot*
INRIA
Email: Laurent.Viennot@inria.fr

***Abstract*—We study the measurement of the Internet according to two graph parameters: treewidth and hyperbolicity. Both tell how far from a tree a graph is. They are computed from snapshots of the Internet released by CAIDA, DIMES, AQUALAB, UCLA, Rocketfuel and Strasbourg University, at the AS or at the router level. On the one hand, the treewidth of the Internet appears to be quite large and being far from a tree with that respect, reflecting some high degree of connectivity. This proves the existence of a well linked core in the Internet. On the other hand, the hyperbolicity (as a graph parameter) appears to be very low, reflecting a tree-like structure with respect to distances. Additionally, we compute the treewidth and hyperbolicity obtained for classical Internet models and compare with the snapshots.**

I. Introduction

A. Motivation

Understanding the structure of the Internet covers several aspects. One is to understand the building process of the Internet, another is to design faithful models for simulation, a third one is to measure its properties to better tune its protocols. We are mostly concerned by this last point. Our approach was guided by finding properties that make Internet tractable, but we are also interested in properties measuring how robust is the topology. Indeed, many graph problems have better algorithmic solutions when the input is a graph with some known properties rather than an arbitrary graph. Better knowing the Internet thus enables better algorithms. We focus on two graph parameters known to bring tractability: treewidth and hyperbolicity. These two parameters measure how far is a graph from a tree from the point of view of connectivity and distances, respectively.

The treewidth parameter is interesting for two reasons. First a low treewidth is known to enable linear algorithms for many NP-hard problems [1]. Second it is a measure of connectivity. The treewidth of a graph is related to the number of nodes required to significantly reduce the connectivity of the graph as described later on. For example, a tree (which always has treewidth equal to one) can be disconnected by removing a single node. A higher treewidth is the sign of a better connectivity.

On the other hand, the hyperbolicity parameter is related to distances: a graph is close to a tree if routes between vertices behave like in a tree. The reason is twofold. First, similar properties have already been observed on the Internet [2]–[4]. Second, graphs with low hyperbolicity offer more tractability, enabling efficient algorithms for routing related problems like compact routing, diameter estimation among others [5]–[8].

*INRIA Project-Team "GANG" joint with Université Paris Diderot and CNRS.

The tractability issues of Internet growth mainly concern the AS-level routing. The global connectivity of the Internet is also mainly concerned with this graph: how many ASes a cyber-attacker has to infect in order to significantly disconnect the Internet?

The Internet can be seen either as a network of networks (the AS-level Internet) or as a collection of networks. This fact results in three types of graphs that are interesting with respect to our parameter measures: the AS level graph where each AS is viewed like a node and two different Ases are connected if they interchange data traffic, the router level graph modeling the connections of all Internet routers and finally the induced graphs within an individual AS. These structures are not publicly available and represent and important research field.

The difficulty of obtain an accurate view of Internet graph is well know and widely studied. We use existing different sources of data with varied techniques. Nevertheless strategies are mainly based in the analyse of BGP routing tables or traceroute probes.

This kind of snapshots are known to be incomplete and lack accuracy [12], [15]. However, this is the best data available as far as we know. On the other side, our measures have some robustness properties with respect to missing nodes and edges. More precisely, the treewidth of a graph is at least the treewidth of any of its subgraphs.

The data comes from six different sources (detailed in Section II), namely: CAIDA [9], DIMES [10], AQUALAB [11], UCLA (Internet Topology Collection) [12], Rocketfuel project [13] and Strasbourg University (MRINFO Project) [14].

B. Contribution

Our first conclusion is that the treewidth of the Internet is high: all the snapshots have a treewidth comparable to that of a square grid with same number of nodes (a grid is a classical example of sparse graph with high treewidth). From that point of view, Internet is far from a tree: it is much better connected. This result gives an estimation of the connectivity of the Internet: there exists a core that cannot be broken in two parts by removing less than $tw/2$ ASes where tw is the treewidth (see Theorem 1). This result holds independently of the accuracy of the measure of the graph: discovering more nodes or links can only increase tw.

Our second conclusion is that the hyperbolicity of the Internet is low, confirming previous results [2]–[4]. Conversely to this previous work, the hyperbolicity definition we use has algorithmic applications [5]–[7] such as enabling compact routing with small additive stretch [8]. From that point of view, Internet is close to a tree and far from a grid.

Additionally, we observe an important churn at the AS level. This can explain why the set of observed IP addresses increases with time in previous work [16].

Finally, we compare these measurements with classical Internet models based on various type of random graph generation. These graphs present a slightly higher treewidth and similar hyperbolicity. The closest models are those obtained from random generation with appropriate degree distribution. Surprisingly, the structure of Internet seems to be fairly represented by such a model. This fact was already observed [17]. The fit is particularly true for hyperbolicity. This somehow contradicts a possible interpretation of previous work [2]–[4] stating that low hyperbolicity is a intrinsic property of the Internet. It happens to be a usual property among graphs with similar degree distribution. On the other hand, random graphs appear to have slightly higher treewidth and thus better connectivity than the Internet. Yet the connectivity of the Internet is not so far from that of a random graph.

C. Related Work

Internet Topology was studied by Pansiot and Grad [18] at the router level and Govindan and Reddy [19] at the AS level. Numerous work has focused on the degree distribution observed at both levels. They are heavy tailed and can be modeled with power laws [20] or Weibull distributions [13], [21], [22].

Broido and claffy study the *connectivity* of Internet [21]. More specifically they inspect how the giant strongly connected component behaves with regard to node removal. The treewidth is a theoretical parameter for measuring such global connectivity.

The *treewidth* parameter was introduced by Robertson and Seymour. [23]. It is related to tractability: many NP-complete problems can be solved in linear time via dynamic programming for any class of graphs with bounded treewidth [1].

A large literature concerns the understanding of *Internet delay space*. Embedding Internet delay space in an Euclidean space allows to build virtual coordinate systems [24], [25]. It has also been noted that embedding in an hyperbolic space rather than an euclidean space could give better results [2]. Hyperbolicity can generally be measured on a metric or a graph according to the definition of Gromov [26]. Ramasubramanian et al. show that a relaxed version of this definition called the four point condition matches Internet delays [3]. This property can be used for predicting Internet delays through an embedding in a tree. Other notions of dimension have been tested on Internet delays such as fractal dimension [27] or doubling dimension [28].

Gromov's notion of *hyperbolicity* is defined as a four point condition and applies to graphs as well. Up to our knowledge, this notion of hyperbolicity has not yet been measured on the Internet. Narayan and Saniee study an alternative definition relying on δ-thin triangles [4] and measure it on Rocketfuel data [13]. We obtain similar results based on Gromov's original definition.

The discovery of *heavy tail* in the degree distribution has raised the question of designing adequate models for the Internet. A first goal for modeling is to understand the emergence of such heavy tail distributions. Preferential attachment was proposed by Barabási and R. Albert [29] for explaining the web graph structure, or social networks. Concerning the Internet, a possible explanation concerns the optimization of tradeoffs [30]. Another reason for modeling is to generate large network with similar properties as the Internet for simulation purposes. In that trend, generating a random graph with an appropriate degree distribution surprisingly appears to fit well with regard to structural properties [17]. Our work gives also credit to that point.

D. Roadmap

Section II describes the snapshots of the Internet we have used for our measurements and makes a comparison of the AS snapshots over time. Section III gives the treewidth definition and its relationship with connectivity. Upper and lower bounds of the treewidth of the snapshots are given. Section IV introduces the hyperbolicity definition and describes the behavior of the snapshots. Section V compares these results to what we obtain for classical graph models.

II. Data Sets

Real Internet topology is unknown. Different techniques have been developed in order to obtain realistic snapshots. In the aim of bypassing the several deficiencies of data collection, we use heterogeneous data sources with diverse inference techniques. For each data source we use snapshots collected at different times to capture the graph dynamic.

There exits different levels of granularity for Internet. First, the *router level* corresponds to IP interconnection between routers. Second, the *AS level* corresponds to the interconnections between ASes (such links can be observed at the BGP routing level or at IP level). Finally, we look at the inner topology (at router level) of certain ASes.

Data sets combine passive and active measurement techniques. BGP data collected passively (by dumping BGP routing tables at some routers) or actively through looking glasses is a main source of inter AS connectivity. For router level, *traceroute* is the widely used tool to discover router interconnections. Traceroute discover IP paths followed by probe packets sent from monitot routers to a list of destination. Additionally, the IP interconnections between two IP addresses can be used to infer an AS interconnection between the ASes who advertise IP prefixes. Internet snapshots used in this work come from the sources of data described below. The Table I shows, for each graph, its size, average degree, the size

			Graphs parameters			Largest bi-connected component		
			$\lvert V\rvert$	Avg. Degree	β	$\lvert V\rvert$	Avg. Degree	β
AS graphs	AQUALAB	12/2007-09/2008	31847	9.00	2.18	25341	10.80	2.18
	CAIDA	12/2010	29797	5.31	2.16	17559	7.49	2.16
	DIMES	12/2010	29542	6.84	2.12	21296	8.72	2.12
	UCLA	12/2010	37450	6.65	2.14	25271	8.73	2.14
router graphs	CAIDA router	04/2003	192,244	6.36	(2.93)	132,367	8.17	(2.96)
	CAIDA router	07/2010	3,360,982	2.93	2.27	1,644,761	3.84	2.18
	mrinfo	09/2008	8,636	2.72	3.26	1705	3.60	3.71
routers within AS #	1221 Telstra (2669)		Australia	2.38	2.45	246	6.07	2.46
	1239 Sprintlink (US)		7337	2.70	2.37	1054	6.70	2.77
	1755 Ebone (Europe)		295	3.68	2.86	178	4.76	3.24
	2914 Verio (US)		4670	3.26	2.59	1644	5.54	2.76
	3257 Tiscali (Europe)		411	3.18	2.77	166	4.81	2.97
	3356 Level3 (US)		1620	8.32	2.39	729	16.03	2.43
	3967 Exodus (US)		375	4.53	2.85	254	5.33	3.22
	4755 VSNL (India)		41	3.32	2.29	22	3.91	2.33
	7018 AT&T (US)		9430	2.48	2,65	1199	5.54	2.89

TABLE I: Basic statistics of snapshots.

of largest bi-connected component (the graph part sufficient for computing treewidth and δ-hyperbolicity, see below), and the exponent β of a regression of the degree distribution on a power law. This exponent is obtained via a linear regression on the complementary cumulative distribution function of degree distribution.

A. AS graphs

Vertex-set is the AS-set and edges are inter-AS links. We use:

- CAIDA, the Cooperative Association for Internet Data Analysis. Its infrastructure consists in about twenty monitors that daily collect traceroute probes to destination in full routed address space subdivided into /24's. The *AS level data set* [9], periodically constructed from probes and IP to AS mapping, gives us snapshots.
- AQUALAB [11] uses peer-to-peer clients to collect traceroute paths which are use to infer AS interconnections. Probes were made between December 2007 and September 2008 from approximate 992,000 P2P user IPs placed in 3,700 ASes.
- DIMES [10] project performs traceroute from a volunteer community of about 1000 agents. A weekly AS snapshot is available.
- UCLA (Internet Topology Collection[1]) [31] collects inter AS links by combining different BGP sources both passive (Route Views, RIPE-RIS, Abilene, CERNET BGP View) and active (Packet Clearing House, UCR, traceroute.org, Route Server Wiki).

B. router graphs

The vertex-set is a set of routers and edges are their known links at the IP level. We use data from:

- Rocketfuel project [13], use `traceroute` tool. Probes are made from public servers and alias resololution is performed from BGP tables.
- *MRINFO project*[2] from Strasbourg University. Using `mrinfo`, an IGMP multicast tool, the topology is discovered by sending IGMP `ASK_NEIGHBORS` messages, which are replied with the list of interfaces of a router. This method discovers all interfaces of replying hosts and avoids alias resolution process. However, replies are obtain only in routers with IPv4 multicast enabled which reduce the set of probed nodes. For details see [14].
- CAIDA also provides *Router level* graphs from the Internet topology data kit[3]. Two snapshots are analyzed here, made in April 2003 and July 2010 respectively.

C. Comparison of AS level snapshots

As illustrated by Figure 1(a), the size of the AS level graph slowly increases over time. Each point corresponds to the data collected during one month. The DIMES curve seems more erratic, this may come from changes in the monitor set used.

To see the influence of the aggregation window, we plot in Figure 1(b) the number of ASes $(|V|)$ and links $(|E|)$ collected between January 1st 2010 and the end of each month of that year. The size of the aggregated AS graph increases surprisingly denoting a high churn of reconnection at the AS level. A similar linear increase was already noticed at the IP level [16] where new IP addresses are continuously discovered when probing the same destination set from the same monitor with traceroute-like measurements. The churn we observe here at AS level provides an explanation for this : as connections between ASes change, new routers become exposed in the IP probes. A possible explanation for this AS churn is that BGP routing policies are frequently updated (at the pace of commercial interactions).

From now on, we use snapshots aggregated over one month except for AQUALAB for which we only have a snapshot aggregated over 10 months.

III. Treewidth

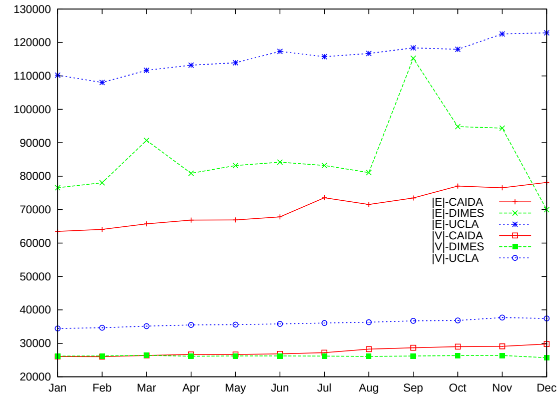
A. Definition

A *tree decomposition* of $G = (V, E)$ consists of a tree T (on a different node set than G), and a subset $V_t \subseteq V$ associated with each node t of T (called a "bag".) The tree T and the

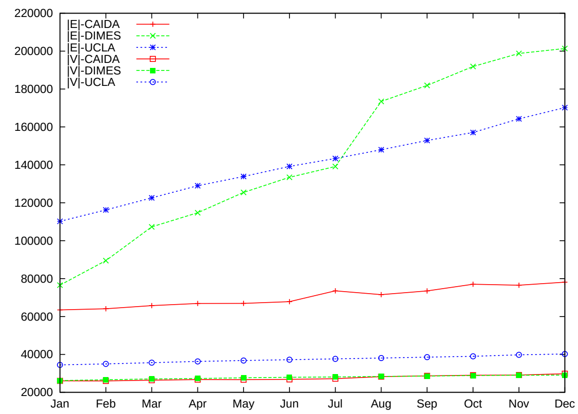
[1] http://irl.cs.ucla.edu/topology/

[2] http://svnet.u-strasbg.fr/mrinfo/index.html

[3] http://www.caida.org/data/active/internet-topology-data-kit/



(a) *Monthly snapshots.*



(b) *Aggregated snapshots since January 2010.*

Fig. 1: Size evolution of AS snapshots in 2010.

collection of bags $\{V_t \mid t \in T\}$ must satisfy the following three properties:

- Every node of G belongs to at least one bag of T
- For every edge e of G, there is some bag V_t containing both ends of e.
- The collection of bags containing a given node of G induces a connected subtree of T.

The *width* of the tree decomposition $(T, \{V_t\})$ is defined to be one less than the maximum size of a bag: $\text{width}(T, \{V_t\}) = \max_t |V_t| - 1$. The **treewidth** of G is the minimum width of a tree decomposition of G (taken over all possibles trees). A connected graph has treewidth 1 if and only if it is a tree. A tree decomposition of G naturally induces a tree decomposition of any subgraph H of G. This implies that $\text{treewidth}(H) \leq \text{treewidth}(G)$. In particular, any lower bound on the treewidth of a subgraph of G is also valid for G.

Sparse graphs (n vertices but $O(n)$ edges) may have a treewidth either low (it is 1 for a tree) or high (a square grid has treewidth $\sqrt{n}$).

Note that the removal of all vertices from the same internal bag of T disconnects the graph G. Efficient algorithms use this property by removing a bag and working recursively on the remaining connected components. The treewidth can also be seen as a measure of global connectivity in the graph (see III-E).

Computing the treewidth of a graph is NP-hard [32]. However there exists heuristics for computing lower and upper bounds of the treewidth [4]. We used them on our data: the results are given below. The treewidth of a graph is the maximum over biconnected components of treewidth. As our graphs have a large biconnected component and many pending trees, we work only on the giant biconnected component to increase computation speed.

B. Treewidth of the AS Graphs (snapshots)

We have computed treewidth lower bounds for the Internet snapshots we have of the inter-AS links (the *AS graph*). Figure 2 plot the lower bound computed for each December snapshot (except in 2011 where the February snapshot is used). We have two special graphs: INTER is the links that appear in *all* sources and UNION the links that appear in *some* source (at a given date, any AS graph is thus a supergraph of INTER and a subgraph of UNION). As seen in Table I, UCLA has more ASes and so without surprise has higher treewidth. Most notably, we see that the treewidth increases over time whatever source we consider (UCLA, DIMES or CAIDA). This may be explained by the fact that the AS graph itself increases as shown by Figure 1(a).

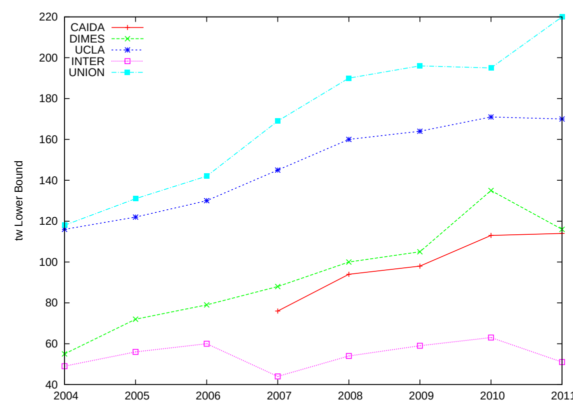


Fig. 2: Treewidth of the AS graph over time.

C. Treewidth of the AS Graphs (long time measurement)

AS links change over time and some data, especially from AQUALAB, are aggregate of a long period of measurement. In order to show how different snapshots are from long-period aggregations, we computed, for each month of 2010, the number of ASes ($|V|$) and of links ($|E|$) collected between January 1st and the end of that month, and the corresponding treewidth (Figure 3). The number of discovered edges seems to linearly increase (but the number of actually used edges evolves more slowly) so treewidth also increases. Therefore in the previous section we used aggregation of at most one month on time whenever possible (ie, not with AQUALAB).

Table II shows the lower bound of treewidth for the various snapshots gathered. We compare with the square root of the

[4]We use software from Bodlaender team available at http://www.treewidth.com/

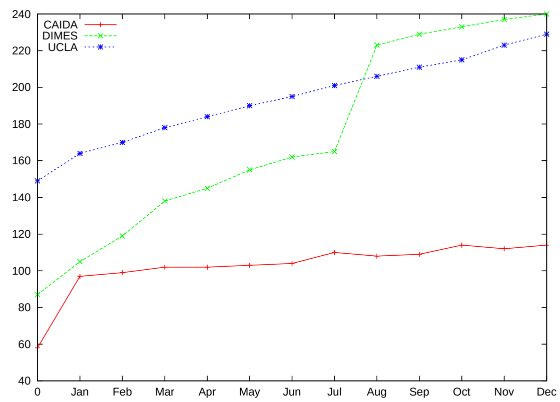


Fig. 3: Evolution of treewidth for different aggregations: $y = treewidth$(links known from January 1st to Month x)

number of vertices, ie the treewidth of a grid with the same number of vertices. The two figures appear to have comparable value. The ratio $treewidth(G)/\sqrt{n}$ especially remains a constant (about 1) for each data source independently from time.

Graph		tw	$\sqrt{\lvert V \rvert}$
AQUALAB	12/2007-09/2008	236	178
CAIDA	12/2010	113	162
DIMES	12/2010	135	156
UCLA	12/2010	171	177
UNION	12/2010	195	179
INTERS.	12/2010	63	146

TABLE II: Treewidth lower bound for AS graphs.

D. Treewidth of the Router Graphs

Figure 4 shows treewidth bounds for router level graphs. We compute the values for both, inner ASes topologies (Rocketfuel data) and complete router views (MRINFO Project and CAIDA). The table compares the bounds with the treewidth of a grid with the same number of nodes (a grid of n nodes has treewidth $\sqrt{n}$. Notice that X-axis is $\sqrt{n}$).

Almost all our snapshots appear to have a treewidth which is close to the treewidth of a square grid with same number of nodes as shown by Figure 4. This proves a high degree of connectivity of Internet. A noticeable exception is AS 3356 (Level3), which appears to have a very high treewidth. This certainly comes from virtual circuits enabled through Multiprotocol Label Switching (MPLS), a technology in which this AS is a leader. Another exception is the MRINFO snapshot which appears to have a lower treewidth than the other snapshots. This certainly comes from the sparsity of this graph. This can be explained by the fact that we keep in this graph only nodes responding IGMP queries and only edges between these nodes. It happens that few neighbors of a responding node do respond to IGMP queries. Although snapshots represent induced subgraph of the Internet at router level graph, treewidth observed is still reasonably high.

Graph	$lb \leq tw \leq ub$		Grid
router level 03/2004	372	-	363
mrinfo (router) 09/2008	11	48	41
1221 Telstra (Austr.)	9	10	15
1239 Sprintlink (US)	29	55	32
1755 Ebone (Europe)	7	8	13
2914 Verio (US)	24	35	40
3257 Tiscali (Europe)	11	14	27
3356 Level3 (US)	54	137	11
3967 Exodus (US)	8	11	15
4755 VSNL (India)	4	4	4
7018 AT&T (US)	18	26	34

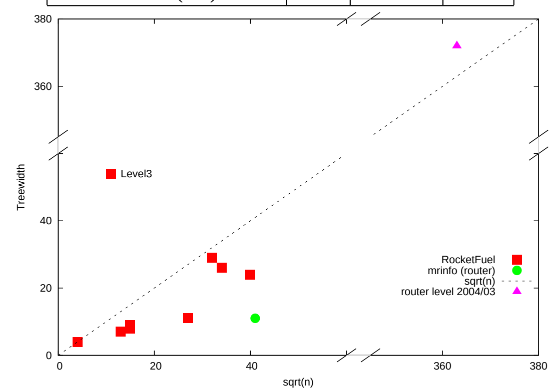


Fig. 4: Treewidth bounds for router graphs. Bottom, plot with comparison to $=\sqrt{n}$

E. Conclusion: a Core of the Network

The treewidth parameter captures how well the graph is globally connected in the following sense. It is linked to the number of nodes required to significantly reduce the connectivity of G. More precisely, a subset S of nodes in G is k-linked if for any subset X with fewer than k nodes, some connected component of $G - X$ contains more than half of the nodes of S. In other words, the removal of less than k nodes cannot drastically disconnect S. The linkedness of a graph is the maximum k such that there exists a k-linked set.

The set S can be seen as a core in the network which cannot be globally disconnected by the removal of fewer than k nodes: more than half of the core always remains connected. The linkedness is a measure of fault tolerance of the core. Both parameters are linked according to the following theorem.

Theorem 1: [33] For any graph G,
$\text{linkedness}(G) \leq \text{treewidth}(G) + 1 \leq 2\,\text{linkedness}(G)$.

We have chosen to estimate the treewidth rather than the linkedness because there exists efficient heuristics for bounding the former. Another reason is that our lower bounds of the treewidth of our snapshots hold for the Internet as long as our snapshots are subgraphs of it.

Lower bounds on the on the linkedness of our snapshots can be deduced for the table of Figure 4 (divide by 2 the lower bound of treewidth ans add one). For example, the linkedness of the AS snapshot of UCLA (2011) is at least 86. This means that there exists a core of ASes that globally remains connected even under the failure of 85 ASes or less. This gives a measure of the fault tolerance of the Internet.

IV. HYPERBOLICITY

A. Definition

Mikhail Gromov introduced and developed hyperbolic groups in the 1980's. In a seminal paper from 1987 [26] he proposed a wide-ranging research program. Aiming at studying groups through their Cayley graph, he defined δ-hyperbolicity. This notion may however be used apart from its group theory context.

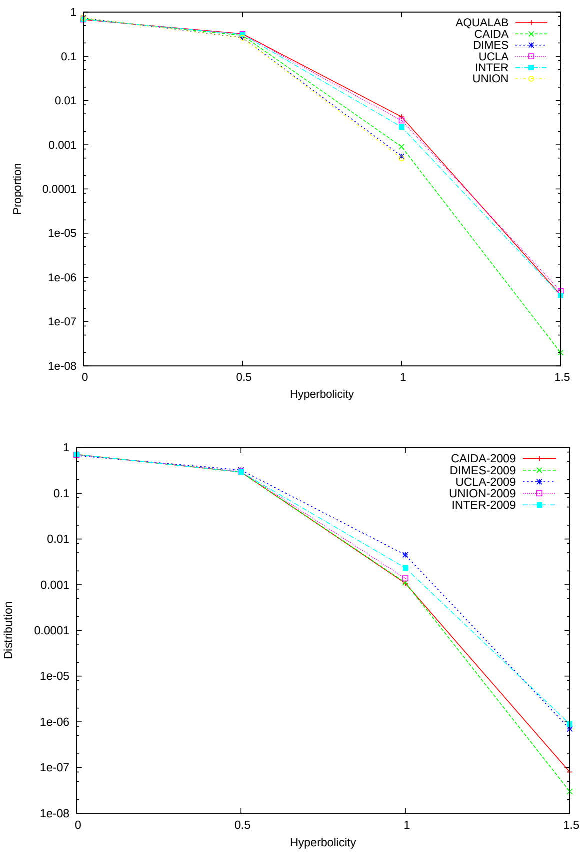


Fig. 5: Distribution of the value of the hyperbolicity $\delta(x, y, z, t)$ of quadruplets for the AS graphs (Top: 10-month aggregation. bottom: 2010/12 month snapshot)

Let x, y, z, t be four vertices. Let d_1, d_2 and d_3 be the three sums $dist(x, y) + dist(z, t)$, $dist(x, z) + dist(y, t)$ and $dist(x, t) + dist(y, z)$ non-increasingly sorted: $d_1 \geq d_2 \geq d_3$. Then define $\delta(x, y, z, t) = \frac{d_1 - d_2}{2}$. The **hyperbolicity** (or δ-hyperbolicity) of a graph G, denoted $\delta(G)$ or just δ when not ambiguous, is $\max_{x,y,z,t \in V(G)} \delta(x, y, z, t)$.

Like treewidth, δ-hyperbolicity indeed measures how far a graph is from a tree. Trees has hyperbolicity equal to zero and conversely any graph with hyperbolicity equal to zero can be isometrically embedded into a tree. Chordal graphs, which have a tree-like structure, has hyperbolicity equal to one. However, there is no relationship between the treewidth and hyperbolicity distance-to-tree measures. For instance a complete graph has large treewidth but is 0-hyperbolic. Conversely a n-cycle is $n/4$ -hyperbolic but has treewidth equal to two. A $n \times m$-grid has both treewidth and δ-hyperbolicity $\min(n, m)$ and is far from a tree in both measures.

It follows from its definition that δ-hyperbolicity can be computed in polynomial time. But the $O(n^4)$ naive implementation is slow and we use heuristics to obtain faster (and exact) computation.

Many problems can be solved efficiently for *δ-hyperbolic* graphs (classes of bounded δ-hyperbolicity). Let us cite fast diameter and center heuristics approximation (using two BFSs), and small stretch spanning tree computation [5]. Covering by balls and k-center, two NP-complete problems for general graphs, are addressed in [6]. A compact distance labeling (enabling to compute the distance between two nodes as a function of their labels) is given in [7]. Finally small-stretch additive spanners and compact routing computation are treated in [8].

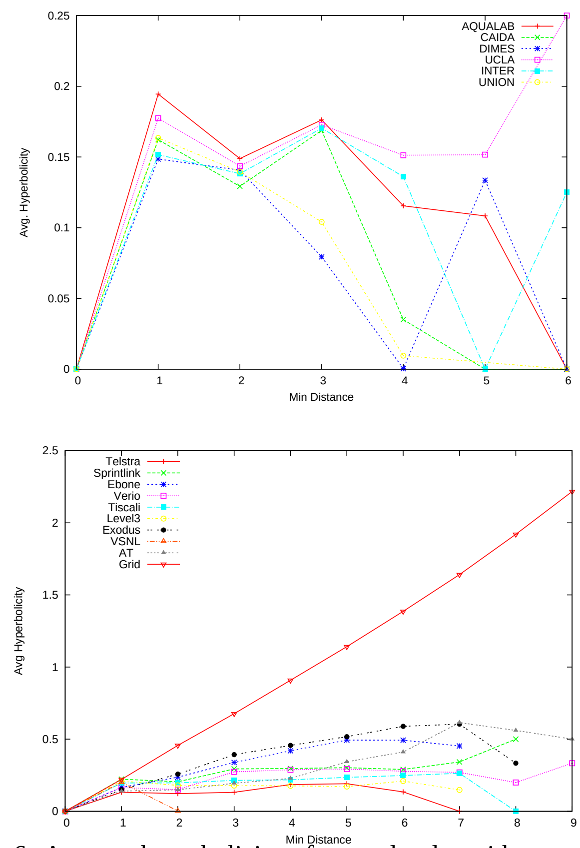


Fig. 6: Average hyperbolicity of a quadruplet with respect to the minimum distance d_3 in the quadruplet (d_3 is defined in Section IV) Top: AS graphs. Bottom: router graphs.

B. Results

We measured δ-hyperbolicity on various data. The δ-hyperbolicity of a graph is maximum over biconnected components of δ-hyperbolicity so we also worked only on the giant biconnected component to increase computation speed.

While classical definition involves only giving the *maximum* hyperbolicity over all quadruplets, we found that the *distribution* also is interesting, since it appears to exponentially decrease with δ and maximum can be deduced easily (see

Figure 5). One can observe that, for almost all quadruplets (x, y, z, t), we have $\delta(x, y, z, t) \leq 1$. Notice that if $\delta(x, y, z, t) = 0$ the shortest routes between these four points can be mapped on a tree. The ASes from Rocketfuel have all a similar behavior.

The distribution of hyperbolicity of the quadruplets seems to be relatively independent from the minimum distance between two points of the quadruplet for most of the snapshots studied (Figure 6), while in a grid it linearly depends from that distance. This is however consistent with a tree-like distance.

While observed maximum δ-hyperbolicity is never more than 2 for other graphs, MRINFO data behave very differently: 0.8% of quaduplets have $\delta = 2.5$ and the maximum of the graph is $\delta = 5$. This imply the existence of long *isometric cycles* [5] but we can not explain why.

The conclusion is then that the distances (in hop count) between any four vertices (routers or ASes) in the Internet is, on average, like in a tree, exactly or with and additive error of 1. Furthermore, as hyperbolicity is low, for any four vertices, their distances are always like in a tree with a small error. This allows efficient routing schemes to be used [8]. Krioukov et al. [34], [35] already noticed it was possible in the Internet, thanks to its scale-free structure (degrees and distances distribution), using a modified Thorup-Zwick scheme.

V. Internet models

We compare our results with various generated graphs, aiming at modeling the AS graphs. They have the same number of nodes than the largest bi-connected component of the CAIDA AS graph. They are:

- An Erdös-Rényi (ER) random graph with same average degree (6.3) than CAIDA AS graph.
- A random graph whose degree distribution follows a power law with the same exponent that in the CAIDA AS graph (it can be generated using [36]).
- A random graph (called "AS Degree Dist") with exactly the same degree distribution than the CAIDA AS graph [36].
- A graph generated using the Barabási-Albert (BA) model [29].

Table III shows the value of parameters studied for each model. The lower bounds for the **treewidth** of these graphs are slightly bigger than the bound of the AS graphs and still comparable to the treewidth of a square grid of the same size (107). That leads us to the conclusion that the treewidth of the Internet is slightly lower than the treewidth of a random graph of the same size and density. A graph generated to follow exactly the same distribution (by random matching) has a treewidth a bit nearer, but it is not relevant enough. The BA graph has treewidth of the same magnitude as the other random graphs.

Figure 8 gives the observed **hyperbolicity** for these generated graphs compared with the AS graphs. One can notice that randoms graphs with the same distribution (exactly, or power-law of same exponent) are very close to the AS graph. BA graph is however closer to ER graph.

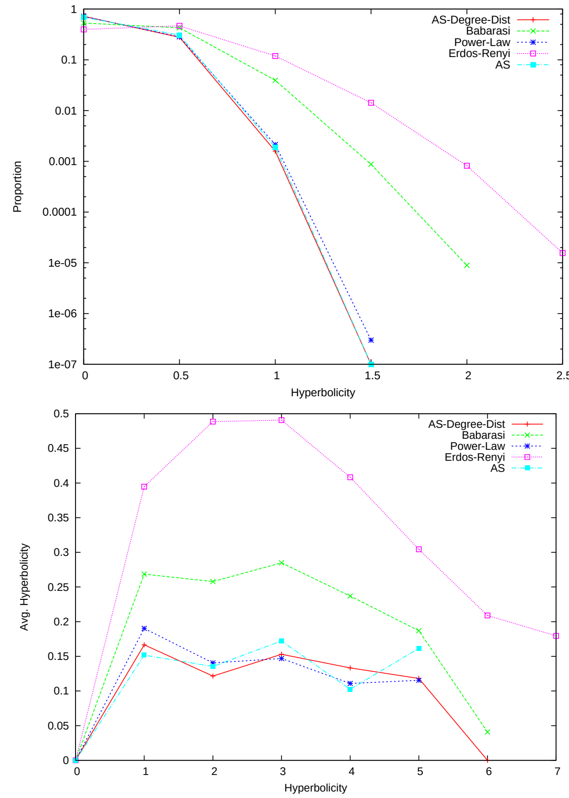


Fig. 8: *Internet models. Top: percentage of quadruplets having a given hyperbolicity. Bottom: mean hyperbolicity of a quadruplet with respect to the minimum distance between vertices in the quadruplet.*

Our conclusion meets Tangmunarunkit et al. work [17]. From both tree-likeness measurements point of view, the simpler (random) "degree based" generators produce even more accurate simulation than the more sophisticated "structural" generators like Barabási-Albert. Especially, for the hyperbolicity, random graphs are so close to the Internet data than they are a valid model of the Internet with respect to that parameter.

Graph	Avg. deg.	Max deg.	β	Hyp.	tw
CAIDA AS 01/09	6.31	1,815	2.19	2.0	82 - 473
Erdös-Rényi	6.34	18	-	2.5 -	135 -
Barabási	6.00	283	2.92	2.0 -	130 -
AS degree dist.	6.31	1,815	2.19	1.5 -	110 -
Power Law	8.97	1,507	2.19	1.5 -	150 -

TABLE III: Statistics for Internet models.

VI. Conclusion

We have observed that AS level snapshots exhibit an important churn over time. This can explain previous observations at the IP level.

We have established that the Internet has a high treewidth both at the AS level and the router level, i.e. a treewidth comparable to a square grid. This result holds independently of the accuracy of the data available. A better accuracy can only result in a higher treewidth.

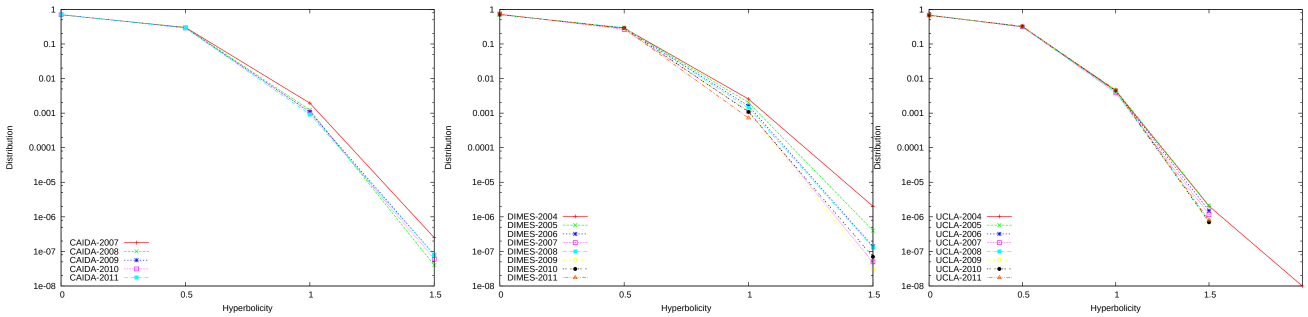


Fig. 7: Variation (or the lack of!) from year to year of hyperbolicity distribution of AS graphs (CAIDA, DIMES and UCLA).

As seen in [4], we observe that the hyperbolicity of almost all Internet snapshots is low. As we use a different hyperbolicity measure, our work comes as a confirmation of this fact. All AS graphs have roughly the same hyperbolicity distribution: for almost all quadruplets it is 0 or 1. However, we point out that this is not the case for MRINFO data, but we cannot conclude whether it is an artifact of the measurement method. Additionally, we observe that low hyperbolicity is a natural property of power law random graphs which appear again as a simple model capturing many structural properties of the Internet.

REFERENCES

[1] H. L. Bodlaender, "Dynamic programming on graphs with bounded treewidth," in *ICALP*, 1988, pp. 105–118.
[2] Y. Shavitt and T. Tankel, "Hyperbolic embedding of internet graph for distance estimation and overlay construction," *IEEE/ACM Trans. Netw.*, vol. 16, no. 1, pp. 25–36, 2008.
[3] V. Ramasubramanian, D. Malkhi, F. Kuhn, M. Balakrishnan, A. Gupta, and A. Akella, "On the treeness of internet latency and bandwidth," in *ACM SIGMETRICS*, New York, NY, USA, 2009, pp. 61–72.
[4] O. Narayan and I. Saniee, "The large scale curvature of networks," http://arxiv.org/abs/0907.1478v1, 2009.
[5] V. Chepoi, F. F. Dragan, B. Estellon, M. Habib, and Y. Vaxès, "Diameters, centers, and approximating trees of delta-hyperbolicgeodesic spaces and graphs," in *Symposium on Computational Geometry*, 2008.
[6] V. Chepoi and B. Estellon, "Packing and covering *delta* -hyperbolic spaces by balls," in *APPROX-RANDOM*, 2007, pp. 59–73.
[7] C. Gavoille and O. Ly, "Distance labeling in hyperbolic graphs," in *ISAAC*, 2005.
[8] V. Chepoi, F. F. Dragan, B. Estellon, M. Habib, Y. Vaxès, and Y. Xiang, "Additive spanners and distance and routing labeling schemes for delta-hyperbolic graphs," To appear in Algorithmica.
[9] Y. Hyun, B. Huffaker, D. Andersen, E. Aben, M. Luckie, kc claffy, and C. Shanno, "The ipv4 routed /24 as links dataset," 2009.
[10] Y. Shavitt and E. Shir, "Dimes: let the internet measure itself," *SIGCOMM Comput. Commun. Rev.*, vol. 35, pp. 71–74, October 2005. [Online]. Available: http://doi.acm.org/10.1145/1096536.1096546
[11] K. Chen, D. R. Choffnes, R. Potharaju, Y. Chen, F. E. Bustamante, D. Pei, and Y. Zhao, "Where the sidewalk ends: extending the internet as graph using traceroutes from p2p users," in *CoNEXT 2009 (Conf. on Emerging Networking Experiments and Tech.)*, 2009, pp. 217–228.
[12] R. Oliveira, D. Pei, W. Willinger, B. Zhang, and L. Zhang, "The (in)completeness of the observed internet AS-level structure," *IEEE/ACM Transac. on Networking*, vol. 18, no. 1, pp. 109–122, 2010.
[13] N. Spring, R. Mahajan, D. Wetherall, and T. Anderson, "Measuring isp topologies with rocketfuel," *IEEE/ACM Trans. Netw.*, vol. 12, no. 1, pp. 2–16, 2004.
[14] P. Mérindol, V. V. den Schrieck, B. Donnet, O. Bonaventure, and J.-J. Pansiot, "Quantifying ases multiconnectivity using multicast information," in *Internet Measurement Conference*, 2009, pp. 370–376.
[15] W. Willinger, D. Alderson, and J. C. Doyle, "Mathematics and the internet: A source of enormous confusion and great potential," *Notices of the AMS*, vol. 56, no. 5, pp. 586–599, 2009.
[16] C. Magnien, F. Ouédraogo, G. Valadon, and M. Latapy, "Fast dynamics in internet topology: preliminary observations and explanations," *CoRR*, vol. abs/0904.2716, 2009.
[17] H. Tangmunarunkit, R. Govindan, S. Jamin, S. Shenker, and W. Willinger, "Network topology generators: Degree-based vs. structural," in *ACM SIGCOMM*, 2002, p. 159.
[18] J.-J. Pansiot and D. Grad, "On routes and multicast trees in the internet," *Computer Communication Review*, vol. 28, no. 1, pp. 41–50, 1998.
[19] R. Govindan and A. Reddy, "An analysis of internet inter-domain topology and route stability," in *IEEE INFOCOM*, 1997, pp. 850–857.
[20] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On power-law relationships of the internet topology," in *ACM SIGCOMM*, 1999, pp. 251–262.
[21] A. Broido and k. claffy, "Internet topology: Connectivity of IP graphs," in *Proceedings of SPIE International Symposium on Convergence of IT and Communication*, 2001.
[22] Q. Chen, H. Chang, R. Govindan, S. Jamin, S. Shenker, and W. Willinger, "The origin of power-laws in internet topologies revisited," in *IEEE INFOCOM*, 2002.
[23] N. Robertson and P. D. Seymour, "Graph minors. i. excluding a forest," *J. Comb. Theory, Ser. B*, vol. 35, no. 1, pp. 39–61, 1983.
[24] T. S. E. Ng and H. Zhang, "Predicting internet network distance with coordinates-based approaches." in *IEEE INFOCOM*, 2002.
[25] F. Dabek, R. Cox, M. F. Kaashoek, and R. Morris, "Vivaldi: a decentralized network coordinate system." in *ACM SIGCOMM*, 2004.
[26] M. Gromov, "Hyperbolic groups," in *Essays in Group Theory*, ser. Math. Sciences Research Inst. Springer, 1987, vol. 8.
[27] B. D. Abrahao and R. D. Kleinberg, "On the internet delay space dimensionality," in *Internet Measurement Conference*, 2008, pp. 157–168.
[28] P. Fraigniaud, E. Lebhar, and L. Viennot, "The inframetric model for the internet," in *IEEE INFOCOM*, Phoenix, 2008, pp. 1085–1093.
[29] A. L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, 1999.
[30] A. Fabrikant, E. Koutsoupias, and C. H. Papadimitriou, "Heuristically optimized trade-offs: A new paradigm for power laws in the internet," in *ICALP*, 2002, pp. 110–122.
[31] W. Willinger, R. Oliveira, and B. Zhang, "Quantifying the completeness of the observed internet as-level structure abstract," UCLA, East Lansing, Michigan, Tech. Rep. TR 080026, September 2008.
[32] S. Arnborg, D. G. Corneil, and A. Proskurowski, "Complexity of finding embeddings in a k-tree," *SIAM Journal on Algebraic and Discrete Methods*, vol. 8, no. 2, pp. 277–284, 1987.
[33] B. Reed, "Tree width and tangles: A new connectivity measure and some applications," *Surveys in combinatorics*, vol. 241, pp. 87–162, 1997.
[34] D. V. Krioukov, K. R. Fall, and X. Yang, "Compact routing on internet-like graphs." in *IEEE INFOCOM*, 2004.
[35] D. V. Krioukov, K. Claffy, K. Fall, and A. Brady, "On compact routing for the internet," *ACM SIGCOMM*, vol. 37, no. 3, 2007.
[36] F. Viger and M. Latapy, "Efficient and simple generation of random simple connected graphs with prescribed degree sequence," in *Computing and Combinatorics*, ser. LNCS, 2005, vol. 3595, ch. 45, pp. 440–449–449.