

Exact and approximate algorithms for computing the hyperbolicity of large-scale graphs

David Coudert, Aurélien Lancin

▶ To cite this version:

David Coudert, Aurélien Lancin. Exact and approximate algorithms for computing the hyperbolicity of large-scale graphs. [Research Report] RR-8074, 2012. hal-00735481v1

HAL Id: hal-00735481 https://inria.hal.science/hal-00735481v1

Submitted on 25 Sep 2012 (v1), last revised 15 Mar 2013 (v4)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Exact and approximate algorithms for computing the hyperbolicity of large-scale graphs

David Coudert, Aurélien Lancin

RESEARCH REPORT N° 8074 September 2012 Project-Team MASCOTTE



Exact and approximate algorithms for computing the hyperbolicity of large-scale graphs

David Coudert^{*}, Aurélien Lancin[†]

Project-Team MASCOTTE

Research Report n° 8074 — September 2012 — 17 pages

Abstract: Let G be a connected graph, and let d(a, b) denotes the shortest path distance between vertices a and b of G. The graph G is δ -hyperbolic if for any vertices a, b, c, d of G, the two largest of the three sums $S_1 = d(a, b) + d(c, d)$, $S_2 = d(a, c) + d(b, d)$, and $S_3 = d(a, d) + d(b, c)$ differ by at most 2δ . This can be determined in time $O\left(\binom{n}{4}\right)$ which could be prohibitive for large graphs.

In this document, we propose an exact algorithm for determining the hyperbolicity of a graph that is scalable for large graphs. The time complexity of this algorithm is a function of the size of the largest bi-connected component of the graph, of the shortest path distance distribution in this componenant and of the value of the hyperbolicity. In the worst case, the time complexity is in $O\left(\binom{n}{2} \cdot \binom{n}{2} - 1\right)/2$, so higher than $O\left(\binom{n}{4}\right)$, but it is much faster in practice. We also propose both a multiplicative factor and an additive constant approximation algorithms. We then analyze further the time complexity of our exact algorithm for several class of graphs, and present some computational results on large-scale graphs.

Key-words: Hyperbolicity, algorithm, graph, approximation.

* MASCOTTE Project, INRIA/I3S(CNRS/UNSA), France. david.coudert@inria.fr † MASCOTTE Project, INRIA/I3S(CNRS/UNSA), France. aurelien.lancin@inria.fr

RESEARCH CENTRE SOPHIA ANTIPOLIS – MÉDITERRANÉE

2004 route des Lucioles - BP 93 06902 Sophia Antipolis Cedex

Algorithmes exact et approché pour le calcul de l'hyperbolicité d'un graphe

Résumé : Soit G un graph connexe et soit d(a, b) la distance entre les sommets a et b dans le graphe. Le graphe G est dit δ -hyperbolic si pour tout quadruplet a, b, c, d de sommets de G, les deux plus grandes des sommes $S_1 = d(a, b) + d(c, d), S_2 = d(a, c) + d(b, d)$, et $S_3 = d(a, d) + d(b, c)$ diffèrent d'au plus 2δ . Cette valeur peut-être déterminé en temps $O\left(\binom{n}{4}\right)$, ce qui est souvent inaccessible pour les grands graphes.

Nous proposons un nouvel algorithme exact pour calculer l'hyperbolicité d'un graphe, permettant de traiter des grands graphes. La complexité temporelle de cet algorithme est fonction de la taille de la plus grande composante bi-connexe du graphe, de la distribution des plus courts chemins dans cette composante, et de la valeur de l'hyperbolicité. Au pire, cet algorithme prendra un temps en $O\left(\binom{n}{2} \cdot \binom{n}{2} - 1\right)/2$, soit plus que $O\left(\binom{n}{4}\right)$. Cependant, l'algorithme est bien plus efficace en pratique. Nous proposons également un algorithme approché à un facteur multiplicatif ou à une constante additive donné en entrée. Nous analysons la complexité temporelle de notre algorithme pour des classes particulères de graphes et présentons des résultats expérimentaux sur des grands graphes.

Mots-clés : Hyperbolicité, algorithme, graphe, approximation.

1 Introduction

The (Gromov) hyperbolicity of a graph reflects how the metric (distances) of the graph is close to the metric of a tree. Gromov [7] defines the notion of δ -hyperbolic metric spaces using the notion of δ -thin triangles. Given any three points x, y, and z of a hyperbolic metric space, the triangle (x, y, z) is δ -thin if any point of the geodesic joining x and y is at distance at most δ of one of the geodesics joining x to z or y to z. A δ -hyperbolic space is a geodesic metric space in which every geodesic triangle is δ -thin. In other words, a graph has hyperbolicity $\leq \delta$ if, for any $u, v, w \in V(G)$ and for any shortest paths P_{uv}, P_{vw}, P_{uw} between these three vertices, any vertex in P_{uv} is at distance at most δ from $P_{vw} \cup P_{uw}$ [7]. Intuitively, in a graph with small hyperbolicity, any two shortest paths between the same pair of vertices are close to each other. For instance, a graph has hyperbolicity 0 if and only if it is acyclic.

An alternative definition given by Gromov [7], and called the 4-points condition, is the following. A metric space is δ -hyperbolic if for any four points u, v, w, x the two larger of the distance sums d(u, v) + d(w, x), d(u, w) + d(v, x), d(u, x) + d(v, w) differ by at most 2δ . These notions extend to connected graphs, and we say that a connected graph G = (V, E) equipped with its standard graph metric d_G (shortest path distance) is δ -hyperbolic if the metric space (V, d_G) is δ -hyperbolic. In other words, a connected graph G = (V, E) is δ -hyperbolic if it satisfies the 4-points condition. The hyperbolicity of a graph measures its tree-likeness. The less the value of δ is, the more the graph looks like a tree.

Determining the hyperbolicity δ of a graph of order N can therefore be done in time $O(N^4)$. In fact, since all 4-tuples have to be checked, the time complexity is in $O(\binom{N}{4})$. An implementation of the naive algorithm for determining the hyperbolicity of a graph as been included into the distory (Distance Between Phylogenetic Histories) package [6] of the CRAN (The Comprehensive R Archive Network) project [4]. This package is devoted to the study of geodesic distance between phylogenetic trees and associated functions. The implementation uses the "revolving doors Gray code" principle [8] for visiting all 4-tuples of the input graph. Unfortunately, the proposed implementation is impracticle for large and dense graphs (size of the largest biconnected component). For instance, for a graph with a biconnected component of order 10000, the number of 4-tuples to consider is higher than $4 \cdot 10^{14}$, which requires approximately 100 hours of computation assuming that 10^9 4-tuples are evaluated per second. A heuristic algorithm for determining the hyperbolicity of CAIDA AS maps has been used in [5]. Other heuristics based on sampling were used in [11, 10] for random Erdös-Renyi GNP graphs generated using probability p = c/N, where c is a small constant (1.5 or 2), and so the resulting graphs are almost trees having very small biconnected components. In addition, a 2-approximation algorithm, with running time in $O(\binom{N}{3})$, is obtained fixing one vertex and evaluating all possible 4-tuples containing that vertex [3].

The need for more efficient algorithms has been expressed by many authors, for instance it has recently been mentionned in the conclusion of [2] that "exact computation of δ by its definition takes $O(N^4)$ time, which is not scalable to large graphs, and thus the design of more efficient exact or approximation algorithms would be of interest".

The algorithm proposed below is the first exact algorithm scalable for large graphs.

Outline of this document. We start in Section 2 with some definitions and notations. Then we formally describe our algorithm for computing the hyperbolicity of a graph in Section 3 and give hints on its time complexity. We also we explain how to turn this algorithm into an approximation algorithm. In Section 4, we evaluate experimentally the performances of our exact algorithm. Next, in Section 5, we report on the computation of the hyperbolicity of large-scale graphs, namely CAIDA maps of the Internet autonomous systems topology. Last, we list some

simple results and proofs on the hyperbolicity of particular graph classes in Section A.

2 Definitions and notations

Let G be a connected graph, and let d(a, b) denotes the shortest path distance between vertices a and b of G. The graph G is δ -hyperbolic if for any vertices a, b, c, d of G, the two largest of the three sums $S_1 = d(a, b) + d(c, d)$, $S_2 = d(a, c) + d(b, d)$, and $S_3 = d(a, d) + d(b, c)$ differ by at most 2δ .

Some examples:

- *Block* graphs (i.e., connected graph in which every 2-connected subgraph is a clique) are 0-hyperbolic, and so are trees and cliques;
- Cycles of order $n = 4p + \varepsilon$, with $p \ge 1$ and $\varepsilon \in \{0, 1, 2, 3\}$, are (p 1/2)-hyperbolic when $\varepsilon = 1$, and p-hyperbolic otherwise (see proof in Appendix A);
- $n \times m$ grids, with $2 \le n \le m$, are (n-1)-hyperbolic (see proof in Appendix A).
- *d*-dimensional grids of side s are $(s-1)\lfloor d/2 \rfloor$ -hyperbolic (see proof in Appendix A);
- The hypercube $\mathcal{H}_2(n)$ of dimension n and order 2^n is $\lfloor n/2 \rfloor$ -hyperbolic (see proof in Appendix A);
- Chordal graphs are \leq 1-hyperbolic [1, 9].

3 Exact algorithm for computing the hyperbolicity

In this section, we formally describe a new exact algorithm for computing the hyperbolicity of graphs. We then give some hints on its time-complexity. Last but not least, we explain how to turn this algorithm into an approximation algorithm.

Our new algorithm is based on the following lemma, in which d(x, y) denotes the graph distance between nodes x and y.

Lemma 1. Let G = (V, E) be a connected graph, let $a, b, c, d \in V$, let $S_1 = d(a, b) + d(c, d)$, $S_2 = d(a, c) + d(b, d)$, and $S_3 = d(a, d) + d(b, c)$, and assume w.l.o.g. that $S_1 \ge \max \{S_2, S_3\}$. We have $h(a, b, c, d) \le \min \{d(a, b), d(c, d)\}$.

Proof. We have $S_2+S_3 = d(a,c)+d(b,d)+d(a,d)+d(b,c) = (d(a,c)+d(b,c))+(d(a,d)+d(b,d))$. Using the triangular inequality, we deduce $S_2 + S_3 \leq 2 \cdot d(a,b)$. Since S_1 is the largest sum, we have $h(a,b,c,d) = S_1 - \max\{S_2, S_3\} \leq S_1 - \max\{S_2, S_3\}/2 = S_1 - d(a,b) = d(c,d)$. We obtain similarly that $h(a,b,c,d) \leq d(a,b)$.

To make use of Lemma 1, we construct in Algorithm 1 the list SLL of 3-tuples ($S_1 = \ell_1 + \ell_2, \ell_1, \ell_2$), for $1 \leq \ell_2 \leq \ell_1 \leq D$, sorted in decreasing lexicographic order (lines 3-4). For instance, if D = 4, the 3-tuples are sorted in the following order:

$$(8,4,4), (7,4,3), (6,4,2), (6,3,3), (5,4,1), (5,3,2), (4,3,1), (4,2,2), (3,2,1), (2,1,1), (3,2,1), (2,1,1), (3,2,1), (3,2,1), (3,2,1), (3,3,1), ($$

We then compute for each 3-tuple $(S_1 = \ell_1 + \ell_2, \ell_1, \ell_2)$ the values $h(a, b, c, d) = S_1 - \max\{d(a, c) + d(b, d), d(a, d) + d(b, c)\}$ for all 4-tuple (a, b, c, d) such that $d(a, b) = \ell_1$ and $d(c, d) = \ell_2$. Thus, the algorithm considers the 4-tuples (a, b, c, d) in such a way that either S_1 is the largest sum, or, in case S_2 (resp. S_3) is larger than S_1 we know that the 4-tuple has previously been considered

4

with S_2 (resp. S_3) as maximum value, and so we have $h(a, b, c, d) \leq 0$. Then, thanks to Lemma 1 we know that at any step of the algorithm the value of ℓ_2 is an upper bound on the value of h(a, b, c, d). Also, if the current lower bound is h^* , none of the 4-tuples such that $d(c, d) \leq h^*$ can be used to improve the lower bound. We can thus cut exploration (lines 6-8).

Algorithm 1 Hyperbolicity

Require: G = (V, E) is a 2-connected graph. **Ensure:** δ , the hyperbolicity of G (observe that $\delta = 2 h^*$). 1: Let $P[\ell]$ be the list of pairs $(a, b) \in V \times V$ with a < b such that $d(a, b) == \ell$ 2: Sort $P[\ell]$ in increasing lexicographic order 3: Let *SLL* be the list of triples $(\ell_1 + \ell_2, \ell_1, \ell_2)$ for $1 \le \ell_2 \le \ell_1 \le D$ 4: Sort *SLL* by decreasing lexicographic order 5: Let $h^* := 0$ 6: for all (S_1, ℓ_1, ℓ_2) in SLL such that $\ell_2 > h^*$ do for all $(a,b) \in P[\ell_1]$, if $\ell_2 > h^*$ do 7: for all $(c,d) \in P[\ell_2]$, if $\ell_2 > h^*$ do {When $\ell_1 == \ell_2$, we ensure that (a,b) < (c,d) in 8: $P[\ell_1]$ $h^* := \max \{h^*, S_1 - \max \{d(a,c) + d(b,d), d(a,d) + d(b,c)\}\}$ 9: 10: **return** $h^*/2$

Since we have $\binom{n}{2}$ pairs, and that Algorithm 1 considers pairs of pairs, the worst case time complexity of the algorithm is in $O\left(\binom{n}{2} \cdot \binom{n}{2} - 1\right)/2$. However, with a more carefull analysis we observe that we can parameterized the time complexity with the optimal value of the hyperbolicity and the distribution of the path lengths. That is,

Proposition 2. Given a δ -hyperbolic graph G of diameter D and the sets $P[\ell]$ of pairs of vertices at distance ℓ from each other, the time complexity of lines 6-9 of Algorithm 1 is in

$$O\left(\sum_{\ell_1=2\delta}^{D} |P[\ell_1]| \left(\frac{|P[\ell_1]|-1}{2} + \sum_{\ell_2=\max\{1, 4\delta-\ell_1\}}^{\ell_1-1} |P[\ell_2]|\right)\right)$$

Proof. Let $P[\ell]$ be the list of pairs $(a, b) \in V \times V$ with a < b such that $d(a, b) = \ell$. For a δ -hyperbolic graph, Algorithm 1 will consider in the worst case all the triples $(\ell_1 + \ell_2, \ell_1, \ell_2)$ such that $\ell_1 + \ell_2 \ge 4\delta$ and $D \ge \ell_1 \ge \ell_2 \ge 1$. Furthermore, when $\ell_1 = \ell_2$, line 8 ensures that pair (a, b) < (c, d) in $P[\ell_1]$. Since we are not interested in the computation time of lines 1-4 (which is clearly dominated by the computation of the distances between each pairs of vertices), the result follows.

For instance, if the input graph is a $n \times n$ grid, with diameter 2n-2 and hyperbolicity $\delta = n-1$ (so $h^* = 2n-2$), the value of the hyperbolicity will be obtained with the first considered 4-tuple. Lines 6-9 of Algorithm 1 will thus be executed in constant time. On the other hand, if the input graph is a $n \times 2$ grid, with diameter n and hyperbolicity $\delta = 1$ (so $h^* = 2$), almost all 4-tuples will be considered and so the running time will be in $O\left(\binom{n}{2} \cdot \binom{n}{2} - 1\right)/2$). Chordal graphs, which have hyperbolicity at most 1 [1], are worst case instances for this algorithm since its running time increases with the gap between the diameter and the hyperbolicity.

Algorithm 1 requires to compute the distances better all pairs of vertices (line 1), to sort pairs of vertices at equal distances in lexicographic order (line 2), and also some negligeable operations (lines 3-4). Therefore we have

RR n° 8074

Corollary 3. Given a δ -hyperbolic graph G of diameter D and the sets $P[\ell]$ of pairs of vertices at distance ℓ from each other, the time complexity of Algorithm 1 is in

$$O\left(\max\left\{n(n+m), \ n^2\log n, \ \sum_{\ell_1=2\delta}^{D} |P[\ell_1]| \left(\frac{|P[\ell_1]|-1}{2} + \sum_{\ell_2=\max\{1, \ 4\delta-\ell_1\}}^{\ell_1-1} |P[\ell_2]|\right)\right\}\right)$$

Since the computational complexity of Algorithm 1 depends on the distance distribution of the graph and of the computed value of the hyperbolicity, we have not been able yet to obtain closed formula for its time complexity on particular graph classes such as Internet-like graphs. However, we will present in Section 4 some experimental results.

Turning Algorithm 1 into an approximation algorithm. Observe that at any step of the algorithm, the value of ℓ_1 is an upper bound for the hyperbolicity of G. Therefore, for large graphs inducing an excessive running time (some instances may require weeks of computations), Algorithm 1 can be turned into an approximation algorithm with approximation ratio $\frac{\ell_1}{h^*}$. It suffices for instance to stop the execution of the algorithm after a given time (for instance one hour) and to return the values h^{*} and ℓ_1 . More precisely, we can insert one of the following test between line 6 and line 7 of the algorithm:

- "If computation time is larger than allowed computation time, then stop computations and return h^{*} and ℓ_1 ". We get $\frac{h^*}{2} \leq \delta \leq \frac{\ell_1}{2}$;
- "If $\frac{\ell_1}{h^*} \leq apx$, then stop computations and return $\frac{h^*}{2}$." We get an approximation of the value δ of the hyperbolicity with proven approximation factor apx (i.e.., $\frac{h^*}{2} \leq \delta \leq apx \cdot \frac{h^*}{2}$);
- "If $\frac{\ell_1}{2} \frac{h^*}{2} \le apx$, then stop computations and return $\frac{h^*}{2}$." We get an approximation of the value δ of the hyperbolicity with proven additive approximation constant apx (i.e., $\frac{h^*}{2} \le \delta \le \frac{h^*}{2} + apx$).

As we show in the next section, the main part of the running time of the algorithm consists in closing the small gap between lower and upper bounds that are generally found very quickly. Therefore, depending on the expected result, it may be appropriate to use the above rules that may allow to save a lot of time while preserving a sufficient precision.

4 Experimental performances

In this section, we evaluate the performances of Algorithm 1 described above. This evaluation is done by comparing the running time of Algorithm 1 with the time of the naive algorithm $O\binom{N}{4}$. Our objective is to determine the topologies for which our algorithm is better than the naive one. Indeed, Algorithm 1 appears to be worse than the naive one for some pathological topologies such as grids. The good news are that its running-time outperforms the one of known algorithms in the case of Internet-like graphs. In particular, Algorithm 1 allows us to compute the hyperbolicity of the last CAIDA AS maps (N = 25815 for the largest bi-connected component of the CAIDA map of September 2012) in two weeks.

We have reported in Table 1 the evolution of the computation time of Algorithm 1 on $N \times M$ grids such that $N * M = 2^2 * 3^2 * 4^2 = 576$ and $N * M = 2^2 * 3^2 * 5^2 = 900$. The running times are averages over 10 executions of the algorithm. The hyperbolicity of a $N \times M$ grid is $\delta_{N \times M} = \min \{N, M\} - 1$. As expected, the computation time on square grids is way smaller than for grids with sides of very different sizes. In Section 3, we said that rectangular grids are some of the worst case instances for Algorithm 1. To verify this claim, we have plotted in Figure 1 the relative number of visited 4-tuples with Algorithm 1 compared to the total number of 4-tuples in the graph. Recall that a 4-tuple may be visited up to three times by the algorithm. In these plots, we observe that as soon as the sides of the grids differ by a factor at least 6, then the computation time of Algorithm 1 is larger than the naive algorithm, but it is much faster when the sides of the grid differ by a factor less than 6.

N	M	δ	time	
24	24	23	0.08	
18	32	17	0.16	
16	36	15	1.13	
12	48	11	15.01	
9	64	8	40.33	
8	72	7	50.79	
6	96	5	71.80	
4	144	3	90.60	
3	192	2	97.08	
2	288	1	101.82	
(a) $N * M = 576$				

N	M	δ	time	
30	30	29	0.16	
25	36	24	0.18	
20	45	19	5.98	
18	50	17	23.10	
15	60	14	86.94	
12	75	11	201.71	
10	90	9	297.54	
9	100	8	349.51	
6	150	5	496.64	
5	180	4	537.42	
4	225	3	574.37	
3	300	2	599.19	
2	450	1	622.15	
(b) $N * M = 900$				

Table 1: Computation time in secondes of the hyperbolicity of $N \times M$ grids such that $N * M = 2^2 * 3^2 * 4^2 = 576$ (Table 1a) and $N * M = 2^2 * 3^2 * 5^2 = 900$ (Table 1b).



Figure 1: Relative performance of Algorithm 1 vs. the naive algorithm on $N \times M$ grids.

Next, we have reported in Figure 2 the running time of Algorithm 1 on Barabasi-Albert (BA) graphs. We have generated graphs with number of nodes in the range $[1\ 000..10\ 000]$ and with different values of the degree k of newly added nodes. We have generated 100 graphs for each couple number of nodes and parameter k, and reported the average values. Figure 2a reports the average values of the hyperbolicity of these graphs. We observe that the hyperbolicity of BA graphs decreases with the degree of newly added nodes. Figure 2b represents the average

RR n° 8074

computation time of Algorithm 1 on these graphs. We observe a slow increase of the computation time with the increase of the parameter k. This slow increase was expected since the time complexity of both the all-pairs shortest path algorithm and the decomposition into bi-connected components algorithm depend on the number of edges of the graph and so on the parameter k. We also observe in Figure 2b some pics in the computation times for particular combinations of the number of nodes and the parameter k. However, at the time of writing this deliverable we have not been able to explain this surprizing behavior. Last, we have reported in Figure 2c the ratio of the average number of 4-tuples visited by Algorithm 1 over the total number of 4-tuples of the graphs. This highlights the drastic running time improvement of Algorithm 1 over the naive algorithm to compute the hyperbolicity of BA graphs. More precisely, Algorithm 1 is between 10^3 and 10^{11} times faster than the naive algorithm.



(c) Ratio visited/total number of 4-tuples

Figure 2: Computation time of the hyperbolicity of Barabasi-Albert graphs.

Last, we have reported in Table 2 the size of the largest bi-connected component (BCC), the hyperbolicity, the running time and the number of visited 4-tuples by Algorithm 1 when computing the hyperbolicity of some CAIDA maps. We have selected CAIDA maps with different values of the hyperbolicity to highlight the running time improvement of Algorithm 1. As expected with the time complexity expressed in Proposition 2, the running time improvement is correlated to the hyperbolicity of the CAIDA maps. Indeed, Algorithm 1 is more than 100 times faster than the naive algorithm for these maps. Nevertheless, the running times remains large. However, as explained in Section 3, at each step of the execution of Algorithm 1 we get proven lower and upper bounds for the hyperbolicity of the considered graph. Also, we have plotted in Figure 3 the time at which new lower and upper bounds are obtained when computing the

hyperbolicity of the maps of Table 2. We observe that a few minutes of computations is sufficient for the lower bound to reach the optimal value but that the time to decrease the upper bounds to the optimal value could be very long. In fact, for maps with low hyperbolicity, almost all computation time is spent to prove the optimality of the lower bound.

AS map name	Largest BCC	Hyperbolicity	Time	Visited 4-tuples	Tot. 4-tuples
2004/01/05	10 424	2.5	336s	$3.8 \cdot 10^{10}$	$\simeq 5 \cdot 10^{14}$
2004/06/07	11 100	2.0	18h50m	$8.8 \cdot 10^{12}$	$\simeq 6.3 \cdot 10^{14}$
2005/09/05	12957	3.0	58s	$1.2 \cdot 10^{8}$	$\simeq 1.2 \cdot 10^{15}$
2012/06/01	25 815	2.0	15d	$1.8 \cdot 10^{15}$	$\simeq 1.8 \cdot 10^{16}$

1000 000,0 05/01/2004 -- LE 05/01/2004 -- UB 100 000.0 07/06/2004 --- I B 07/06/2004 -- UB 10 000,0 05/09/2005 -- LB 05/09/2005 -- UB 1 000,0 01/06/2012 -- LB Time in sec. (log. scale) 01/06/2012 -- UB 100,0 10,0 1,0 0.1 0,5 1 2,5 3 3,5 1,5 Hyperbolicity

Table 2: Computation time of the hyperbolicity of some CAIDA maps.

Figure 3: Time to reach lower and upper bounds of the hyperbolicity of CAIDA maps since 2004. Computation times for lower bounds are plotted from left to right, and from right to left for upper bounds.

5 Hyperbolicity of large-scale graphs

Thanks to Algorithm 1, we have been able to compute the hyperbolicity of all available CAIDA maps since 2004 (173 maps). Although the new algorithm improves upon the naive one, the overall computation time represents several months of computations. Indeed, the determination of the hyperbolicity of the 2011 CAIDA map only took 12 days of computations (but maps with hyperbolicity 3 requires only a few minutes of computations). The computed values are reported in Figure 4 where the first axis corresponds to the dates the maps where produced. We have one dot per CAIDA map and we have also reported the linear interpolation of the hyperbolicity.

Figure 4 provides several informations. First, we observe some measurement bias, for instance for the map of 07/06/2004 which has hyperbolicity 2 while maps produced in the months before and after have hyperbolicity 2.5. The same holds for the maps with hyperbolicity 3 (05/09/2005and 06/02/2006). A frequent variation is also observed between consecutive maps in the period from 2007 till 2009. We still do not know whether this behavior is due to some bias of the measurement or whether it comes from some hidden fact.



Figure 4: Evolution of the hyperbolicity of CAIDA maps since 2004.

The main observation resulting from this experimet is that the hyperbolicity of CAIDA maps has decreased in average from 2.5 to 2 and the hyperbolicity is stable since 2009. This is not surprising and it is clearly due to the fact that the AS network has become bigger during the last decade. In particular, many new links have appeared and have broken large cycles that made 4-tuples with bigger hyperbolicity.

Acknowledgments

This work has been partially supported FP7 STREP EULER, and ANRs AGAPE and GRATEL.

References

- G. Brinkmann, J. H. Koolen, and V. Moulton. On the hyperbolicity of chordal graphs. Annals of Combinatorics, 2001.
- [2] W. Chen, W. Fang, G. Hu, and M. W. Mahoney. On the hyperbolicity of small-world and tree-like random graphs. Technical Report arXiv:1201.1717, September 2012.
- [3] V. Chepoi, F. F. Dragan, B. Estellon, M. Habib, and Y. Vaxès. Notes on diameters, centers, and approximating trees of delta-hyperbolic geodesic spaces and graphs. *Electronic Notes* in Discrete Mathematics, 31:231–234, 2008.
- [4] CRAN: The Comprehensive R Archive Network project. http://cran.r-project.org/.
- [5] F. de Montgolfier, M. Soto, and L. Viennot. Treewidth and hyperbolicity of the internet. In 10th IEEE International Symposium on Networking Computing and Applications (NCA), pages 25–32. IEEE Comp. Soc., 2011.
- [6] distory: Distance between phylogenetic histories. http://cran.r-project.org/web/ packages/distory/index.html.
- [7] M. Gromov. Hyperbolic groups. Essays in Group Theory, 8:75-263, 1987.
- [8] D. Knuth. *The art of computer programming*, volume 4, Fascicle 2, chapter Generating All tuples and permutations. Addison-Wesley, February 2005.

- [9] J. H. Koolen and V. Moulton. Hyperbolic bridged graphs. Eur. J. Comb., 23(6):683–699, 2002.
- [10] O. Narayan, I. Saniee, and G. H. Tucci. Lack of spectral gap and hyperbolicity in asymptotic Erdös-Renyi sparse random graphs. In 5th International Symposium on Communications, Control and Signal Processing (ISCCSP), pages 1–4, Rome, Italy, May 2012. IEEE.
- [11] Y. Shang. Lack of gromov-hyperbolicity in colored random networks. *PanAmerican Mathematical Journal*, 21(1):27–36, 2011.

A Simple results

Lemma 4. Cycles of order $n = 4p + \varepsilon$, with $p \ge 1$ and $\varepsilon \in \{0, 1, 2, 3\}$, are (p - 1/2)-hyperbolic when $\varepsilon = 1$, and p-hyperbolic otherwise.

Proof of Lemma 4. Let C be a cycle of order $n = 4p + \varepsilon$, with $p \ge 1$ and $\varepsilon \in \{0, 1, 2, 3\}$, and let the vertices be indexed $\{0, 1, \dots, 4p + \varepsilon - 1\}$. Let a, b, c, d be four vertices of C. In order to break symmetries and w.l.o.g. we assume that $a = 0 < b \le p < c \le 2p + \lfloor \varepsilon/2 \rfloor < d \le 3p + \lfloor \varepsilon/2 \rfloor + \lfloor \varepsilon/3 \rfloor$, and so we have

$$1 \le \mathbf{d}(a, b) \le p \tag{1a}$$

$$p+1 \le d(a,c) \le 2p + |\varepsilon/2| \tag{1b}$$

$$p + \varepsilon - \lfloor \varepsilon/2 \rfloor - \lfloor \varepsilon/3 \rfloor \le d(a, d) \le 2p + \varepsilon - \lfloor \varepsilon/2 \rfloor - 1$$
 (1c)

$$1 \le d(b,c) \le 2p + \lfloor \varepsilon/2 \rfloor - 1 \tag{1d}$$

$$p + \lfloor \varepsilon/2 \rfloor + 1 \le d(b, d) \le 2p + \lfloor \varepsilon/2 \rfloor$$
(1e)

$$1 \le d(c,d) \le 2p + \lfloor \varepsilon/2 \rfloor + \lfloor \varepsilon/3 \rfloor - 1 \tag{1f}$$

and,

$$2 \le S_1 = d(a,b) + d(c,d) \le 3p + \lfloor \varepsilon/2 \rfloor + \lfloor \varepsilon/3 \rfloor - 1$$
(1g)

$$2p + \lfloor \varepsilon/2 \rfloor + 2 \le S_2 = d(a,c) + d(b,d) \le 4p + 2\lfloor \varepsilon/2 \rfloor$$
(1h)

$$p + \varepsilon - \lfloor \varepsilon/2 \rfloor - \lfloor \varepsilon/3 \rfloor + 1 \le S_3 = d(a, d) + d(b, c) \le 4p + \varepsilon - 2$$
(1i)

For any values of b, c, d, we have $S_1 + S_3 = d(a, b) + d(c, d) + d(a, d) + d(b, c) = 4p + \varepsilon$, and so maximizing S_1 decreases S_3 and conversely.

Observe that the maximum value of S_1 is obtained with b = p, c = p + 1, and $d = 3p + \lfloor \varepsilon/2 \rfloor + \lfloor \varepsilon/3 \rfloor$. For these values we have $S_2 = 3p + \lfloor \varepsilon/2 \rfloor + \lfloor \varepsilon/3 \rfloor + 1 > S_1$, $S_3 = p + \varepsilon - \lfloor \varepsilon/2 \rfloor - \lfloor \varepsilon/3 \rfloor + 1 < S_1$, and so h(a, b, c, d) = 1. Similarly, the maximum value of S_3 is obtained with b = 1, $c = 2p + \lfloor \varepsilon/2 \rfloor$, and $d = 2p + \lfloor \varepsilon/2 \rfloor + 1$, which gives $S_1 = 2$, $S_2 = 4p + 2 \lfloor \varepsilon/2 \rfloor$, and so $h(a, b, c, d) = \lfloor \varepsilon/2 \rfloor + 1$, which gives $S_1 = 2$, $S_2 = 4p + 2 \lfloor \varepsilon/2 \rfloor$, and so $h(a, b, c, d) = \lfloor \varepsilon - 2 - 2 \lfloor \varepsilon/2 \rfloor \rfloor \leq 2$.

Let us now assume that $S_2 = \max \{S_1, S_2, S_3\}$. Since $S_1 + S_3 = 4p + \varepsilon$, the best we can do to maximize h(a, b, c, d) is to choose values of a, b, c, d such that the maximum among S_1 and S_3 is the closest possible to $2p + \varepsilon/2$, and such that S_2 is as large as possible. This can be achieve choosing $b = p, c = 2p + \lfloor \varepsilon/2 \rfloor$, and $d = 3p + \lfloor \varepsilon/2 \rfloor + \lfloor \varepsilon/3 \rfloor$. Indeed, with these values, we get $S_2 = 4p + 2 \lfloor \varepsilon/2 \rfloor$, so the maximum possible value, and $S_1 = 2p + \lfloor \varepsilon/3 \rfloor$ and $S_3 = 2p + \varepsilon - \lfloor \varepsilon/3 \rfloor \ge S_1$ which are the closest possible of $2p + \varepsilon/2$. We obtain $h(a, b, c, d) = 2p + 2 \lfloor \varepsilon/2 \rfloor - \varepsilon + \lfloor \varepsilon/3 \rfloor$.

Since $\varepsilon \in \{0, 1, 2, 3\}$, it follows that cycles of order $n = 4p + \varepsilon$, with $p \ge 1$, are (p - 1/2)-hyperbolic when $\varepsilon = 1$, and p-hyperbolic otherwise.

Lemma 5. $n \times m$ grids, with $2 \le n \le m$, are (n-1)-hyperbolic.

Proof of Lemma 5. Let $\{v_{i,j}, 1 \leq i \leq n, 1 \leq j \leq m\}$ the vertices of the $n \times m$ grid.

Firstly, suppose that n = m, and consider the four vertices $v_{1,1}, v_{1,n}, v_{n,1}, v_{n,n}$. We have $S_1 = d(v_{1,1}, v_{1,n}) + d(v_{n,1}, v_{n,n}) = 2n - 2$, $S_2 = d(v_{1,1}, v_{n,n}) + d(v_{1,n}, v_{n,1}) = 4n - 4$, $S_3 = d(v_{1,1}, v_{n,1}) + d(v_{1,n}, v_{n,n}) = 2n - 2$, and so $h(v_{1,1}, v_{1,n}, v_{n,1}, v_{n,n}) = 2n - 2$. Since h(a, b, c, d) is less or equal to the diameter of the graph for any 4-tuple (a, b, c, d), this value is optimal.

Secondly, let $n \leq x \leq m$ and consider the four vertices $v_{1,1}, v_{1,x}, v_{n,1}, v_{n,x}$. We have $S_1 = d(v_{1,1}, v_{1,x}) + d(v_{n,1}, v_{n,x}) = 2x - 2$, $S_2 = d(v_{1,1}, v_{n,x}) + d(v_{1,x}, v_{n,1}) = 2n + 2x - 4$, $S_3 = d(v_{1,1}, v_{n,1}) + d(v_{1,x}, v_{n,x}) = 2n - 2$, and since $S_2 \geq S_1 \geq S_3$ we obtain $h(v_{1,1}, v_{1,x}, v_{n,1}, v_{n,x}) = 2n - 2$.

Let now $v_{i_1,j_1}, v_{i_2,j_2}, v_{i_3,j_3}, v_{i_4,j_4}$ be four vertices of the grid such that $1 < i_1, i_2 \le n/2 < i_3, i_4 < n$, and $1 < j_1, j_3 \le n/2 < j_2, j_4 < n$, and let us show that $h(v_{i_1,j_1}, v_{i_2,j_2}, v_{i_3,j_3}, v_{i_4,j_4}) \le h(v_{i_1-1,j_1-1}, v_{i_2,j_2}, v_{i_3,j_3}, v_{i_4,j_4})$. We have $S_1 = d(v_{i_1,j_1}, v_{i_2,j_2}) + d(v_{i_3,j_3}, v_{i_4,j_4}), S_2 = d(v_{i_1,j_1}, v_{i_3,j_3}) + d(v_{i_2,j_2}, v_{i_4,j_4})$, and $S_3 = d(v_{i_1,j_1}, v_{i_4,j_4}) + d(v_{i_2,j_2}, v_{i_3,j_3})$.

According the shortest path distance in the grid, we have

$$d(v_{i_1,j_1}, v_{i_2,j_2}) + 1 \le d(v_{i_1-1,j_1-1}, v_{i_2,j_2}) \le d(v_{i_1,j_1}, v_{i_2,j_2}) + 2$$
(2a)

$$d(v_{i_1,j_1}, v_{i_3,j_3}) + 1 \le d(v_{i_1-1,j_1-1}, v_{i_3,j_3}) \le d(v_{i_1,j_1}, v_{i_3,j_3}) + 2$$

$$d(v_{i_1,j_1}, v_{i_3,j_3}) + 1 \le d(v_{i_1-1,j_1-1}, v_{i_3,j_3}) \le d(v_{i_1,j_1}, v_{i_3,j_3}) + 2$$

$$(2b)$$

$$d(v_{i_1-1,j_1-1}, v_{i_4,j_4}) = d(v_{i_1,j_1}, v_{i_4,j_4}) + 2$$
(2c)

and so

$$S_1 + 1 \le S'_1 = d(v_{i_1 - 1, j_1 - 1}, v_{i_2, j_2}) + d(v_{i_3, j_3}, v_{i_4, j_4}) \le S_1 + 2$$
(2d)

$$S_2 + 1 \le S'_2 = \mathrm{d}(v_{i_1 - 1, j_1 - 1}, v_{i_3, j_3}) + \mathrm{d}(v_{i_2, j_2}, v_{i_4, j_4}) \le S_2 + 2 \tag{2e}$$

$$S'_{3} = d(v_{i_{1}-1,j_{1}-1}, v_{i_{4},j_{4}}) + d(v_{i_{2},j_{2}}, v_{i_{3},j_{3}}) = S_{3} + 2$$
(2f)

From these values, we get $|S'_1 - S'_2| \leq |S_1 - S_2| + 1$, $|S'_1 - S'_3| \leq |S_1 - S_3| + 1$, and $|S'_2 - S'_3| \leq |S_2 - S_3| + 1$. Therefore, we have $h(v_{i_1,j_1}, v_{i_2,j_2}, v_{i_3,j_3}, v_{i_4,j_4}) \leq h(v_{i_1-1,j_1-1}, v_{i_2,j_2}, v_{i_3,j_3}, v_{i_4,j_4})$, and we can prove the same result when decreasing some of the values of i_1, i_2, j_1, j_3 or increasing some of the values of i_3, i_4, j_2, j_4 .

Altogether, we conclude that $n \times m$ grids, with $2 \le n \le m$, are (n-1)-hyperbolic.

Lemma 6. The d-dimensional grid of side s is $(s-1)\lfloor d/2 \rfloor$ -hyperbolic.

Proof. The vertices of the *d*-dimensional grid of side *s* are labeled $(u_0, u_1, \cdots, u_{d-1})$, with $u_i \in [0..s-1]$ for $i \in [0..d-1]$, and there is an edge between $u = (u_0, u_1, \cdots, u_{d-1})$ and $v = (v_0, v_1, \cdots, v_{d-1})$ if $u_i = v_i$ for $i \in [0..j-1] \cup [j+1..d-1]$ and either $v_j = u_j + 1$ if $u_j + 1 < s$ or $v_j = u_j - 1$ if $u_j - 1 \ge 0$.

Consider now the four vertices u, v, w, x such that

$$u = (0, 0, \cdots, 0)$$
 contains only 0's (3a)

 $v = (s - 1, s - 1, \cdots, s - 1) \quad \text{contains only } s - 1\text{'s}$ (3b)

 $w = (w_0, w_1, \cdots, w_{d-1}) \quad \text{contains } \lfloor d/2 \rfloor \text{ 0's and } \lceil d/2 \rceil s - 1\text{'s}$ (3c)

 $x = (x_0, x_1, \dots, x_{d-1})$ obtained from w replacing each 0 with a s-1 and each s-1 with a 0 (3d)

We have for instance $d(u, w) = (s - 1) \lfloor d/2 \rfloor$, and so

$$S_1 = d(u, v) + d(w, x) = 2(s - 1)d$$
(3e)

$$S_2 = d(u, w) + d(v, x) = 2(s - 1) \left\lceil d/2 \right\rceil$$
(3f)

$$S_3 = d(u, x) + d(v, w) = 2(s - 1) \lfloor d/2 \rfloor$$
(3g)

We obtain $h(u, v, w, x) = S_1 - S_3 = 2(s - 1) \lfloor d/2 \rfloor$.

The optimality proof is similar to the proof of Lemma 5.
$$\Box$$

Corollary 7. The hypercube $\mathcal{H}_2(n)$ of dimension n and order 2^n is |n/2|-hyperbolic.

Proof of Corollary 7. The proof follows from Lemma 6 since the *n*-dimensional grid of side 2 is the hypercube $\mathcal{H}_2(n)$.

RR n° 8074

B Hyperbolicity of some CAIDA AS maps

Table 3 lists the hyperbolicity of all CAIDA AS maps since 2004. We have also reported in this table the number of 4-tuples with given hyperbolicity among a large set of randomly chosen 4-tuples. We can derive from these values the statistical distribution of the hyperbolicity of the 4-tuples. Last, we have reported some hyperbolicity *certificates*, i.e., a 4-tuple with maximum hyperbolicity.

AS map date	δ	Distribution sampling over 10 ⁷ 4-tuples	Certificate
2004/01/05	2.5	$\{0: 6715039, 0.5: 3219787, 1: 65133, 1.5: 41\}$	[3233, 8338, 8923, 13268]
2004/02/02	2.5	$\{0: 6861885, 0.5: 3089540, 1: 48558, 1.5: 17\}$	
2004/03/01	2.5	$\{0: 6958090, 0.5: 3005122, 1: 36783, 1.5: 5\}$	
2004/04/05	2.5	$\{0: 7046925, 0.5: 2928389, 1: 24683, 1.5: 3\}$	
2004/05/03	2.5	$\{0: 7130375, 0.5: 2831755, 1: 37851, 1.5: 19\}$	
2004/06/07	2	$\{0: 7123845, 0.5: 2842636, 1: 33503, 1.5: 16\}$	
2004/07/05	2.5	$\{0: 7201729, 0.5: 2759975, 1: 38285, 1.5: 11\}$	
2004/08/02	2.5	$\{0: 7303943, 0.5: 2668190, 1: 27859, 1.5: 8\}$	
2004/09/06	2.5	$\{0: 7289703, 0.5: 2683638, 1: 26647, 1.5: 12\}$	
2004/10/04	2.5	$\{0: 7335555, 0.5: 2640181, 1: 24256, 1.5: 8\}$	
2004/11/01	2.5	$\{0: 7307127, 0.5: 2670440, 1: 22426, 1.5: 7\}$	
2004/12/06	2.5	$\{0: 7324065, 0.5: 2653489, 1: 22436, 1.5: 10\}$	
2005/01/03	2.5	$\{0: 7224372, 0.5: 2742389, 1: 33221, 1.5: 18\}$	
2005/02/07	2.5	$\{0: 7359925, 0.5: 2620657, 1: 19414, 1.5: 4\}$	
2005/03/07	2.5	$\{0: 7328803, 0.5: 2650637, 1: 20553, 1.5: 7\}$	
2005/04/04	2.5	$\{0: 7356271, 0.5: 2624514, 1: 19213, 1.5: 2\}$	
2005/05/02	2.5	$\{0: 7328310, 0.5: 2652281, 1: 19404, 1.5: 5\}$	
2005/06/06	2.5	$\{0: 7333564, 0.5: 2645964, 1: 20467, 1.5: 5\}$	
2005/07/04	2.5	$\{0: 7350693, 0.5: 2628664, 1: 20640, 1.5: 3\}$	
2005/08/01	2.5	$\{0: 7333700, 0.5: 2647182, 1: 19116, 1.5: 2\}$	
2005/09/05	3	$\{0: 7321286, 0.5: 2658325, 1: 20382, 1.5: 7\}$	[12739, 12764, 16010, 34645]
2005/10/03	2.5	$\{0: 6552966, 0.5: 3409964, 1: 37067, 1.5: 3\}$	
2005/11/07	2.5	$\{0: 6562109, 0.5: 3416873, 1: 21011, 1.5: 7\}$	
2005/12/05	2.5	$\{0: 6603889, 0.5: 3376486, 1: 19623, 1.5: 2\}$	
2006/01/02	2.5	$\{0: 6249800, 0.5: 3683136, 1: 67058, 1.5: 6\}$	
2006/01/09	2.5	$\{0: 6277033, 0.5: 3629123, 1: 93815, 1.5: 29\}$	
2006/01/16	2.5	$\{0: 6430631, 0.5: 3503733, 1: 65607, 1.5: 29\}$	
2006/01/23	2.5	$\{0: 6259049, 0.5: 3647352, 1: 93504, 1.5: 95\}$	
2006/01/30	2.5	$\{0: 6259226, 0.5: 3644761, 1: 95918, 1.5: 95\}$	
2006/02/06	3	$\{0: 6251194, 0.5: 3648303, 1: 100398, 1.5: 105\}$	[1955, 13210, 20722, 34416]
2006/02/13	2.5	$\{0: 6438592, 0.5: 3495437, 1: 65921, 1.5: 50\}$	
2006/02/20	2.5	$\{0: 6280456, 0.5: 3621250, 1: 98186, 1.5: 108\}$	
2006/02/27	2.5	$\{0: 6258051, 0.5: 3637521, 1: 104291, 1.5: 137\}$	
2006/03/06	2.5	$\{0: 6277239, 0.5: 3622292, 1: 100368, 1.5: 101\}$	
2006/03/13	2.5	$\{0: 6467799, 0.5: 3463946, 1: 68227, 1.5: 28\}$	
2006/03/20	2.5	$\{0: 6289738, 0.5: 3610075, 1: 100106, 1.5: 81\}$	
2006/03/27	2.5	$\{0: 6668540, 0.5: 3271859, 1: 59522, 1.5: 79\}$	
, ,	1		Continued on next page

Table 3: Hyperbolicity of CAIDA AS maps since 2004.

Inria

AS map date	δ	Distribution sampling over 10 ⁷ 4-tuples	Certificate
2006/04/03	2.5	$\{0: 6550807, 0.5: 3375904, 1: 73127, 1.5: 162\}$	
2006/04/10	2.5	$\{0: 6719065, 0.5: 3219717, 1: 61145, 1.5: 73\}$	
2006/04/17	2.5	$\{0: 6413929, 0.5: 3558977, 1: 27082, 1.5: 12\}$	
2006/04/24	2.5	$\{0: 6414858, 0.5: 3556266, 1: 28861, 1.5: 15\}$	
2006/05/01	2.5	$\{0: 6332056, 0.5: 3619895, 1: 48032, 1.5: 17\}$	
2006/05/08	2.5	$\{0: 6411851, 0.5: 3557981, 1: 30160, 1.5: 8\}$	
2006/05/15	2.5	$\{0: 6857882, 0.5: 3104832, 1: 37271, 1.5: 15\}$	
2006/05/22	2.5	$\{0: 6398173, 0.5: 3563320, 1: 38500, 1.5: 7\}$	
2006/05/29	2.5	$\{0: 6409381, 0.5: 3557696, 1: 32914, 1.5: 9\}$	
2006/06/05	2.5	$\{0: 6836190, 0.5: 3117207, 1: 46583, 1.5: 20\}$	
2006/06/12	2.5	$\{0: 6399370, 0.5: 3568385, 1: 32240, 1.5: 5\}$	
2006/06/19	2.5	$\{0: 6855270, 0.5: 3099098, 1: 45600, 1.5: 32\}$	
2006/06/26	2.5	$\{0: 6842193, 0.5: 3113002, 1: 44788, 1.5: 17\}$	
2006/07/03	2.5	$\{0: 6842883, 0.5: 3107914, 1: 49168, 1.5: 35\}$	
2006/07/10	2.5	$\{0: 6460342, 0.5: 3488446, 1: 51168, 1.5: 44\}$	
2006/07/17	2.5	$\{0: 6295688, 0.5: 3628291, 1: 75996, 1.5: 25\}$	
2006/07/24	2.5	$\{0: 6273429, 0.5: 3643831, 1: 82708, 1.5: 32\}$	
2006/07/31	2.5	$\{0: 6405024, 0.5: 3531128, 1: 63813, 1.5: 34, 2: 1\}$	
2006/08/07	2.5	$\{0: 6938393, 0.5: 3013454, 1: 48143, 1.5: 10\}$	
2006/08/14	2.5	$\{0: 6932566, 0.5: 3016186, 1: 51238, 1.5: 10\}$	
2006/08/21	2.5	$\{0: 6444189, 0.5: 3521115, 1: 34692, 1.5: 4\}$	
2006/08/28	2.5	$\{0: 6419793, 0.5: 3537662, 1: 42540, 1.5: 5\}$	
2006/09/04	2.5	$\{0: 6925202, 0.5: 3025368, 1: 49417, 1.5: 13\}$	
2006/09/11	2.5	$\{0: 6949629, 0.5: 3001455, 1: 48903, 1.5: 13\}$	
2006/09/18	2.5	$\{0: 6955874, 0.5: 2996263, 1: 47854, 1.5: 9\}$	
2006/09/25	2.5	$\{0: 6944121, 0.5: 3008660, 1: 47204, 1.5: 15\}$	
2006/10/02	2.5	$\{0: 6972967, 0.5: 2981858, 1: 45166, 1.5: 9\}$	
2006/10/09	2	$\{0: 6962105, 0.5: 2992007, 1: 45879, 1.5: 9\}$	
2006/10/16	2.5	$\{0: 6949064, 0.5: 3004384, 1: 46536, 1.5: 16\}$	
2006/10/23	2.5	$\{0: 6945509, 0.5: 3008070, 1: 46411, 1.5: 10\}$	
2006/10/30	2.5	$\{0: 6987522, 0.5: 2970515, 1: 41950, 1.5: 13\}$	
2006/11/06	2.5	$\{0: 7026357, 0.5: 2938867, 1: 34765, 1.5: 11\}$	
2006/11/13	2.5	$\{0: 6872877, 0.5: 3083727, 1: 43381, 1.5: 15\}$	
2006/11/20	2.5	$\{0: 7006072, 0.5: 2954138, 1: 39783, 1.5: 7\}$	
2006/11/27	2.5	$\{0: 7005084, 0.5: 2952690, 1: 42217, 1.5: 9\}$	
2006/12/04	2.5	$\{0: 6843850, 0.5: 3130423, 1: 25724, 1.5: 3\}$	
2006/12/11	2.5	$\{0: 6840067, 0.5: 3134004, 1: 25921, 1.5: 8\}$	
2006/12/18	2.5	$\{0: 6839307, 0.5: 3134617, 1: 26071, 1.5: 5\}$	
2006/12/25	2.5	$\{0: 6838367, 0.5: 3133685, 1: 27945, 1.5: 3\}$	
2007/01/01	2.5	$\{0: 6854612, 0.5: 3119413, 1: 25973, 1.5: 2\}$	
2007/01/08	2.5	$\{0: 6851704, 0.5: 3122760, 1: 25535, 1.5: 1\}$	
2007/01/15	2.5	$\{0: 6927249, 0.5: 3023527, 1: 49212, 1.5: 12\}$	
2007/01/22	2.5	$\{0: 6856404, 0.5: 3117985, 1: 25606, 1.5: 5\}$	
2007/01/29	2.5	$\{0: 6953842, 0.5: 3003236, 1: 42912, 1.5: 10\}$	
2007/02/05	2.5	$\{0: 6932677, 0.5: 3020205, 1: 47110, 1.5: 8\}$	
2007/02/12	2.5	$\{0: 6832169, 0.5: 3142025, 1: 25797, 1.5: 9\}$	
		·	Continued on next page

Table 3 – continued from previous page

15

AS map date	δ	Distribution sampling over 10^7 4-tuples	Certificate
2007/02/19	2.5	$\{0: 6374287, 0.5: 3553405, 1: 72279, 1.5: 29\}$	
2007/02/26	$\frac{-10}{25}$	$\{0: 6371092, 0.5: 3555698, 1: 73178, 1.5: 32\}$	
2007/03/05	2.5	$\{0: 6399288, 0.5: 3529179, 1: 71507, 1.5: 26\}$	
2007/03/12	$\frac{-10}{25}$	$\{0: 6404282, 0.5: 3523699, 1: 71982, 1.5: 37\}$	
2007/03/19	$\frac{2.0}{2.5}$	$\{0: 6390005, 0.5: 3537871, 1: 72092, 1.5: 32\}$	
2007/03/26	2.0 2.5	$\{0: 6506308, 0.5: 3428145, 1: 65516, 1.5: 31\}$	
2007/04/02	2.0 2.5	$\{0: 6520975, 0.5: 3415957, 1: 63037, 1.5: 31\}$	
2007/04/02	2.0 2.5	$\{0, 6535495, 0.5, 3403074, 1, 61408, 1.5, 23\}$	
2007/04/05 2007/04/16	2.5	$\{0, 6230874, 0.5, 3603857, 1, 75234, 1.5, 25\}$	
2007/04/10 2007/04/23	2.0 2.5	$\{0, 6442512, 0.5, 3093037, 1, 15234, 1.5, 55\}$	
2007/04/23	2.5	$\{0, 6647201, 0.5, 3494013, 1, 02043, 1.5, 26\}$	
2007/04/30 2007/05/07	2.0	$\{0, 0047691, 0.0, 0299926, 1, 02140, 1.0, 00\}$	
2007/05/07	2.0	$\{0, 640062, 0.5, 5525595, 1, 59097, 1.5, 20\}$	
2007/05/14	2.0	$\{0, 0400014, 0.0, 05000, 1, 00900, 1.0, 21\}$	
2007/05/21	2.0	$\{0: 0309730, 0.5: 3535210, 1: 55010, 1.5: 22\}$	
2007/05/28	2.0	$\{0: 0433451, 0.5: 3500950, 1: 59509, 1.5: 30\}$	
2007/06/04	2.5	$\{0: 6440919, 0.5: 3500753, 1: 58297, 1.5: 31\}$	
2007/06/11		$\{0: 6440151, 0.5: 3501987, 1: 57838, 1.5: 24\}$	
2007/06/18	2.5	$\{0: 6380476, 0.5: 3557678, 1: 61814, 1.5: 32\}$	
2007/06/25	2.5	$\{0: 6427726, 0.5: 3513950, 1: 58297, 1.5: 27\}$	
2007/07/02	2.5	$\{0: 6408167, 0.5: 3531634, 1: 60169, 1.5: 30\}$	
2007/07/09	2.5	$\{0: 6421170, 0.5: 3522066, 1: 56731, 1.5: 33\}$	
2007/07/16	2.5	$\{0: 6443330, 0.5: 3502183, 1: 54460, 1.5: 27\}$	
2007/07/23	2.5	$\{0: 6564480, 0.5: 3378525, 1: 56972, 1.5: 23\}$	
2007/07/30	2.5	$\{0: 6527236, 0.5: 3414461, 1: 58274, 1.5: 29\}$	
2007/08/06	2.5	$\{0: 6534131, 0.5: 3409880, 1: 55957, 1.5: 31, 2: 1\}$	
2007/08/13	2.5	$\{0: 6562985, 0.5: 3381653, 1: 55327, 1.5: 35\}$	
2007/08/20	2.5	$\{0: 6385632, 0.5: 3553272, 1: 61065, 1.5: 31\}$	
2007/08/27	2.5	$\{0: 6396236, 0.5: 3545887, 1: 57857, 1.5: 20\}$	
2007/09/03	2.5	$\{0: 6259583, 0.5: 3667659, 1: 72733, 1.5: 25\}$	
2007/09/10	2.5	$\{0: 6273355, 0.5: 3655273, 1: 71326, 1.5: 46\}$	
2007/09/17	2	$\{0: 6721884, 0.5: 3236887, 1: 41214, 1.5: 15\}$	
2007/09/24	2.5	$\{0: 6309020, 0.5: 3635421, 1: 55531, 1.5: 28\}$	
2007/10/01	2.5	$\{0: 6308654, 0.5: 3636417, 1: 54898, 1.5: 31\}$	
2007/10/08	2.5	$\{0: 6294556, 0.5: 3647818, 1: 57594, 1.5: 32\}$	
2007/10/15	2.5	$\{0: 6297088, 0.5: 3650644, 1: 52243, 1.5: 25\}$	
2007/10/22	2.5	$\{0: 6322781, 0.5: 3626490, 1: 50709, 1.5: 20\}$	
2007/10/29	2.5	$\{0: 6307057, 0.5: 3642778, 1: 50140, 1.5: 25\}$	
2007/11/05	2.5	$\{0: 6334022, 0.5: 3616051, 1: 49897, 1.5: 29, 2: 1\}$	
2007/11/12	2.5	$\{0: 6474339, 0.5: 3479180, 1: 46452, 1.5: 29\}$	
2007/11/19	2.5	$\{0: 6471634, 0.5: 3483304, 1: 45039, 1.5: 23\}$	
2007/11/26	2.5	$\{0: 6471207, 0.5: 3481323, 1: 47445, 1.5: 25\}$	
2007/12/03	2.5	$\{0: 6268234, 0.5: 3654378, 1: 77367, 1.5: 21\}$	
2007/12/10	2.5	$\{0: 7193646, 0.5: 2776470, 1: 29880, 1.5: 4\}$	
2007/12/17	2.5	$\{0: 6211603, 0.5: 3704146, 1: 84222, 1.5: 29\}$	
2007/12/24	2.5	$\{0: 6203681, 0.5: 3711907, 1: 84393, 1.5: 19\}$	
2007/12/31	2.5	$\{0: 6152455, 0.5: 3748686, 1: 98575, 1.5: 284\}$	
, ,	1		Continued on next page

Table 3 – continued from previous page

Continued on next page

Inria

AS map date	δ	Distribution sampling over 10^7 4-tuples	Certificate
2008/01/07	2.5	$\{0: 6181559, 0.5: 3721697, 1: 96460, 1.5: 283, 2: 1\}$	
2008/01/14	2.5	$\{0: 6181805, 0.5: 3724102, 1: 93805, 1.5: 288\}$	
2008/01/21	2.5	$\{0: 6091114, 0.5: 3823553, 1: 85302, 1.5: 31\}$	
2008/01/28	2.5	$\{0: 6148914, 0.5: 3773245, 1: 77813, 1.5: 28\}$	
2008/02/04	2.5	$\{0: 6110081, 0.5: 3804218, 1: 85664, 1.5: 37\}$	
2008/02/11	2.5	$\{0: 6153417, 0.5: 3769283, 1: 77263, 1.5: 37\}$	
2008/02/18	2.5	$\{0: 6494425, 0.5: 3469038, 1: 36533, 1.5: 4\}$	
2008/02/25	2.5	$\{0: 6508414, 0.5: 3453013, 1: 38561, 1.5: 12\}$	
2008/03/03	2.5	$\{0: 6074445, 0.5: 3835576, 1: 89948, 1.5: 31\}$	
2008/03/10	2.5	$\{0: 6116647, 0.5: 3796576, 1: 86751, 1.5: 26\}$	
2008/03/17	2.5	$\{0: 6469621, 0.5: 3494777, 1: 35598, 1.5: 4\}$	
2008/03/24	2.5	$\{0: 6680713, 0.5: 3279135, 1: 40128, 1.5: 24\}$	
2008/03/31	2.5	$\{0: 6760962, 0.5: 3200027, 1: 38986, 1.5: 25\}$	
2008/04/07	2.5	$\{0: 6146697, 0.5: 3772221, 1: 81061, 1.5: 21\}$	
2008/04/14	2	$\{0: 6678617, 0.5: 3280468, 1: 40884, 1.5: 31\}$	
2008/04/21	2.5	$\{0: 6754821, 0.5: 3205187, 1: 39968, 1.5: 24\}$	
2008/04/28	2.5	$\{0: 6747599, 0.5: 3221263, 1: 31123, 1.5: 15\}$	
2008/05/05	2.5	$\{0: 6705534, 0.5: 3253230, 1: 41224, 1.5: 12\}$	
2008/05/12	2	$\{0: 6736779, 0.5: 3224340, 1: 38866, 1.5: 15\}$	
2008/05/19	2.5	$\{0: 6744533, 0.5: 3223722, 1: 31733, 1.5: 12\}$	
2008/05/26	2.5	$\{0: 6676541, 0.5: 3283593, 1: 39854, 1.5: 12\}$	
2008/06/02	2.5	$\{0: 6263034, 0.5: 3629604, 1: 107150, 1.5: 212\}$	
2008/06/09	2.5	$\{0: 6119294, 0.5: 3801752, 1: 78928, 1.5: 26\}$	
2008/06/16	2.5	$\{0: 6140937, 0.5: 3784952, 1: 74087, 1.5: 24\}$	
2008/06/23	2	$\{0: 6117524, 0.5: 3804487, 1: 77964, 1.5: 25\}$	
2008/06/30	2	$\{0: 6108712, 0.5: 3809083, 1: 82182, 1.5: 23\}$	
2008/07/07	2	$\{0: 6121446, 0.5: 3796503, 1: 82028, 1.5: 23\}$	
2008/07/14	2	$\{0: 6125452, 0.5: 3792550, 1: 81974, 1.5: 24\}$	
2008/07/21	2.5	$\{0: 6104044, 0.5: 3803429, 1: 92504, 1.5: 23\}$	
2008/07/28	2.5	$\{0: 6172496, 0.5: 3746211, 1: 81267, 1.5: 26\}$	
2008/08/04	2	$\{0: 6714605, 0.5: 3241415, 1: 43963, 1.5: 17\}$	
2008/08/11	2.5	$\{0: 6227344, 0.5: 3689657, 1: 82979, 1.5: 20\}$	
2008/08/18	2	$\{0: 6107882, 0.5: 3802076, 1: 90023, 1.5: 19\}$	
2008/08/25	2.5	$\{0: 6102253, 0.5: 3806062, 1: 91670, 1.5: 15\}$	
2008/09/01	2.5	$\{0: 6869610, 0.5: 3090027, 1: 40351, 1.5: 12\}$	
2008/09/10	2	$\{0: 6858274, 0.5: 3100938, 1: 40786, 1.5: 2\}$	
2008/10/20	2.5	$\{0: 6016429, 0.5: 3894959, 1: 88587, 1.5: 25\}$	
2008/11/10	2.5	$\{0: 6018567, 0.5: 3901161, 1: 80259, 1.5: 13\}$	
2008/12/01	2.5	$\{0: 6117362, 0.5: 3804621, 1: 77255, 1.5: 762\}$	
2008/12/22	2	$\{0: \ 6125964, \ 0.5: \ 3798219, \ 1: \ 75002, \ 1.5: \ 815\}$	
2009/01/05	2		$[8643, 15416, 17641, 296\overline{63}]$
2010/01/20	2	$\{0: 6980969, 0.5: 3003824, 1: 15206, 1.5: 1\}$	
2011/01/16	2	$\{0: 6884018, 0.5: 3059911, 1: 56070, 1.5: 1\}$	
2012/06/01	2		

Table 3 – continued from previous page

17



RESEARCH CENTRE SOPHIA ANTIPOLIS – MÉDITERRANÉE

2004 route des Lucioles - BP 93 06902 Sophia Antipolis Cedex Publisher Inria Domaine de Voluceau - Rocquencourt BP 105 - 78153 Le Chesnay Cedex inria.fr

ISSN 0249-6399