# Hyperbolic Trees in Complex Networks 

Zalán Heszberger ${ }^{1,2}$ András Gulyás ${ }^{1,2}$ András Biró ${ }^{1}$ László Balázs ${ }^{1}$ József Bíró ${ }^{1}$<br>${ }^{1}$ Budapest University of Technology and Economics<br>${ }^{2}$ MTA-BME Information Systems Modelling Research Group Budapest, Hungary, heszi@tmit.bme.hu


#### Abstract

The two-dimensional hyperbolic space turned out to be an efficient geometry for generative models of complex networks. The networks generated with this hyperbolic metric space share their basic structural properties (like small diameter or scale-free degree distribution) with several real networks. In this paper, we present a new model for generating trees in the two-dimensional hyperbolic plane. The generative model is not based on known hyperbolic network models: the trees are not inferred from the existing links of any network; instead, the hyperbolic tree is generated from scratch purely based on the hyperbolic coordinates of nodes. We show that these hyperbolic trees have scale-free degree distributions and are present to a large extent both in synthetic hyperbolic complex networks and real ones (Internet autonomous system topology, US flight network) embedded in the hyperbolic plane.


Index Terms-hyperbolic trees, complex networks, scale-free distribution

## I. Introduction

Trees play an essential role in network operation and management. Carefully chosen ones can be considered skeletons of more complex, real networks, acting as a scaffold for certain vital functions like routing [9], navigation [8], cluster analysis, or broadcasting [11]. For example, the so-called minimum (weight) spanning tree can efficiently be used for designing routing algorithms and protocols in computer and communication networks. The spanning tree of a network contains all the nodes with only a subset of the edges. That is, it is algorithmically generated from the whole original network. In this paper, we follow a completely different approach to generating hyperbolic trees. Only hyperbolic plane coordinates of network nodes are used, and a simple rule is applied to establish tree edges. The rule significantly differs from the ones used for hyperbolic complex network generation [7], and no other structural properties of the network are considered in the hyperbolic tree generation. First, it is shown, both analytically and numerically, that the degree distribution of the hyperbolic trees generated are scale-free, that is, they follow a power-law function. Second, it is also demonstrated that the hyperbolic trees are highly present in synthetic hyperbolic networks as well as in real networks embedded in the hyperbolic plane.

[^0]
## II. The Hyperbolic Plane

The hyperbolic plane (the two-dimensional hyperbolic space) is a metric space, hence there is a well-defined distance calculation such as (hyperbolic cosine theorem)

$$
\begin{equation*}
\cosh d(u, v)=\cosh r_{u} \cosh r_{v}-\sinh r_{u} \sinh r_{v} \cos \phi \tag{1}
\end{equation*}
$$

where $r_{u}$ and $r_{v}$ are the radial components of the polar coordinates of points $u$ and $v$, and $\phi=\phi_{u}-\phi_{v}$ is the difference of the angular components of the polar coordinates [2], [5]. Note that it significantly differs from the Euclidean cosine theorem. The fundamental nature of hyperbolic space is its constant negative curvature, which is not specified here because we do not directly need it. From now on, we assume unit negative curvature, which corresponds to the distance calculation formula above. The direct consequence of negative curvature is the exponential behavior; for example, the circumference and area of circles are exponential functions of the radii instead of polynomials like in the Euclidean space. The area of a hyperbolic disk with radius $R$ is

$$
\begin{equation*}
A_{R}=2 \pi(\cosh (R)-1) \tag{2}
\end{equation*}
$$

A further exciting and counter-intuitive property of the hyperbolic plane is that the area of triangles is bounded above by $\pi$ and equals to (in case of unit negative curvature)

$$
\begin{equation*}
A_{\text {triangle }}=\pi-\alpha-\beta-\gamma \tag{3}
\end{equation*}
$$

where $\alpha, \beta, \gamma$ are the angles of the triangle. The immediate consequence is that the triangle areas can usually be neglected, which can significantly simplify the calculations. For example, the area of a circle sector with $R=5$ and $\phi=\pi / 2$ (a quarter circle) is $\frac{\pi}{2}(\cosh 5-1) \approx 115$ while the area of the rightangled triangle with sides 5 a is less than $\pi / 2 \approx 1.57$. This means that the area of a circle sector is mainly concentrated to the circle segment. The above equations represent the three main properties, the careful use of which provides the basis of our calculations and derivations.

## III. Hyperbolic Trees

For generating hyperbolic random trees, let the model be the following. Let $N$ points having uniform distribution random coordinates be generated on a hyperbolic disk of radius $R$. The uniform distribution means in this context that the node density (i.e., the number of nodes per unit area) on the $R$-disk is constant. From this and (1), it follows that the probability
density of the angle coordinates of the points is uniform between 0 and $2 \pi$, while the radial coordinate density is given by

$$
\begin{equation*}
\rho(r)=\frac{\sinh r}{\cosh R-1} \tag{4}
\end{equation*}
$$

In a generated point set, the coordinates of the points are randomly sampled from the distributions above. The construction of the hyperbolic tree is given by the following rules. Each point in the set generated corresponds to a node in the tree. First, the points are ordered by increasing radial coordinates. The first node (the point with the smallest radial coordinate) is the root of the tree. Then the $i^{\text {th }}$ node $(i=2, \ldots, N)$ is connected to the closest one among $(1, \ldots, i-1)$, that is to the closest one having smaller radial coordinate. An example can be seen in Fig. 1 in which $N=200$ nodes are placed uniformly on a disk of radius $R=9$. Here, we should note that for illustration purposes, the native representation of hyperbolic space is used; that is, in the drawings, we use hyperbolic coordinates as if they were Euclidean. Although this may cause some strange phenomena in the figures, the textual descriptions of the algorithms and examples become more comfortable to interpret. As the calculations and derivations do not rely on the properties of any specific representations, there is no reason to use more sophisticated models like the Poincarè disk or the Klein model in our investigations. In


Fig. 1. Points generated on a radius $R=9$ hyperbolic disk according to uniform distribution (left side) and the hyperbolic tree inferred from this point (right side). One can observe that several links in the tree are oriented towards radial directions due to the explanation of hyperbolic distance characteristics.

Fig. 2 a more complex example can be seen with $R=14$, $N=1000$. One can immediately notice that the nodes seem to be non-uniformly distributed on the disk. This is due to the non-isometric nature of the native representation. Note that all other representation models in the Euclidean plan lack isometry, the hyperbolic plane simply "too big", e.g., contains exponentially larger amount of space than the Euclidean one. In the case of the disk with a radius of 14 , the total area is $2 \pi(\cosh 14-1) \approx 3.78 \times 10^{6}$. The half of this area is $1.8910^{6}$, which can be covered by a disk with radius 13.3 ! Due to the uniform node distribution, half of the nodes are expected within the disk of 13.3. In the particular realization in Fig. 2, exactly 500 nodes are inside the 13.14 radius disk (shown by red dashed line in the figure) while
the other 500 ones have radial coordinate values within 13.14 and 14 . This clearly illustrates the exponential behavior of a hyperbolic plane: areas and node densities are not what they seem in the figures. Another strange phenomenon in Fig. 2


Fig. 2. A hyperbolic tree with 1000 nodes. One can observe the strong hierarchy and the wide spectrum of node degrees. There are exactly half of the nodes (500) inside the red dashed circle. Areas and node densities are not what they seem. Distances also significantly differ from Euclidean one, e.g. node 936 and 999 closer to node 1 than to each other (green dots), see explanation in the text.
and Fig. 1 is that the orientation of most of the connections lies close to the radial direction. It seems that lots of node pairs (especially in the outer rim of the disk) should have been connected because they look to lie close enough to each other; further evidence that the hyperbolic distance calculation significantly differs from the Euclidean one. We demonstrate that with simple examples. In Fig 2, the green connections highlight the first node (with the smallest radial coordinate) connected to nodes 936 and 999. The polar coordinates are as follows: $\left(\phi_{1}, r_{1}\right)=(1.07836,1.88073),\left(\phi_{936}, r_{936}\right)=$ $(0.93668,13.8972),\left(\phi_{999}, r_{999}\right)=(0.829746,13.9966)$. According to the hyperbolic cosine law (1) the distances are $d(1,936)=12.2075, d(1,999)=12.6142 d(936,999)=$ 22.0355 . Thus nodes 936 and 999 turn out to be much farther from each other than it seems in the native representation, becoming the first node the closest neighbor to both. Hence they should not be connected. In fact, if the radial coordinates of two nodes are close to each other, the hyperbolic distance is rapidly increasing with increasing angular difference, especially in the range of small values. Fig. 3 illustrates this behaviour, in which the angular coordinate $\phi_{999}$ of node 999 is varying from $\phi_{936}$ to $\phi_{936}-\pi$.

One can observe that node 999 would be closer to node 936 only in an extremely small range of $0<\Delta \phi<0.00076$; for all larger $\Delta \phi d(936,999)>d(1,999)$ holds.


Fig. 3. Hyperbolic distances between nodes 936,999 and nodes 1,999 are presented. In case of similar radial coordinates $\left(r_{936}, r_{999}\right)$ the distance is quickly increasing with angular difference, while in case of large radial coordinate difference ( $r_{1}, r_{999}$ ) the distance function is quite flat.

## IV. Degree Distribution of Hyperbolic Trees

In this section the degree distribution of the nodes in the hyperbolic tree is investigated. First approximate analytical derivations are performed, than numerical experiments are presented. The derivation of the approximate degree distribution formula consists of three main steps. In the first step, the connection probability $p(u, v)$ for a node pair $u, v\left(r_{u}<r_{v}\right)$ is derived. In the second step, the conditional expected degree $\bar{k}\left(r_{u}\right)^{1}$ is computed based on $p(u, v)$ and the probability density of node $v$. In the third step, the degree distribution is inferred from $\bar{k}\left(r_{u}\right)$ with appropriate deconditioning.

Assume that $N$ points are randomly placed on a disk of radius $R$ evenly distributed over its area and a tree is generated applying the previously described algorithm. We can establish the probability $p(u, v)$ that an arbitrary node pair $(u, v)$ is connected. More specifically, $p(u, v)$ is the probability that $u$ and $v$ are connected under the condition that $r_{u}<r_{v}{ }^{2}$ In Fig. 4 a node pair $u, v$ are shown with $r_{u}<r_{v}$. According

[^1]

Fig. 4. Geometric illustration for the calculation of the connection probability $p(u, v)$ : Node $v$ is connected to node $u\left(r_{u}<r_{v}\right)$ if no nodes are contained in the intersection of the $O$-centered disk (colored in red) with radius $r_{v}$ and the $v$-centered disk (green line) with radius $d(u, v)$.
to the tree generation rule, $v$ is connected to the closest node which has a smaller radial coordinate than $r_{v}$. In a geometric interpretation, $v$ connects to the closest node inside the $O$-centered disk of radius $r_{v}$ (red disk in the figure). The closest node will be exactly $u$ when no points lie inside the intersection of the $O$-centered disk with radius $r_{v}$ and the $v$ centered disk with radius $d(u, v)$ (green disk in the figure). The probability that all the $N-2$ points fall outside the intersection area is $1-\frac{A_{Q P}}{A_{R}}$, where $A_{Q P}$ denotes the area of the intersection, and $A_{R}^{A_{R}}=2 \pi(\cosh R-1)$ is the area of the whole disk. It follows that
$p(u, v)=\left(1-\frac{A_{Q P}}{A_{R}}\right)^{N-2}=\left(1-\frac{A_{Q P}}{A_{R}}\right)^{A_{R} \frac{N-2}{A_{R}}} \approx \mathrm{e}^{-\delta A_{Q P}}$,
where $\delta$ is the node density: $\delta=\frac{N}{A_{R}} \approx \frac{N-2}{A_{R}}$ What remains to be done to calculate the probability $p(u, v)$ is to establish $A_{Q P}$. One can observe in Fig. 4 that $A_{Q P}$ is equal to the sum of the two circle sectors $P O Q$ (red, sector angle is $2 \alpha$ ) and $P v Q$ (green, sector angle is $2 \beta$ ) minus the area of the two triangles $P Q O$ and $P Q v$ :

$$
\begin{equation*}
A_{Q P} \approx 2 \alpha\left(\cosh r_{v}-1\right)+2 \beta(\cosh d(u, v)-1) \tag{6}
\end{equation*}
$$

where the area of the two triangles $2(\pi-\alpha-2 \beta)$ is already neglected. The angles $\alpha$ and $\beta$ can be determined by applying the hyperbolic cosine law (1) to the triangle $v O Q$ with edges $O Q$ and $v Q$. From these

$$
\begin{equation*}
\alpha\left(d_{u v}, r_{v}\right)=\arccos \frac{\cosh ^{2} r_{v}-\cosh d_{u v}}{\sinh ^{2} r_{v}} \tag{7}
\end{equation*}
$$

and

$$
\begin{equation*}
\beta\left(d_{u v}, r_{v}\right)=\arccos \frac{\cosh r_{v}\left(\cosh d_{u v}-1\right)}{\sinh r_{v} \sinh d_{u v}} \tag{8}
\end{equation*}
$$

Now the equations (6),(7) and (8) can be combined to get $A_{Q P}$ as a function of $d(u, v)$ and $r_{v}$. Because the formula is a bit lengthy and not very expressive, it is not repeated here. However, based on a sequence of approximations a much simpler and useful expression can be obtained as

$$
\begin{equation*}
A_{Q P} \approx 4 \mathrm{e}^{\frac{d(u, v)}{2}} \tag{9}
\end{equation*}
$$

by which we get a simple formulae for the approximation of the connection probability:

$$
\begin{equation*}
p(u, v) \approx \mathrm{e}^{-\delta 4 \mathrm{e}^{\frac{\mathrm{d}(\mathrm{u}, \mathrm{v})}{2}}} \tag{10}
\end{equation*}
$$

At this point, it is practical to postpone the in-depth analysis of the details and quality of these approximations, as two more weighing functions are to appear in the computation of $\bar{k}\left(r_{u}\right)$. Instead, an illustrative example is shown in Fig. 5 to study the performance of the different approximations of $A_{Q P}$, where $r_{v}=7$ is fixed and $d(u, v)$ is running from 1 to 14 (note that $d(u, v)<2 r_{v}$ holds for all $\left.(u, v), r_{u}<r_{v}\right)$. One can


Fig. 5. Different approximations of $A_{Q P}$ and their relative difference.
observe that for a wide range of $d(u, v)$ values, the simpler approximation (9) is quite close to the more tedious one; the relative difference is below $5 \%$ when $3<d(u, v)<13$. Because the distance $d(u, v)$ is also a function of $r_{u}$ and $r_{v}$ itself, it is worth expanding it exploiting (1) as

$$
\begin{equation*}
\cosh d(u, v)=\cosh r_{u} \cosh r_{v}-\sinh r_{u} \sinh r_{v} \cos \phi \tag{11}
\end{equation*}
$$

where $\phi$ is the difference between the angle coordinates of $u$ and $v$. Using the approximations $\cosh x \approx \frac{\mathrm{e}^{x}}{2}$ and $\sinh x \approx \frac{\mathrm{e}^{x}}{2}$ when $x$ is not too small, one can write

$$
\begin{equation*}
\frac{\mathrm{e}^{d(u, v)}}{2}=\frac{\mathrm{e}^{r_{u}+r_{v}}}{4}(1-\cos \phi) . \tag{12}
\end{equation*}
$$

Based on this equation we get

$$
\begin{equation*}
\mathrm{e}^{\frac{d(u, v)}{2}} \approx \mathrm{e}^{\frac{r_{u}+r_{v}}{2}} \sqrt{\frac{1-\cos \phi}{2}} \tag{13}
\end{equation*}
$$

Using the remarkable identity $\sqrt{\frac{1-\cos \phi}{2}}=\sin \frac{\phi}{2}$ a further, more useful formula can be obtained for $p(u, v)$ for subsequent derivation as

$$
\begin{equation*}
p(u, v) \approx \mathrm{e}^{-4 \delta \mathrm{e}^{\frac{r_{u}+r_{v}}{2}} \sin \frac{\phi}{2}} \tag{14}
\end{equation*}
$$

Now we can continue with the derivation of $\bar{k}\left(r_{u}\right)$. Roughly speaking, $\bar{k}\left(r_{u}\right)$ is the expected number of neighbors of $u$ when all nodes $v,\left(r_{u}<r_{v}<R\right)$ are counted for which $u$ is the closest. There is surely one more "downward" connection of $u$ to a node with an even lower radial coordinate due to the tree generation rule (except when $u$ has the smallest $r_{u}$, but this exception does not influence the results). More formally, $p(u, v)$ is to be integrated with the probability densities of $r_{v}$ and $\phi(\phi):=\phi_{v}-\phi_{u}$ (and a constant 1 should be added), that is

$$
\begin{equation*}
\bar{k}\left(r_{u}\right)=1+N \int_{r_{v}=r_{u}}^{R} \int_{\phi=0}^{2 \pi} p(u, v) \frac{1}{2 \pi} \frac{\sinh r_{v}}{\cosh R-1} \mathrm{~d} \phi \mathrm{~d} r_{v} \tag{15}
\end{equation*}
$$

First, consider the integral with respect to the angle difference $\phi$ :

$$
\begin{equation*}
\int_{\phi=0}^{2 \pi} p(u, v) \mathrm{d} \phi \tag{16}
\end{equation*}
$$

Interestingly enough, if we use approximation (14), it can be expressed exactly with two distinguished functions: the zero order modified Bessel function $I_{0}(x)$ and the zero order modified Struve function $L_{0}(x)$

$$
\begin{equation*}
\int_{\phi=0}^{2 \pi} \mathrm{e}^{-x \sin \frac{\phi}{2}} \mathrm{~d} \phi=2 \pi\left(\boldsymbol{I}_{0}(x)-\boldsymbol{L}_{0}(x)\right) \tag{17}
\end{equation*}
$$

where $x=4 \delta \mathrm{e}^{\frac{r_{u}+r_{v}}{2}}$. It is known that for large $x, \boldsymbol{I}_{0}(x)-$ $\boldsymbol{L}_{0}(x)$ quickly tends ${ }^{3}$. to $\frac{2}{\pi x}$, therefore the right hand side of the equation above is approximated by

$$
\begin{equation*}
\frac{4}{x}=\frac{1}{\delta} \mathrm{e}^{-\frac{r_{u}+r_{v}}{2}} \tag{18}
\end{equation*}
$$

Putting it back to equation (15), we get

$$
\begin{equation*}
1+N \int_{r_{v}=0}^{R} \frac{1}{\delta} \mathrm{e}^{-\frac{r_{u}+r_{v}}{2}} \frac{1}{2 \pi} \frac{\sinh r_{v}}{\cosh R-1} \mathrm{~d} r_{v} \tag{19}
\end{equation*}
$$

[^2]Using the definition of node density $\delta=\frac{N}{2 \pi(\cosh R-1)}$, we get

$$
\begin{equation*}
1+\int_{r_{v}=0}^{R} \mathrm{e}^{-\frac{r_{u}+r_{v}}{2}} \sinh r_{v} \mathrm{~d} r_{v} \approx 1-1+\mathrm{e}^{\frac{R-r_{u}}{2}} \tag{20}
\end{equation*}
$$

Finally, we get the stunningly simple formula

$$
\begin{equation*}
\bar{k}\left(r_{u}\right) \approx \mathrm{e}^{\frac{R-r_{u}}{2}} \tag{21}
\end{equation*}
$$

The final formula reflects the property observed in the previous numerical examples, namely that nodes with smaller radial coordinates tend to have much higher number of connections. There is an interesting and easy way to perform a sanity check on the formula. All trees with $N$ nodes have exactly $N-1$ links, that is the grand average node degree is $\bar{k}=2(N-$ 1) $/ N \approx 2$. This means that the result of integrating $\bar{k}\left(r_{u}\right)$ with respect to the density $r_{u}$ should be very close to 2 :

$$
\begin{equation*}
\bar{k}=\int_{r_{u}=0}^{R} \mathrm{e}^{\frac{R-r_{u}}{2}} \mathrm{e}^{r_{u}-R} \mathrm{~d} r_{u}=2\left(1-\mathrm{e}^{-\frac{R}{2}}\right) \tag{22}
\end{equation*}
$$

which differs from 2 with a negligible factor when $R$ is not too small (for example, in case of $R=10 \bar{k}=1.98652$ ).

As a third step, the complement cumulant degree distribution is derived, i.e. the probability that an arbitrary node degree is larger than a given value $k, P($ degree $>k)$ is computed. Here we present an approximate and more perceptible reasoning instead of a sophisticated but tedious one. If we accept the approximation (21), than from its monotonicity if follows, that nodes with higher expected degrees than a given $k$ are exactly inside the circle of radius $r_{u}(k)$, where $r_{u}(k)$ is the inverse function of (21). One can also state that nodes with degrees exactly higher than a given $k$ are expectedly inside the circle with radius $r_{u}(k)$. These two statements are approximately equivalent, therefore $P($ degree $>k)$ can be approximated by the ratio of the area of $r_{u}(k)$-disk and the $R$-disk

$$
\begin{equation*}
P(\text { degree }>k) \approx \frac{2 \pi\left(\cosh r_{u}(k)-1\right)}{2 \pi(\cosh R-1)} \approx \frac{\mathrm{e}^{r_{u}(k)}}{\mathrm{e}^{R}} \tag{23}
\end{equation*}
$$

Substituting the inverse function $r_{u}(k)=R-2 \log k$ into this equation

$$
\begin{equation*}
P(\text { degree }>k) \approx k^{-2} \tag{24}
\end{equation*}
$$

is obtained, that is approximately a power law distribution with parameter 2.

## V. Numerical Analysis

In this section numerical results are presented in connection with synthetic and real networks embedded on the hyperbolic plane. First, the generative model for synthesizing complex networks on the hyperbolic plane is considered [6] . The network generation rule in the simplest case is as follows: distribute $N$ nodes uniformly on a hyperbolic disk with radius $R$ then connect every node pair which are closer than $R$. It is analytically shown and numerically confirmed that the degree distribution of such networks follow power law and the ccdf is proportional to $k^{-2}$ [6], [10].

Two numerical examples are presented for $N=1000$, $R=13$ and for $N=3000, R=14$. In Fig. 6 the hyperbolic
tree generated and the synthetic network can be seen. It is observable that in spite of the completely different generation rule the structure of the synthetic network is strongly hierarchical and resembles to the tree network. Moreover, 860 of the total 999 links can be found in the original network too. In the second larger example (not shown in the figure) 2857 of the total 2999 tree links are also present in the synthetic network.


Fig. 6. Hyperbolic tree and a synthetic complex network with $N=1000$, $R=13$ for the same set of nodes.

The structural similarity is confirmed by the histograms for the degree distributions which are shown on Fig. 7. Note that the theoretical function $k^{-2}$ well fits to the measured degree distribution of the tree and is in accordance with the similar decay in synthetic network for larger values of $k$.


Fig. 7. Degree distributions.
The hyperbolic tree networks are also generated and tested on two types of real networks embedded in the hyperbolic plane. Here we do not outline the embedding process, just note that the embedding process do not use algorithms based on any trees inferred from the original network. For more details of the embedding see [3], [6], [12]. Thus our hyperbolic tree generation method does not "know" directly anything about any trees contained in the original real networks.

The first network embedded is the US flight network. This network was downloaded from the Bureau of Transportation Statistics http://transtats.bts.gov/ on 5 November 2017. It consists of 283 nodes and 1973 edges. In the network the nodes are US airports. Two nodes are linked if they are connected by
a direct flight. The hyperbolic tree and the original network are drawn in Fig. 8. The hyperbolic tree has 282 links of which 246 are also in the original network. It corresponds to a $87.2 \%$ tree inclusion ratio. The degree distributions of the hyperbolic tree and the original flight network can be seen in Fig. 9. One can observe, that the decay of the histograms are different, however, both of them are power law functions which confirms our theoretical results. The flight network is embedded into a 11.3 radius disk, which is determined by the embedding algorithm, so the average node density on this disk is about 0.0035 . The other network is the much larger Internet


Fig. 8.


Fig. 9. US flight network and its tree degree distribution. The tree is generated based only on the hyperbolic coordinates.

AS level topology. The corresponding data set representing the global Internet structure at the autonomous system (AS) level is from [4]. In this network we use the inner core of 10000 nodes (out of the total 23748), the number of edges is 40605 and the network is embedded into an 26.87 radius disk. In this case the node density is much lower, $6.83 \times 10^{-9}$. The degree distribution can be seen in Fig. 10. Similar power law behaviour also appears here, the difference between the decay rates is smaller than in the previous case. The inclusion ratio of the hyperbolic tree in the original network is also as high as $84.7 \%$. In both cases, the difference between the exponents of the power functions can be attributed to the fact that real network embeddings usually cause non-uniform node distributions, and according to our numerical experiments this non-uniformity results in smaller change in the tree degree distribution than in that of the original network.


Fig. 10. Internet AS level topology (inner core of 10000 nodes) and its tree degree distribution.

## VI. Conclusion

A new hyperbolic tree generation method has been presented. The generating algorithm is based on the presumption that successful navigability of many real networks is key to their evolution. Both analytically and by numerical experiments on synthetic and real data, it is shown that the degree distribution of the trees generated is approximately a power function. It has also been demonstrated that the trees are present to a large extent in real and synthetic networks. Our results open further research on the deeper relation between navigable hyperbolic trees and the structural evolution of networks.

## REFERENCES

[1] M. Abramovitz and I. Stegun. Handbook of Mathematical Functions. Courier Dover Publication, 1965.
[2] James W Anderson. Hyperbolic geometry. Springer Science \& Business Media, 2006.
[3] M. Boguna, D. Krioukov, and K. C. Claffy. Navigability of complex networks. Nat Phys, 5(1):74-80, 2009.
[4] Marián Boguná, Fragkiskos Papadopoulos, and Dmitri Krioukov. Sustaining the internet with hyperbolic mapping. Nature communications, 1(1):1-8, 2010.
[5] J. Bolyai. APPENDIX - The Theory of Space (with Introduction, Comments and Addenda by F. Karteszi). North-Holland, 1987.
[6] Krioukov D. et al. Hyperbolic geometry of complex networks. Physical Review E, 82(3):036106, 2010.
[7] Luca Gugelmann, Konstantinos Panagiotou, and Ueli Peter. Random hyperbolic graphs: degree sequence and clustering. In International Colloquium on Automata, Languages, and Programming, pages 573585. Springer, 2012.
[8] András Gulyás, József J Bíró, Attila Kőrösi, Gábor Rétvári, and Dmitri Krioukov. Navigable networks as nash equilibria of navigation games. Nature communications, 6:7651, 2015.
[9] Attila Kőrösi, András Gulyás, Zalán Heszberger, József Bíró, and Gábor Rétvári. On the memory requirement of hop-by-hop routing: Tight bounds and optimal address spaces. IEEE/ACM Transactions on Networking, 2020.
[10] F. Papadopoulos, D. Krioukov, M. Bogua, and A. Vahdat. Greedy forwarding in dynamic scale-free networks embedded in hyperbolic metric spaces. In Proc. of IEEE Infocom, pages 1-9. IEEE, 2010.
[11] János Tapolcai, József Bíró, Péter Babarczi, András Gulyás, Zalán Heszberger, and Dirk Trossen. Optimal false-positive-free bloom filter design for scalable multicast forwarding. IEEE/ACM Transactions on Networking, 23(6):1832-1845, 2014.
[12] Ivan Voitalov, Rodrigo Aldecoa, Lan Wang, and Dmitri Krioukov. Geohyperbolic routing and addressing schemes. ACM SIGCOMM Computer Communication Review, 47(3):11-18, 2017.


[^0]:    Thanks to OTKA FK17 123957, FK 128233, KH18 129589, K17 124171 for funding. Heszberger Zalán is also supperted by MTA Bolyai János Research Grant and UNKP-20-5 Bolyai+ Research Grant. László Balázs and András Biró are supported by the UNKP-20-1 New National Excellence Program of the Ministry of Innovation and Technology from the source of the national research, development and innovation fund.

[^1]:    ${ }^{1}$ We introduce the short notation $\bar{k}\left(r_{u}\right)$ for $\mathrm{E}\left[k(u) \mid r_{u}\right]$, that is, the conditional expectation of the degree of a node $u$, whose radial coordinate is $r_{u}$
    ${ }^{2}$ The notation $p\left(u\right.$ connected to $\left.v \mid r_{u}, r_{v}, r_{u}<r_{v}\right)$ could be more appropriate, but for brevity we use $p(u, v)$.

[^2]:    ${ }^{3}$ The asymptotic series expansion of $\boldsymbol{I}_{0}(x)-\boldsymbol{L}_{0}(x)$ at $x=\infty$ starts with $\frac{2}{\pi x}$, see [1]. The asymptotic series expansion in this case means that $\lim _{x \rightarrow \infty} \frac{\boldsymbol{I}_{0}(x)-\boldsymbol{L}_{0}(x)}{\frac{2}{\pi x}}=1$

