

The Skeleton of Hyperbolic Graphs for Greedy Navigation

Zalán Heszberger József Bíró András Gulyás

Dept. of Telecommunications and Media Informatics, Budapest University of Technology and Economics

Budapest, Hungary, biro@tmit.bme.hu

Abstract—Random geometric (hyperbolic) graphs are important modeling tools in analyzing real-world complex networks. Greedy navigation (routing) is one of the most promising information forwarding mechanisms in complex networks. This paper is dealing with greedy navigability of complex graphs generated by using a metric (hyperbolic) space. Greedy navigability means that every source-destination pairs in the graph can communicate in such a way that every node passes the information towards that neighboring node which is "closest" to the destination in terms of node coordinates in the metric space. A set of compulsory links in greedy navigable graphs called Greedy Skeleton is identified. Because the two-dimensional hyperbolic plane (\mathbb{H}^2 , also known as the two dimensional Bolyai-Lobachevsky Space [2]) turned out to be extremely useful in modelling and generating real-like networks, we deal with the statistical properties of the Greedy Skeleton when the metric space is \mathbb{H}^2 . Some examples of numerical studies and simulation results supporting the analytical formulae are also performed. The significance of the results lies in that every (either artificial or natural) network formation process aiming at greedy navigability must contain this Greedy Skeleton. Furthermore, this could be an important step towards the formal argumentation of the very high greedy navigability of some models observed only experimentally for the time being, and also to analyze equilibrium of greedy network navigation games on \mathbb{H}^2 .

Index Terms—Random Hyperbolic Graphs, Greedy Routing, Scale-free Networks

I. INTRODUCTION

A graph embedded in a metric space (nodes of the graphs are labeled with coordinates of the metric space) is said to be greedy navigable (or equivalently said greedy routing is applicable) if greedy path exists between any source-destination pair. Greedy path is a set of consecutive links between a source and a destination, in which the next hop of a node along this path is the neighbor of this node which is closest to the destination in terms of distance measured in the metric space. Clearly, greedy path does not necessarily exist in a graph with an underlying metric space. Informally, if a piece of information (e.g. a packet) arrives at a node, this node should calculate first which of its neighboring nodes is closest to the destination (called also greedy next hop), and then pass the information to this next hop node. In [3] R. Kleinberg showed in his seminal paper that for each node of an arbitrary graph a virtual coordinate in the hyperbolic plane can be assigned, and greedy routing can be performed with

respect to these virtual coordinates. The assignment is based on his fundamental theorem that every connected finite graph has a greedy embedding in the hyperbolic plane. Nevertheless, this theorem does not provide that this embedding has the desirable properties of low congestion (the number of distinct pairs s, t whose greedy route uses a given node v) and low stretch (the ratio of the number of hops on a greedy route to the number of hops on the shortest route between the same pair of nodes)

In what follows we expose that a skeleton (a set of links in the graph) must exist in *every* 100% greedy navigable graph. Let a node u be supposed to reach a node v through a greedy path. Let also be the v -centered ball with $d(u, v)$ radius considered. If there is no other node connected to v within this ball, then the (u, v) directed link should exist, otherwise u can not reach v in a greedy manner. In other words, u will be the closest node to v . Eventually, for all nodes these directed links can be identified, we call them together the Greedy Skeleton. Every node in such a skeleton has only one incoming link (every node has one closest neighbour). From this it immediately follows that the average degree of the Greedy Skeleton is 1. (This could be a good checking point when the degree distribution is derived). A node can have more than one outgoing links (a node can act as a closest node of more than one other nodes).

II. THE SKELETON OF GREEDY NAVIGABLE HYPERBOLIC RANDOM GRAPHS ON \mathbb{H}^2

A very simple and appealing model using \mathbb{H}^2 as metric space is introduced in [4] for generating scale-free graph, which is as follows: First, place N points (nodes) on a hyperbolic disk with radius R on \mathbb{H}^2 , with generating their polar coordinates (r, ϕ) according to the following density functions ¹ on the whole disk

$$\rho(r) = \frac{\alpha \sinh(r)}{\cosh(\alpha R) - 1}, \quad \rho(\phi) = \frac{1}{2\pi}, \quad (1)$$

where $\alpha \geq 1/2$ is a parameter controlling the heterogeneity of the layout. If $\alpha = 1$, the nodes are distributed uniformly over the hyperbolic disk because the area element at coordinates (r, ϕ) is $dA = \sinh(r) dr d\phi$. Next, for every node pair

Thanks OTKA FK17 123957, KH18 129589, K17 124171 for funding. Heszberger Zalán is also supported by MTA Bolyai János Research Grant and UNKP-19-4 Bolyai+ Research Grant.

¹Note, that the desired node scattering is achieved in simulations by placing nodes u at polar coordinates $r_u = (1/\alpha) \operatorname{arccosh} \{1 + [\cosh(\alpha R) - 1]U\}$ and $\phi_u = 2\pi U$ where U for each u is a random number drawn from the uniform distribution on $[0, 1]$.

u, v establish a link between them, if they are not farther than R from each other, i.e. $d(u, v) \leq R$, where $d(u, v)$ is the hyperbolic distance between u, v on \mathbb{H}^2 . This link generation process can also be viewed that for every node u , a u -centered circle is drawn with radius R , then every point within this circle is connected to u .²

The distance calculation on \mathbb{H}^2 by polar coordinates can be performed by the hyperbolic law of cosine:

$$\cosh(d_{uv}) = \cosh(r_u) \cosh(r_v) - \sinh(r_u) \sinh(r_v) \cos(\phi) \quad (2)$$

where r_u, r_v are the radial coordinates of nodes u, v and ϕ is the difference of their angle coordinates.

Now, similarly to the previous discussion, one can identify the greedy skeleton of any greedy navigable graph generated on \mathbb{H}^2 in the following way: Recall that the link (u, v) is contained in GS if and only if the $d(u, v)$ -disk centered at v does not contain any node other than u (see Fig. 1). This means that u cannot reach v by greedy routing through any other nodes than v , and so a link must exist towards v to keep greedy navigability of the graph. Note again, that the in-degree of each node in GS will be exactly one.

In the next section, we analyze the statistical properties of GS .

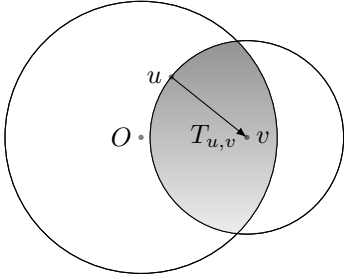


Fig. 1. An edge in the GS

III. STATISTICAL PROPERTIES OF GS

A. The expected out degree and degree distribution of GS for uniform node density ($\alpha = 1$)

Because the area T_{uv} of the intersection of the R -disk and the $d(u, v)$ -disk at origin v plays central role in the statistical properties of GS , first we give an approximation for this. An acceptable approximation for T_{uv} is as follows: T_{uv} is apparently equals to $2\pi(\cosh d_{uv} - 1)$ ($\approx \pi e^{\frac{d_{uv}}{2}}$ for not so small d_{uv}) when the d_{uv} -disk is completely inside the R -disk. On contrary, if $R - r_v < d(u, v)$ (there is real intersection) then much less evidently T_{uv} is approximately $T_{uv} \approx 4e^{\frac{d_{uv}}{2}} e^{\frac{R-r_v}{2}}$. In Fig. III-A two characteristic cases are depicted when there is real intersection of the d_{uv} -disk and the R -disk. Let the polar coordinates of node v be (r_v, ϕ_v) ,

²It can also be interpreted that the link establishment rule is a Heaviside step function on $d(u, v)$ as $H(R - d(u, v))$. In [4] a continuous relaxation of H is also presented, nevertheless, in this paper we follow the model using H .

and of node u be (r_u, ϕ_u) . Let $\phi = |\phi_u - \phi_v|$. The area T_{uv} is the function of r_u, r_v, ϕ , and R , and can be calculated as the sum of the two circle sectors with angle 2α , radius d_{uv} and angle 2β radius R , and minus the area of the two triangles with angles α, β, γ . That is

$$T_{uv} = 2\beta (\cosh(R) - 1) + 2\alpha (\cosh(d_{uv}) - 1) - 2(\pi - \alpha - \beta - \gamma) \quad (3)$$

where the angles and d_{uv} are given by the hyperbolic law of cosines, however, here the following simpler approximations are used (which are accurate enough for larger r_u and d_{uv}):

$$d_{uv} \approx R + r_v + 2 \ln \frac{\beta}{2} \Rightarrow \beta \approx 2e^{\frac{d_{uv}}{2} - \frac{R+r_v}{2}} \quad (4)$$

$$R \approx d_{uv} + r_v + 2 \ln \frac{\alpha}{2} \Rightarrow \alpha \approx 2e^{-\frac{d_{uv}}{2} + \frac{R+r_v}{2}} \quad (5)$$

Applying (3) with neglecting the triangle areas, and using $\cosh(R) - 1 \approx e^R/2$, $\cosh(d_{uv}) - 1 \approx \frac{e^{d_{uv}}}{2}$ we get $T_{uv} \approx 4e^{\frac{d_{uv}}{2}} e^{\frac{R-r_v}{2}}$. In summary:

$$T_{uv} \approx \begin{cases} \pi e^{d_{uv}} & , \text{ if } 0 \ll d_{uv} < R - r_v \\ 4e^{\frac{d_{uv}}{2}} e^{\frac{R-r_v}{2}} & , \text{ if } d_{uv} > R - r_v \end{cases} \quad (6)$$

The expected *out-degree* as the function of r_u can be expressed by the integral:

$$k_{out}(r_u) = (N - 1) \int_{R-disk} \left(\frac{T_R - T_{uv}}{T_R} \right)^{N-2} dT \approx \delta \int_0^R \int_0^{2\pi} e^{-T_{uv}\delta} d\phi \sinh(r_v) dr_v \quad (7)$$

where T_{uv} (see Fig. 1) is the area of the intersection of the R -disk and the $d(u, v)$ -disk at origin v and $\delta = N/T_R$ is the density of the points (T_R denotes the area of the R -disk).

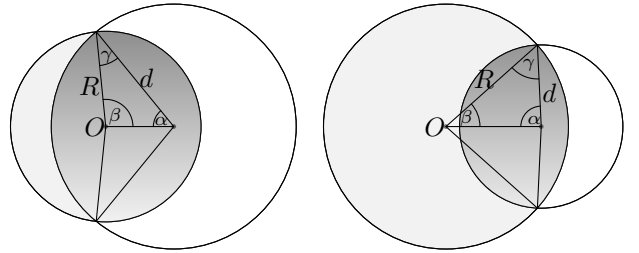


Fig. 2. Illustration for T_{uv}

These approximations are illustrated in Fig. 3 for $R = 12$, $r_v = 6$ and $r_v = 8$. Solid lines are the exact T_{uv} calculations based on (3) and exact computations of angles. Note that there is a sharp change on logarithmic scale between the d_{uv} -slope and $d_{uv}/2$ -slope around $R - r_v$. The dashed lines are the T_{uv} approximations when $d_{uv} > R - r_v$.

The joint asymptotic expansion of the double integral with respect to r_v and ϕ reveals that the dominant terms will be those in which $d_{uv} > R - r_v$. Hence, considering the first integral by ϕ as

$$\delta \int_0^{2\pi} e^{-\delta T_{uv}} d\phi \quad (8)$$

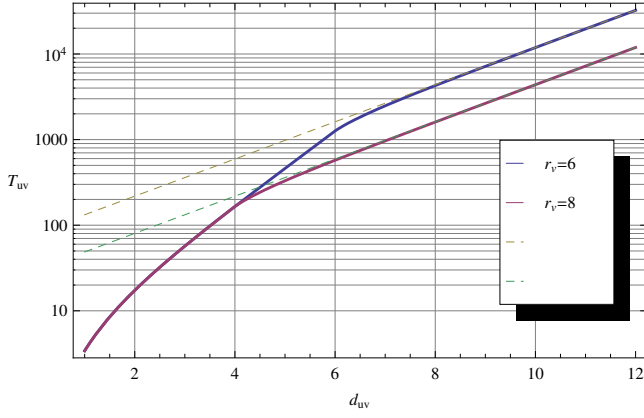


Fig. 3. $T_{uv} \approx 4e^{\frac{d_{uv}}{2}} e^{\frac{R-r_v}{2}}$ when there are real intersections (that is when $d(u, v) > 6$ and 4 , respectively).

can be approximated when T_{uv} is substituted by its approximation, that is by

$$\delta \int_0^{2\pi} e^{-\delta 4e^{\frac{d_{uv}}{2}} e^{\frac{R-r_v}{2}}} d\phi. \quad (9)$$

Applying the approximation $e^{\frac{d_{uv}}{2}} \approx e^{\frac{r_u+r_v}{2}} \sqrt{\frac{1-\cos\phi}{2}}$ the integral can be approximated as

$$\delta \int_0^{2\pi} e^{-\delta 4e^{\frac{d_{uv}}{2}} e^{\frac{R-r_v}{2}}} d\phi \approx 2\pi\delta(I(0, x) - S(0, x)) \approx e^{-\frac{R+r_u}{2}} \quad (10)$$

where $x = 4\delta e^{\frac{R+r_u}{2}}$ and the last wave due to that $I(0, x) - S(0, x)$ (difference of the BesselI and the modified Struve functions) quickly tends to $\frac{2}{\pi}x^{-1}$. Now the second integration by r_v gives the expected degree approximation in the frame topology, that is

$$k_{out,GS}(r_u) \approx \int_0^R e^{-\frac{R+r_u}{2}} \sinh(r_v) dr_v = e^{-\frac{R+r_u}{2}} (\cosh(R) - 1) \approx \frac{1}{2} e^{\frac{R}{2}} e^{-\frac{r_u}{2}}. \quad (11)$$

The quality of this approximation can be checked with the followings. The average in-degree in the frame topology is exactly 1, therefore the average out-degree should also be 1, that is

$$\bar{k}_{out,GS} = \int_0^R k_{out,GS}(r_u) \frac{\sinh(r_u)}{\cosh(R) - 1} dr_u \approx \frac{1}{3} (3 + e^{-2R} - 4e^{-R/2}) \approx 1. \quad (12)$$

Let us recall that in case of uniform distribution of points on an R -disk of the hyperbolic plane, the density of the radial coordinates of the points is

$$\rho(r) = \frac{\sinh r}{\cosh R - 1} \quad (13)$$

Note that the expected degree of node u is exponential in the radial coordinate r_u as in [4]. Because of this and the fact

that equilibrium graph of NNG is also sparse [1] the degree distribution can be calculated in the same way as in [4] :

$$P(k) = \int_0^R g(k, k_{out}(r_u)) \rho(r_u) dr_u = \frac{1}{2} \frac{\Gamma(k-2, \frac{1}{2})}{k!} \quad (14)$$

where $g(k, k_{out}(r_u))$ is the conditional distribution of the degree of a node with radial coordinate u , and it is Poissonian with mean $k_{out}(r_u)$ in case of sparse graphs. It can also be shown that for larger k

$$P(k) \approx \frac{1}{2k^3}. \quad (15)$$

The direct derivation of the complement cumulant degree distribution from $P(k)$ seems to be intangible, however, from its approximation it can be computed as

$$\bar{F}(k) \approx 1 - \left(\int \frac{1}{2k^3} dk + C \right) \quad (16)$$

where the constant C is 1, and $k \geq \frac{1}{2}$ (in order to have distribution function), that is

$$\bar{F}(k) \approx \frac{\bar{k}^2}{4} k^{-2}, \quad k \geq \frac{1}{2}. \quad (17)$$

It is interesting to show that this approximation can also be obtained as the *exact* ccdf of the conditional expected node degrees $k_{out}(r_u)$. This approximation can be computed as

$$\bar{F}(k) \approx \int_{r=0}^{r_u(k)} \rho(r) dr \approx e^{r_u(k)-R} \quad (18)$$

where $r_u(k)$ is the inverse function of $k_{out}(r_u, R)$ w.r.t. r_u , i.e.

$$r_u(k) = R - 2 \ln(2k). \quad (19)$$

Applying this one can obtain the same before as

$$\bar{F}(k) \approx \frac{\bar{k}^2}{4} k^{-2}, \quad k \geq \frac{1}{2\bar{k}}. \quad (20)$$

Note that the average degree can also be computed as

$$\int_{k=\frac{1}{2}}^{\infty} \left(k \frac{\partial(1 - \bar{F}(k))}{\partial k} \right) = 1 \quad (21)$$

and it is indeed 1.

We have studied the quality of the approximations above also by extensive numerical investigations and eventually found that the exponential decay of the expected degree of nodes ($k_{out}(r_u)$) is extremely good approximation of the numerically evaluated expected degree function for a wide range of node density δ .

This is illustrated by the following examples. First consider the setup $R = 16.5$, $n = 10000$. In this case $\delta = 2.17 \cdot 10^{-4}$. Fig. 4 shows how the expected degree decay is matching the exponential decay. One can observe that for larger values of r_u the coincidence is very good, for smaller values of r_u there are larger errors of the approximation.

To quantify this observation, note that 99.9% of points have $r_u > 10$ (that is in case of uniformly distributed points on the $R(=16.5)$ -disk, only about 10 points of the 10000 is inside

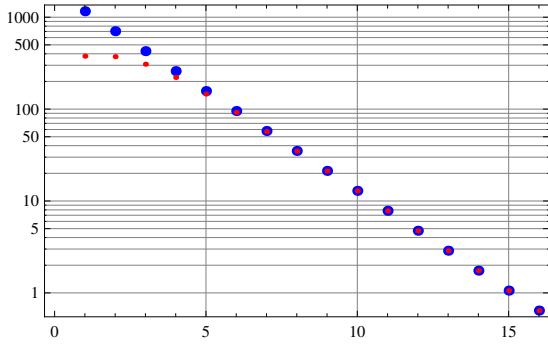


Fig. 4. Comparison of the exponential decay ($C(R)e^{-\frac{r_u}{2}}$, larger blue dots) and the numerically evaluated exact decay (smaller red dots) of the expected degree in the function of r_u , $R = 16.5$, $n = 10000$.

the disk with radius 10). If we consider the relative errors of the matching one can reveal that for $r_u > 10$ it is smaller than 0.15%, that is for 99.9% of points the expected degree approximation has smaller than 0.15% relative error, as shown by Fig. 5. To increase the number of points to $n = 30000$ and $n = 50000$ ($\delta = 6.54 \cdot 10^{-4}$, $\delta = 1.08 \cdot 10^{-3}$), the relative error is increasing, especially for smaller values of r_u , but still for 99.9% of the points the relative error smaller than 0.25% and 1%, respectively. If we dramatically decrease the node-density, for example $n = 500$, the relative errors also increase (compared to the $n = 10000$ case), however, it still remains under 0.2% for 99.9% of the points.

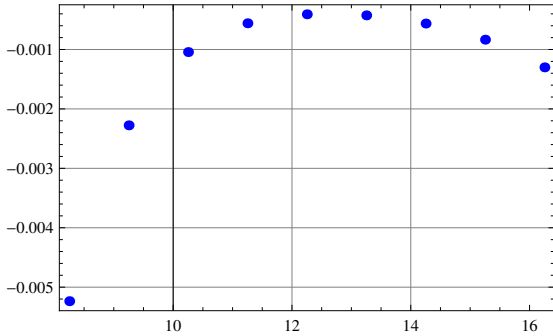


Fig. 5. Relative error of matching the exponential decay in the function of r_u , $R = 16.5$, $n = 10000$.

B. The expected out degree and degree distribution of GS for quasi-uniform node density

The radial coordinate density in case of quasi-uniform node density is

$$\rho(r, \alpha) := \frac{\alpha \sinh(\alpha r)}{\cosh(\alpha R) - 1} \approx \alpha e^{\alpha(r-R)} \quad (22)$$

while the angle density remains uniform ($\frac{1}{2\pi}$) over the range $[1, 2\pi]$. Given a point pair (u, v) , first we determine the probability $p(r_u, \alpha)$ that the $u \rightarrow v$ link exists, then based on this the average out degree $k(r_u, \alpha)$ of u is calculated, and finally $\bar{F}(k, \alpha)$ is also given.

The probability $p(r_u, \alpha)$ is equal to that none of the remaining $N - 2$ points fall in the intersection of the v -centered d_{uv} circle and the R -disk. Let us denote by p_1 the probability that a point whose coordinates generated by randomly according to the densities above falls inside the intersection. Using p_1 the probability $p(r_u, \alpha)$ can be calculated and approximated as

$$p(r_u, \alpha) = (1 - p_1)^{N-2} \approx e^{-Np_1} \quad (23)$$

The calculation of p_1 can be performed by using the node density function in the following way [4]

$$p_1 = \int_0^{\max(0, d-r_v)} \rho(r, \alpha) dr + \frac{1}{2\pi} \int_{|d-r_v|}^{\min(R, d+r_v)} \rho(r, \alpha) 2\theta(r) dr \quad (24)$$

where

$$\theta(r) = \arccos \frac{\cosh r_v \cosh r - \cosh d}{\sinh r_v \sinh r} \quad (25)$$

In [4] a useful approximation is presented for quite similar integrals, based on which one can write

$$p_1 \approx \frac{4e^{\frac{1}{2}(d-R-r_v)} \alpha}{\pi(-1+2\alpha)} \quad (26)$$

for $0.5 < \alpha \leq 1$.

Now the expected out-degree of u can be written as

$$k_{out}(r_u, \alpha) \approx \frac{N}{2\pi} \int_0^R \int_0^{2\pi} e^{-Np_1} d\phi \rho(r_v) dr_v \quad (27)$$

Using the approximation of p_1 and $\cosh(d/2) \approx e^{\frac{r_u+r_v}{2}} \sin \frac{\phi}{2}$ one can formulate

$$\int_0^{2\pi} e^{-Np_1} d\phi \approx \int_0^{2\pi} e^{-x \sin \frac{\phi}{2}} d\phi \approx 2\pi(I(0, x) - S(0, x)) \approx \frac{4}{x} \quad (28)$$

where

$$x = 4 \frac{N}{\pi} \frac{\alpha}{2\alpha - 1} e^{\frac{r_u - R}{2}} \quad (29)$$

Note, that x does not depend on r_v , therefore the second integration by r_v results

$$k_{out}(r_u, \alpha) \approx \frac{N}{2\pi} \frac{4}{x} \int_0^R \rho(r_v, \alpha) dr_v = \frac{2\alpha - 1}{2\alpha} e^{\frac{R}{2}} e^{-\frac{r_u}{2}} \quad (30)$$

Note, that for $\alpha = 1$ we get back the result for the uniform density case.

Now the (approximation of the) complement cumulative distribution function $\bar{F}(k)$ can be derived as,

$$\bar{F}(k) = \int_0^{r_u(k)} \rho(r, \alpha) dr \approx e^{\alpha(r_u(k) - R)} = \left(\frac{1 - \frac{1}{2\alpha}}{k} \right)^{2\alpha} \quad (31)$$

where $r_u(k)$ is the inverse function of $k_{out}(r_u)$.

Here we note that further investigations are needed for deriving analytical formulae for the case $0 < \alpha \leq 0.5$.

C. Comparing analytical and simulation results

For comparing the analytical formula and simulation results of the degree distribution of GS, 20 random graphs have been generated each with 5000 nodes distributed uniformly ($\alpha = 1$) on a hyperbolic disk with radius $R = 15$. For each graph the pdf of node degrees is generated and "averaged" them over the 20 graphs. Fig 6 shows the results. One can observe good matching for lower degrees, however, in this simulation setup nodes having larger degree than 20 are very rare resulting less reliable statistics for this region.

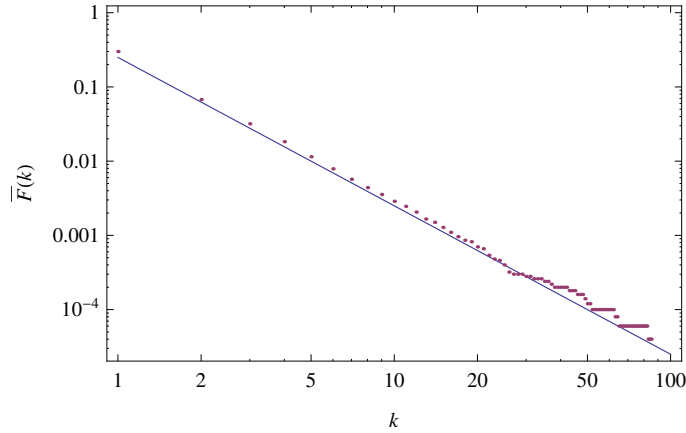


Fig. 6. CCDF of the node degrees, analytical formula (continuous line) and the simulation results on a log-log plot. $N = 5000$, $R = 15$.

IV. CONCLUSION

We have presented the Greedy Skeleton as a set of mandatory links in every 100% greedy navigable graphs. Statistical properties of GS of greedy navigable graphs generated by the metric space \mathbb{H}^2 have also been performed. As the main result of this paper, it has been shown that the expected degree of a node in the Greedy Skeleton is proportional to $\exp(-\frac{r}{2})$ (where r is the radial coordinate of the node in \mathbb{H}^2) and the degree distribution of the Greedy Skeleton is scale-free that is proportional to k^{-3} . Numerical examples and simulation results highlighted the accuracy of the analytical formulae.

REFERENCES

- [1] M. Boguna. Class of correlated random networks with hidden variables. *Physical Review E*, 68:1–13, 2003.
- [2] J. Bolyai. *APPENDIX - The Theory of Space (with Introduction, Comments and Addenda by F. Karteszi)*. North-Holland, 1987.
- [3] R. Kleinberg. Geographic Routing Using Hyperbolic Space. In *Proc. of IEEE Infocom 2007*, volume 3, pages 1902 – 1909. IEEE, 2007.
- [4] 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.

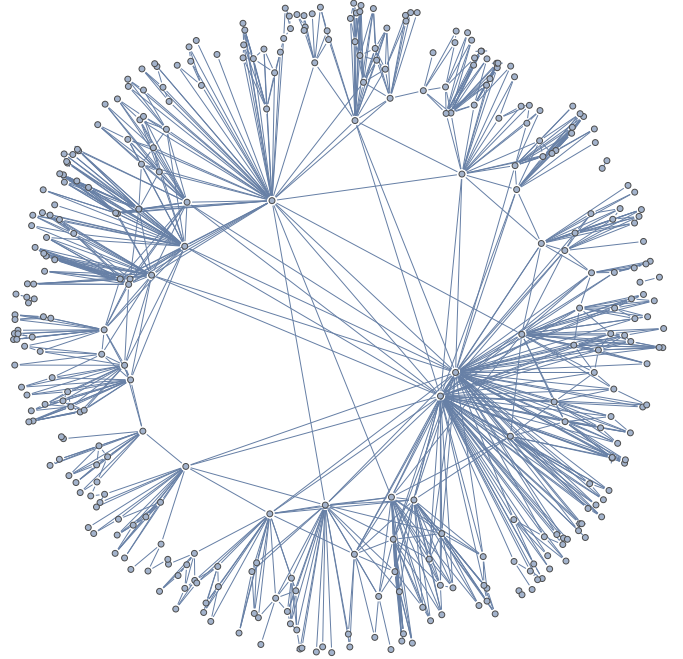


Fig. 7. The full graph $N = 350$, $R = 10$ to include the mandatory links for navigation. Note that for the ease of illustration, hyperbolic polar coordinates used as Euclidean ones, however, the link establishment is in full accordance of the hyperbolic distance calculation.

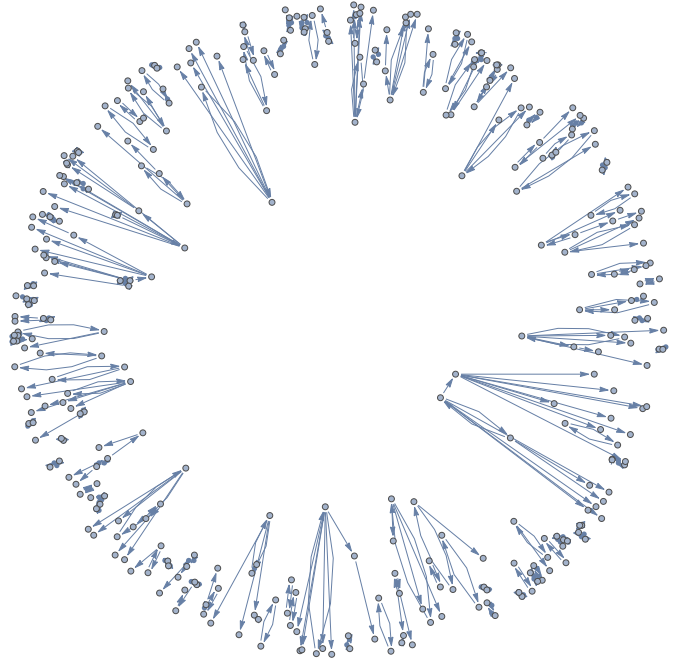


Fig. 8. Skeleton necessary for 100% greedy navigable graphs, $N = 350$, $R = 10$.