Proximity in the Brain

Zalán Heszberger[1,2] András Gulyás[1,2] András Majdán[1] András Biró[1]
László Balázs[1] Szabolcs Mezei[1] József Bíró[1]
[1]*Budapest University of Technology and Economics*
[2]*MTA-BME Information Systems Modelling Research Group*
Budapest, Hungary,
heszi@tmit.bme.hu

***Abstract*—The structural navigability of complex networks is an important question in the function-structure perspective of complex network analysis. This may provide hints on the underlying mechanisms that have been forming the structure of networks for a desirable level of navigation. It has been already discovered that greedy navigational cores as minimalistic networks with 100% greedy navigability considerably present in many real networks, including the structural networks of the human brain. Because the greedy navigational core is not unique, the connection between the level of its presence in a network and the structural navigability of that network is far from clear. In this paper, we deal with a special subset of the greedy navigational core, the so-called greedy proximity links (GPL), whose presence is necessary for 100% greedy navigability of a network. We show that the greedy proximity links are highly present in the brain networks, and the presence is consistent throughout the the individual subjects.**

***Index Terms*—structural brain networks, greedy navigation, proximity links**

I. Introduction

Thanks to the advancement of the last decade in MRI technologies and biomedical imaging, mammalian brains (including the human brain) is getting more and more explored [5]. This exploration has been relied and performed on a new prospering discipline, the network science. Networked neuroscience as a subdiscipline at the cross-roads of neuroscience and network science deals with functional and structural characteristics of brain networks using mathematical tools and methodologies like graph theory, probability and statistics [2], [5], [12]. One of the largest measurement projects in this area was the Human Connectome Project, whose goal was to build a "network map" that would show structural (anatomical) and functional connectivity within the healthy human brain. This project induced many activities in scanning and imaging brains and investigating them. The resulted data are often networks (graphs) which show interconnections or interrelations between different parts of the brain. In case of structural brain networks the gray matter is divided into brain parcels (RoI, region of interest) providing the vertices of the network, and the neuronal pathways through the white matter connecting these parcels form the edges of the networks.

Thanks to OTKA FK17 123957, FK 128233, KH18 129589, K17 124171 for funding. Zalán Heszberger is also supperted by MTA Bolyai János Research Grant and UNKP-19-4 Bolyai+ Research Grant. László Balázs and András Biró are supported by the ÚNKP-20-1 New National Excellence Program of the Ministry of Innovation and Technology from the source of the national research, development and innovation fund.

Navigation in complex networks has been a widely studied function from operational and structural point of view [1], [9], [10]. In technological networks this path exploring, finding and coordination mechanism is often referred to as routing. Routing protocols are set up and operate in these networks to provide efficient information/vehicle/energy transport. In natural networks like in the brain, communication between the nodes seems to be important as well, hence the navigation may also play central role in the operation of these networks [1], [11], [13].

Greedy navigation is a widely studied concept of routing in networks due to its simplicity and local way of operation [3], [7]–[10]. This hop-by-hop routing can work in geometrically embedded networks, i.e. the nodes of the networks have to possess coordinates chosen from some kind of metric space. Routing decisions are performed at every node based on distance calculations in the underlying metric space. More specifically, at a node the information parcel (packet) is forwarded to that neighbor which is closest to the destination node by a certain distance metric. Apparently, geometrically embedded networks are not necessarily 100% navigable by this mechanism. Hence, in this context, a very important question is the following: Given a set of nodes with coordinates without any connection (empty network). What is the minimal network (in terms of the number of links) which can provide 100% navigability between the nodes?

In [6] it has been shown that the minimalistic 100% greedy navigable core (GNC) of the networks are the Nash equilibrium of a network navigation game played by the nodes. The GNC can be considered as a navigation skeleton of real networks that inherits many structural properties. It was also demonstrated in [6] that the greedy navigable cores are highly present in a human brain structural network averaged over 5 subjects. A follow up study can be found in [4] which contains a more detailed analysis on the precision (i.e., the level of presence) of the greedy navigational core in human brain structural networks. Because the GNC as the Nash equilibrium of the network navigation game is not unique, there is no direct connection between the level of greedy navigability and the presence of GNC in a network. However, there exists a set of links which shows up in all possible greedy navigational cores, we call the Greedy Proximity Links (GPL). In this paper, we investigate the presence of GPL in 40 subject based human brain structural networks with five different resolutions. We

show that GPL presence is high in the brain networks and consistent in all scales, even in case of network pruning when anatomically weak or spurious links have been cut out from the networks.

II. Brain Networks

In this paper we use a high quality dataset on MRI measurements on human subjects in five different resolutions resulting in 200 structural brain networks. The dataset used in our investigations contains 40 healthy human subjects who underwent an MRI measurement procedure, where Diffusion Spectrum Imaging (DSI) data were obtained for each subject. The DSI data was processed according to the methods described in [2], resulting in 40 weighted, undirected structural connectivity maps comprising 83, 129, 233, 463, 1015 nodes in five different scales, respectively. Each node represents a region of cortical or sub-cortical gray matter, and the links represent white matter streamlines connecting the brain regions. Connection weights measure the average density of white matter streamlines. A connection is identified only if the density is above 10^{-8}, resulting structural networks containing an average of 1119, 1976, 3799, 7246 and 14254 connections per subject. The geometric centres of the brain regions are identified and are presented with their 3D Euclidean coordinates, leading to an embedding of the brain networks in the 3D Euclidean space.

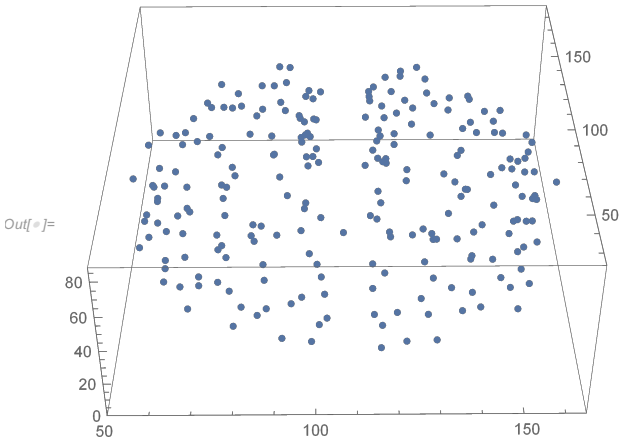

Fig. 1. Brain parcel centres of a scale 3 brain network as drawn by their 3D Euclidean coordinates. One can easily identify the two hemispheres. The points represent the centres of brain parcels ROI's (region of interest). Due to anatomical variations the 3D Euclidean coordinates of the centres of ROI's varies from network/subject to network/subject.

The DSI technology, as many other non-invasive imaging methods, has limitations due to measurement errors and artifacts of data post-processing. It can happen that spurious links appear in the brain networks by inference or existing weak connections remain hide from imaging. One possible way of handling this problem at some extent is to score somehow the links in the brain network and then a sequence of pruned networks can be generated by using the link scores. In our data set, in every scale, 40 networks are inferred from measurements carried out on 40 individuals. Because the anatomical parcellations of the individuals were consistent, a node in the brain networks represents the same brain parcel in all subjects. Therefore, a link (connection) can have a prevalence score, which simply means the number of networks among the 40 ones containing that link. The link prevalence score (LPS) can be used to prune the brain networks and then the robustness of any results or considerations on them.

III. Greedy Proximity Links

As mentioned earlier, greedy navigation is not always successful in geometrically embedded networks. Greedy routing can get stuck at a node, when all neighbours of the node is farther from the destination node than itself, that is, there is no way (no route) to get closer to the destination. However, there is always a common subset of all maximally navigable subnetworks, which we called Greedy Proximity Links (GPL). The presence of GPL links is necessary but still not a sufficient condition for maximal greedy navigability. This set of links can be identified as follows. Let us take two nodes in the network, u and v. If v is the closest node to u then for ensuring 100% greedy navigability the (directed) link $v \to u$ must exist. Otherwise there would be no greedy path from v (or through v) to u. Note that this kind of proximity is not a symmetric property. If v is the closest node to u this does not necessarily imply that u is the closest node to v, nevertheless it can happen, typically when u and v are close enough to each other and both of them are far from all the other nodes. Based on the node coordinates, Greedy Proximity Links can also be identified in the brain networks. Because in the structural brain networks the connections are undirected, we have also made the GPL links undirected. This means that if both $u \to v$ and $v \to u$ are proximity links they are presented as a single undirected connection $u - v$. Otherwise, if there is only one proximity link between a node pair, then this directed link is converted to an undirected link. Apparently, there are N directed proximity links in an N-node network, from which it follows that the number of undirected proximity links is less than N. Here it is worth emphasizing that the Greedy Proximity Links can be determined in the brain networks by using purely the 3D coordinates of the parcels; no other anatomical information or consideration is needed.

Now we turn to the detailed analysis of the GPL precision in the brain networks. The GPL links are generated for all the nodes of the networks in all scales and then they are tested against the brain networks. Adjacency matrices are inferred for GPL links in all networks and used together with the original brain network adjacency matrices to calculate the required parameters. First some basic properties of the GPL networks are presented. The average number of proximity links (in brackets the number of nodes) in the different scales are:

$$58.6(83), 90.2(129), 163.8(234), 324.4(463), 716.7(1015)$$

As mentioned earlier, in case on N nodes, the number of undirected proximity links is between $N/2$ and N depending on how many double directed proximity links are between the node pairs. It follows that the average degree of nodes

$\bar{k}_{GPL}$ with respect to the undirected proximity links should be between 1 and 2. The average degree of GPL networks in the five scales (from scale 1 to scale 5 respectively) are as follows:

$$1.41, 1.39, 1.40, 1.40, 1.41,$$

which are obtained by

$$\bar{k}_{GPL} = \frac{2 * \texttt{Mean}(\#\texttt{GPL links})}{N}. \quad (1)$$

One can not only observe that the previous statement $1 < \bar{k}_{GPL} < 2$ is confirmed but also the average degree of nodes in GPL networks are quite stable across the different scales. In other words, it is insensitive to the resolution of the brain parcellation. This also means that the ratio of the number of double directed proximity links [1] and the number of single directed proximity links is insensitive to the size of the brain networks. Because the greedy proximity links depend only on the 3D euclidean coordinates of the nodes this insensitivity property may relate to the consistency of the anatomical parcellation. The GPL precisions have been measured in all networks and all scales and the results are as follows: The mean values (and standard deviation) of the GPL precisions (without pruning) in Scale 1 to Scale 5 are

$$0.94(0.022), 0.98(0.024), 0.95(0.026), 0.83(0.031), 0.59(0.034).$$

One can observe that these inclusion ratios in the first three scales are exceptionally high and decreases significantly only in scale 5. The low variability of the precision values are valid within all scales. In Fig. 2 all the GPL precisions are plotted. On the figure one can also recognize the low deviations of the precisions within the scales.

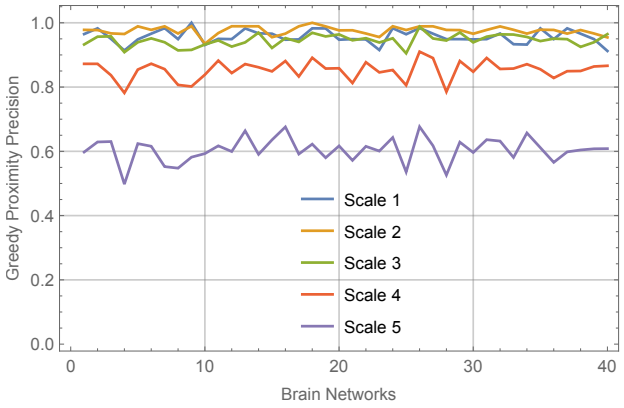

Fig. 2. Individual GPL precisions in all brain networks and all scales.

An illustration of the undirected proximity links of scale 5 brain parcels can be seen in the left part of Fig 3. On the right, the proximity links are shown, which are also part of the original brain network. One can easily identify the two hemispheres and other major parts in both cases. In this particular case the total number of links in the scale 5 brain network is 14695, while the number of proximity links is 728. Among these 728 proximity links there are 458 ones, which are present in the original brain network. This corresponds to a 0.63 precision, which is much lower than in the first four scales, however, still amazingly high.

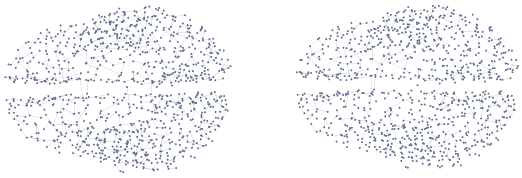

Fig. 3. Greedy proximity links in a scale 5 brain network. The left drawing represents all the proximity links (728) inferred from the coordinates of the nodes. The right part of the figure shows only those proximity links (458) which are also present in the original brain network.

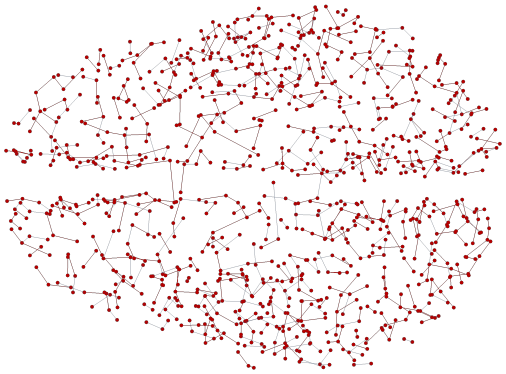

Fig. 4. Greedy proximity links in a scale 5 brain network are shown with highlighting the ones which are also present in the original brain network.

Now, results are presented on the precision of GPL in pruned brain networks. The pruning process is the following. In case of pruning threshold parameter $p, p = 1 \ldots 40$ all links having scores less than p are cut out from all networks. In case of $p = 1$ all links are kept, thus we have the original set of networks. In case of $p = 40$, we have only those networks which contain only the links with $LPS = 40$, all other links are cut out. In this way, in every scale, 1600 networks have been generated. GPL precisions are measured for every set of pruned networks for all possible pruning threshold values. For example, in a scale 2 brain network (with 2073 connections) 86 of 88 greedy proximity links are present. This corresponds to a GPL precision of 97.7 % . The average values of precisions are shown in Fig. 5, which underlines that pruning out links up to $p = 10$ precisions do not change significantly in any of the scales. In case of lower resolutions (the first three scales) it is even true for $p = 30$. In the highest resolution (scale 5) the GPL precision is remarkably lower than in others, but the decrease is more flat between $p = 20$ and $p = 40$. Based on the

[1] A proximity link between a node pair (u, v) is double directed if u is closest to v and also v is closest to u.

results one can deduce that precisions of the Greedy Proximity Links in the pruned brain networks are robust against pruning. This means that most of the proximity links contained by the brain networks have high link prevalence scores; cutting out the possibly non-existent low score links does not affect significantly the GPL precisions in a wide range of LPS.

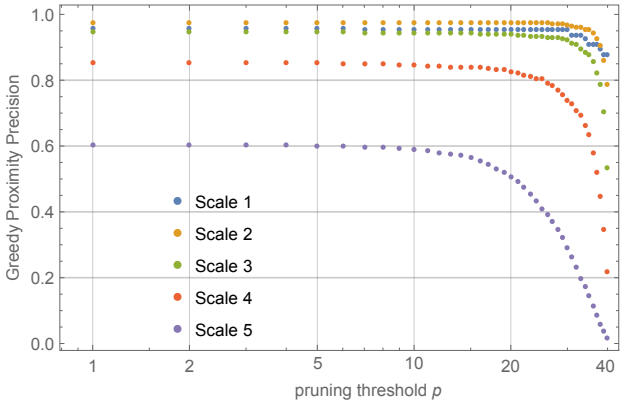

Fig. 5. Average greedy proximity precisions against link prevalence score thresholds.

These statements can also be supported by the distributions of the LPS values of the greedy proximity links present in the brain networks. The empirical probability distribution functions (PDF) of link prevalence score values are generated for all greedy proximity links in every scales, which can be seen in Fig. 6. In the first three scales the greedy proximity

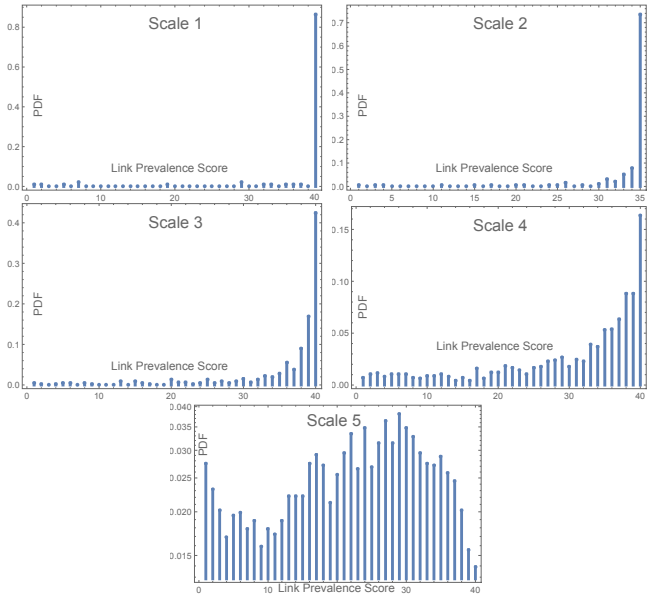

Fig. 6. LPS distributions in all scales.

links have very high link prevalence scores, for example, in the particular scale 2 network there are $70, 6, 5, 2, 1, 1, 1$ links with LPS values $40, 39, 38, 37, 36, 35, 29$, respectively. The LPS distribution of GPL is more widened in scale 4 and scale 5, i.e. they show a higher variance, however, they concentrate around higher LPS values as well. The mean LPS values of proximity links are $37.55, 38.32, 36.00, 30.23, 21.57$ in scale 1 to 5, respectively.

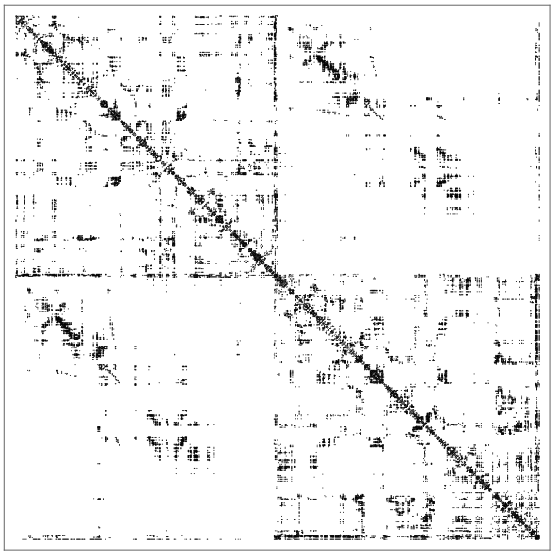

Fig. 7. Array plot of a scale 5 structural brain network. In this plot a dot at (i, j) represents the connection between nodes i and j. One can recognize higher connectivity inside the hemispheres, and lower density for inter-hemispheric connections.

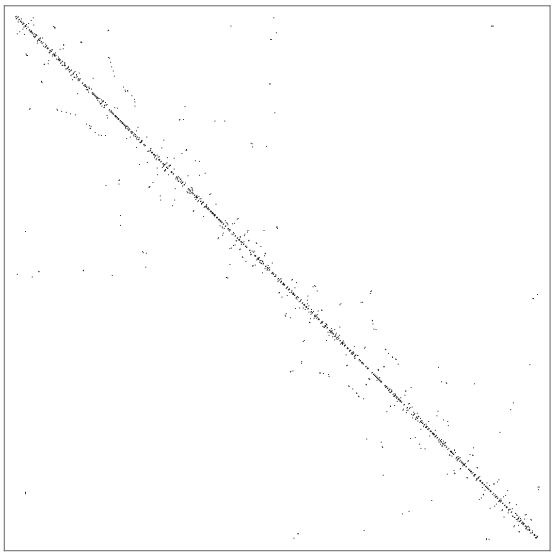

Fig. 8. Array plot for scale 5 greedy proximity links. If we take a closer look at the pattern, we can recognize that the morphology resembles to the previous full brain network, however, the density is much lower because it contains only 1015 links.

IV. Conclusion

Greedy Proximity Links have been identified whose existence are necessary (but not sufficient) for 100% greedy navigability in networks. GPL precision (inclusion ratio) is tested in a high-quality data set on structural networks of the human brain. The precision is very high in lower resolutions

of brain parcellations, lower (but still significantly high) at higher resolutions and consistent with low variability in all scales. The results are remarkable in light of that Greedy Proximity Links are inferred only from the 3D Euclidean coordinates of brain parcels; no other information is used from the underlying anatomy or functions of the brains. One may speculate that the strong universal consistency of greedy proximity link presence in structural brain networks can relate to the navigation functionalities of the brain, nevertheless, there is still no direct evidence for that.

REFERENCES

[1] Andrea Avena-Koenigsberger, Bratislav Misic, and Olaf Sporns. Communication dynamics in complex brain networks. *Nature Reviews Neuroscience*, 19(1):17, 2018.

[2] Richard F Betzel, Andrea Avena-Koenigsberger, Joaquín Goñi, Ye He, Marcel A De Reus, Alessandra Griffa, Petra E Vértes, Bratislav Mišic, Jean-Philippe Thiran, Patric Hagmann, et al. Generative models of the human connectome. *Neuroimage*, 124:1054–1064, 2016.

[3] Karl Bringmann, Ralph Keusch, Johannes Lengler, Yannic Maus, and Anisur Rahaman Molla. Greedy routing and the algorithmic small-world phenomenon. In *Proceedings of the ACM Symposium on Principles of Distributed Computing*, pages 371–380, 2017.

[4] Z. Heszberger et al. Greedy navigational cores in the human brain. In *Proceedings of the CSCE'20 - The 2020 World Congress in Computer Science, Computer Engineering, Applied Computing*. ACSE, 2020.

[5] Alex Fornito, Andrew Zalesky, and Edward Bullmore. *Fundamentals of brain network analysis*. Academic Press, 2016.

[6] 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.

[7] Robert Kleinberg. Geographic routing using hyperbolic space. In *IEEE INFOCOM 2007-26th IEEE International Conference on Computer Communications*, pages 1902–1909. IEEE, 2007.

[8] Dmitri Krioukov, Fragkiskos Papadopoulos, Marián Boguñá, and Amin Vahdat. Greedy forwarding in scale-free networks embedded in hyperbolic metric spaces. *ACM SIGMETRICS Performance Evaluation Review*, 37(2):15–17, 2009.

[9] Alessandro Muscoloni and Carlo Vittorio Cannistraci. Navigability evaluation of complex networks by greedy routing efficiency. *Proceedings of the National Academy of Sciences*, 116(5):1468–1469, 2019.

[10] Ioannis Pappas, Michael M Craig, David K Menon, and Emmanuel A Stamatakis. Structural optimality and neurogenetic expression mediate functional dynamics in the human brain. *Human Brain Mapping*, 41(8):2229–2243, 2020.

[11] Caio Seguin, Martijn P Van Den Heuvel, and Andrew Zalesky. Navigation of brain networks. *Proceedings of the National Academy of Sciences*, 115(24):6297–6302, 2018.

[12] Olaf Sporns. *Networks of the Brain*. MIT press, 2010.

[13] Dale Zhou, Christopher W Lynn, Zaixu Cui, Rastko Ciric, Graham L Baum, Tyler M Moore, David R Roalf, John A Detre, Ruben C Gur, Raquel E Gur, et al. Efficient coding in the economics of human brain connectomics. *arXiv preprint arXiv:2001.05078*, 2020.