

The Budapest Reference Connectome Server Ver. 2.0

Balázs Szalkai^{a,**}, Csaba Kerepesi^{a,**}, Bálint Varga^{a,**}, Vince Grolmusz^{a,b,*}

^a*PIT Bioinformatics Group, Eötvös University, H-1117 Budapest, Hungary*

^b*Uratim Ltd., H-1118 Budapest, Hungary*

Abstract

The connectomes of different human brains are pairwise distinct: we cannot talk about an abstract "graph of the brain". Two typical connectomes, however, have quite a few common graph edges that may describe the same connections between the same cortical areas. The Budapest Reference Connectome Server Ver. 2.0 (<http://connectome.pitgroup.org>) generates the common edges of the connectomes of 96 distinct cortexes, each with 1015 vertices, computed from 96 MRI data sets of the Human Connectome Project. The user may set numerous parameters for the identification and filtering of common edges, and the graphs are downloadable in both csv and GraphML formats; both formats carry the anatomical annotations of the vertices, generated by the Freesurfer program. The resulting consensus graph is also automatically visualized in a 3D rotating brain model on the website.

The consensus graphs, generated with various parameter settings, can be used as reference connectomes based on different, independent MRI images, therefore they may serve as reduced-error, low-noise, robust graph representations of the human brain.

1. Introduction

Several large-scale projects for brain-mapping are being executed [1, 2, 3], but the neuron-scale graph of the human brain, where the nodes are the neurons, and two neurons are connected by an edge if they are joined through a synapse, is out of reach today [4]. The difficulties come from the number of the neurons to be mapped, and also from the lack of the high-throughput methods for mapping their connections. The neuron-scale graphs were constructed only for very simple organisms with a very small number of neurons [5, 6, 7] or for just small cortical areas of more complex organisms [8, 9].

*Corresponding author

** Joint first authors

Email addresses: szalkai@pitgroup.org (Balázs Szalkai), kerepesi@pitgroup.org (Csaba Kerepesi), balorkany@pitgroup.org (Bálint Varga), grolmusz@pitgroup.org (Vince Grolmusz)

The application of magnetic resonance imaging (MRI) offers numerous methods for mapping physical and functional connections between subdivided anatomical areas of the brain (called "Regions of Interests", ROIs), each consisting of millions of neurons. The vertices are the ROIs, and two ROIs are connected by an edge if connections are detected between them by an MRI-based method. This method can either be diffusion MRI imaging, depicting the Brownian motion of water molecules in axons, consequently, mapping the axons between different cortical areas; or functional MRI (fMRI) imaging, depicting brain areas of elevated blood flow while the subject rests or performs different mental tasks.

In this note we present a web-server, which, starting from the diffusion MRI data published as a result of the Human Connectome Project [3], compiles differently parametrized reference graphs from the common edges of the graphs describing 96 different 1015-vertex graphs of 96 human subjects. Additionally, a default, single graph, the Budapest Reference Connectome Ver. 2.0 is also presented in two downloadable formats.

The resulting graphs may be used for identifying more robust, more error-free connections between the cortical areas, represented by ROIs: for example, in the default reference graph if an edge is present then it is present in at least 14 different source graphs. In general, one may set the "Minimum edge confidence" to value k anywhere between $k = 1$ (where an edge is included if it is present in at least one source graph) through $k = 96$ (where an edge is present in the reference graph if it can be found in all the 96 source graphs).

Therefore, the resulting graphs contain *common*, *consensus* edges (i.e., Fig. 1) originated from multiple graphs with user-specified parameters, computed from the diffusion MRI data of different subjects.

Version 2.0 of the Budapest Reference Connectome Server is described here in detail. Choosing Version 1.0 is also possible on the website: Version 1.0 applies the source data from the classical article of [10] describing six connectomes of five subjects, each with 998 vertices. Version 1.0 of the webserver computes the consensus edges from those six graphs only.

By filtering edges with very few occurrences or those with small weights, one may get a connectome with more reliable edges and weights than in the case of any single dataset in the input. Therefore, we may get robust edges and weights in the consensus graphs generated by the server.

2. Results and Discussion

The webserver is available at <http://connectome.pitgroup.org>. The default, "canonical" Budapest Reference Connectome Ver. 2.0 can be downloaded by simply hitting the "Download graph" button without changing the default options. This default graph has 1015 vertices, 8507 edges.

The following options can be set after choosing the "Show options" button:

- (i) Version 1.0 or 2.0. The default choice is 2.0, using the graphs of 96 subjects, computed from the Human Connectome Project [3]. The user may

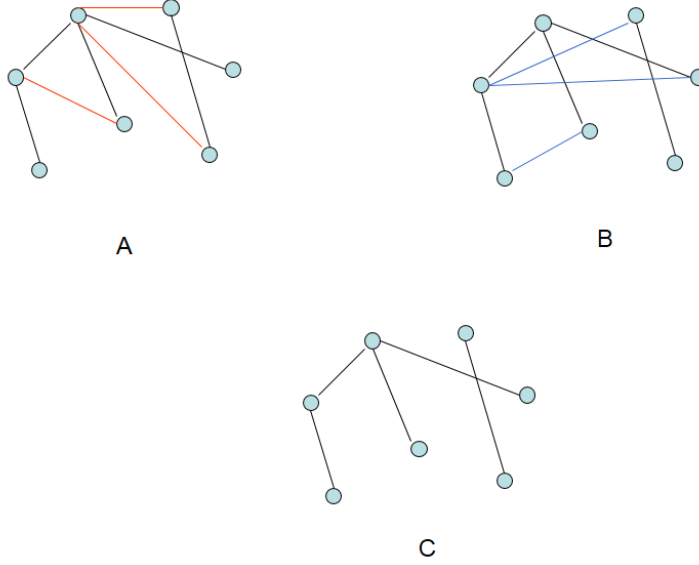


Figure 1: The black edges of graphs A and B are common edges; they form graph C, the consensus graph.

alternatively choose Version 1.0, that applies only six graphs computed and described by the classical article of [10].

- (ii) Minimum edge confidence: The graph to be constructed will contain all the edges that are present in at least k graphs, between the very same vertices in each graph. The valid choices for $k = 1, \dots, 96$. The last choice means that each source graph needs to contain the edge in order to be presented in the resulting consensus graph.

For each edge $\{u, v\}$, the weight of that edge is a fraction n/L , where n is the number of fibers connecting u and v , and L is the average length of the fibers.

- (iii) Minimum edge weight: One may set a slider to a value of minimum weight required. The returned graph will contain edges whose mean or median weights are larger than or equal to this value. The mode of computation (mean or median) can be set by the next option.
- (iv) Weight calculation mode: There are two choices: Median or Mean. Choice "Median" means that from the list of weights appearing as the weights of

the same edge in different graphs, we use the "central" element, that is, first we sort the weights, and next the element is chosen that separates the upper half of the weights from the lower half of the weights. "Mean" means the arithmetic average of the weights. The default choice is the median, since the median is more robust than the mean: typically extreme large or small strengths have less impact to the median than to the mean.

The resulting graph can be downloaded in CSV or GraphML formats, or can readily be visualized on the webpage. The downloaded file-names inherit the parameter-settings as follows: e.g., the Budapest Reference Connectome Ver. 2.0 is given as the file "budapest_connectome_2.0_14.0_median.csv", that is, the csv file contains the graph generated by Version 2.0 of the server, with a minimum confidence of 14 (i.e., each edge of the graph is contained in at least 14 input graphs), and the minimum edge weight is 0 and the weights of the edges of the reference graph are computed as the median of the weights of the corresponding edges of the input graphs.

The format of the CSV file is demonstrated on Table 1.

v_1	v_2	name of v_1	name of v_2	conf.	weight-med.	weight-mean
501	504	rh.insula_9	Right-Putamen	92	0.883	0.974

Table 1: The structure of the result file in csv format. Each row corresponds to an edge of the graph. The first four columns identify the vertices, connected by the edge; the first two entries describe the vertex numbers, the third and fourth columns the anatomical name according to the mapping generated by Freesurfer. The fifth column contains the confidence, the sixth and seventh columns the weights computed according to the median and the mean of the source edges.

The number of the common edges in at least n graphs ($n=1,2,\dots,96$) are given on Figure 2.

3. Methods

The main server, denoted as "Version 2.0", was created as follows:

The dataset is a subset of Human Connectome Project 500 Subjects Release (<http://www.humanconnectome.org/documentation/S500/>), containing MRI images of healthy adult males and females between the ages of 22 and 35. The data was downloaded in October, 2014.

Partitioning, tractography, and graph construction were done by the Connectome Mapper Toolkit (<http://cmtk.org>).

Partitioning of the grey matter was done by the Lausanne2008 method [10] into 1015 ROIs of surface area of about 1.5 mm^2 .

For tractography, the deterministic streamline method was applied.

The graphs were constructed as follows: Two ROIs were connected by an edge if there exists at least one fiber, determined by the tractography step, that connects these two ROIs. The number and the length of the fibers are taken care of by computing the weights of the edges: For each edge $\{u, v\}$, the weight

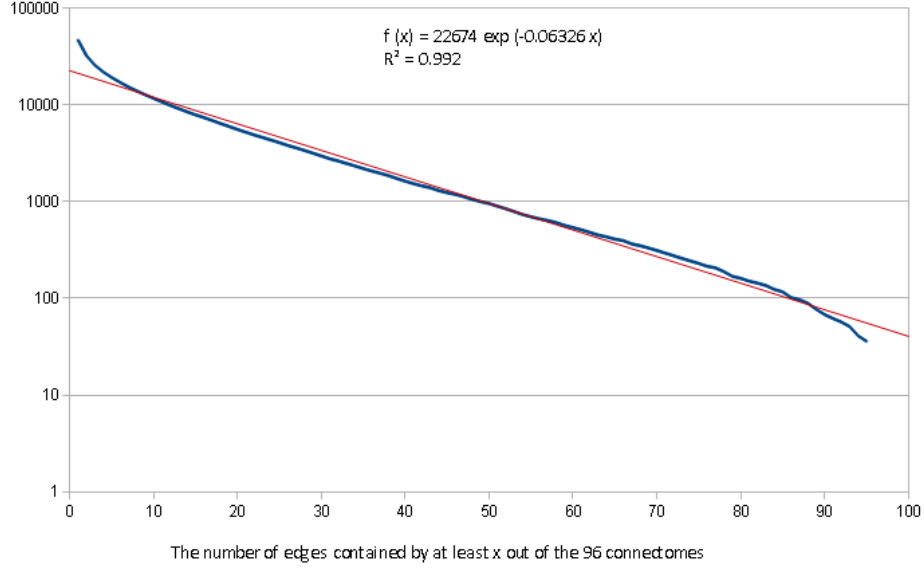


Figure 2: The plot of the number of common edges.

of that edge is a fraction n/L , where n is the number of fibers connecting u and v , and L is the average length of the fibers.

After 96 graphs were computed, each with 1015 vertices, we identified the common edges, their confidence, and weights, computed according to their median and mean. The large, pre-computed tables were integrated into the web-server.

Version 1.0 of the webserver applies the six graphs that were described in [10]. The definition of weight (called strength) and its computation, and also the parcellation of the cortex used are described in [10]. The six connectomes were downloaded from http://www.cmtk.org/datasets/homo_sapiens_01.cff in September, 2014.

The visualization component applies a modified version of the WebGL Brain Viewer [11].

4. Acknowledgments

The authors declare no conflicts of interest.

References

- [1] Heidi Johansen-Berg. Human connectomics - what will the future demand? *Neuroimage*, 80:541–544, Oct 2013. doi: 10.1016/j.neuroimage.2013.05.082. URL <http://dx.doi.org/10.1016/j.neuroimage.2013.05.082>.

- [2] Yaniv Assaf, Daniel C. Alexander, Derek K. Jones, Albero Bizzi, Tim E J. Behrens, Chris A. Clark, Yoram Cohen, Tim B. Dyrby, Petra S. Huppi, Thomas R. Knoesche, Denis Lebihan, Geoff J M. Parker, Cyril Poupon, C. O. N. N. E. C. T consortium, Debbie Anaby, Alfred Anwander, Leah Bar, Daniel Barazany, Tamar Blumenfeld-Katzir, Silvia De-Santis, Delphine Duclap, Matteo Figini, Elda Fischì, Pamela Guevara, Penny Hubbard, Shir Hofstetter, Saad Jbabdi, Nicolas Kunz, Francois Lazeyras, Alice Lebois, Matthew G. Liptrot, Henrik Lundell, Jean-François Mangin, David Moreno Dominguez, Darya Morozov, Jan Schreiber, Kiran Seunarine, Simone Nava, Cyril Poupon, Till Riffert, Efrat Sasson, Benoit Schmitt, Noam Shemesh, Stam N. Sotiropoulos, Ido Tavor, Hui Gary Zhang, and Feng-Lei Zhou. The CONNECT project: Combining macro- and micro-structure. *Neuroimage*, 80:273–282, Oct 2013. doi: 10.1016/j.neuroimage.2013.05.055. URL <http://dx.doi.org/10.1016/j.neuroimage.2013.05.055>.
- [3] Jennifer A. McNab, Brian L. Edlow, Thomas Witzel, Susie Y. Huang, Himanshu Bhat, Keith Heberlein, Thorsten Feiweier, Kecheng Liu, Boris Keil, Julien Cohen-Adad, M Dylan Tisdall, Rebecca D. Folkerth, Hannah C. Kinney, and Lawrence L. Wald. The Human Connectome Project and beyond: initial applications of 300 mT/m gradients. *Neuroimage*, 80:234–245, Oct 2013. doi: 10.1016/j.neuroimage.2013.05.074. URL <http://dx.doi.org/10.1016/j.neuroimage.2013.05.074>.
- [4] Chaehyun Yook, Shaul Druckmann, and Jinhyun Kim. Mapping mammalian synaptic connectivity. *Cell Mol Life Sci*, 70(24):4747–4757, Dec 2013. doi: 10.1007/s00018-013-1417-y. URL <http://dx.doi.org/10.1007/s00018-013-1417-y>.
- [5] Dmitri B. Chklovskii, Shiv Vitaladevuni, and Louis K. Scheffer. Semi-automated reconstruction of neural circuits using electron microscopy. *Curr Opin Neurobiol*, 20(5):667–675, Oct 2010. doi: 10.1016/j.conb.2010.08.002. URL <http://dx.doi.org/10.1016/j.conb.2010.08.002>.
- [6] Siemon C. de Lange, Marcel A. de Reus, and Martijn P. van den Heuvel. The Laplacian spectrum of neural networks. *Front Comput Neurosci*, 7:189, Jan 2014. doi: 10.3389/fncom.2013.00189. URL <http://dx.doi.org/10.3389/fncom.2013.00189>.
- [7] Emma K. Towlson, Petra E. Vértes, Sebastian E. Ahnert, William R. Schafer, and Edward T. Bullmore. The rich club of the c. elegans neuronal connectome. *J Neurosci*, 33(15):6380–6387, Apr 2013. doi: 10.1523/JNEUROSCI.3784-12.2013. URL <http://dx.doi.org/10.1523/JNEUROSCI.3784-12.2013>.
- [8] Cole Gilbert. Brain connectivity: revealing the fly visual motion circuit. *Curr Biol*, 23(18):R851–R853, Sep 2013. doi: 10.1016/j.cub.2013.08.018. URL <http://dx.doi.org/10.1016/j.cub.2013.08.018>.

- [9] James R. Anderson, Bryan W. Jones, Carl B. Watt, Margaret V. Shaw, Jia-Hui Yang, David Demill, James S. Lauritzen, Yanhua Lin, Kevin D. Rapp, David Mastronarde, Pavel Koshevoy, Bradley Grimm, Tolga Tasdizen, Ross Whitaker, and Robert E. Marc. Exploring the retinal connectome. *Mol Vis*, 17:355–379, 2011.
- [10] Patric Hagmann, Leila Cammoun, Xavier Gigandet, Reto Meuli, Christopher J. Honey, Van J. Wedeen, and Olaf Sporns. Mapping the structural core of human cerebral cortex. *PLoS Biol*, 6(7):e159, Jul 2008. doi: 10.1371/journal.pbio.0060159. URL <http://dx.doi.org/10.1371/journal.pbio.0060159>.
- [11] Daniel Ginsburg, Stephan Gerhard, John Edgar Congote, and Rudolph Pienaar. Realtime visualization of the connectome in the browser using webgl. *Frontiers in Neuroinformatics*, 2011. URL http://www.frontiersin.org/10.3389/conf.fninf.2011.08.00095/event_abstract.