The Eurekometric Connectome: Discovering unexplored areas of neuroscience research

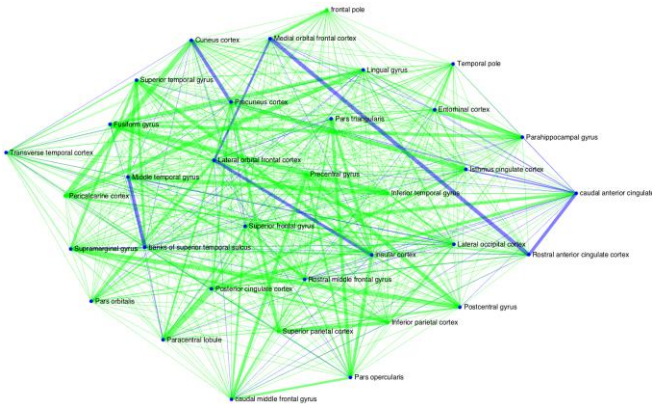
Malhar Jere, Ravi Kiran Raman, and Lav R. Varshney
University of Illinois at Urbana Champaign, IL, USA

In this study, we build a *eurekometric connectome* (representing how often brain regions are studied together) using papers from the Public Library of Science. The brain regions are nodes and edge weight is the number of papers in which two regions are co-studied. This connectome is compared with structural and functional connectomes to find potentially understudied brain regions.

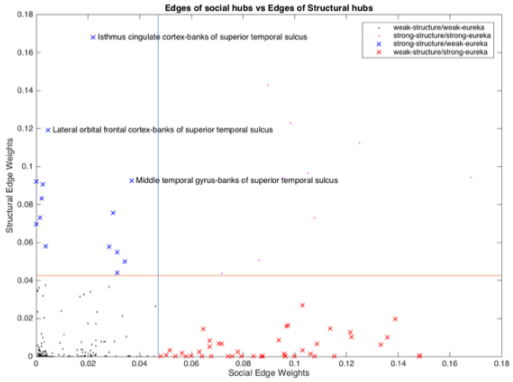
In building a statistical methodology for multilayer networks we ran tests to determine such regions. The first test found the highest strength nodes common to the eurekometric and structural connectomes, plotted these nodes' edges against each other and used *k*-means clustering to divide them into quadrants (Fig. b). The second and fourth quadrants represent over- or understudied edges that are of primary interest.

The second test partitions the structural graph on its edge weights using a fixed threshold, determines corresponding edges in the eurekometric graph and runs the Kolmogorov-Smirnov test on both datasets. After plotting this test value for different thresholds and against several null random-graph ensembles, we observed that the *eurekometric* graph is derived from the structural graph. We also observed that among the top 35 structural edges, 6 are missing from the *eurekometric* graph, implying they haven't been studied. (Fig. a).

This work uses text mining to extend scientometrics from using papers and their citation patterns as atomic units of study to using actual concepts of science (here brain regions), and uses novel multilayer network science methods to suggest new groupings of brain regions for future study in systems neuroscience.



(a)



(b)

Eurekometric Connectomes. (a) Structural connectome in green with edges that do not exist in the eurekometric connectome highlighted in blue (b) Edges of top hubs (by outgoing edge weight of each node) that appear in both connectomes partitioned into structure/eureka sets. The top three edges that are strong-structure/weak-eureka are labeled.