



THE UNIVERSITY  
of NORTH CAROLINA  
at CHAPEL HILL

# Identifying Individuals from Brain Connectivity Networks Using Graph Measures of Similarity

Oscar Fawcett, Dr. Nicolas Fraiman, Dr. Eran Dayan  
Department of Statistics and Operations Research, Department of Radiology, UNC-Chapel Hill

## SUMMARY

- Functional magnetic resonance imaging (fMRI) of the brain can be used for accurately identifying, or “fingerprinting”, individuals.
- This process of identification is done by calculating how similar brain network are across different sessions, and predicting a subject’s identity based on the chosen similarity measure.
- However, the common similarity measure used is based on simple correlation analysis between each vectorized network, which lacks mathematical foundation.
- We investigate here the integration of other network similarity measures by methods found in traditional graph theory such as edit distance and spectral distance.

## INTRODUCTION

- Functional magnetic resonance imaging (fMRI) of the brain can be used for accurately identifying individuals [1].
- Referred to as “connectome fingerprinting”, this process is done by comparing how similar each functional connectivity network from one imaging session was to another.
- However, the similarity measure used thus far is based on simple correlation analysis between each vectorized functional connectivity network, which lacks mathematical foundation.

**Here, we investigate the integration of other network similarity measures found in traditional graph theory to the connectome fingerprinting problem.**

## DATA PROCESSING

- Due to the complexity of raw fMRI data, we transform the data into a network representation, which is known as a functional connectivity network. In this way, analytical methods found in network science and graph theory can be applied.
- A network, also known as a graph, is a mathematical object containing a collection of nodes, edges, and weights. This is often represented as the adjacency matrix  $A$ . Each element of this matrix  $A$  can be expressed as follows:

$$A_{ij} = \begin{cases} w_{ij} & \text{if edge } (i, j) \text{ exists} \\ 0 & \text{otherwise} \end{cases}$$

Where  $w_{ij}$  is the weight given for edge  $(i, j)$ .

- The processing pipeline used is illustrated in figure 1..

## ANALYSIS & RESULTS

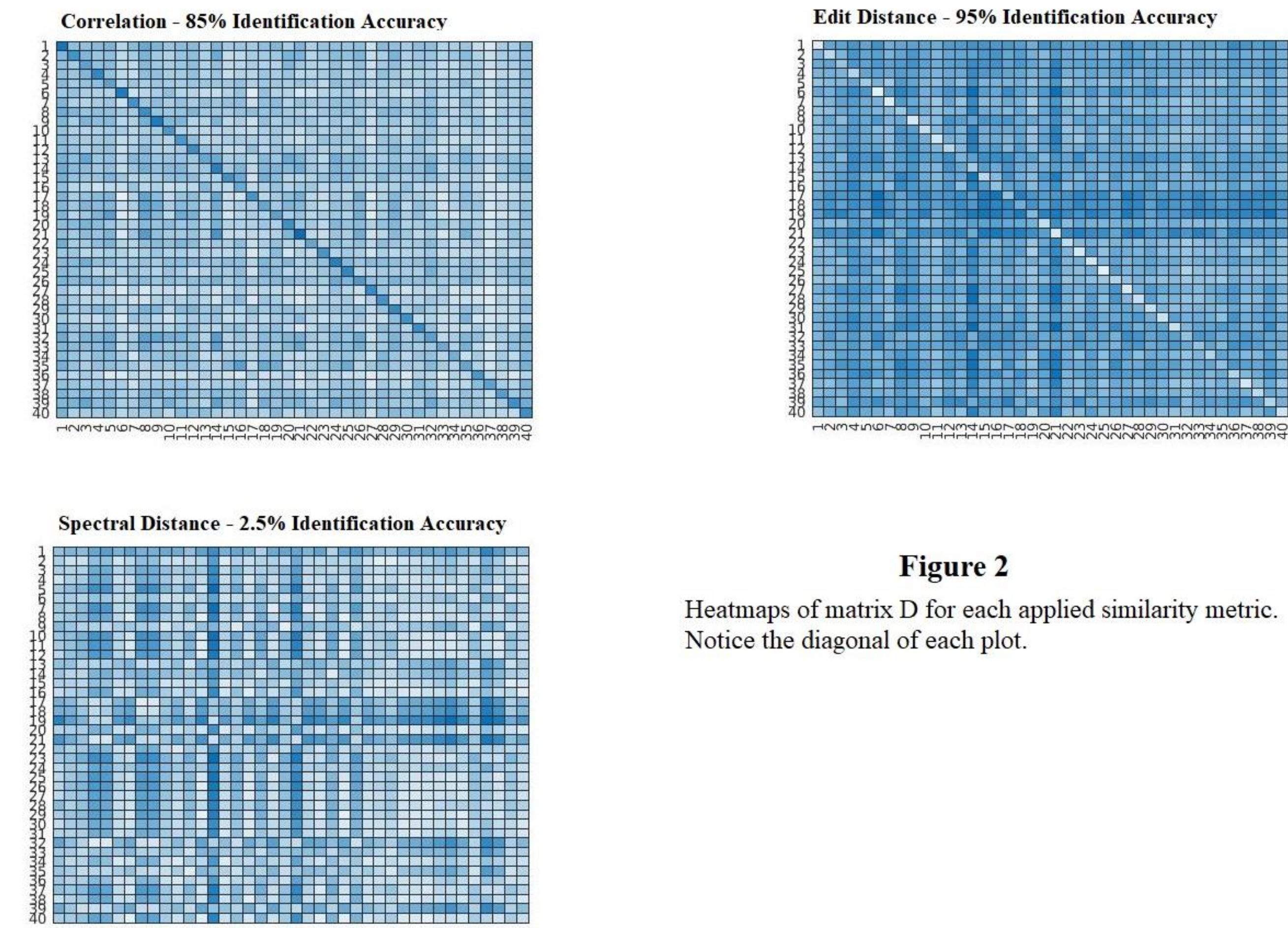
- For two sets of adjacency matrices  $G^1$  and  $G^2$  respectively, we define the distance matrix  $D$  such that  $D_{ij} = d(G_i^1, G_j^2)$ , for subjects  $i$  and  $j$  and similarity function  $d$ . We then consider a successful subject identification if for some subject  $k$ ,  $D_{kk}$  is the maximum or minimum value (depending on the similarity function  $d$ ) of the  $k^{th}$  row of  $D$ .
- As stated before, the common similarity measure used is simple correlation of the vectorized adjacency matrix.
- Here, we utilize measures of similarity as outlined in this paper [2]. Two methods that have been applied thus far are edit distance and spectral distance. Both are defined as follows:

$$d_{edit}(A, A') = \sum_{i=1}^n \sum_{j=1}^n |A_{ij} - A'_{ij}|$$

$$d_{spectral}(A, A') = \sqrt{\sum_{i=1}^n (\lambda_i - \lambda'_i)^2}$$

Where  $n$  is the number of nodes,  $A$  and  $A'$  are adjacency matrices, and  $\lambda_i$  is the  $i^{th}$  eigenvalue of the corresponding adjacency matrix.

**Our preliminary analysis includes 40 of the unrelated 100 dataset from the Human Connectome Project [3]. We found that edit distance outperformed correlation in subject identification.** Our results are shown in figure 2.



**Figure 2**  
Heatmaps of matrix D for each applied similarity metric. Notice the diagonal of each plot.

## DISCUSSION

**From the results found in the previous section, there are a few things of note:**

- Since in our current dataset edit distance is outperforming simple correlation, there is potential for this measure of similarity to be a better choice for the task of subject identification.
- Spectral distance performed very poorly on our dataset. A possible explanation for this is that while edit distance measures local structural differences between two networks, spectral distance measures more global structural differences between two networks [2]. So, it could be the case that functional connectivity networks have very similar global structures to them, resulting in more local measures of similarity performing better in distinguishing them apart.

Moving forward, processing more fMRI data for our analysis will be prioritized, with the target goal being to have the full unrelated 100 dataset from the Human Connectome Project [3] processed. Simultaneously, more measures of similarity will be tested, including NETSIMILE, resistance distance, and DELTACON. There will also be opportunity to study the robustness of the predictions each similarity measure makes, as well as which nodes are most useful in subject identification.

## REFERENCES

- [1] Finn, E., Shen, X., Scheinost, D. *et al.* Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nat Neurosci* **18**, 1664–1671 (2015). <https://doi.org/10.1038/nn.4135>
- [2] Willis, Peter and Meyer, Francois G.. Metrics for Graph Comparison: A Practitioner’s Guide. *PLoS ONE* **15**(2): e0228728, 2022. <https://doi.org/10.1371/journal.pone.0228728>
- [3] Van Essen, D.C. *et al.* The WU-Minn Human Connectome Project: an overview. *Neuroimage* **80**, 62–79 (2013).
- [4] Shen X, Tokoglu F, Papademetris X, Constable RT. Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. *Neuroimage*. 2013 Nov 15;82:403-15. doi: 10.1016/j.neuroimage.2013.05.081. Epub 2013 Jun 4. PMID: 23747961; PMCID: PMC3759540.

**Figure 1**

Data processing pipeline. For this project, the Shen parcellation was used [4].

