
Grafos en neurociencias (pre TP2)

Víctor A. Bettachini

Datamining en ciencia y tecnología 2023
Especialización en Explotación de Datos y Descubrimiento del Conocimiento
bettachini@gmail.com

Resumen

?

1. Materiales y métodos

Datos Se hace uso de datos producto de la medición de la señal de resonancia magnética funcional (fMRI). Definidos 116 volúmenes de interés del cerebro en términos de su activación [1], se publicaron coeficientes de correlación lineal entre sus medias en distintos segmentos temporales [2]. Con estos datos se generó una matriz de correlación, y a partir de la misma los grafos analizados en este trabajo.

Recurso informático Un cuaderno (notebook) Jupyter provisto por los docentes en el sitio web denominado “Campus” [3] es la plantilla donde se escribió código en lenguaje Python. Este explotó funciones de las biblioteca NetworkX [4].

1.1. Preprocesamiento de los datos

Carga del conjunto de datos Los archivos provistos corresponden a los estadíos de sueño N1, N2, N3 y despierto (W) para 18 sujetos. Estos estuvieron acompañados de una tabla que describe la denominación y ubicación espacial las regiones en que se parcializó el cerebro.

2. Resultados

2.1. Manipulación de datos

Matriz de correlación La matriz de adyacencia pesada que muestra la figura 1a corresponde a la condición de despierto para el sujeto número 2.

Una medida que caracteriza un grafo es la densidad, δ , definida como la razón entre los enlaces presentes sobre todos los posibles en el grafo. Para convertir la matriz en una de adyacencia binaria con una densidad de enlaces $\delta = 0,8$ se discriminaron sus pesos con un umbral 0,77997 obteniéndose la matriz que muestra la figura 1b.

Grafo de la matriz de adyacencia binaria No es totalmente conectado. Hay tres componentes conectadas de 92, 5 y 2 nodos en tanto que los 17 restantes están aislados. La distancia mínima entre nodos contabiliza cuantos intermedios deben atravesarse para ir de un nodo a otro. Esta medida para los aislados no tiene sentido, por lo en este grafo no puede obtenerse una distancia media d a lo fines de tener una media de que tan “conectado” se presenta el grafo. Una medida numérica alternativa la provee la densidad. Para el conjunto, no conectado, este valor parece bajo, $\approx 8\%$. Y en el componente mayoritario, el de 92 nodos, este valor apenas se incrementa hasta un $\approx 12\%$.

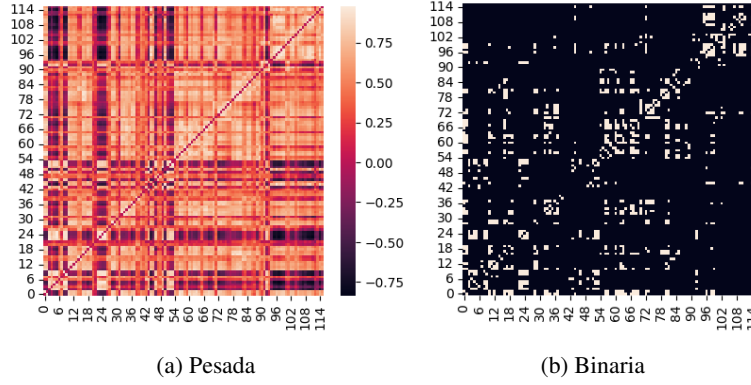


Figura 1: Matrices de correlación de las medias de las señales de las regiones parcializadas para el sujeto número 2 despierto.

Eficiencia de conectividad La distancia mínima entre nodos contabiliza cuantos intermedios deben atravesarse para ir de un nodo a otro. Un promedio de la inversa de esta distancia es una medida de la eficiencia de conectividad global del grafo, que para este caso resultó ser $\approx 0,389$.

Distribución de grado El número de enlaces por nodo, o grado k , se distribuye en forma dispar. De un total de 534 enlaces el grado mayor resultó $k_{\text{máx}} = 30$, y un relativamente alto promedio $\langle k \rangle \approx 9,207$ aunque no hay que olvidar que no participan aquí los nodos aislados. El histograma de k que muestra la figura 2a que este $\langle k \rangle$ es representativo de la distribución.

En una medida similar a la densidad que cuenta la proporción de enlaces sobre los posibles puede hacerse algo similar calculando la proporción de cuantos de los posibles enlaces entre primeros vecinos efectivamente se realizan. Esto se denomina coeficiente de agrupamiento (clustering) por nodo C_i , cuyo promedio para este grafo es $\langle C_i \rangle \approx 0,527$. Coloreando cada nodo según su C_i y ubicandole por las coordenadas (y,z) en el cerebro se obtiene la figura 2b.

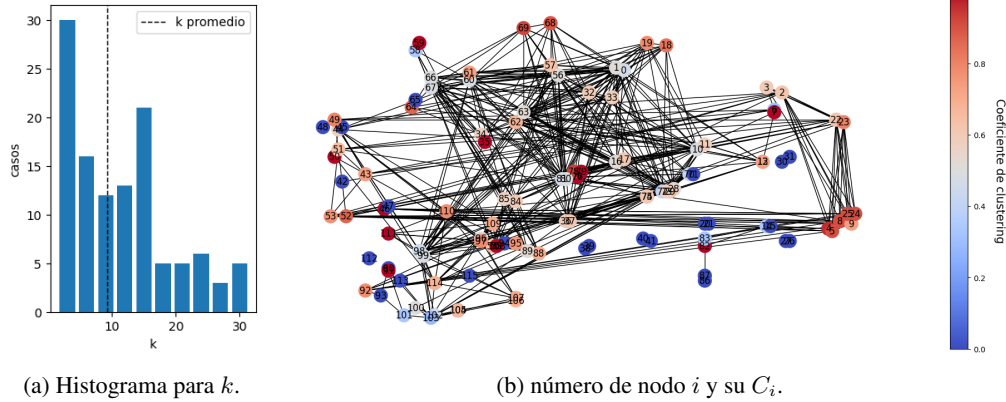


Figura 2: Distribución de grado y coeficiente de agrupamiento.

Grafos prototípicos Se generaron grafos aleatorios con la misma cantidad de nodos, $n = 116$ que el grafo de la matriz de adyacencia binaria a los fines de encontrar semejanzas y determinar cuál de estos explicaría la la estructura del grafo producto de las mediciones.

El algoritmo de Erdős–Rényi enlaza un par de nodos si una probabilidad al azar supera un umbral [5, sección 3.2]. Estableciendo en función del $\langle k \rangle$ de la red original tal probabilidad como $p = \frac{\langle k \rangle}{n-1}$ se obtuvo una densidad prácticamente idéntica $\delta = 0,0808$ a la de la red original, 0,0801, con una suma de enlaces, 539, también muy cercana, 534. El coeficiente de agrupamiento $\langle C_i \rangle = 0,7986$ resulta bastante disímil con el 0,5271 de la red original. Esta red azarosa, a diferencia de la original, mostró

estar totalmente conectada, como se aprecia en la figura 3a. En contrapartida, las redes reales suelen estar particionadas en múltiples componentes conectadas [5, sección 3.7].

Si una red fuera perfectamente regular y todos los nodos tuvieran idéntica k , la distancia mínima entre dos nodos d , es decir cuantos intermedios debe atravesarse para enlazar uno con otro, seguiría una dependencia polinomial con n . En las redes reales d presenta una dependencia mucho menor con n , con $\log(n)$ de hecho. Este fenómeno recibe el nombre de “mundo pequeño” (small-world) por la sorpresiva poca distancia entre dos cualesquiera nodos de la red. Asimismo, en las reales, el coeficiente de agrupamiento C_i suele ser más alto que en una red azarosa [5, sección 3.9]. En respuesta a estas observaciones, Watts y Strogatz extendieron el modelo ordenado con un k regular con un parámetro adicional que determina la probabilidad de que un enlace cambie su enlace hacia otro nodo, generando redes intermedias entre una regular y una azarosa. El número de enlaces total más cercano al de la red original, 580, se obtuvo con $k = \text{int}(\langle k \rangle) + 1$. Puesto que la $\delta = 0,870$ es insensible al parámetro de reconexión se lo ajustó buscando un $\langle C_i \rangle$ similar al de la red original. Este fluctúa por efectos azarosos pero se logró que este oscilara en torno al $\approx 0,527$ con una probabilidad de reconexión de 0,815. Pero la red generada se presenta totalmente conectada y con un aspecto muy diferente a la original, como muestra la figura 3b. Se pudo generar con el algoritmo de Watts-Strogatz una red con aspecto más similar incrementando la reconexión y así haciendola más azarosa y por tanto a la original, pero el resultante $\langle C_i \rangle$ se aleja del de esta.

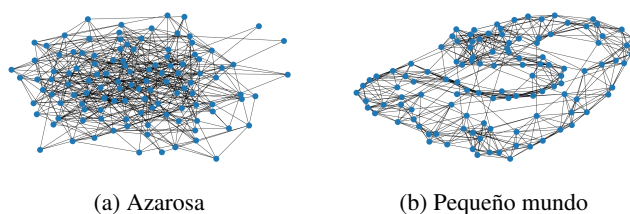


Figura 3: Grafos prototípicos con número n igual y k similar a los de la red producto de las mediciones.

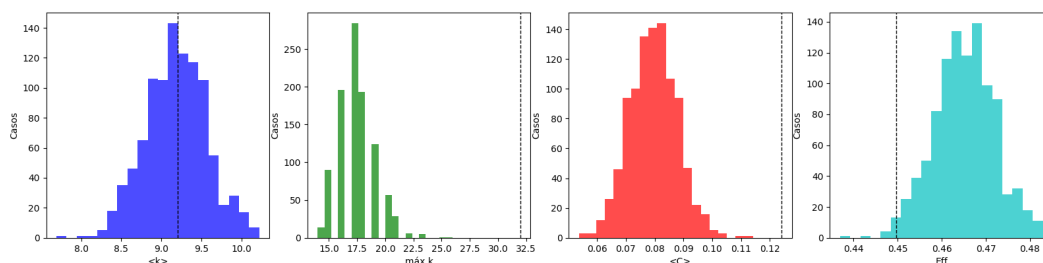


Figura 4: Distribución de magnitudes que caracterizan las mil redes azarosas generadas.

Coefficientes de grafos prototípicos

Referencias

- [1] N. Tzourio-Mazoyer et al. «Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain». En: *NeuroImage* 15.1 (ene. de 2002), págs. 273-289. URL: <https://linkinghub.elsevier.com/retrieve/pii/S1053811901909784>.
- [2] Enzo Tagliazucchi et al. «Large-scale brain functional modularity is reflected in slow electroencephalographic rhythms across the human non-rapid eye movement sleep cycle». En: *NeuroImage* 70 (abr. de 2013), págs. 327-339. URL: <https://linkinghub.elsevier.com/retrieve/pii/S1053811913000128>.
- [3] Juan A. Kamienkowski. *Curso: Data Mining en Ciencia y Tecnología*. 2023. URL: <https://datamining.dc.uba.ar/campus/course/view.php?id=37>.

- [4] Aric A Hagberg, Daniel A Schult y Pieter J Swart. «Exploring Network Structure, Dynamics, and Function using NetworkX». En: *Proceedings of the 7th Python in Science Conference (SciPy 2008)*. SciPy 2008. Pasadena, CA USA, ago. de 2008, págs. 11-15. URL: https://conference.scipy.org/proceedings/SciPy2008/paper_2/.
- [5] Albert-László Barabási. *Network Science*. Cambridge University Press, jul. de 2016. 456 págs. URL: <http://networksciencebook.com/>.