

TP2: Redes en el Cerebro

Juan E Kamienkowski, Pablo Riera
Data Mining en Ciencia y Tecnología

2 de diciembre de 2019

La aplicación del análisis de redes en general, y del análisis de grafos en particular, en registros de actividad cerebral en humanos es un campo en constante ebullición [Sporns, 2010, Barabási et al., 2016]. El primer paso del análisis es tomar una medida de similaridad entre la señal tomada de los electrodos o de las distintas regiones del cerebro. Estas medidas pueden estar enfocadas sobre una frecuencia particular o no, y suele tomarse en consideración variaciones de una sincronía de fase o a partir de las amplitudes [Sporns, 2010, Cohen, 2014].

Este trabajo está basado en el trabajo de Tagliazucchi y colaboradores (2013) que busca relacionar cambios en la modularidad de las redes construidas a partir de la señal de resonancia magnética funcional (fMRI) con los distintos estadios del sueño, detectados a partir de los ritmos de onda lenta en la señal del electroencefalograma (EEG) [Tagliazucchi et al., 2013]. Los estadios de sueño explorados no incluyen sueño REM (rapid eye movements) y fueron determinados manualmente por un experto según el criterio de la American Academy of Sleep Medicine (AASM) [Berry et al., 2012].

1. Objetivo

Este segundo Trabajo Práctico tiene tres objetivos:

- Explorar los cambios en la red en función de la profundidad del sueño.
- Explorar, en particular, los cambios asociados a la modularidad en función de la profundidad del sueño.
- Identificar los nodos en los cuales se producen estos cambios.

2. Estructura de los datos:

En la carpeta *DataSujetos* se encuentran los archivos separados por cada sujeto y estadio del sueño bajo la siguiente notación *[Estadio del sueño]_[Número de sujeto].csv*.

Además se incluyen los nombres de las 116 regiones en un archivo aparte: *aal_extended.csv*. Estas regiones están definidas a partir del atlas Automatic Anatomical Labeling (AAL) [Tzourio-Mazoyer et al., 2002].

Ejemplos de los procedimientos para comenzar el análisis pueden encontrarse en https://github.com/jkamienkowski/dmcyt_tp2/.

3. Tarea 1: Visualización

A partir de los datos individuales, visualizar la estructura de las redes pesadas a partir de los datos promedio para cada estadio del sueño (despierto - W - y $N1$, $N2$, $N3$). Transformar el grafo pesado de los datos promedio en uno no pesado manteniendo constante la densidad de aristas (δ). Sobre los datos promedio, extraer medidas de centralidad, grado, camino mínimo, y coeficiente de clustering en función de la densidad de aristas en función de δ .

Visualizar el grafo no pesado para distintos valores de δ que resulten interesantes.

4. Tarea 2: Comunidades y coeficiente de modularidad

Para los distintos valores de la densidad de aristas (δ), determinar comunidades a partir de algoritmo de Louvain [Blondel et al., 2008] y estimar el coeficiente de modularidad (Q). Graficar Q y el número de comunidades (N_C) en función de δ , y comparar los resultados con el comportamiento de una red random que preserve la distribución de grados de los nodos.

4.1. Opcional 1 (Tarea 2):

Algoritmos de detección de comunidades Comparar el resultado del algoritmo de Louvain con otros, como Girvan-Newman. Discutir las diferencias para el caso particular.

5. Tarea 3: Estadística

Repetir las curvas de modularidad (Q) y número de comunidades (N_C) para cada sujeto y estadio del sueño, en función de δ , y comparar cada estadio $N1$, $N2$ y $N3$, con el estadio despierto W . Para cada valor de δ incluir una noción de significancia entre estadios del sueño.

5.1. Opcional 2 (Tarea 3):

Corrección por comparaciones múltiples En cada caso discuta si resulta necesario corregir por comparaciones múltiples y cómo lo haría. De ser necesario, implementar una corrección adecuada.

5.2. Opcional 3 (Tarea 3):

Corrección por comparaciones múltiples Repetir el procedimiento sobre las medidas de centralidad, grado, camino mínimo, y coeficiente de clustering ¿Cómo interpretaría estos resultados a la luz de los obtenidos para la modularidad?

6. Tarea 4: Diferencias en la membresía para los diferentes estadios

Para identificar diferencias significativas globales en la membresía de los nodos entre los distintos estadios del sueño ($N1$, $N2$ y $N3$), con el estadio despierto (W) se propone seguir el procedimiento propuesto por Alexander-Bloch y colaboradores [Alexander-Bloch et al., 2012]. Se quiere saber si las comunidades encontradas en un estadio de sueño NX y las encontradas en el cerebro despierto W son similares.

1. Se toman para todos los pares de sujetos en NX el Índice de Rand ajustado (adjusted-for-chance Rand index).
2. Se toman para todos los pares de sujetos en W el Índice de Rand ajustado (adjusted-for-chance Rand index).
3. Se promedian y se obtiene el Índice de Rand ajustado "*within-group*" observado (RI_o). Se lo graficar junto al error estandar de la media.
4. Se aleatorizan las etiquetas de los parece $NX-W$ de forma que siempre esten todos los sujetos en cada grupo pero con las etiquetas cambiadas, y se repiten los pasos 1-3 de forma de obtener un Índice de Rand ajustado "*within-group*" permutado (RI_p). Este paso se repite N_p veces generando una distribución empirica de valores de valores RI_p .
5. El p-valor se calcula como la cantidad de permutaciones que dieron $RI_p > RI_o$ dividido N_p .

7. Opcional 5: Rol de nodos, y cambios en el rol de los nodos

Finalmente, se busca definir los roles que cumple cada nodo dentro de las comunidades, e identificar cuales son los nodos que, o bien están cambiando de comunidad o bien están cambiando de rol. Se deberán clasificar los nodos para cada estadio según el coeficiente de participación (P_i) y el z-score del grado intra-comunidad (z_i),

$$z_i = \frac{k_i - \langle k \rangle}{\sigma_k} \quad (1)$$

donde k_i es el grado intra-comunidad, y $\langle k \rangle$ y σ_k son el promedio y el desvío estándar del grado intra-modular.

$$P_i = \sum_j^{N_M} \left(\frac{k_i^{U_j}}{k_i} \right)^2 \quad (2)$$

donde $k_i^{U_j}$ es el número de aristas en el nodo i y la comunidad j , N_M es el número de comunidades, y k_i es el grado total del nodo i .

Esto se realizará siguiendo los criterios propuestos en Tagliazucchi y colaboradores (2013) [Tagliazucchi et al., 2013]:

- Hubs: $P_i > P_C$ & $z_i > z_C$
- Provincial Hubs: $P_i < P_C$ & $z_i > z_C$
- Provincial Nodes: $P_i < P_C$ & $z_i < z_C$
- Connector Nodes: $P_i > P_C$ & $z_i < z_C$

en principio, para los umbrales propuestos (P_C y z_C) en el mismo trabajo. Graficar el número de nodos por cada clase en función de δ , y comparar estadísticamente cada estadio $N1$, $N2$ y $N3$, con el estadio despierto W .

Para un valor de δ que resulte particularmente interesante a partir de los análisis anteriores, visualizar el grafo coloreando la membresía de cada nodo y su rol. Identificar en este grafo cuáles son los nodos que cambian entre los distintos estadios del sueño.

Opcional 4 (Tarea 5): Criterios para definir los roles Revisar los criterios definidos por Tagliazucchi y colaboradores (2013) a partir de los histogramas del coeficiente de participación (P_i) y del z-score del grado intra-comunidad (z_i).

7.1. Opcional 6:

Visualizaciones Usar el archivo *aal_extended.csv* para generar visualizaciones coloreando por regiones anatómicas y hemisferios. ¿Qué puede concluir a partir de las diferencias entre agrupamientos funcionales y anatómicos para los diferentes estadios del sueño?

8. Formato

Repetir el formato anterior incorporando los comentarios de la devolución del TP1.

9. Nota final

El TP se realizará en grupos de tres a cinco personas. El TP consiste de una serie de tareas, que pueden consistir en un análisis o contestar una pregunta. Algunas de estas preguntas o tareas están indicadas como optativas. Realizar estas tareas suma puntos pero no son obligatorias. Se puede usar cualquier herramienta de análisis o combinación de herramientas, debiendo indicarlas en el informe. El lenguaje en el que se desarrolle el TP no es excluyente.

La fecha límite de entrega es el día Domingo 15/12/2019 a las 00.00hs, para ambos turnos, vía email a juank@dc.uba.ar o pablo.riera@gmail.com.

Tabla de puntos: Cantidad máxima de puntos que se pueden obtener por ...

- ... la tarea obligatoria 1: 1.5
- ... la tarea obligatoria 2: 1.5
- ... la tarea obligatoria 3: 2.5
- ... la tarea obligatoria 4: 2.5
- ... la tarea opcional 1: 0.5
- ... la tarea opcional 2: 1.0
- ... la tarea opcional 3: 0.5

- ... la tarea opcional 4: 0.5
- ... la tarea opcional 5: 1.5
- ... la tarea opcional 6: 0.5

Puntaje máximo posible: diez

Referencias

- [Alexander-Bloch et al., 2012] Alexander-Bloch, A., Lambiotte, R., Roberts, B., Giedd, J., Gogtay, N., and Bullmore, E. (2012). The discovery of population differences in network community structure: new methods and applications to brain functional networks in schizophrenia. *Neuroimage*, 59(4):3889–3900.
- [Barabási et al., 2016] Barabási, A.-L. et al. (2016). *Network science*. Cambridge university press.
- [Berry et al., 2012] Berry, R. B., Brooks, R., Gamaldo, C. E., Harding, S. M., Marcus, C., Vaughn, B. V., et al. (2012). The aasm manual for the scoring of sleep and associated events. *Rules, Terminology and Technical Specifications, Darien, Illinois, American Academy of Sleep Medicine*, 176.
- [Blondel et al., 2008] Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10):P10008.
- [Cohen, 2014] Cohen, M. X. (2014). *Analyzing neural time series data: theory and practice*. MIT press.
- [Sporns, 2010] Sporns, O. (2010). *Networks of the Brain*. MIT press.
- [Tagliazucchi et al., 2013] Tagliazucchi, E., Von Wegner, F., Morzelewski, A., Brodbeck, V., Borisov, S., Jahnke, K., and Laufs, H. (2013). Large-scale brain functional modularity is reflected in slow electroencephalographic rhythms across the human non-rapid eye movement sleep cycle. *Neuroimage*, 70:327–339.
- [Tzourio-Mazoyer et al., 2002] Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., and Joliot, M. (2002). Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain. *Neuroimage*, 15(1):273–289.