

Modelos matemáticos para ayudar a entender las enfermedades neurodegenerativas

por

Sergio Ardanza-Trevijano, Universidad de Navarra

Introducción

Vamos a recorrer parte un trabajo interdisciplinar fruto de la colaboración entre un grupo de investigación en neurociencias del Centro de Investigación Médica Aplicada de la Universidad de Navarra y el departamento de Física y Matemática Aplicada de dicha universidad. Los resultados que se exponen han sido obtenidos por el Dr. Joaquín Goñi como parte de su tesis doctoral que fué codirigida por el Dr. Pablo Villoslada y quien escribe estas letras. Para más detalles sobre esta investigación, el lector puede consultar los trabajos ya publicados [1] y [2]. Entre los dos trabajos suman más de 11 investigadores entre los que se encuentran Médicos especialistas en neurología, psicólogos, biólogos, bioquímicos, ingenieros informáticos y matemáticos.

El grupo de enfermedades neurodegenerativas estudia enfermedades de tanto calado social como la Esclerosis Múltiple o el Alzheimer y tiene entre otros objetivos, el diagnóstico precoz y la prognosis (saber a priori como va a evolucionar un enfermo). Para esto se dispone de numerosas herramientas como los tests psicotécnicos, la imagen médica, el estudio de rutas metabólicas y vías de señalización, etc... Todas estas herramientas están llenas de oportunidades para un trabajo interdisciplinar en los que las matemáticas (y la física) tienen un papel importante (ver [3], [4], [5] para otros ejemplos de estas colaboraciones con neurociencias en los que nuestro departamento ha

estado involucrado). Hay muchos otros grupos tanto en nuestro país como en el extranjero que realizan este tipo de colaboraciones y el objeto de este texto es sobre todo animar a los matemáticos y físicos que no lo hayan hecho aún, a establecer contacto con colegas científicos de otras disciplinas, ya que la colaboración interdisciplinar es una experiencia sumamente reconfortante (aunque no exenta de dificultades).

En esta exposición nos centraremos en un tipo concreto de tests que se realizan a los enfermos: los tests de fluencia verbal.

0.1 Tests de Fluencia verbal

Aislamos una tarea simple en un sistema complejo. En nuestro caso la tarea fue:

En 90 segundos, diga todos los animales que pueda

Los test de fluencia verbal están relacionados con la memoria semántica, que es aquella parte de la memoria que acumula conocimiento de conceptos o significados. Veamos las primeras palabras de un test de fluencia verbal típico:

perro, gato, león, tigre, elefante, tiburón, ballena, delfín, ...

Es fácil apreciar en este test un doble mecanismo que utilizamos para evocar la información almacenada en nuestra memoria: por un lado la exploración de una categoría semántica determinada (que llamamos clustering) y por otro lado saltos o cambios de categoría (que llamaremos switching). Así perro-gato sería una exploración en profundidad (clustering) y elefante-tiburón sería un salto de categoría semántica (switching). En este ejemplo apreciamos los cambios de categoría semántica de animales domésticos – animales de la selva– animales marinos.

Estos dos mecanismos ya habían sido descritos por el grupo de Troyer en 1998 ([8]) y se hablaba entonces de una modulación fronto temporal en los tests de fluencia verbal donde el lóbulo frontal del cerebro se ocuparía de los cambios de categoría y el lóbulo temporal de las exploraciones en profundidad. Sin embargo este doble mecanismo no se había utilizado con fines diagnósticos y de estos tests de fluencia verbal, solo se tenía en cuenta el número de palabras citadas por cada paciente.

0.2 Nuestra muestra

Realizamos el test a 200 personas sanas. Obtuvimos una media de 31.57 ± 6.99 palabras por tests. Al realizar estos tests notamos por una parte que cada uno de los participantes lograba evocar tan solo una pequeña parte

de su conocimiento (en el campo semántico de los animales). Además era notable la variabilidad tanto en el número de palabras de cada test como en las frecuencias con que cada animal había sido mencionado en el total de los tests.

0.3 Objetivos del estudio

1. Crear un modelo de red semántica de los animales e investigar su estructura.
2. Crear un modelo de exploración de esta red.
3. Estudiar las diferencias en switching y clustering que hay entre personas sanas y personas enfermas en distintas etapas de la enfermedad.

Los dos primeros objetivos están relacionadas con una mejor comprensión del mecanismo de evocación (switching-clustering) hasta el punto de replicarlo con un modelo matemático adecuado, y el tercero está relacionado con las aplicaciones clínicas en diagnóstico y prognosis de la enfermedad.

1. De los test de fluencia verbal a una red de conceptos

Para entender mejor los mecanismos de evocación necesitamos crear un modelo de la red semántica de los animales. Queremos obtener un grafo que re-presente dicha red semántica. Será un grafo no dirigido, donde los vértices serán los distintos animales y los enlaces entre ellos las posibilidades de exploraciones en profundidad. Utilizaremos el vocabulario usual en sistemas complejos donde se denomina nodo de una red, al vértice de un grafo.

1.1 Inferencia de la red de conceptos

Suponer que la palabra perro ha sido dicha en tercera posición en un test y la palabra ratón ha sido dicha en la quinta posición en un test. Entonces decimos que las palabras perro y ratón están a una distancia 2 en este test (5-3). En general si tenemos dos palabras en un mismo test, una ocupando la posición i y otra la posición j (con $i < j$) diremos que estas palabras están a distancia $j - i$, y diremos que dos palabras que aparecen simultáneamente en una ventana de tamaño k si están a distancia $\leq k$. Un método utilizado frecuentemente para inferir una red de conceptos en situaciones similares consiste en añadir enlaces entre aquellas palabras que aparecen en una ventana pequeña en un mismo test. Sin embargo con este método no se explica suficientemente el mecanismo de cambio de categoría semántica (switching) que se observa en cada test. Podríamos refinar este método borrando “a mano” los

enlaces que consideremos saltos en categorías semánticas, sin embargo queremos un método de inferencia no supervisado. Conseguiremos este algoritmo no supervisado introduciendo un test de hipótesis.

Sean w_1, \dots, w_n todas las palabras distintas mencionadas en el conjunto de los test de fluencia verbal realizados. Llamamos f_{w_i} al número de veces que la palabra w_i aparece en el conjunto de todos los tests y M el número total de tests (en nuestro caso 200). La probabilidad de que un test contenga a la palabra w_i se puede estimar por:

$$\hat{P}_{w_i} = \frac{f_{w_i}}{M}$$

Por otra parte, la probabilidad de que dos palabras aparezcan en el mismo test, si consideramos ambos sucesos independientes, es:

$$P_{w_i, w_j}^{test} = \hat{P}_{w_i} \hat{P}_{w_j} = \frac{f_{w_i}}{M} \frac{f_{w_j}}{M},$$

Dadas dos palabras del mismo test, aproximamos la probabilidad de que estén a una distancia l por

$$2 \frac{N-l}{\binom{N}{2}} = 2 \frac{N-l}{N(N-1)}, \quad 1 \leq l < N.$$

donde N es la longitud media de los tests. Notar que estamos haciendo un simplificación en nuestro modelo al asumir en este cálculo de probabilidad teórica que los tests tienen la misma longitud (la media N de las longitudes de los tests).

Dadas dos palabras del mismo test, la probabilidad de que estén en una ventana de tamaño l es

$$P_{w_i, w_j}^{lventana} = 2 \sum_{i=1}^l \frac{N-i}{\binom{N}{2}} = \frac{2}{N(N-1)} (lN - \frac{l(l+1)}{2}), \quad 1 \leq l < N.$$

Entonces, podemos aproximar la probabilidad de que dos palabras ocurran en el mismo test y ventana, P_{w_i, w_j}^{juntas} , por

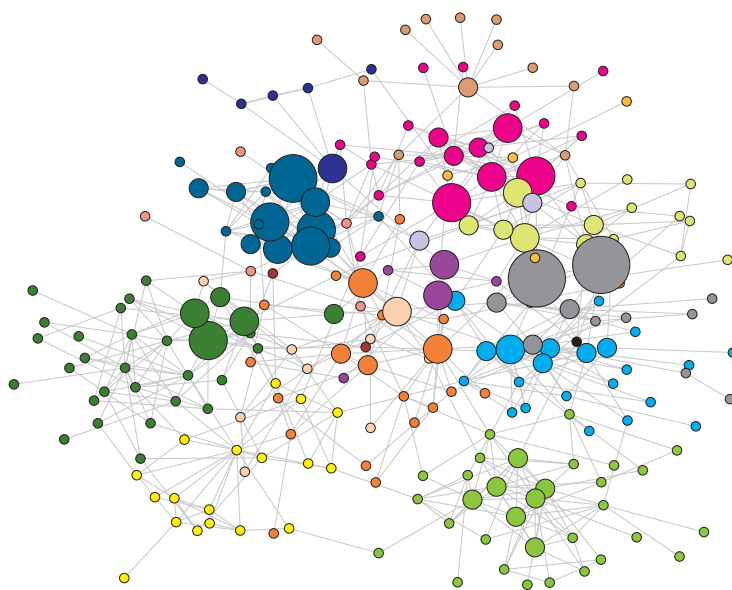
$$P_{w_i, w_j}^{juntas} = P_{w_i, w_j}^{test} P_{w_i, w_j}^{lventana} = \frac{f_{w_i}}{M} \frac{f_{w_j}}{M} \frac{2}{N(N-1)} (lN - \frac{l(l+1)}{2}).$$

Para cada par de palabras, calculamos el intervalo de confianza al 95% de confianza, para el parámetro p de una distribución binomial donde el número de éxitos es el número de co-ocurrencias en una ventana de tamaño ($l = 2$) y el número de intentos es el número de tests ($M = 200$) y lo

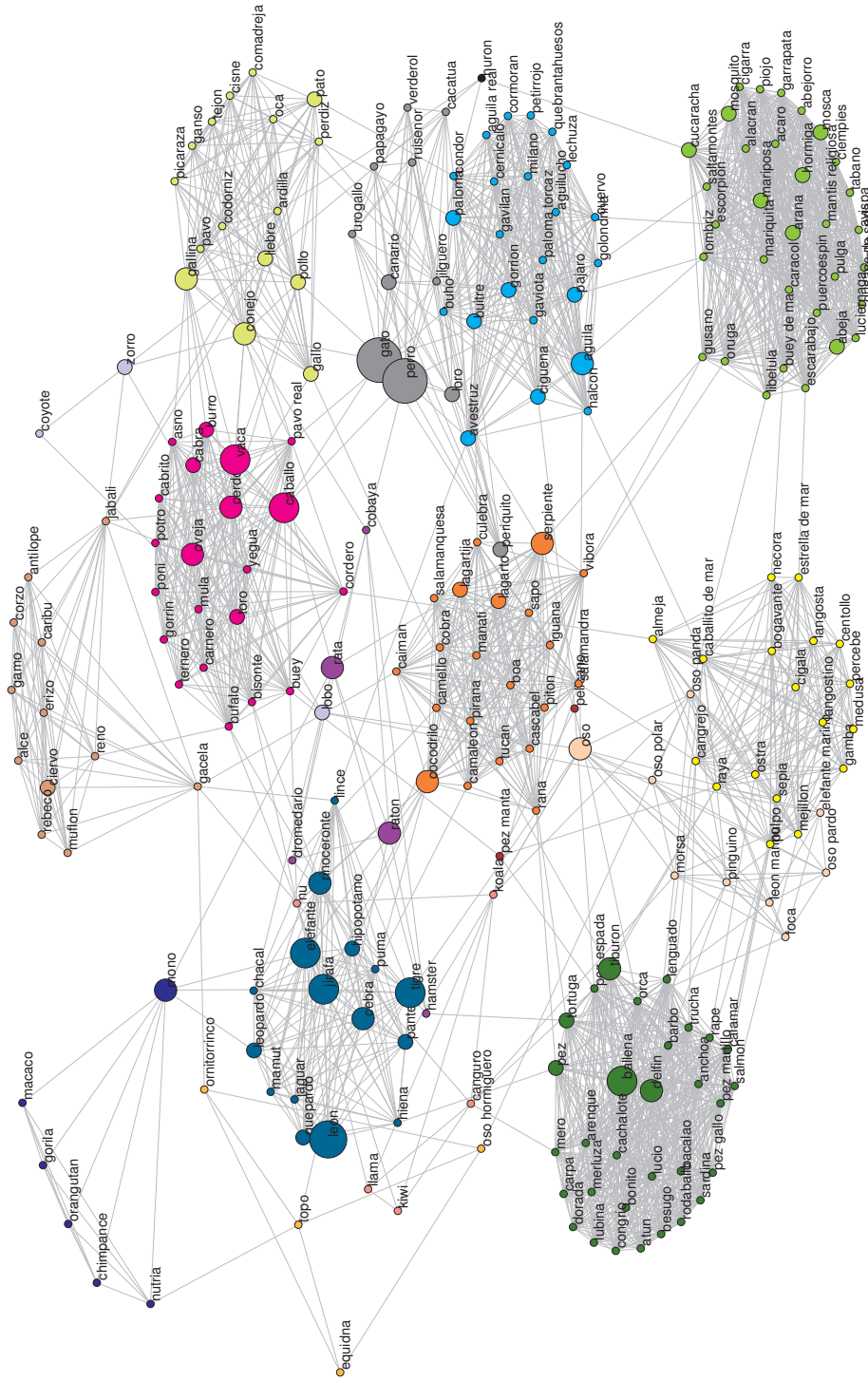
donde $J(i, j)$ es el número de vecinos compartidos por los nodos i y j , $\min(k_i, k_j)$ es el mínimo de los grados de k_i y k_j .

Yip y sus colaboradores obtuvieron una generalización GTOM k , que expande la lista de los vecinos de cada nodo antes de obtener la matriz de superposición. Así GTOM1=TOM, mientras que en GTOM2, añadiría a $J(i, j)$ el número de vecinos de vecinos que comparten los nodos i y j y en general para GTOM k definiría $J(i, j)$ como el número de vecinos a distancia $\leq k$ que comparten los nodos i y j . Notar que en cualquier caso $0 \leq \text{GTOM}k \leq 1$.

Para nuestro experimento resultó más adecuado tomar $k = 2$, en un test de fluencia verbal es raro que un individuo haga clustering muy largos, la mayor parte tienen dos o tres individuos. Una vez obtenida la matriz GTOM2, se divide el conjunto de nodos en módulos usando un agrupamiento jerárquico, donde la disimilaridad entre nodos viene dada por $d(i, j) = 1 - \text{GTOM}2(i, j)$. Obtenemos la siguiente red, donde distinguimos los módulos por colores.



Para terminar nuestra inferencia automática vamos a hacer un último paso motivado porque algunas transiciones de clustering no aparecen como tales en la red de conceptos obtenida posiblemente por no disponer de suficientes tests. Por ejemplo tenemos los enlaces paloma – gorrión – cuervo. En esta situación esperaríamos tener también un enlace entre paloma y cuervo que sin embargo no aparece.



Una manera de solventar este problema, siempre de modo no supervisado consiste en enriquecer la red añadiendo todos los enlaces que falten entre animales del mismo módulo, convirtiendo así cada módulo en un clique (una red con todos los nodos conectados entre sí). Así finalizamos el proceso y obtenemos nuestra red de conceptos enriquecida. Recomendamos al lector explorar alguno de los módulos que allí aparecen, recordando que ha sido obtenida de manera no supervisada.

Recordamos que los colores de los nodos indican los distintos módulos y el tamaño la frecuencia de la palabra en los 200 tests.

Si examinamos los módulos obtenidos de esta manera no supervisada, observamos que salvo tres excepciones, todos los animales han quedado clasificados en grupos reconocibles.

Descripción	tamaño	% tests en los que aparece	Más frec.
Domésticos	11	95	perro
Sabana y felinos	16	93	león
Aves salvajes	23	86	águila
Granja-pequeño	16	85	gallina
Peces	31	84	ballena
Granja-grande	21	83	caballo
Reptiles	21	80	culebra
Insectos	32	69	mosca
Roedores	5	55	ratón
Osos y Polares	9	47	oso
Simios	6	41	mono
Crustáceos y moluscos	18	39	cangrejo
Cervátidos	12	35	ciervo
Canidos salv.	3	27	lobo
Australianos	5	26	canguro
Ovomamíferos	4	17	ornitorrinco
Sin clasificar	2	09	manta raya
Sin clasificar	1	05	verderón

Con este proceso de enriquecimiento termina la inferencia no supervisada de la red a partir de los tests. Recordamos el proceso en tres pasos que hemos hecho:

1. A partir de 200 tests creamos una red con nodos las palabras distintas dichas en los tests y donde ponemos un enlace según el resultado de un test de hipótesis acerca de la coocurrencia de dos palabras en una misma ventana de tamaño 2.

2. Usando GTOM2 y agrupamiento jerárquico hemos detectado distintos módulos en la red.
3. Hemos enriquecido la red original haciendo que cada módulo tenga enlaces conectando todos sus nodos entre sí.

1.2 Evaluadores *in-silico* de cambios de categoría semántica y exploración en profundidad

Veamos como nuestra red enriquecida nos sirve para evaluar en cada test cuando ocurren los cambios de categoría y cuando hay exploraciones en profundidad así si decimos dos palabras consecutivas y hay un enlace entre ellas en la red diremos que hay una exploración en profundidad y si no hay enlace que hemos hecho un salto. Denotamos CN a la red de conceptos antes del enriquecimiento, y a ECN a la red de conceptos después del enriquecimiento.

Ahora clasificamos el paso de una palabra a otra en un test como exploración en profundidad (clustering) cuando las dos palabras estan en el mismo módulo de la red enriquecida y como cambio de categoría semántica (switching), en caso contrario. Para comprobar la bondad de nuestra clasificación pedimos a dos jueces que clasifiquen los cambios de palabras como switching o clustering y evaluamos la concordancia entre distintos clasificadores utilizando el coeficiente Kappa, un estadístico que toma valores entre -1 y 1 usado frecuentemente para medir concordancias entre variables cualitativas. Se obtuvieron unos resultados sorprendentes ya que la concordancia entre nuestra clasificación y la de los jueces era casi tan buena como la que obtuvieron los jueces entre sí (un coeficiente Kappa de 0.82 y 0.83 con cada uno de los jueces frente a 0.88 que tuvieron los jueces entre sí. Así se ha logrado el primer objetivo, inferir de manera no supervisada una red de conceptos con la que se pueda explicar el mecanismo exploración en profundidad–cambio de categoría semántica, en los resultados de un test de fluencia verbal.

2. Exploración automática de una red. Caminos aleatorios con saltos

En el test de fluencia verbal que nos ocupa en este trabajo, intentamos evocar los animales que tenemos en la memoria. Ya que hemos construido un modelo de la organización de los conceptos en nuestra memoria como una red, podemos hacer una analogía del proceso de evocación con el de explorar esta red. Una primera aproximación es un camino aleatorio. Este proceso consiste en recorrer la red pasando de un nodo a cualquiera de sus vecinos que seleccionemos al azar. Como veremos podemos hacer un modelo de nuestra exploración en profundidad (clustering) usando caminos aleatorios, pero ¿qué

aporta el cambio de categoría semántica (switching)? Vamos a comprobar que el switching hace más eficaz la exploración de una red. Una manera de medir la eficacia en la exploración de una red, consiste en calcular el tiempo medio que se tarda en llegar a un nodo por primera vez. Veamos como un modelo de exploración de redes basado en la dualidad switching-clustering descrita anteriormente optimiza este tiempo medio.

2.1 Exploración de la red como una cadena de Markov

Un proceso de Markov finito está determinado por un conjunto de estados $\{s_1, \dots, s_n\}$ y una matriz de transición P de tal modo que p_{ij} nos da la probabilidad de ir del estado i al estado j . Así en cada paso del tiempo, para calcular las probabilidades de a que estado puedo ir, solo importa el estado en el que estoy y no por qué estados haya podido pasar. Notar que las potencias P^k de la matriz de transición nos dan las probabilidades de ir de un nodo a otro en k pasos. Diremos que el proceso es ergódico si se puede ir de cualquier nodo a cualquier nodo en un número finito de pasos y que es regular, si además existe un N tal que P^k tiene todas sus entradas no nulas para $k \geq N$.

Recorrer una red consiste en ir pasando de un nodo a otro nodo a lo largo del tiempo. Vamos a hacer un modelo basado en switching y clustering que será una cadena de Markov finita regular, con estados los nodos de nuestra red. Necesitamos definir entonces la matriz de transición.

Proponemos tres modelos alternativos de switching. Uno en el que la probabilidad de ir a cada nodo es la misma y otros que favorecen o penalizan el ir a nodos con muchas conexiones (grado alto).

Para definir matriz de probabilidades de transición P^{sw} , denotamos por $|V|$ el número de nodos de nuestra red, por K_{max} al máximo grado entre todos los nodos y por SK a la suma de los grados de todos los nodos. Así la matriz P^{sw} tiene todas las filas iguales y viene dada por:

$$p_{ij}^{sw} = \lambda_j$$

donde λ_j toma, según el modelo elegido, uno de estos valores:

- $\frac{1}{|V|}$ cada nodo tiene la misma probabilidad
- $\frac{k(v_j)}{SK}$ gradiente positivo por grado.
- $\frac{K_{max} - k(v_j) + 1}{SK}$ gradiente negativo por grado.

Notar que en los tres casos se cumple la condición de que las filas de la matriz de transición suman 1. Así hemos obtenido tres posibles modelos para el cambio de categoría semántica (switching) como una cadena de Markov.

Para la exploración en profundidad o clustering nos conformamos con un camino aleatorio que puede ser descrito también como una cadena de Markov con matriz de transición P^{cl} definida por:

$$p_{ij}^{cl} = \frac{a_{ij}}{k(v_i)}$$

donde $A = [a_{ij}]$ es matriz de adyacencia del grafo ($a_{ij} = 1$ si i y j son vecinos, y $a_{ij} = 0$ si no lo son) y $k(v_i)$ es el grado del vértice i .

Vamos ahora a combinar los dos procesos para crear nuestro explorador que llamaremos camino aleatorio con Switching o con la abreviatura SWR. Para conseguir que siga siendo un proceso de Markov simplemente consideramos una combinación lineal convexa de probabilidades de transición (sw y cl)

$$P^{srw} = q P^{sw} + (1 - q) P^{cl} \quad 0 \leq q \leq 1,$$

donde las probabilidades de switching y clustering son q y $1 - q$ respectivamente.

El tiempo de primer paso medio mide el número medio de pasos necesarios para ir del nodo i al nodo j en un proceso de Markov y lo presentamos como una matriz $\langle T \rangle = [\langle t_{ij} \rangle]$.

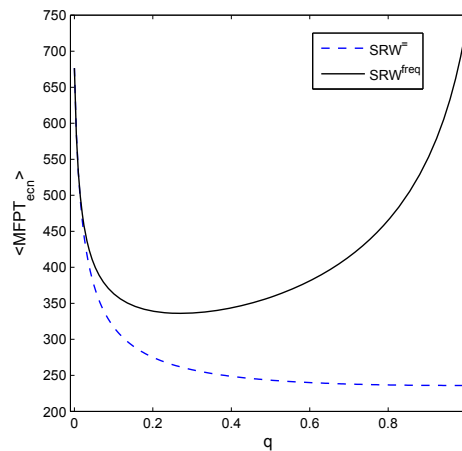
Esta matriz se puede calcular analíticamente cuando la matriz de transición de probabilidades (P^{srw} en nuestro caso) es *ergódica* y *regular* y el grafo es *finito* (ver la excelente exposición de Laurie Snell en [7] para más detalles sobre las cadenas de Markov finitas y como calcular este resultado analíticamente usando un poco de álgebra lineal). Notar que la condición de regularidad necesita que la matriz de transición tenga todas sus entradas no nulas, pero el switching nos asegura esta condición.

Denotamos MFPT de un grafo al promedio de los tiempo de primer paso medio $\langle t_{ij} \rangle$ para todos los $i \neq j$ (notar que T no va a ser simétrica en general). Vamos a evaluar, utilizando el parámetro MFPT, la eficiencia de nuestros SRW (en sus tres tipos según consideramos en la matriz que representa el switching P^{sw} probabilidades iguales para todos los nodos (en cuyo caso denotaremos SWR=), gradiente positivo por grado (SWR+) o gradiente negativo por grado (SWR-).

Realizamos la evaluación en cuatro tipos de redes distintas que comparten las características de tener el mismo número de nodos (500) y de enlaces (2000), aunque difieren en la manera de conectar los nodos.

1. Red libre de escala (scale-free). Esta red se puede generar con el siguiente procedimiento: construimos un anillo con todos los nodos de tal forma que cada nodo esté conectado con otros 8 nodos. Una vez construido este anillo, fijamos una probabilidad pequeña y reconectamos los nodos al azar con esa probabilidad.
2. Un grafo de Erdős-Rényi. Dado un conjunto de 500 nodos, escogemos al azar y sin repetición 2000 pares de nodos distintos que serán nuestros enlaces.
3. Red modular aleatoria (random modular). Construimos 10 componentes de 50 nodos y 200 enlaces entre ellos elegidos al azar (como en el modelo de Erdős-Rényi). Una vez construidas las componentes, fijamos una probabilidad pequeña y reconectamos los nodos al azar con esa probabilidad.
4. Red libre de escala con conexión preferencial (Preferential Attachment). A partir de una semilla formada por nueve nodos conectados, se conecta cada nodo nuevo con 8 de los nodos existentes con probabilidad proporcional al número de enlaces que tenga cada nodo.

y medimos la eficiencia de la exploración de los tres tipos de SRW en cada una de las cuatro redes mediante el MFPT según los valores del parámetro q que nos da una medida de la frecuencia del switching o cambio de categoría semántica. Se observó (ver figura 3 en [1]) que con $SRW=$ y $SWR+$, aumentar switching mejora la exploración de la red para todos los tipos de redes (disminuye el MFPT). En el caso de una red modular y $SWR-$ hay un valor de q que nos da un mínimo para el MFPT. Es decir hay un parámetro de frecuencia de switching óptimo.



La red de conceptos enriquecida (ECN) que hemos construido es una red modular y también encontramos que nuestro explorador aleatorio con cambios de categoría semántica tiene también un valor óptimo de switching para la exploración de la red, si consideramos la matriz de switching P^{sw} con un gradiente que favorece hacer switching a los animales más frecuentes.

Este procedimiento de exploración eficaz de una red usando una combinación lineal convexa de procesos de Markov es similar al utilizado por Google al hacer su PageRank aunque allí el objetivo sea crear un ranking de las páginas web “interesantes” para la persona que busca. Este algoritmo de Google se basa en el comportamiento de una persona surfeando la red, que resulta ser similar al de nuestro explorador aleatorio con cambios de categoría semántica, ya que la persona va siguiendo enlaces de páginas hasta que se cansa y va a una página no enlazada. En este caso la web es un grafo dirigido, pero también hay una exploración local (siguiendo enlaces) y saltos (introducir una nueva dirección).

3. Problemas de acceso al léxico en enfermedades neurodegenerativas

Nos gustaría poder utilizar el conocimiento adquirido para mejorar la diagnosis de diversas enfermedades neurodegenerativas. Tradicionalmente se ha utilizado la longitud de los tests como único parámetro de estudio. Sin embargo, comparando la longitud de los tests de los pacientes no se aprecian diferencias significativas entre los individuos sanos y aquellos enfermos sin deterioro cognitivo que son difíciles de diagnosticar (aunque sí son patentes las diferencias entre ambos grupos citados y los enfermos con deterioro cognitivo). Sin embargo estudiando las distribuciones de switching y clustering si se han detectado diferencias significativas entre enfermos sin deterioro cognitivo aparente y controles. Se aprecia que los individuos enfermos sin deterioro cognitivo aparente hacen más switching y menos clustering que los controles. Esto sugiere un mecanismo de compensación ante la dificultad que pudieran encontrar estos enfermos con las exploraciones en profundidad.

Referencias bibliográficas

- [1] Goñi et al: *Switcher-Random-Wals: A cognitive-Inspired Mechanism for network Exploration*, IJBC 30:3, 913-922, 2010.
- [2] Goñi et al: *The semantic organization of the animal category: evidence from semantic verbal fluency and network theory*, Cogn. Process 12, 183-196, 2011.
- [3] Goñi et al: *A computational analysis of the protein-protein interaction*

networks in neurodegenerative diseases, BMC Systems Biology.

[4] Esteban F.J. et al: *Fractal dimension analysis of grey matter in multiple sclerosis*, Journal of the Neurological Sciences 282, 67-71, 2009.

[5] Martínez-Forero, I. et al: *Steady State Detection of Chemical Reaction Networks Using a Simplified Analytical Method*, Plos One, 5(6): e10823, 2010.

[6] Ravasz et al: *Hierarchical organization of modularity metabolic networks*, Science 297 (5586): 1551-1555.

[7] Snell, J.L: *Finite markov chains and its applications*, American Mathematical Monthly 66, 99-104, 1956.

[8] Troyer et al: *Clustering and switching on verbal fluency: the efect of focal frontal and temporal lobe sessions*, Neuropsychologia 36(6): 499-504.

[9] Yip A, Horvath S: *Gene network interconnectedness and the generalized topological overlap measure*, BMC Bioinformatics 8(22), 2007.

Sergio Ardanza-Trevijano

Universidad de Navarra

Departamento de Física y Matemática Aplicada

Edificio de los Castaños

C/Irunlarrea s/n

31080 Pamplona

e-mail: sardanza@unav.es

