



Análisis exploratorio de datos

Profesor: Miguel Serrano Reyes

Proyecto final: Análisis de
redes de dos conetomas
distintos

Grupo: 04

Zuleta Ceja Juan Manuel
No. de cuenta 315227969

SEMESTRE 2020-2
24/JUNIO/2020

ÍNDICE

Introducción	2
Investigación previa	2
Análisis del conectoma del gato	3
Análisis del conectoma del macaco	12
Referencias	18

Introducción

El presente proyecto trata del procesamiento y análisis de datos que se le da un par de conectomas de dos animales distintos: un gato y un macaco.

El desarrollo se muestra a la par de las conclusiones obtenidas, sustentadas en la teoría de grafos como en una breve investigación de las estructuras cerebrales sobre las cuales se hizo el estudio. Se debe aclarar que la inmersión en los artículos de los que se extrajo los datos de los conectomas no fue posible debido al poco conocimiento que se tiene en el área. Sin embargo, basándose en lo comentado más arriba, se hicieron especulaciones sobre la interpretación que se le pudo dar a la información dispuesta y tratada de manera cualitativa y cuantitativa.

Investigación previa

Tálamo

Estructura cerebral que tiene una ubicación central en el cerebro, se caracteriza por su simetría bilateral, pero sus mitades anterior y posterior difieren. Anatómicamente, se habla de núcleos laterales o ventrales, anteriores o posteriores y de núcleos pulvinares, que están ubicados en la zona lateral posterior y los núcleos geniculados, ubicados en el extremo posterior.¹ La función principal del tálamo es retransmitir señales sensoriales y motoras a la corteza cerebral, se le reconoce como la estructura que integra la información proveniente del sistema nervioso periférico. Además, vincula la información que procesa a otras estructuras permitiendo vinculaciones, por ejemplo, entre percepción y emoción.²

En investigaciones recientes se ha visto que una parte fundamental del tálamo, es que determina a cual estímulo sensorial debemos de enfocar la atención.¹

De manera sencilla, la manera en que los estímulos externos detectados por la vista, oído y sistema somato-sensorial, incluyendo la piel, articulaciones y músculo; con la excepción de los estímulos del olfato, entran al tálamo por medio de un circuito, llegando al tálamo, sirve su función de integración de la información y es necesario este paso antes de llegar a los centros de procesamiento que se encuentra en la **corteza cerebral**.³

La manera en la que el tálamo se estructura es por medio de núcleos, estos núcleos son un conjunto de cuerpos neuronales, es decir, sustancia gris que se dividen en núcleos de conexión específica, que son los que mandan la información a zonas concretas de la corteza; núcleos de conexión inespecífica, que como su nombre lo dice, lo mandan a zonas más amplias de la corteza sin discriminar y por último los núcleos de asociación, que comunican la corteza con estructuras subcorticales.⁴

Siendo su principal función la de conectar los estímulos externos con la corteza cerebral, no se reduce a ser un simple puente de comunicación. El tálamo regula las funciones de la corteza asociativa y es importante en funciones como el lenguaje, el habla y las funciones cognitivas, medidas por la corteza.⁴ Existen tres regiones importantes de la corteza: parieto-temporo-occipital, prefrontal y la límbica, a estas zonas son a donde son proyectados los núcleos talámicos.¹

Corteza cerebral

En los grandes mamíferos, la corteza cerebral se pliega, proporcionando una mayor área superficial en el volumen confinado del cráneo. Además de minimizar el cerebro y el volumen craneal, plegable cortical es crucial para el cableado del cerebro y su organización funcional. En los mamíferos con un pequeño cerebro no hay plegado y la corteza es suave.¹

La corteza cerebral es una fina capa de neuronas. Se encarga de asociar la información que le es demandada, principalmente por el tálamo, con el fin de generar patrones de reconocimiento y conducta. Referente a la motricidad, la corteza se encarga también de controlar los movimientos del cuerpo.

Muchas áreas de la corteza cerebral se especializan en funciones concretas. Algunas áreas, denominadas corteza primaria, presentan conexiones directas con la médula debido a que llevan la función de movimiento. Otras regiones primarias reciben las aferencias sensitivas de los núcleos talámicos (excepto el olfato).

- La corteza parieto-temporo-occipital (siendo las áreas 39 y 40 respectivamente en el cerebro) se relacionan con funciones perceptivas, la visión y la lectura y recibe la información de los núcleos pulvinares.⁴
- La corteza asociativa prefrontal es importante para la planificación, la conducta y los movimientos, la cognición, el aprendizaje, la memoria y el pensamiento.⁴
- La corteza límbica, relacionada con el aprendizaje, la memoria y la emoción, recibe eferencias del núcleo anterior talámico.⁴

Análisis del conectoma del gato

CARACTERIZACIÓN DE RED

En primera instancia, al ser su matriz adyacente ponderada y dirigida, se reconfiguró a no dirigida y binaria, sobre la que se pudo obtener la gráfica mostrada en la figura 1.

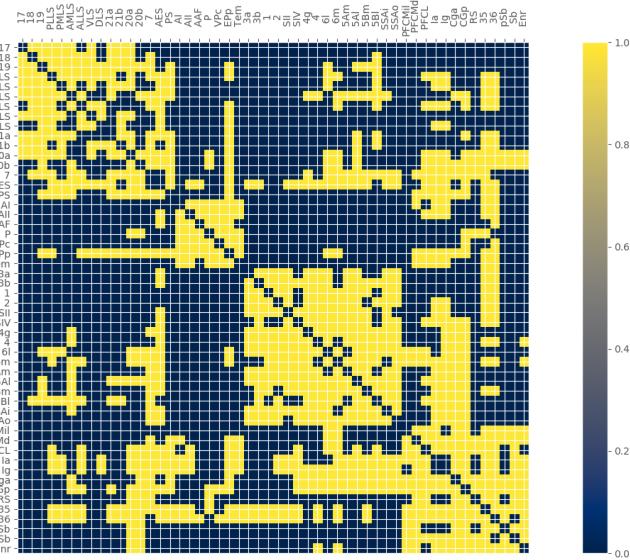


Figura 1. Matriz adyacente del conectoma del gato

A primera de vista se pueden apreciar regiones densas en conectividad, lo que se podría atribuir a la relación de las estructuras cerebrales estudiadas, pues el tálamo se conoce por integrar información proveniente del sistema nervioso periférico para después comunicar dicha información pre-procesada a la corteza, lo que daría pie a pensar que deben existir varias conexiones de una zona con las demás y viceversa dada la tarea de traducir las señales en bruto a unas manipulables para la corteza.

Para una mayor apreciación en cuanto a las regiones con mayor número de conexiones, se obtuvo la distribución de grado, mostrada en la figura 2, en la que se ve que las regiones como la 35 y 36, según la investigación original, están relacionadas con las estaciones más elevadas del sistema auditivo⁵ además de estar relacionadas con la PS (Posterior Suprasylvian area) relacionada con corteza visual del gato⁶.

Además, se tiene que el mínimo de conexiones es de 7 y el máximo de 37, con una media de 19.807 conexiones siendo esto, un posible indicador de la mucha actividad que se tiene en la comunicación que hay entre el tálamo y la corteza.

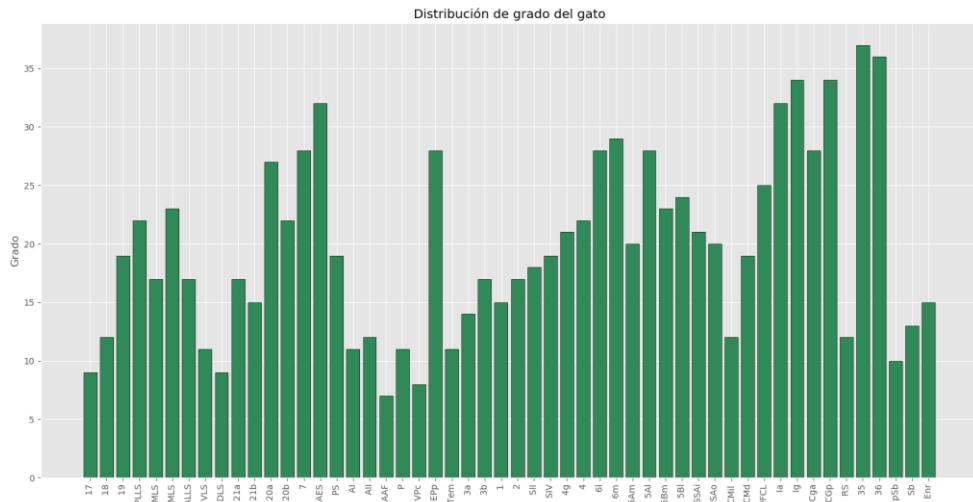


Figura 2. Distribución de grado para las regiones estudiadas en el cerebro del gato

En adición a la información procesada se obtuvo la matriz de distancias, cuya visualización permite conocer que tan próximas se encuentran las zonas de las diversas zonas relacionadas. En la barra de color de la matriz de distancias, mostrada en la figura 3 se señalan las magnitudes de las distancias, las cuales son relativamente cortas, pues obteniendo el promedio de éstas (directamente de Gephi) siendo de 1.85, implica que de una de una región se deben pasar casi 2 nodos para llegar a otra.

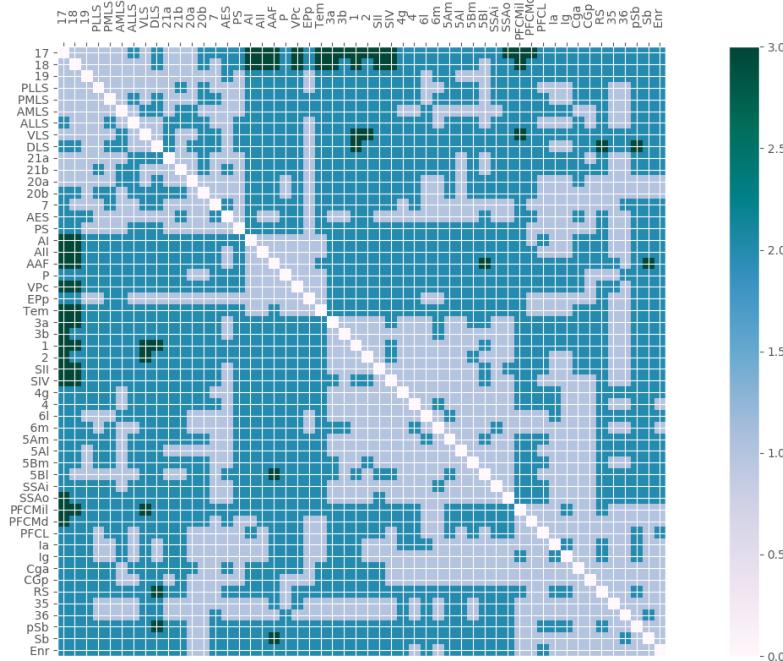


Figura 3. Matriz de distancias entre regiones del conectoma del gato.

En añadidura a la visualización cualitativa de los datos, se creó una red del grafo, mostrada a continuación.

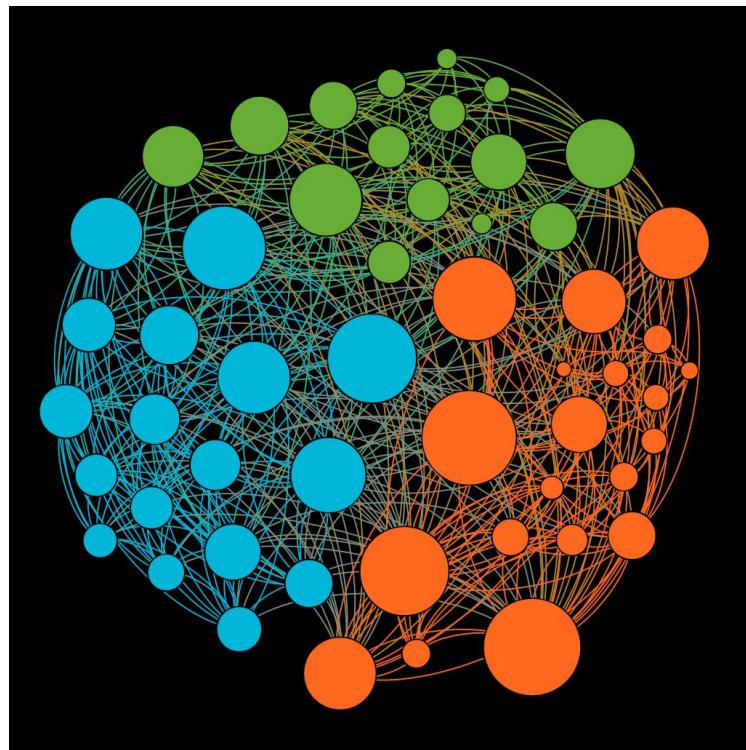


Figura 4. Red del grafo del conectoma del gato.

De la figura 4, se observa la existencia de 3 comunidades, determinadas por las clases de modularidad, lo que quiere decir que esas regiones están muy relacionadas para las actividades que colectivamente hacen en el procesamiento y transferencia de información. Además, el tamaño que tiene cada nodo representa el grado, haciendo evidente así las conexiones que guardan respecto a su comunidad y las otras. Cabe recalcar que ningún nodo es de naturaleza solitaria.

Así mismo, se calcularon los parámetros de red siguientes:

- *Densidad = 0.388*: esto indica que del 100% de las conexiones que deberían estar para que el grafo fuera completo, sólo hay un 38.8%, lo cual en realidad está bien, pues si estuvieran todas no habría una división en el trabajo que realiza cada una de las regiones observadas.
- *Eficiencia global=0.69*: señala que la manera en la que fluye información en toda la red se pasa sin pérdidas o retrasos es de un 69% lo cual puede señalar imperfecciones poco relevantes pero existentes o bien el estado en el que se encuentra el gato sólo permite dicho porcentaje.

- Coeficiente de agrupamiento global=0.6641: la red tiende a agruparse en un 66.41%, señalando también la proximidad de las conexiones.

De este último parámetro, se muestra a continuación la distribución del coeficiente de agrupamiento local, de donde se puede percibir específicamente que tanto los nodos de cada región están de cerca o agrupados con sus nodos vecinos.

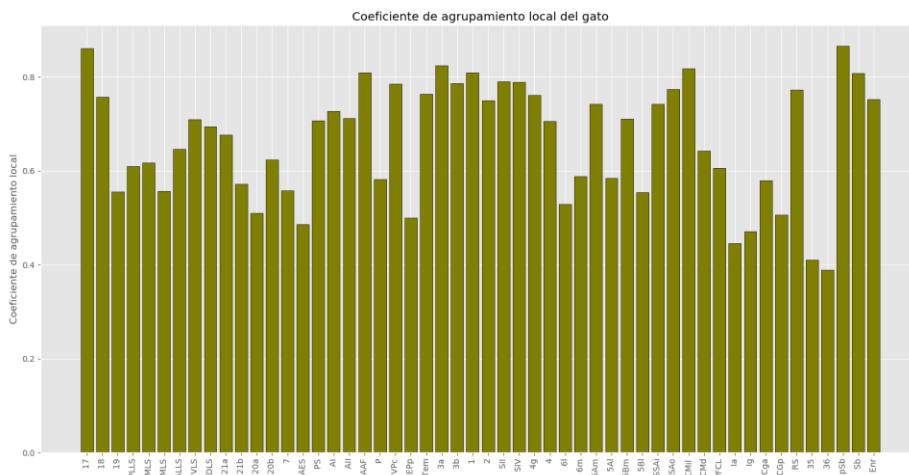


Figura 5. Distribución del coeficiente de agrupación por región en el conectoma del gato.

Esta figura permite visualizar cómo es que el coeficiente de agrupamiento global toma el valor de 66.41 %, indicando con ello que en promedio, más de la mitad de las regiones se encuentran condensadas en sí mismas.

AGRUMPAMIETO POR MEAN SHIFT⁷

A pesar de que en la figura 1 se muestra un agrupamiento de la red, se buscó otro método para ejecutarlo mediante código. El algoritmo seleccionado es el llamado **Mean Shift** que se caracteriza por ser no paramétrico que calcula, según la información que se le provea para los ejes de la gráfica, el número de grupos óptimo para ser mostrado.

En la figura siguiente se muestra el resultado del procesamiento del algoritmo, que consideró como 4 los grupos a formar, marcando con una estrella roja el centroide de cada uno.

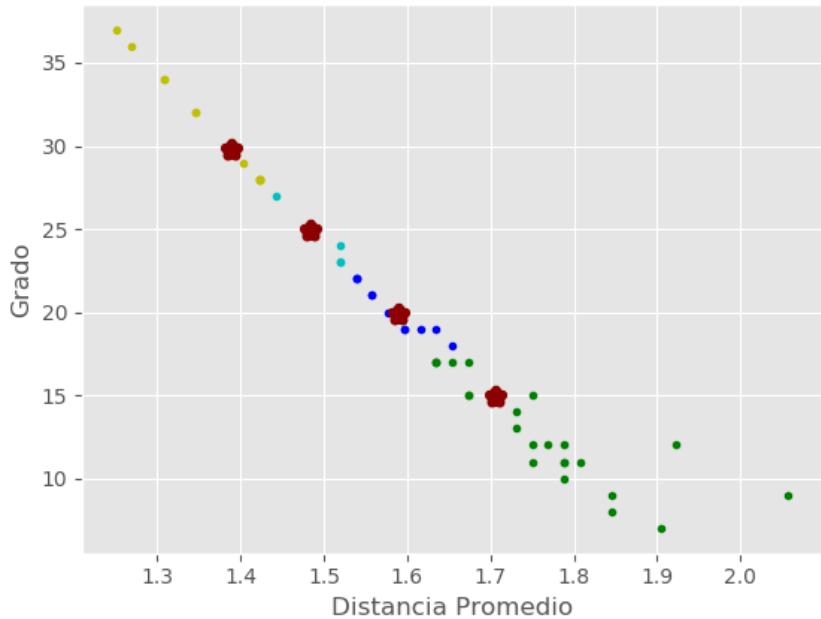


Figura 6. Agrupamiento por Mean Shift

Del gráfico mostrado se observa que abundan las regiones que tienen conexiones con una distancia relativamente lejana con un grado que va desde las 7 conexiones hasta un margen entre 15 y 20. Después del primer, lo demás cuentas con pocas conexiones siendo las de menor cantidad las que son cercanas y de grado alto (mayor a 25 conexiones). Todas estas observaciones se relacionan con que los centros de procesamiento robustos no se encuentran tan alejados como los que parecieran ser intermedios, sin embargo, esto es especulativo.

Es importante mencionar que la diferencia entre los grupos hallados en la figura 4 y los de la figura 6, varían debido a las variaciones que tienen los respectivos algoritmos, aunque nos es útil comparar sus gráficos para ubicar las regiones más conectadas y la cercanía entre ellas.

INTERVALOS DE CONFIANZA

Posteriormente, se elaboró el establecimiento de los intervalos de confianza referidos a la media de simulaciones tanto completamente aleatorias como aleatorias conservando el grafo del conectoma. Para el primer caso se hicieron 1000 simulaciones a las cuales se calculó su media, para después visualizar la gráfica mostrada en la figura 7. Es esta figura está la distribución de medias y el intervalo de confianza que se generó a partir de la suma y resta de dos veces la desviación estándar de la media, que se grafican, envolviendo a la primera curva.

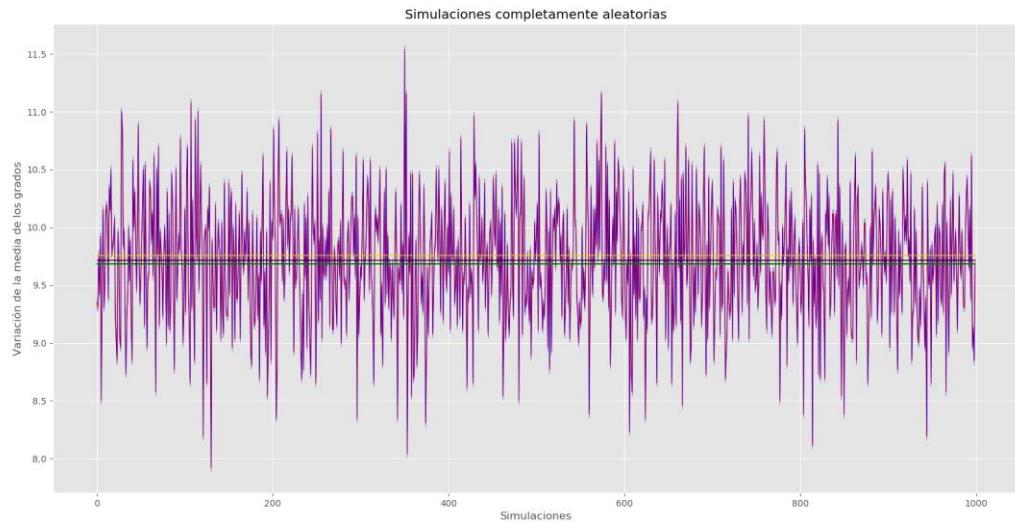


Figura 7. Intervalos de confianza con simulaciones completamente aleatorias

Otro rasgo que se presentó es el movimiento de la media de las 3 cuervas, que remarcan el movimiento de la curva de las medias.

Referente al caso de las otras simulaciones, debido a la preservación del grado, se obtuvo una distribución idéntica a la del conectoma original, con lo que se procedió de la misma forma que el caso pasado para el establecimiento de sus respectivos intervalos de confianza.

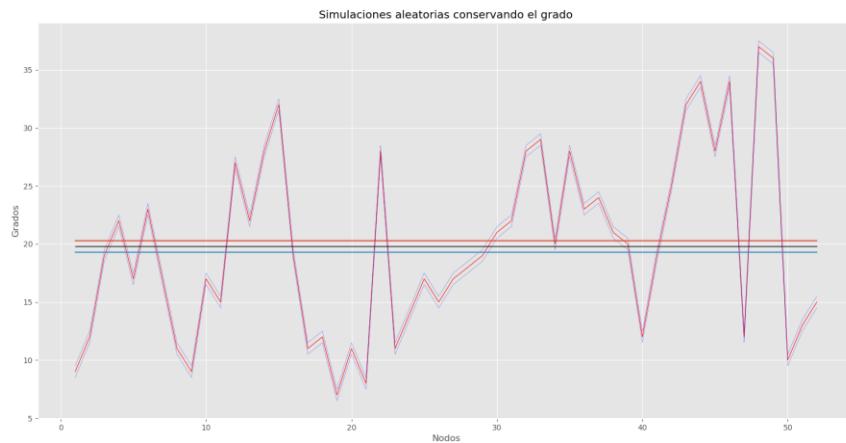


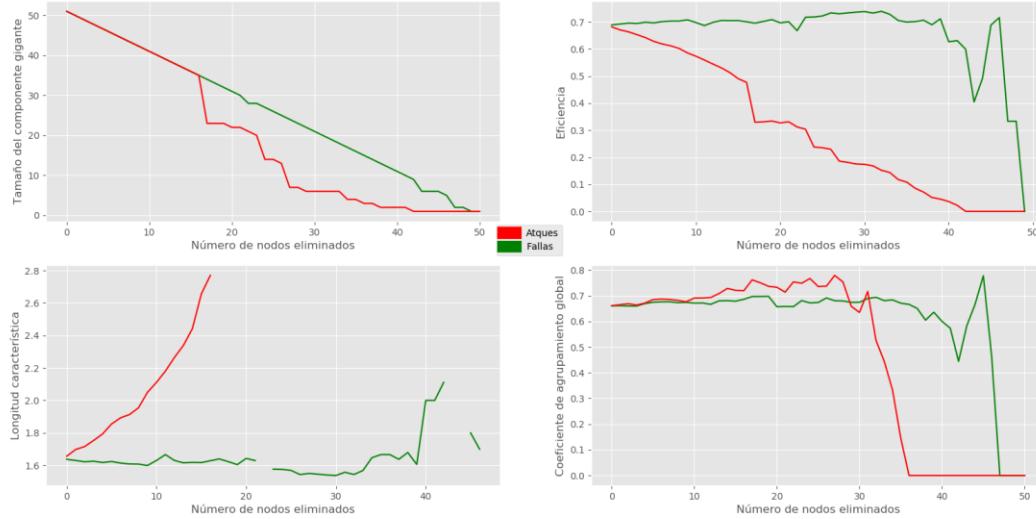
Figura 8. Intervalos de confianza con simulaciones preservando el grado.

ROBUSTEZ DE LA RED

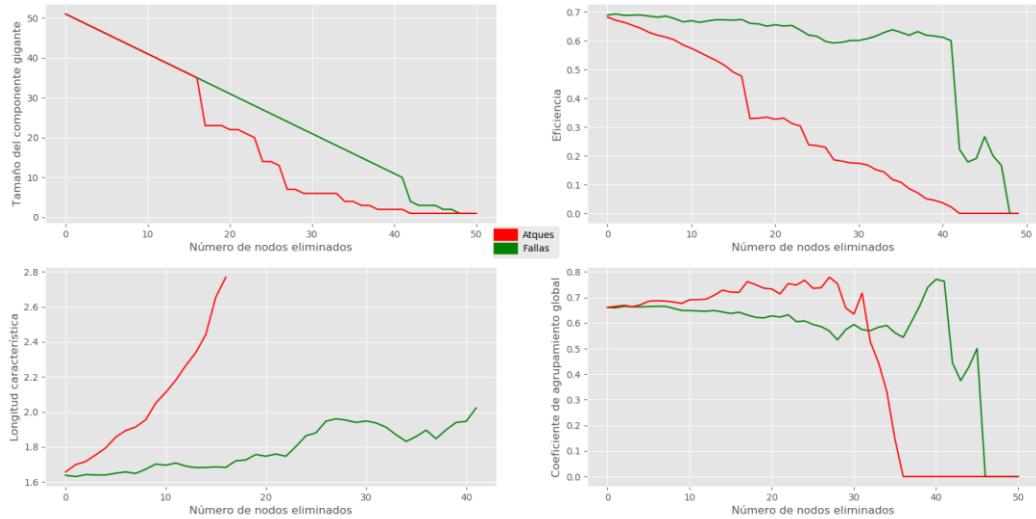
Ahora se verá la resistencia que tiene el conectoma ante la desaparición progresiva de las conexiones, a través de las fallas aleatorias y los ataques dirigidos. Dentro de los parámetros de la red que se analizarán están:

1. Tamaño del componente gigante
2. Eficiencia
3. Longitud característica
4. Coeficiente de agrupamiento global

En una primera ejecución de esta prueba se obtuvieron las siguientes gráficas.



(a)



(b)

Figura 9. Prueba primer (a) y prueba segunda (b)

En lo que respecta al tamaño del componente gigante, la eficiencia y el coeficiente, sus comportamientos son similares en ambas pruebas:

- Para fallas el componente decrece de manera casi lineal, mientras que para los ataques hacia los nodos con mayor grado, varía desde un cierto punto a las fallas con cambios con reducción en las conexiones de manera más abrupta.
- Para la eficiencia, de igual modo, se puede sostener en un rango de nivel hasta que de pronto comienza a reducir drásticamente, aunque con leves “amortiguaciones”. Mientras que los ataques provocan un decaimiento casi lineal.
- Para el coeficiente, los ataques lo reducen de manera no completamente lineal sino que trata de mantenerse para luego caer antes que la reducción provocada por las fallas.

Sin embargo, la longitud presenta cualidades más extravagantes debido a que se trata del cálculo de un recíproco el cual se indetermina cuando su denominador es 0, provocando que se maneje como infinito donde ocurre esto. En los ataques es constante que, a partir de un punto de tienda a infinito, mientras que en las fallas sucede o al final o intermitentemente.

Con el análisis hecho se puede pronosticar lo que sucedería si se desaparecen las conexiones de las diversas regiones, para así señalar posibles daños cerebrales que puedan provocar la interrupción de algunas conexiones y estudiar sus consecuencias. De estas últimas se podrían esperar daños graves tratándose de las estructuras tratadas en el estudio original.

Análisis del conectoma del macaco

CARACTERIZACIÓN DE RED

Para esta red que era dirigida se pasó a ser no dirigida, mostrando en la figura 10, el resultado de su matriz adyacente.

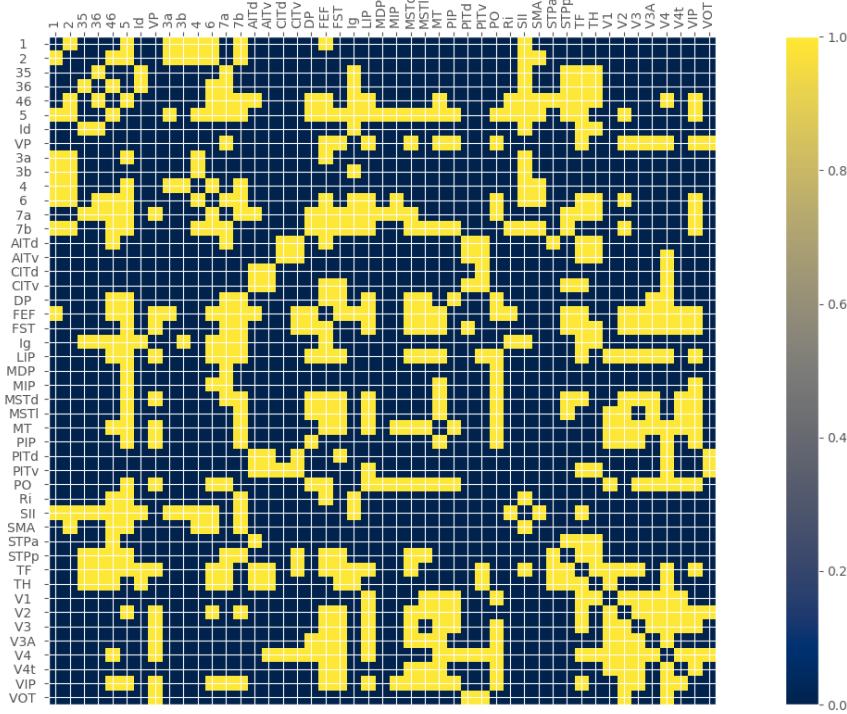


Figura 10. Matriz adyacente del conectoma del macaco

En primera instancia se aprecia una dispersión notoria en cuanto a los cúmulos de conexiones, suponiendo así, que tal vez existe una actividad más distribuida en cuanto la percepción espacial y noción del tiempo, respecto a las regiones de la corteza, señalando con esto la manera en la que opera el procesamiento dentro de la corteza, o sea no centralizado.

La distribución de grado, mostrada en la figura 2, señala las regiones con más conexiones. Donde las áreas de la corteza con mayor número son la 5 y el FEF (Frontal Eye Field). Del primero según el estudio, se señaló como una de las áreas con una elevada centralidad de intermediación⁸ (*Betweenness Cetrality*) lo que representa que dicha región se conecta a grupos que de otro modo estarían inconexos⁹. Por otro lado, respecto a la otra zona, se encarga de que con el movimiento de los ojos y la cabeza ubique objetos ya sea para evitarlos o para acercarse a ellos¹⁰.

La relevancia cuantitativa de estas dos áreas, tratándose en el estudio de aquellas responsables en la corteza del movimiento en el espacio del macaco, tal vez se deba a la importante función de los ojos para el mismo (hablando de la FEF), mientras que el área

35 debido al papel que tiene en el establecimiento de la comunicación a través de sus conexiones con las demás regiones.

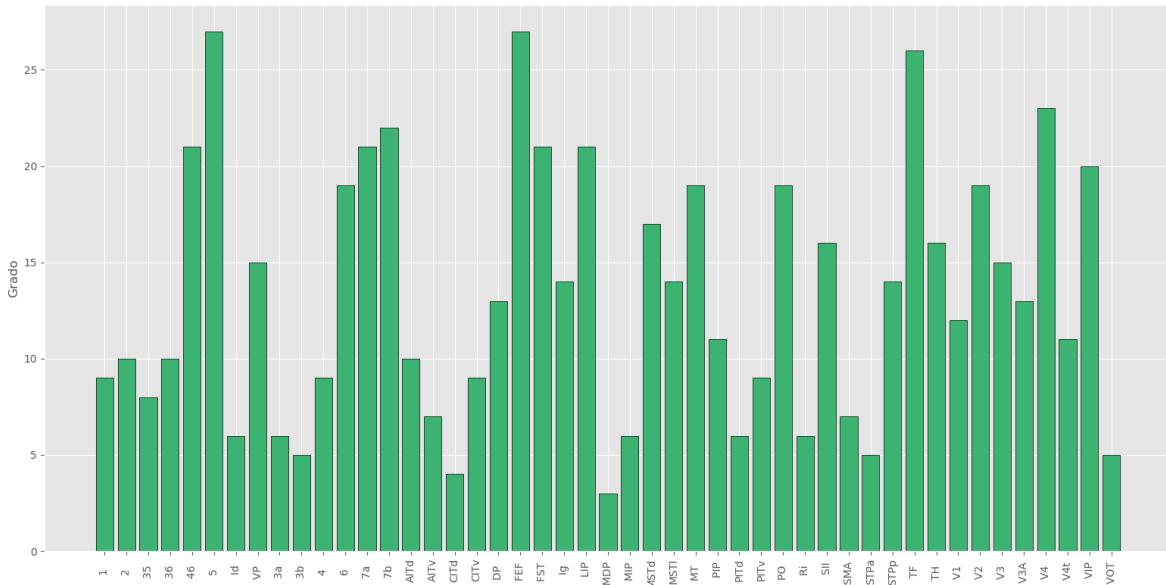


Figura 11. Distribución de grado de las áreas estudiadas en el cerebro del macaco

Para conocer la proximidad de las regiones y las distintas conexiones se recurrió a la gráfica de la matriz de distancias (ver figura 12).

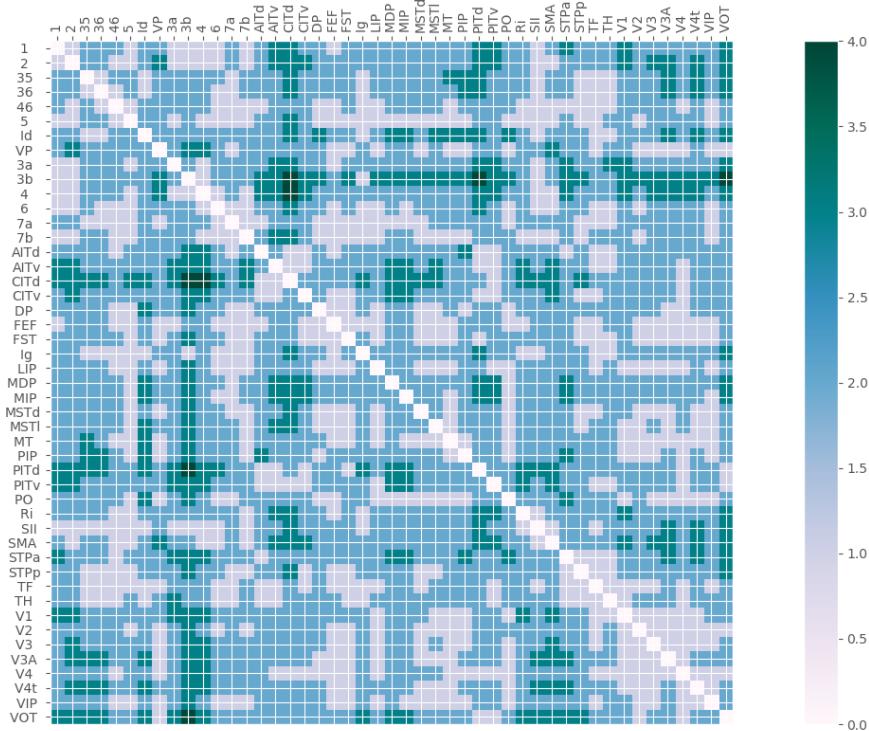


Figura 12. Matriz de distancias para el conectoma del macaco

De la última gráfica, si bien las regiones que están conectadas varían entre 0.5 y 1.5, se puede notar que las regiones están más esparcidas el límite superior de la barra de colores.

Al igual que con el conectoma del gato se recurrió a la visualización mediante Gephi de la red. En la figura 13 se muestran las comunidades halladas (con la modularidad), 4 en total, con sus nodos representando el grado correspondiente con su tamaño.

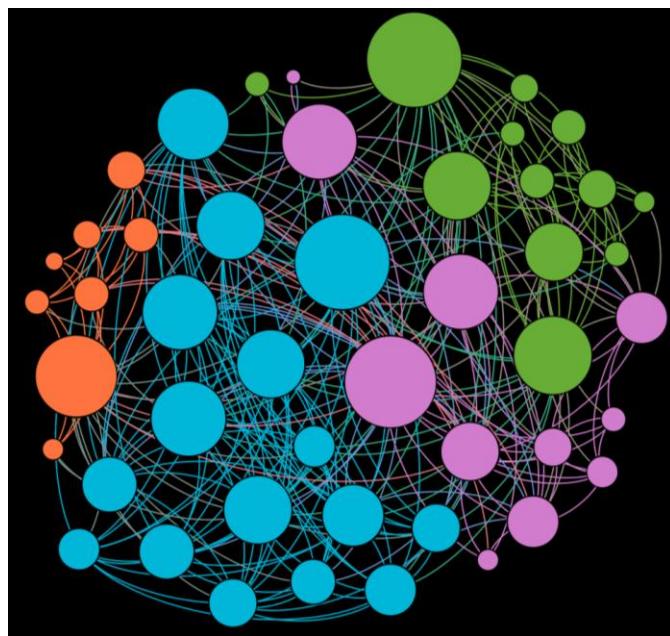


Figura 13. Red del grafo del conectoma del macaco.

En la figura de arriba se puede notar una compactación de la misma puesto que las comunidades y todos sus nodos se ven no muy lejanos entre sí.

De modo cuantitativo se obtuvieron los valores de los siguientes parámetros de la red:

- Densidad = 0.2895: este valor quiere decir que sólo está presente el 28.95% de todas las conexiones posibles en el conectoma, lo cual no es necesario, pues quiere decir que cada área con sus respectivas conexiones con las demás, se dedica a hacer una parte del trabajo.
- Eficiencia global = 0.6218: indica que la información de toda la red fluye a un 62.18% entre todos los nodos. Que también puede estar relacionado con el estado de actividad del macaco.
- Coeficiente de agrupamiento global = 0.6412: los nodos de las regiones tienden a agruparse en un 64.12%, indicando la proximidad que hay entre cada región. Esto se puede relacionar con las distancias mostradas en la figura 12.

En complemento, el coeficiente de agrupamiento local, cuya distribución se muestra en la figura 14, permite delinear mejor aún el agrupamiento que hay entre cada nodo de la red y sus vecinos, además de facilitar la visualización de que, en promedio, cada una de las áreas están agrupadas en un 64.12%.

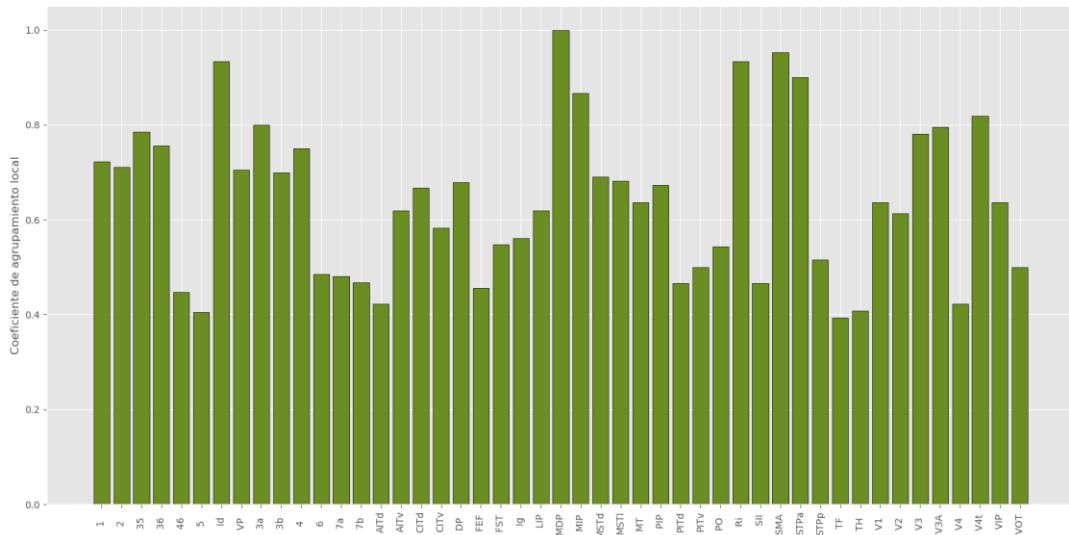


Figura 14. Distribución del coeficiente de agrupamiento local del macaco

AGRUMPA MIETO POR MEAN SHIFT

Como se hizo con el gato, se recurrió a ejecutar código para el agrupamiento mediante el método de **Mean Shift**. El resultado de este procesamiento se muestra a continuación.

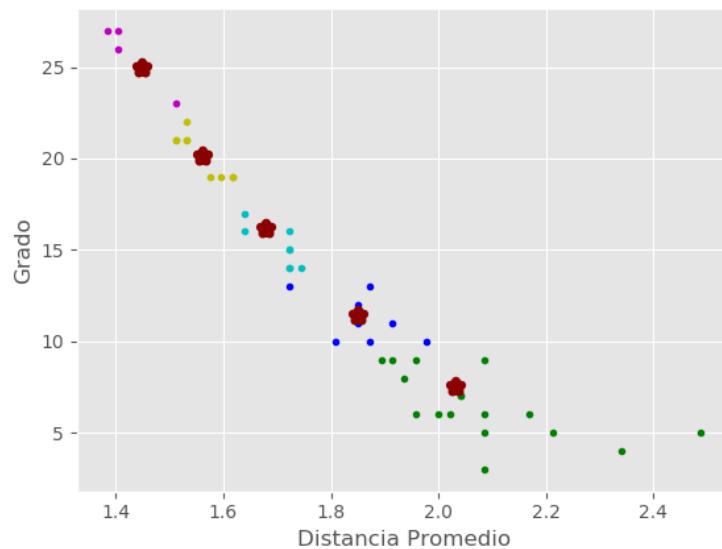


Figura 15. Agrupamiento por Mean Shift

El algoritmo determinó como 5 grupos el óptimo para ser mostrado con base al grado de los nodos y la distancia promedio. Esto de nuevo incide de nuevo en la observación especulativa de que las áreas más conectadas se encuentran próximas unas a las otras mientras que los nodos de menor grado están más distanciados.

INTERVALOS DE CONFIANZA

El procedimiento para establecer los intervalos de confianza es exactamente el mismo que para el primer conectoma, por lo tanto, se procederá a la vista de los gráficos para ambos casos.

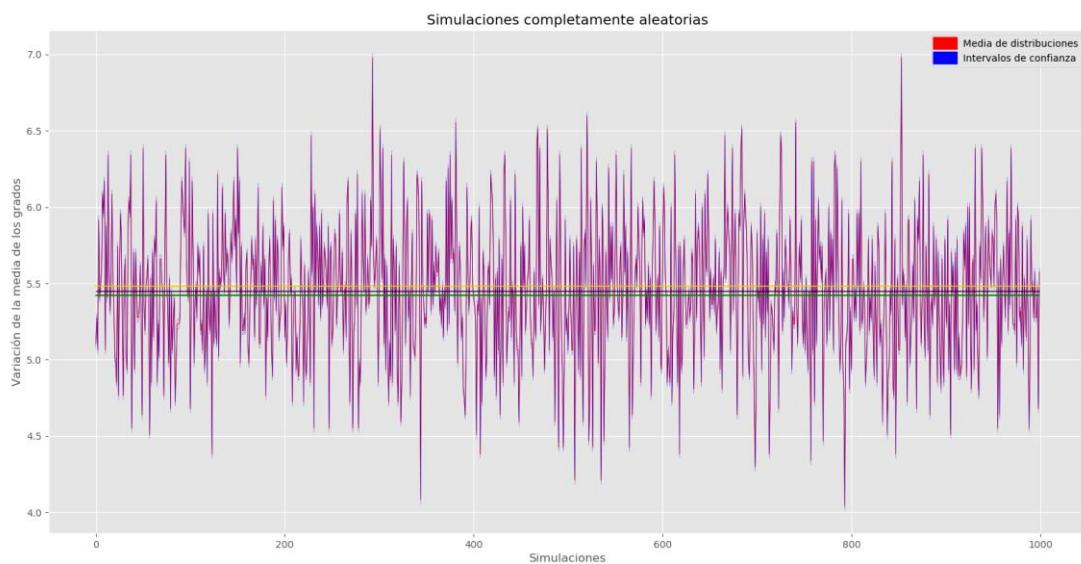


Figura 16. Intervalos de confianza con simulaciones completamente aleatorias

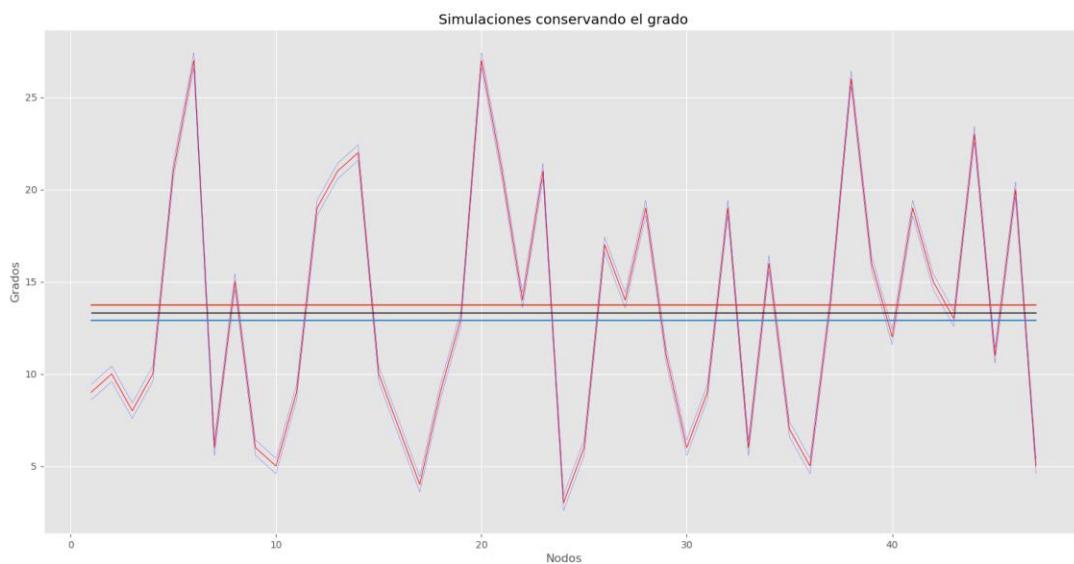
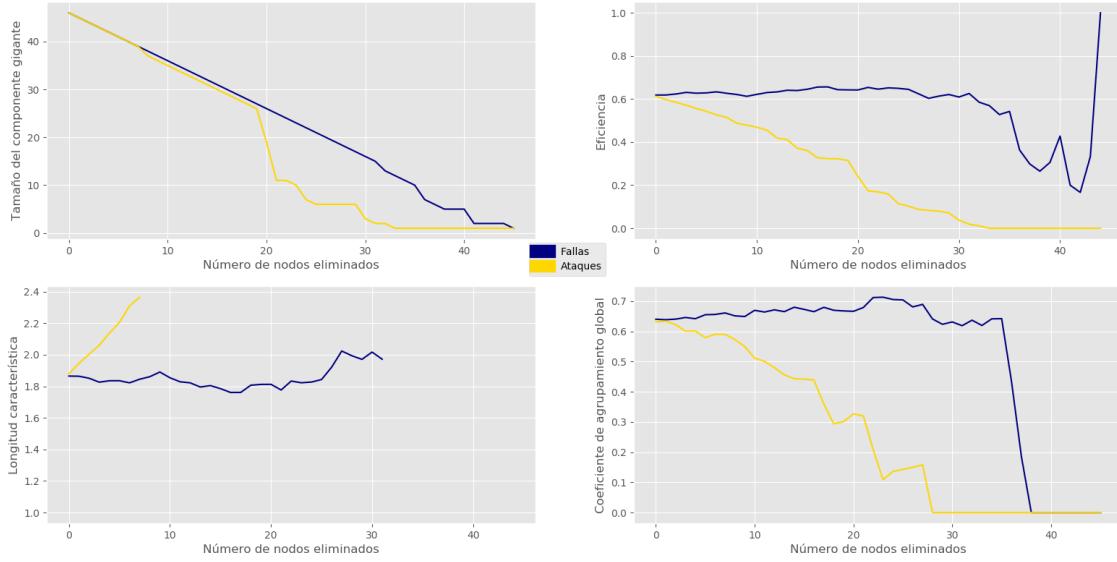


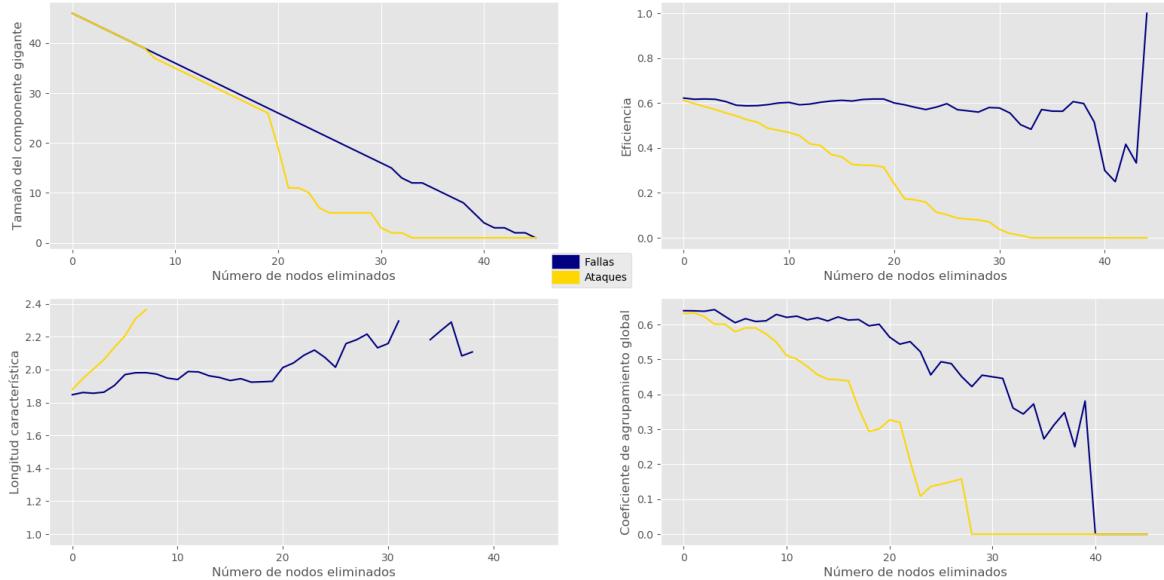
Figura 17. Intervalos de confianza con simulaciones preservando el grado.

ROBUSTEZ DE LA RED

Para el análisis en la resiliencia de las áreas de la corteza relacionadas con este conectoma, se harán do pruebas mostradas en la figura siguiente.



(a)



(b)

Figura 18. Primera prueba (a) y segunda prueba (b)

Como es de apreciarse, los resultados de las fallas tardan más en cobrar un efecto más dañino a la red que los ataques lo cuales hacen decaer tanto la comunicación que hay en

toda la red (con la eficiencia) como la longitud de las rutas geodésicas que rápidamente tiende a indeterminarse por la desaparición o malfuncionamiento de los nodos.

Esto repercutiría, a nivel funcional, el movimiento del macaco, dejándolo inmóvil o con incapacitaciones relevantes si alguna de las áreas llegara a presentar algún tipo de daño o desconexión.

Referencias

1. Guyton A. C., Hall J. E. Tratado de Fisiología Médica III. Editorial Interamericana McGraw-Hill, Nueva York, 1998. (este es mi libro de fisiología)
2. A. Triglia, "Tálamo: anatomía, estructuras y funciones", *Psicología Y Mente.com*, 2020. [Online]. Available: <https://psicologiyamente.com/neurociencias/talamo-anatomia-funciones>. [Accessed: 22 Jun- 2020]
3. Perea, M. Victoria & Ladera, Valentina. (2004). Neurofunctional aspects of the thalamus. *Revista de neurología*. 38. 687-93.
4. Nadal, Marcela & Amarillo, Yimy. (2018). El tálamo en el centro de la atención.
5. J. Scanell, G. Burns, C. Hilgate, M. O'Neil and M. Young, "The Connectional Organization of the Cortico-thalamic System of the Cat", 1999
6. M.Krupa,E.Maire-Lepoivre and M.Imbert, "Visual properties of neurons in the suprageniculate nucleus of the cat", Elsevier Ireland Ltd,1984.
7. "ML - Clustering Mean Shift Algorithm - Tutorialspoint", *Tutorialspoint.com*, 2020. [Online]. Available: https://www.tutorialspoint.com/machine_learning_with_python/machine_learning_with_python_clustering_algorithms_mean_shift.htm. [Accessed: 24- Jun- 2020].
8. C. Honey, R. Kötter, and O. Sporns, "Network structure of cerebral cortex shapes functional connectivity on multiple time scales" *PNAS*, 2007.
9. "Betweenness Centrality - an overview | ScienceDirect Topics", *Sciencedirect.com*, 2020. [Online]. Available: <https://www.sciencedirect.com/topics/computer-science/betweenness-centrality>. [Accessed: 24- Jun- 2020].
10. "Frontal Eye Fields - an overview | ScienceDirect Topics", *Sciencedirect.com*, 2020. [Online]. Available: <https://www.sciencedirect.com/topics/neuroscience/frontal-eye-fields>. [Accessed: 24- Jun- 2020].