

Trabajo Fin de Grado

Ingeniería de Organización Industrial

Análisis del movimiento de residentes en España mediante Redes Complejas

Autora: Ana Domínguez-Adame Palomo

Tutora: Ester Gutiérrez Moya

Dpto. Organización Industrial y Gestión de Empresas I
Escuela Técnica Superior de Ingeniería
Universidad de Sevilla

Sevilla, 2021



Trabajo Fin de Grado
Ingeniería de Organización Industrial

Análisis del movimiento de residentes en España mediante Redes Complejas

Autora:

Ana Domínguez-Adame Palomo

Tutora:

Ester Gutiérrez Moya

Dpto. de Organización Industrial y Gestión de Empresas I

Escuela Técnica Superior de Ingeniería

Universidad de Sevilla

Sevilla, 2021

Trabajo Fin de Grado: Análisis del movimiento de residentes en España mediante Redes Complejas

Autora: Ana Domínguez-Adame Palomo

Tutora: Ester Gutiérrez Moya

El tribunal nombrado para juzgar el Proyecto arriba indicado, compuesto por los siguientes miembros:

Presidente:

Vocales:

Secretario:

Acuerdan otorgarle la calificación de:

Sevilla, 2021

El Secretario del Tribunal

A mi familia

A mis maestros

Resumen

Este trabajo tiene como objetivo el análisis del movimiento de residentes en España en el periodo 2015-2019 a través de la metodología de Redes Complejas. Dicha metodología proporcionará una mejor comprensión de los flujos de movilidad entre las distintas Comunidades Autónomas (CC. AA.). En primer lugar se analizarán las principales características del movimiento nacional a través del análisis de los estadísticos que describen la topología de la red. Los datos analizados proceden de las encuestas anuales realizadas durante el periodo 2015-2019, información publicada por el Instituto Nacional de Estadística (INE) de España.

En segundo lugar, se expondrán los modelos Exponenciales de Grafos Aleatorios (Exponential Random Graph Models, ERGM) mediante los cuales se podrá analizar la no aleatoriedad de la red estudiada a través de un modelo estadístico que considera variables endógenas (arcos, reciprocidad, triangulación) y variables exógenas (conexión con alta velocidad ferroviaria, tipo de destino turístico, festividades, localización geográfica), permitiendo la detección de patrones de comportamiento en la movilidad entre Comunidades Autónomas, además de la posibilidad de poder explicar la relevancia de los flujos y de la red en su conjunto.

Finalmente, se realizará un análisis del impacto que la actual crisis sanitaria, provocada por el virus SARS-CoV-2, ha tenido sobre la movilidad.

Abstract

This work aims to analyze the movement of residents in Spain in the period 2015-2019 through the Complex Networks methodology. This methodology will provide a better understanding of the mobility flows between the different Autonomous Communities (CC. AA.). Firstly, the main characteristics of the national movement will be analyzed through the analysis of the statistics that describe the topology of the network. The data analyzed comes from the annual surveys carried out during the period 2015-2019, information published by the National Institute of Statistics (INE) of Spain.

Secondly, the Exponential Random Graph Models (ERGM) will be exposed through which the non-randomness of the studied network can be analyzed through a statistical model that considers endogenous variables (arcs, reciprocity, triangulation). and exogenous variables (connection with high-speed rail, type of tourist destination, festivities, geographical location), allowing the detection of behavior patterns in mobility between Autonomous Communities, in addition to the possibility of being able to explain the relevance of the flows and the network as a whole.

Finally, an analysis will be made of the impact that the current health crisis, caused by SARS-CoV-2 virus, has had on mobility.

Resumen	9
Abstract	11
Índice	12
Índice de Figuras	14
Índice de Tablas	15
1 Introducción	17
2 Metodología	19
2.1. <i>Introducción a Redes Complejas</i>	19
2.2. <i>Principales Conceptos de las Redes Complejas</i>	19
2.2.1 Matriz de Adyacencia	19
2.2.2 Distribución de grados	19
2.2.3 Grado de Incidencia	20
2.2.4 Peso	20
2.2.5 Longitud de los caminos	20
2.2.6 Red conexa	21
2.3. <i>Modelos de Redes Complejas</i>	21
2.3.1 Redes Regulares	21
2.3.2 Redes Aleatorias	22
2.3.3 Redes de Mundo Pequeño	22
2.3.4 Modelos Libres de Escala	23
2.4. <i>Comunidad de Redes</i>	24
2.5. <i>Camino y camino geodésico</i>	24
2.6. <i>Coficiente de Clustering</i>	24
2.7. <i>Centralidad</i>	25
2.8. <i>Vulnerabilidad</i>	25
2.9. <i>Análisis de Riesgo</i>	26
3 Principales Estadísticos de la Red	27
3.1. <i>Grado Medio</i>	27
3.2. <i>Índice de Centralización de grado</i>	27
3.3. <i>Densidad</i>	27
3.4. <i>Conectividad</i>	28
3.5. <i>Fragmentación</i>	28
3.6. <i>Distancia Media</i>	28
3.7. <i>Diámetro</i>	28
3.8. <i>Compacidad</i>	29
3.9. <i>Amplitud</i>	29
3.10. <i>Page Rank</i>	29
4 Análisis de la Movilidad en España. Características	30
4.1. <i>La Movilidad en España</i>	30
4.2. <i>Movimiento entre Comunidades Autónomas</i>	32
4.3. <i>Motivación de la Movilidad en España</i>	34
4.4. <i>Análisis de la movilidad mediante Ucinet</i>	35
4.4.1 Datos utilizados	35
4.4.2 Visualización de la Red 2015	35
4.4.3 Visualización de la Red 2016	36
4.4.4 Visualización de la Red 2017	36

4.4.5	Visualización de la Red 2018	37
4.4.6	Visualización de la Red 2019	38
4.4.7	Visualización de la Red Promedio (2015-2019)	40
4.5.	<i>Comentarios adicionales</i>	41
4.6.	<i>Estadística de la red de desplazamiento (2015-2019)</i>	42
4.6.1	Resultados de los Estadísticos de la Red 2015	44
4.6.2	Resultados de los Estadísticos de la Red 2016	44
4.6.3	Resultados de los Estadísticos de la Red 2017	45
4.6.4	Resultados de los Estadísticos de la Red 2018	45
4.6.5	Resultados de los Estadísticos de la Red 2019	45
4.6.6	Resultados de los Estadísticos de la Red Promedio (2015-2019)	46
4.7.	<i>Resultados de la distribución Ley de Potencias</i>	46
4.8.	<i>Resultados del Contraste de Bondad de Ajuste</i>	48
5	Modelos ERGM	50
5.1.	<i>Introducción a los Modelos ERGM</i>	50
5.2.	<i>Descripción de los Modelos ERGM</i>	50
5.3.	<i>Estimación y análisis del modelo ERGM</i>	51
5.3.1	Variables endógenas	51
5.3.2	Variables exógenas	52
5.4.	<i>Simulación con RStudio</i>	53
6	Efectos del virus SARS-CoV-2 sobre la movilidad	59
7	Conclusiones	65
	Referencias bibliográficas	66
	Anexo I: Tablas	70
	Anexo II: Código en R	72

Índice de Figuras

Figura 1: Ejemplo de Distribución Ley de Potencias	20
Figura 2: Ejemplo de Redes Regulares	21
Figura 3: Ejemplo de Redes Aleatorias	22
Figura 4: Ejemplo de Redes de Mundo Pequeño	23
Figura 5: Ejemplo de Redes Libres de Escala	23
Figura 6: Vulnerabilidad	26
Figura 7: Clasificación de la movilidad (INE)	30
Figura 8: Mapa Político de España	31
Figura 9: Pernotaciones de residentes por tipo de alojamiento (millones) (2019)	32
Figura 10: Visualización de la Red (2015)	35
Figura 11: Visualización de la Red (2016)	36
Figura 12: Visualización de la Red (2017)	37
Figura 13: Visualización de la Red (2018)	38
Figura 14: Visualización de la Red (2019)	39
Figura 15: Visualización Red 2015 y 2016 (comparativa)	39
Figura 16: Visualización Red 2017, 2018 y 2019 (comparativa)	40
Figura 17: Visualización de la Red (Valores Promedio)	41
Figura 18: Principales rutas por avión (2012)	42
Figura 19: Trayectos de AVE en España	42
Figura 20: Red según PageRank	47
Figura 21: Red según PageRank de la Autonomía más influyente	48
Figura 22: Resultados Power-Law	49
Figura 23: Resultados GOF arcos incidentes	58
Figura 24: Resultados GOF arcos emisores	58
Figura 25: Casos de SARS-CoV-2 en el mundo	59
Figura 26: Casos de SARS-CoV-2 en el mundo (acumulado)	60
Figura 27: Casos de SARS-CoV-2 en España (acumulado)	61
Figura 28: Casos de SARS-CoV-2 en España (desarrollo)	61
Figura 29: Viajes por CC. AA. de residencia de los viajeros	62
Figura 30: Evolución de personas vacunadas en España	63

Índice de Tablas

Tabla 1: Población por Comunidad Autónoma (2019)	31	
Tabla 2: Total viajes por Comunidad Autónoma (2019).	32	
Tabla 3: Total viajes por Comunidad Autónoma desglosado (2019)	33	
Tabla 4: Total viajes por Comunidad Autónoma a otra Autonomía (2019).	33	
Tabla 5: Resultados Red 2015	Tabla 6: Resultados Red 2016	43
Tabla 7: Resultados Red 2017	Tabla 8: Resultados Red 2018	43
Tabla 9: Resultados Red 2019	Tabla 10: Resultados Red Valores Pormedio	44
Tabla 11: PageRank		47
Tabla 12: Ajuste de la ley de Potencias según grado		48
Tabla 13: Resultados del modelo (ERGM) número de arcos		54
Tabla 14: Resultados del modelo (ERGM)		54
Tabla 15: Resultados del modelo (ERGM)		55
Tabla 16: Resultados del modelo (ERGM) con efectos exógenos		55
Tabla 17: Resultados del modelo (ERGM) con efectos exógenos (entrada)		56
Tabla 18: Resultados del modelo (ERGM) con efectos exógenos (salida)		56
Tabla 19: Resultados GOF (p-valor)		57
Tabla 20: Movilidad internacional en España (2020-2021)		60
Tabla 21: Movilidad nacional en España (2019 - 2020)		62
Tabla 22: Datos de movilidad entre CC. AA. (2015)		70
Tabla 23: Datos de movilidad entre CC. AA. (2016)		70
Tabla 24: Datos de movilidad entre CC. AA. (2017)		70
Tabla 25: Datos de movilidad entre CC. AA. (2018)		71
Tabla 26: Datos de movilidad entre CC. AA. (2019)		71
Tabla 27: Datos de movilidad entre CC. AA. (promedio)		71

1 INTRODUCCIÓN

La movilidad de personas se entiende como el conjunto de desplazamientos que se producen en un entorno físico, realizados haciendo uso de diferentes sistemas de transporte, ya sea mediante automóvil, avión, transporte público, etc. en función de la dimensión que tenga el movimiento.

La movilidad ha sido una necesidad a lo largo de la historia humana, incrementándose el volumen de los flujos con los avances industriales y la creación de los diferentes medios de transporte.

Además, los avances tecnológicos han permitido que el movimiento de personas se realice con mayor rapidez y frecuencia, aumentando las distancias de los trayectos y disminuyendo su coste. En consecuencia, esto ha supuesto que los desplazamientos se produzcan en masa pues, a medida que pasa el tiempo, la posibilidad de desplazarse es cada vez más accesible. En efecto, el aumento de los trayectos ha supuesto numerosas ventajas en muchos ámbitos. Sin embargo, no debemos dejar de lado otros aspectos que han tenido un impacto colateral, como la contaminación atmosférica fruto de, entre otros motivos, los gases contaminantes emitidos por diferentes sistemas de transporte.

Así, la movilidad en su conjunto se puede clasificar según su principal motivación: como una fuente de atracción de turismo o una fuente de migración, aunque a lo largo del trabajo se podrá comprobar que el espectro de las principales causas de movilidad es mucho más amplio.

En cuanto al desplazamiento con fines turísticos, se puede afirmar que la presencia del turismo internacional es un pilar básico de la economía española, pues es uno de los motores económicos esenciales de este país. La aportación del turismo a la economía española supone un 12,4% (dato correspondiente al año 2019) sobre el total del porcentaje del PIB, siendo su aportación monetaria de 154.487 millones de euros, generando a su vez 2,72 millones de puestos de trabajo. En cuanto a la caracterización del turismo, las opciones son variadas pues, según la actividad turística a realizar, la clasificación resulta la siguiente: turismo cultural, ecológico, religioso, deportivo, de playa, de nieve, rural y gastronómico. Sin embargo, el turismo que se realiza en España no sólo proviene de los turistas internacionales. Esta misma clasificación se puede aplicar al turismo nacional puesto que casi el 90% de los viajes de residentes se hicieron dentro de España en 2019 haciendo que el número de viajes realizados por los residentes en España ascendiera a 173,75 millones de viajes con destino interno, siendo “únicamente” 20,12 millones de viajes con destino externo. Como se ha señalado, el turismo en España se realiza tanto de forma nacional como internacional. En este sentido, los principales países exportadores de turistas a España son Reino Unido, Alemania y Francia, siendo las principales Comunidades Autónomas de destino en 2019: Cataluña, Baleares y Canarias.

El segundo fin por el que se producen desplazamientos es la migración. Las causas de este concepto son muy variadas, ya sean por una económica, pobreza, hambruna, guerras y conflictos armados, exilios y persecuciones o desastres naturales, todas ellas devastadoras, produciendo consecuencias de mayor o menor envergadura tanto en el lugar de origen como en el destino, creando: cambios demográficos, intercambio cultural y étnico, cambios en la economía e incluso xenofobia. El saldo migratorio desde España hacia otros países fue de 51.250 personas con nacionalidad española durante el año 2019, ascendiendo la inmigración de extranjeros a 661.065 personas. Cabe destacar que son Madrid y Cataluña las Comunidades Autónomas que mayor número de extranjeros recibe, siendo Francia, Alemania y Estados Unidos los países a los que emigran más españoles.

Los flujos de movilidad que se analizan mediante redes complejas en el presente trabajo se verán influidas por variables endógenas y exógenas que afectan a la estructura de las mismas. Por lo tanto, los modelos Exponenciales de Grafos Aleatorios (Exponential Random Graph Models, ERGM) permiten relacionar diversos atributos y características propias y compartidas que tienen las Comunidades Autónomas que afectan al flujo de la movilidad entre las CC. AA., lo cual, en definitiva, permitirá afirmar la presencia de variables que resultan determinantes para la formación del flujo.

Finalmente, es importante precisar que los datos que se analizan en el presente trabajo corresponden al año 2019 y años anteriores pues, a consecuencia del SARS-CoV-2, la situación que se ha establecido en el año 2020 y 2021 es muy dispar a la de años precedentes. La crisis económica y sanitaria, así como las restricciones

asociadas a retener la propagación del SARS-CoV-2, han producido una paralización en el número de movimientos realizados tanto en el territorio español como a nivel internacional. En cuanto a viajes de los españoles realizados al extranjero, las cifras son desproporcionadas: para tener una idea genérica del cambio, partiendo del segundo trimestre del 2019, teniendo 4.933.015 viajes, la cifra ha disminuido a 157.532 viajes (un 96.8%) para el mismo periodo del año 2020. La cifra de los viajes de españoles dentro del territorio español es similar pues se pasó de tener 45.541.375 viajes en el segundo trimestre del 2019 a 8.869.729 viajes en el año 2020 para el mismo periodo (disminuyendo un 80,5%).

El análisis de las redes requiere una aplicación informática apropiada. En este trabajo se utilizará Ucinet (Borgatti, 2002) y el entorno de programación estadística R (R Core Team, 2021), un producto de libre distribución bajo la GNU General Public License. R ofrece un conjunto de paquetes que permiten construir las Redes Complejas, y además, la comunidad de usuarios mantiene constantemente actualizada esta plataforma de computación estadística.

Por tanto, los principales objetivos de este trabajo son: en el Capítulo 2 se presentan algunos conceptos básicos de la metodología a tratar. El Capítulo 3 describe con detalle los estadísticos que caracterizan las redes complejas. El Capítulo 4 analiza el movimiento de residentes en España, a partir de los valores de los datos publicados por el Instituto Nacional de Estadística en sus encuestas anuales en el periodo 2015-2019. En el Capítulo 5 se presentan los modelos ERGM, de forma que se pueda describir las distintas configuraciones que posee la estructura de la red con el fin de poder explicar la relevancia que tiene la misma. A modo de conclusión, en el Capítulo 6 se realizará un análisis de la situación producida por el virus SARS-CoV-2, tanto de forma global como nacional.

2 METODOLOGÍA

En este capítulo se presenta una introducción a la metodología de Redes Complejas, basada en estadísticos que permiten caracterizar la topología de las redes. Finalmente, se van a identificar los tipos de redes complejas existentes en la metodología.

2.1. Introducción a Redes Complejas

Una *red compleja* es una red con propiedades estadísticas y topológicas que no tienen las redes simples. Se pueden observar este tipo de redes en la naturaleza, ejemplo de ello son las redes neuronales, las redes sociales, las redes de metro o de cualquier tipo de transporte.

Este tipo de redes abundan en la naturaleza, pudiéndose dar en todo tipo de situaciones y a diferentes niveles de organización, abarcando disciplinas como la física, biología, sociología, neurología, economía o medicina, entre otras. Los diferentes niveles de organización hace referencia a que una red de proteínas o una red neuronal se da a un nivel de organización mucho menor que una red de comunicación o una red informática o social.

Una red compleja se puede representar gráfica y matemáticamente mediante un grafo. La Teoría de Grafos permite la representación de las relaciones entre objetos de todo tipo. Se denota un grafo $G = (V, E)$ formado por un conjunto V de vértices (nodos) y un conjunto E de arcos entre estos nodos, que más adelante se verá que pueden ser dirigidos o no dirigidos. Se considera un nodo individual $i \in V$, siendo un arco un par (i, j) formado por dos nodos, siendo a su vez i y j nodos vecinos o adyacentes. Un grafo será dirigido si el arco (i, j) es distinto al arco (j, i) (también denominado *digrafo*).

A continuación, se procede a comentar los conceptos que caracterizan a las redes complejas y por los que quedan definidas e identificadas.

2.2. Principales Conceptos de las Redes Complejas

2.2.1 Matriz de Adyacencia

La *matriz de adyacencia* A_G o A_{ij} representa una matriz cuadrada donde la casilla será 1 si $(i, j) \in E$ y será 0 en caso contrario. También se puede expresar como:

$$A_{ij} = \begin{cases} 1 & \text{si } (i, j) \in E \\ 0 & \text{en caso contrario} \end{cases}$$

Cuando A_{ij} es simétrica, se dice que la red es no es dirigida, siendo todos su autovalores reales. En el caso contrario, si A_{ij} es no simétrica, la red es dirigida.

2.2.2 Distribución de grados

El *grado del nodo* i , k_i , representa el número total de conexiones que tiene. Se denomina *grado medio de la red*, \bar{k} , al promedio de los grados de los nodos de la red.

En una red se hace uso de la función de distribución $P(k)$ para representar la distribución de grados en la misma, y representa la probabilidad de que un nodo seleccionado al azar tenga exactamente k enlaces (o conexiones). Para una red completamente aleatoria, la distribución de grados sigue una distribución Poisson. Sin embargo, la mayoría de las redes reales se describen de forma más exacta mediante la *ley de Potencias* (en

inglés, *Power-Law*). La ley de Potencias es una relación funcional entre dos cantidades, de modo que una variación relativa en una cantidad genera un cambio relativo proporcional en la otra cantidad, con independencia del valor inicial de ambas cantidades (ver Figura 1).

$$P(x) = c \cdot x^{-\gamma}$$

donde $P(x)$ es la probabilidad de que un nodo esté conectado a x nodos, c es una constante de normalización y el exponente γ identifica el tipo de red ($2 < \gamma \leq 3$ si la ley de Potencias se ajusta al conjunto de datos).



Figura 1: Ejemplo de Distribución Ley de Potencias

2.2.3 Grado de Incidencia

Como ya se comentó anteriormente, las redes pueden ser de dos tipos: no dirigidas y dirigidas. Se denomina *grado de incidencia* de un nodo para redes no dirigidas como el número de arcos que inciden en el nodo, sin diferenciar de qué extremo sea. Para redes dirigidas, se denomina *grado de entrada* del nodo al número de arcos que lo tienen como extremo final, y *grado de salida* al número de arcos que lo tienen como extremo inicial. Al sumar el grado de entrada y el grado de salida de un nodo se obtiene el grado del nodo. La suma de los grados de entrada y de salida de todos los nodos coincide y es igual al número de arcos que tiene la red.

2.2.4 Peso

Los arcos pueden tener asignado un *peso* (denominado en inglés, *strength*), haciendo referencia a la intensidad de la relación entre nodos conectados. La red en su conjunto pasará a denominarse, en este caso, red ponderada. Del mismo modo que ocurría con el grado de un nodo, el peso puede ser de entrada (en inglés, *in-strength*) o de salida (en inglés, *out-strength*), representando la suma de los pesos de los arcos que inciden o parten del nodo i , respectivamente.

2.2.5 Longitud de los caminos

La longitud de los caminos entre nodos puede definirse como:

- *Paseo*: es la sucesión de nodos conectados entre sí.
- *Camino*: se trata de un paseo donde no se visita ningún nodo en más de una ocasión.
- *Sendero*: es un paseo donde ningún arco se recorre más de una vez.
- *Ciclo*: supone un paseo cerrado de al menos tres arcos, sin recorrer ningún arco más de dos veces.
- *Distancia*, d_{ij} : representa la distancia entre el nodo i y el nodo j , y es el número de enlaces del camino más corto que los conecta.
- *Diámetro*, D : distancia máxima entre cualquier par de nodos de la red.
- *Longitud promedio*, L : es la media de las distancias entre todos los pares de nodos.

2.2.6 Red conexas

Una red será *conexa* si se puede trazar un camino entre cada nodo, es decir, si dos nodos cualesquiera pueden ser conectados mediante un camino con origen en cualquiera de éstos, siendo fuertemente conexo cuando dos nodos cualesquiera están conectados por un arco, o dicho de otra forma, conectados por al menos dos caminos, pudiendo coincidir.

La *componente conexa* representa subconjuntos conexos de nodos y aristas de forma que añadiendo un nuevo nodo, deja de ser conexo. Al igual que se comentó previamente, este concepto se puede ampliar para ser una componente fuertemente conexa.

2.3. Modelos de Redes Complejas

De forma simplificada, un *sistema complejo* es un sistema en el que el estudio de cada parte no puede explicar el comportamiento del conjunto en su totalidad. Las redes complejas se utilizan para representar gráficamente los sistemas complejos, de forma que se puedan analizar y visualizar para un posterior estudio de los mismos.

Existen diferentes algoritmos (tales como algoritmo de Tutte (1963), algoritmo de Kamada-Kawai (1989), entre otros) que se usan para reducir las redes en otras más pequeñas, de forma que se pueda extraer la información de forma más clara y útil. Sin embargo, para el trabajo en cuestión, no se utilizarán, puesto que se parte de una red inicialmente reducida. De forma inversa, existen diferentes modelos de redes (tales como redes de Erdős-Renyi (1960), redes de Watts-Strogatz (1998), redes de Barabási-Albert (1999), entre otros) utilizadas para la generación de los mismos. No obstante, en este trabajo únicamente se comentarán sus características principales, puesto que la red se crea de forma muy sencilla al tratarse de los viajes que se realizan entre las CC. AA., habiendo viajes (conexiones) entre todas ellas.

Para poder entender la estructura de la red que se va a analizar, es necesario seguir ciertos pasos para poder llegar a ese punto. En primer lugar, es necesario medir propiedades básicas mediante los conceptos que se han explicado anteriormente. A continuación, se deberá desarrollar el modelo matemático por el cual se podrá obtener las propiedades estadísticas y así poder aplicar métodos matemáticos para analizar la red. Sin embargo, estas etapas dependen del tipo de red a estudiar, pudiéndose distinguir entre redes regulares, redes aleatorias, redes de mundo pequeño o modelos libre de escala.

2.3.1 Redes Regulares

Una *red regular* es aquella que tiene todos sus nodos con el mismo grado, con una cantidad lineal de enlaces cuando se da el caso de $x \ll N$ (ver Figura 2). Por otro lado, tiene un alto coeficiente de clustering, además de una alta longitud promedio.

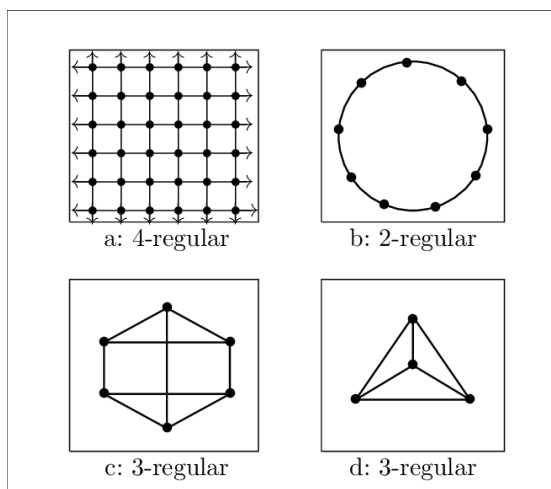


Figura 2: Ejemplo de Redes Regulares

2.3.2 Redes Aleatorias

Las *redes aleatorias* fueron estudiadas en mayor profundidad por Erdős y Rényi a finales de los 50, y resulta ser un tipo de red completamente opuesta a las redes regulares. Este tipo de redes se caracteriza por tener una gran cantidad de nodos, $N \gg 1$, conectando cada par de nodos con un enlace mediante una probabilidad p , en todos los casos, con una longitud de camino pequeña. Se obtendrá así una red aleatoria con N nodos y con $\frac{p \cdot N \cdot (N-1)}{2}$ conexiones.

Erdős y Rényi (1950) descubrieron que si la probabilidad p era superior a cierto umbral, $p_c \sim \frac{\ln(N)}{N}$, la red estaba casi totalmente conectada. En el caso contrario, para valores por debajo de ese umbral, la red resultaba estar casi desconectada (ver Figura 3).

Los estadísticos de las redes aleatorias vienen definidos por las siguientes expresiones:

- Grado medio: $\bar{k} = p \cdot (n - 1) \sim p \cdot N$
- Longitud del camino promedio: $L \sim \frac{\ln(N)}{\bar{k}}$
- Coeficiente de clustering: $C = p = \frac{\bar{k}}{N} \ll 1$. Al tener muchos nodos la red, apenas tendrá clusterización.

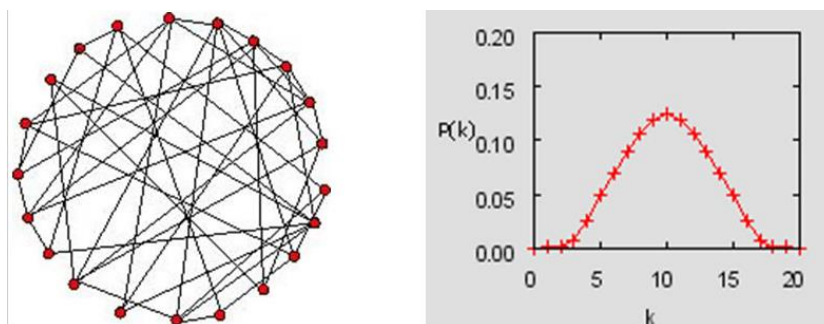


Figura 3: Ejemplo de Redes Aleatorias

Por lo tanto, para valores grandes de N , las redes aleatorias son casi homogéneas. Esto significa que la conectividad se distribuirá según una distribución Poisson.

2.3.3 Redes de Mundo Pequeño

Este tipo de red (denominadas en inglés, *Small-World Network*) combina los conceptos de las redes aleatorias y regulares pues tienen un alto índice de agrupamiento (o clustering) como las redes regulares, además de una longitud de camino pequeña, característico de las redes aleatorias.

Las redes de mundo pequeño fueron estudiadas por Watts y Strogatz (1998), creando el siguiente algoritmo: haciendo uso de una red regular de N nodos, se selecciona cada enlace al azar con probabilidad p y se conecta uno de sus extremos a cualquier nodo de la red. De este modo, se tienen $\frac{p \cdot N \cdot k}{2}$ enlaces que conectan todos los nodos, y que, en el caso de los dos tipos de redes anteriores, estarían posiblemente desconectados.

Las redes de mundo pequeño son las más frecuentes, donde gran parte de los nodos no están conectados pero la mayoría de ellos pueden ser alcanzados desde cualquier nodo de origen a través de un número relativamente reducido de enlaces entre sí (Figura 4). Ejemplo de ello son las redes sociales no virtuales, en las que con un determinado número de amigos se puede llegar a conocer a toda la comunidad.

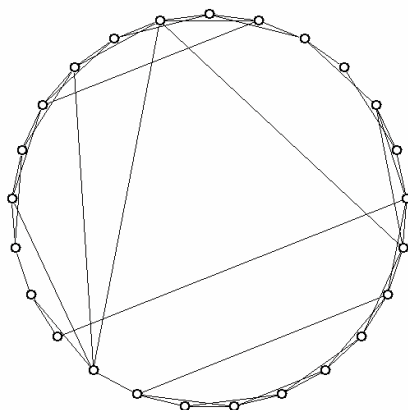


Figura 4: Ejemplo de Redes de Mundo Pequeño

2.3.4 Modelos Libres de Escala

A diferencia de las redes aleatorias y las redes de mundo pequeño, con distribución de grados homogénea, las redes complejas de gran tamaño no siguen el mismo comportamiento, pues tienen una distribución de grado que sigue una ley de potencias.

Fueron Barabási y Albert (1999) quienes estudiaron los *modelos libres de escala* (ver Figura 5), exponiendo los siguientes argumentos:

- Las redes reales añaden de forma continuada nuevos nodos a la red. Sin embargo, otros modelos resultan ser estáticos puesto que al añadir enlaces a la red, el número de nodos es constante.
- En las redes aleatorias y en las redes de mundo pequeño, las probabilidades son uniformes al crear nuevos enlaces, hecho que no es realista. Ejemplo de ello son las citas de artículos: un nuevo artículo será más propenso a citar artículos ya reconocidos que otros menos reconocidos.

Cabe destacar la importancia de los conceptos de crecimiento y enlace preferencial de las redes:

- Crecimiento: partiendo de un número pequeños de nodos (k_0), en cada unidad de tiempo se añaden nuevos nodos y se conectan a los ya existentes ($k \leq k_0$).
- Enlace preferencial: la probabilidad de que un nodo nuevo se conecte al nodo x ya existente depende del grado x_i . Su expresión fraccional es: $\prod_i = \frac{k_i}{\sum_j k_j}$.

Tras t pasos, se obtendrá una red con $N = t + m_0$ nodos y con $(m \cdot t)$ enlaces, de forma que la distribución de grados queda inalterada con el tiempo y sigue una ley de Potencias de exponente $\gamma = 3$. Si el exponente es $\gamma = 3$ equivale a decir que la probabilidad de encontrar un nodo con k enlaces es de k^{-3} . Sin embargo, este modelo tiene ciertas restricciones al compararlo con redes reales.

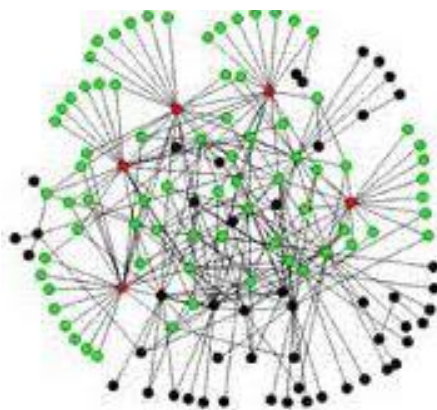


Figura 5: Ejemplo de Redes Libres de Escala

2.4. Comunidad de Redes

En las redes se pueden identificar subgrupos de nodos con una relación más estrecha que el resto, también denominado *comunidades de redes*. Es decir, se trata de un conjunto de nodos con un mayor número de arcos dentro de la comunidad que fuera (la red en su conjunto). Existen diferentes algoritmos o métodos para detectar comunidades dentro de la red:

- *Particionamiento de redes*: se dividen los nodos en g conjuntos de un tamaño predeterminado y se minimiza el número de arcos que conectan nodos de distintos conjuntos. Se denomina tamaño del corte al número de arcos que conectan nodos de distintos conjuntos. Sin embargo, no resulta ser un algoritmo eficiente, pues es necesario predeterminar el número de nodos que tiene cada comunidad.
- *Clustering jerárquico*: se define una distancia o alguna similitud para los nodos que tiene la red y se aplica algún tipo de clustering jerárquico.
- *Clustering particional*: se define una distancia o alguna similitud para los nodos que tiene la red y se aplica algún tipo de clustering divisivo.
- *Clustering espectral*: se hace uso de la matriz de similitud o de distancia, de forma que se pueda transformar la red en un conjunto de puntos para que la distancia entre puntos sea proporcional a la relación entre las distintas redes. Adicionalmente, se hará uso de algoritmos de clustering para agrupar nodos.

Hay un concepto clave para medir la calidad de los algoritmos usados: la medida de *desempeño de la generación de caminos* (denominado en inglés, *performance*), que permite contabilizar el número de arcos en una comunidad y el número de pares de nodos en distintas comunidades, según el número de arcos que se podrían formar con todos los nodos de la red. Su valor está comprendido entre 0 y 1, siendo mayor cuanto mejor es la clasificación que se encuentre.

2.5. Camino y camino geodésico

Hay múltiples definiciones posibles para el concepto de *camino*, siendo una de ellas la siguiente: el camino entre dos nodos supone una secuencia de arcos, estando esos nodos conectados entre sí. El *camino geodésico* entre dos nodos es el camino de longitud mínima entre ellos. Otra definición posible es: el camino de menor coste entre dos nodos, siendo su coste el peso (para redes ponderadas) o en caso contrario, ese peso será la unidad (para redes no ponderadas). La *distancia geodésica* entre esos nodos es la longitud del camino geodésico, o dicho de otro modo, la mínima longitud entre ellos. Cabe la posibilidad que la red tenga más de un camino geodésico entre dos nodos. En el extremo opuesto, si dos nodos no están conectados entre sí, se considera que la distancia geodésica es ∞ .

Existe un algoritmo conocido como Algoritmo de Bellman-Ford (1958) con el que se puede calcular la distancia geodésica entre dos nodos de forma eficiente. Este algoritmo admite pesos negativos. Sin embargo, para el caso en el que todos los pesos sean positivos, el Algoritmo de Dijkstra (1959) es más eficiente que el anterior, aplicado a redes no dirigidas ponderadas positivas formadas por una única componente conexa (concepto explicado más adelante). Este algoritmo proporciona la distancia mínima de cada nodo a un nodo dado k_o .

2.6. Coeficiente de Clustering

El *coeficiente de clustering*, $\langle C \rangle$, representa la proporción media de pares de vecinos de un nodo que son a su vez vecinos entre sí.

Por tanto, el coeficiente de clustering de la red es el promedio de los coeficientes de clustering de todos los nodos, tomando el valor 1 cuando la red es completa, estando todos los nodos conectados entre sí. En caso contrario, para una red aleatoria con N nodos, el valor $\langle C \rangle$ será igual a $1/N$.

$$\langle C \rangle = \frac{1}{N} \sum_{i=1}^p C_i$$

2.7. Centralidad

El concepto de *centralidad* se comenzó a poner en práctica durante los años 40 en el ámbito sociológico en el Grupo Networks Laboratory del Massachusetts Institute of Technology. Sin embargo, con los años se empleó en conceptos dispares: en las áreas geográficas, políticas económicas o industriales. Al tratarse de un concepto complejo que, como se observa, se aplica a varias áreas, se considera que un elemento es central si verifica las siguientes propiedades:

- Los elementos centrales de una red son aquéllos con mayor grado de adyacencia, es decir, los que más relacionados están con otros nodos.
- Son aquéllos que pertenecen a la mayor cantidad de caminos geodésicos posibles entre dos nodos cualesquiera de esa misma red.
- Son los que están más cerca del resto de nodos, minimizando la suma de la distancia geodésica al resto de puntos de la red.

Sin embargo, con estas tres propiedades no se termina de definir la centralidad de un nodo. La centralidad de un nodo es función de, por un lado, la cantidad de nodos a los que está conectado, y por otro, a la centralidad de los nodos a los que está conectado.

Otro concepto relevante es la *centralidad de intermediación* (denominado en inglés, *betweenness*). Expresa el número de caminos geodésicos entre dos nodos cualesquiera que pasan por un nodo dado respecto al número de geodésicas que unen esos dos nodos.

Mediante la *centralidad de cercanía* (en inglés, *closeness*), se puede calcular la distancia geodésica entre un nodo y todos los demás de la red. Es decir, la centralidad de cercanía de un nodo hace referencia a que cuanto más cercano es un nodo al resto, más popular será.

Otra medida en arcos dirigidos es la de *centralidad de centro/autoridad* (en inglés, *hub/authority*) por la que, para cada nodo, se contabilizan los nodos de la red que se conectan a ese nodo a través de un camino, penalizando según la distancia geodésica que tengan. Serán nodos *hub* aquellos nodos con muchos arcos de salida, y nodos *authority* los nodos con muchos arcos de entrada. De este modo, para un mismo concepto se obtienen dos medidas, según sea *hub* o *authority*.

2.8. Vulnerabilidad

El análisis de la *vulnerabilidad* de una red se realiza a partir de una red de gran dimensión con sus nodos conectados. A continuación, se procede a eliminar los nodos de forma aleatoria y sucesiva, uno en cada unidad temporal. Tras eliminar un nodo, también se eliminan todas sus conexiones con otros nodos, lo que supone un posible aumento de la distancia entre ellos, así como la distancia geodésica de la red. Tras varios periodos de tiempo, si se había partido de una red completamente conectada, es habitual que se termine teniendo varias redes más pequeñas conectadas entre sí por pocos, o incluso un solo enlace (ver Figura 6).

Se dice que la conectividad de la red es *robusta*, poco vulnerable, si tras eliminar mucho de los nodos que tiene la red, la componente conexa grande sigue siendo la misma y no ha sido eliminada. Este es el caso que se da en las redes libres de escala. En este sentido, la robustez, o también denominado *resiliencia*, es la capacidad de mantener la red su estructura o funcionalidad ante fallos y ataques. Un fallo es aquel que afecta a nodos y arcos de forma aleatoria, siendo un ataque el fallo intencionado según la criticidad que tenga el nodo seleccionado.

Sin embargo, las redes libres de escala son sensibles, pues al eliminar ciertos nodos específicos, la red queda muy dividada como si fueran pequeñas islas aisladas. Éste es el segundo caso ante el que nos podemos encontrar: que la red sea frágil por la distribución no homogénea de sus grados.

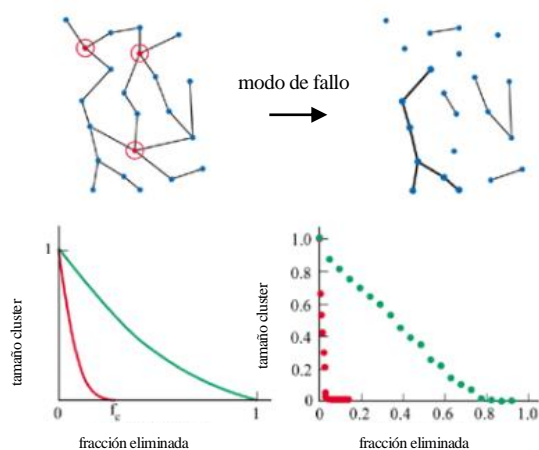


Figura 6: Vulnerabilidad

2.9. Análisis de Riesgo

En el estudio de redes complejas, es relevante el concepto de *análisis de riesgos* pues, se ha visto que si se tiene una red completamente relacionada entre sí, el fallo en un nodo afectará al resto con los que esté conectado de una forma más o menos perjudicial según la relación que tengan. Este concepto se tratará en un capítulo posterior aplicándolo a la situación actual causada por el virus SARS-CoV-2.

3 PRINCIPALES ESTADÍSTICOS DE LA RED

En este capítulo se van a identificar y explicar los principales estadísticos que caracterizan las redes complejas que permitirán comparar las redes mostradas en el Capítulo 4.

3.1. Grado Medio

El *grado* del nodo x , $gr(x_i)$, hace referencia al número de vecinos que tiene. Por tanto, el *grado medio* de un nodo, $\langle gr \rangle$, (denominado en inglés, *average degree*) es:

- En redes dirigidas:

$$\langle gr \rangle = \frac{1}{N} \sum_{i=1}^N gr_i^{entrada} = \frac{1}{N} \sum_{i=1}^N gr_i^{salida} = \frac{L}{N}$$

- En redes no dirigidas:

$$\langle gr \rangle = \frac{1}{N} \sum_{i=1}^N gr(x_i) = \frac{2L}{N}$$

donde N representa el número de nodos de la red y L representa el número de enlaces de la red.

3.2. Índice de Centralización de grado

El *índice de centralización*, IC_g (en inglés, *degree centralization index*) mide el grado de centralización que tiene la red, siendo su valor máximo igual a 1, si un solo nodo está conectado al resto que no tienen conexiones entre sí; su valor mínimo es 0, correspondiendo a una red regular donde todos los nodos tienen el mismo grado.

$$IC_g = \frac{1}{(N-1)(N-2)} \sum_{i=1}^N (gr^{max} - gr(x_i))$$

- Índice de Centralización de grado de entrada (en inglés, *degree in-centralization index*):

$$IC^{entrada} = \frac{\sum_i (gr^{entrada,max} - gr_i^{entrada})}{(N-1)^2}$$

- Índice de Centralización de grado de salida (en inglés, *degree out-centralization index*):

$$IC^{salida} = \frac{\sum_i (gr^{salida,max} - gr_i^{salida})}{(N-1)^2}$$

3.3. Densidad

La *densidad* (ρ) de la red (denominado en inglés, *density*) representa el grado de conectividad a nivel global de la red.

- En redes dirigidas: $\rho = \frac{L}{N(N-1)}$
- En redes no dirigidas: $\rho = \frac{2L}{N(N-1)}$

3.4. Conectividad

La *conectividad* (en inglés, *connectedness*) representa la capacidad de la red para conectar en una red diferentes nodos mediante distintas rutas. Tiene relación con la vulnerabilidad de la red, puesto que ante ataques sucesivos, el tamaño del subgrafo conectado más grande se va reduciendo, tras eliminar los arcos que conectan los nodos. Se puede expresar como:

$$\text{Conectividad} = 1 - \text{Fragmentación}$$

3.5. Fragmentación

La *fragmentación* (en inglés, *fragmentation*) representa la proporción de pares de nodos que son inalcanzables por otros nodos.

$$F = 1 - \frac{2 \sum_{i>j} r_{ij}}{N(N-1)}$$

Donde $r_{ij}=1$ si el nodo i puede alcanzar al nodo j por algún camino de cualquier longitud, siendo $r_{ij}=0$, en caso contrario.

3.6. Distancia Media

La *distancia* entre un nodo i y otro nodo j , d_{ij} hace referencia al número de arcos que hay entre ellos en el camino más corto que los separa. Por lo tanto, la *distancia media* (en inglés, *average distance*) representa la distancia media desde el nodo i al resto de nodos.

- En redes dirigidas:

$$\langle d \rangle = \frac{1}{2 L_{max}} \sum_{j \neq i}^N d_{ij}$$

- En redes no dirigidas:

$$\langle d \rangle = \frac{1}{L_{max}} \sum_{j>1}^N d_{ij}$$

donde L_{max} representa el número máximo de enlaces de la red.

3.7. Diámetro

El *diámetro*, D , (en inglés, *diameter*) supone el número de arcos que separan los dos nodos más alejados, es decir, la mayor distancia que existe entre dos nodos de la red conectados por un arco geodésico. Las redes con mayor diámetro son más densas, existiendo menos caminos entre cada par de nodos.

- En redes dirigidas:

$$D = \max_{j \neq i} d_{ij}$$

- En redes no dirigidas:

$$D = \max_{j>i} d_{ij}$$

3.8. Compacidad

La *compacidad* (denominado en inglés, *compactness*) supone la media de todas las distancias entre dos nodos conectados entre sí, es decir, de distancias recíprocas.

3.9. Amplitud

El estadístico *amplitud* (denominado en inglés, *breadth*) se define como:

$$\textit{Amplitud} = 1 - \textit{Compacidad}$$

3.10. Page Rank

Los modelos PageRank (Brin y Page, 1998) son usados, de forma simplificada, como el comportamiento de los usuarios que conforman una red. Su explicación es la siguiente: partiendo de un “surfero aleatorio” (en inglés, “*random surfer*”) al que se le ha dado una página web aleatoria, éste clickea links de forma continuada y aleatoria, sin tener la posibilidad de ir hacia atrás. Tras un tiempo, se aburrirá y comenzará por otra página aleatoria (damping factor). De esta forma, PageRank, o random surfer de una página web, es la posibilidad de que el surfero clickee en una página concreta de forma aleatoria.

Por lo tanto, PageRank genera un ranking del número de enlaces que recibe la página web, así como la importancia de las páginas web que la enlazan. El algoritmo de PageRank se puede describir como:

$$PR(A) = (1 - d) + d \sum_{i=1}^n \frac{PR(i)}{C(i)}$$

donde:

- $PR(A)$: corresponde al PageRank de la página A .
- d : representa un factor de amortiguación con valor $[0, 1]$, con valor, usualmente, de 0,85.
- $PR(i)$ hace referencia a los valores PageRank que tiene cada página i que enlaza a la página A .
- $C(i)$ representa el número total de enlaces salientes que tiene cada página i , aún no enlazando con la página A .

4 ANÁLISIS DE LA MOVILIDAD EN ESPAÑA. CARACTERÍSTICAS

En el Capítulo 4 se analiza la estructura de los flujos de movilidad en el marco autonómico, a partir de los valores de los datos publicados por el Instituto Nacional de Estadística (INE) en sus encuestas anuales durante el periodo desde el año 2015 a 2019.

Asimismo, dicho análisis proporcionará una visión más normalizada de la situación previa al SARS-CoV-2, que ha provocado la paralización de todos los movimientos a nivel nacional e internacional.

4.1. La Movilidad en España

El movimiento de residentes nacionales se puede clasificar de distintas formas: concepto turístico (viajes, pernoctaciones y/o duración media de los viajes), destino principal (nacional o internacional), alojamiento principal, motivo principal, transporte principal, forma de organización del viaje, duración del viaje, Comunidad Autónoma (C. A) de residencia y periodo (Figura 7).

Seleccione valores a consultar

Concepto turístico

Viajes
Pernoctaciones
Duración media de los viajes

Seleccionados: 0 Total: 3

Destino principal

España
01 Andalucía
02 Aragón
03 Asturias, Principado de
04 Balears, Illes
05 Canarias

Seleccionados: 0 Total: 30

Alojamiento principal

Total alojamientos
De mercado
Hotelero
Vivienda en alquiler
Camping
Casa rural

Seleccionados: 0 Total: 12

Motivo principal

Total motivos
Negocios y otros motivos profesionales
Congresos, ferias y convenciones
Desplazamiento al centro de trabajo
Otros motivos profesionales
Motivos personales

Seleccionados: 0 Total: 22

Transporte principal

Total transporte
Transporte aéreo
Transporte marítimo
Vehículo propio
Transporte por autobús
Transporte por ferrocarril

Seleccionados: 0 Total: 7

Forma de organización del viaje

Total viajes
Con paquete turístico
Sin paquete turístico

Seleccionados: 0 Total: 3

Duración del viaje

Total noches
De 1 a 3 noches
De 4 a 7 noches
De 8 a 15 noches
Más de 15 noches

Seleccionados: 0 Total: 5

Comunidad autónoma de residencia

Total
01 Andalucía
02 Aragón
03 Asturias, Principado de
04 Balears, Illes
05 Canarias
06 Canarias

Seleccionados: 0 Total: 20

Periodo

2020
2019
2018
2017
2016
2015

Seleccionados: 0 Total: 6

Figura 7: Clasificación de la movilidad (INE)

En este trabajo se considera que los nodos de la red están representados por las distintas 17 Comunidades Autónomas, además de las 2 Ciudades Autónomas (Figura 8), y los arcos representarán los flujos de movimiento, emisores o receptores, según el sentido de la flecha del arco.



Figura 8: Mapa Político de España

Como se comentó en secciones anteriores, se puede considerar la red de dos formas distintas, ya sea asignándole un peso a los arcos o no. Una posible versión con pesos considera el número de personas que realizan esos viajes. En la Tabla 1 se muestra el total de población por CC. AA. en el año 2019.

Tabla 1: Población por Comunidad Autónoma (2019)

Posición	Comunidad Autónoma	Total habitantes	% sobre total
1	Andalucía	8.414.240	17,9%
11	Aragón	1.319.291	2,8%
14	Asturias	1.022.800	2,2%
12	Baleares	1.149.460	2,4%
8	Canarias	2.153.389	4,6%
16	Cantabria	581.078	1,2%
6	Castilla y León	2.399.548	5,1%
9	Castilla-La Mancha	2.032.863	4,3%
2	Cataluña	7.675.217	16,3%
4	Valencia	5.003.769	10,6%
13	Extremadura	1.067.710	2,3%
5	Galicia	2.699.499	5,7%
3	Madrid	6.663.394	14,2%
10	Murcia	1.493.898	3,2%
15	Navarra	654.214	1,4%
7	País Vasco	2.207.776	4,7%
17	La Rioja	316.798	0,7%
19	Ceuta	84.777	0,2%
18	Melilla	86.487	0,2%
	Total población	47.026.208	

Fuente: Instituto Nacional de Estadística

Se aprecia como la C. A con un mayor número de habitantes es Andalucía, seguido de Cataluña y Madrid. En el caso contrario, las CC. AA. con menor número de habitantes son las Ciudades Autónomas de Ceuta y Melilla, además de la Comunidad Autónoma de La Rioja.

4.2. Movimiento entre Comunidades Autónomas

Considerando el año 2019 como el último periodo completo con ausencia de cambios bruscos relativos a los datos en la movilidad, se puede afirmar que en ese año, la mayoría de los viajes de los españoles se realizó dentro de nuestro país (casi el 90% del total de desplazamientos). Como es habitual, en gran medida los viajes se realizaron durante periodo estival, principalmente durante el mes de agosto, representando un 25% del total de desplazamientos (Figura 9).

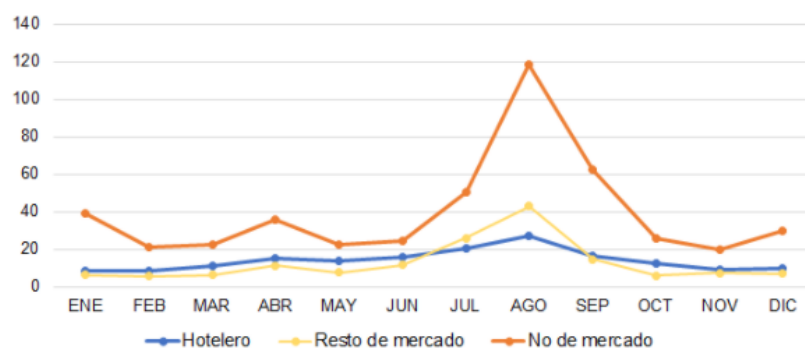


Figura 9: Pernotaciones de residentes por tipo de alojamiento (millones) (2019)

Fuente: Instituto Nacional de Estadística

En la Tabla 2 se muestra los desplazamientos que tuvieron como destino cada una de las Comunidades Autónomas durante el año 2019. La C. A. que recibió mayor número de desplazamientos fue Madrid (18,2)%, seguida de la Cataluña (17,3)% y Andalucía (15,3)%; con un peso algo menor, Valencia (9,2)%. La Comunidad Autónoma con menor entrada de turistas españoles fueron: La Rioja (0,7)%, seguido de las Ciudades Autónomas de Ceuta (0,2)% y Melilla (0,1)%.

Tabla 2: Total viajes por Comunidad Autónoma (2019).

Posición	Comunidades Autónomas	Viajes	% sobre total
3	Andalucía	29.728.070	15,3%
10	Aragón	6.690.201	3,5%
12	Asturias	4.441.749	2,3%
14	Baleares	3.622.222	1,9%
9	Canarias	6.931.553	3,6%
16	Cantabria	2.133.001	1,1%
5	Castilla y León	11.297.345	5,8%
8	Castilla-La Mancha	8.422.734	4,3%
2	Cataluña	33.448.057	17,3%
4	Valencia	17.838.521	9,2%
13	Extremadura	3.842.266	2,0%
7	Galicia	9.854.554	5,1%
1	Madrid	35.343.853	18,2%
11	Murcia	4.691.885	2,4%
15	Navarra	2.973.847	1,5%
6	País Vasco	10.742.903	5,5%
17	La Rioja	1.310.083	0,7%
18	Ceuta	325.917	0,2%
19	Melilla	235.956	0,1%
	Total	193.874.716	

Fuente: Instituto Nacional de Estadística

En la Tabla 3 se detallan los desplazamientos según el trimestre del año 2019. De acuerdo a lo comentado previamente, además de lo observado en la Figura 9, se comprueba como la mayoría de las Comunidades Autónomas y Ciudades Autónomas reciben un mayor número de desplazamientos en el tercer trimestre del año 2019, que transcurre desde el 1 de julio hasta finales de septiembre.

Tabla 3: Total viajes por Comunidad Autónoma desglosado (2019)

Comunidad Autónoma	2019T4	2019T3	2019T2	2019T1
Andalucía	6.514.810	9.438.764	6.860.177	6.914.318
Aragón	1.420.515	2.146.013	1.649.399	1.474.274
Asturias	752.366	1.330.060	1.296.103	1.063.221
Baleares	945.883	878.690	950.381	847.269
Canarias	1.447.914	2.295.291	1.617.000	1.571.347
Cantabria	515.067	576.134	534.243	507.558
Castilla-León	2.492.339	3.257.439	2.952.066	2.595.502
Castilla la Mancha	1.840.757	2.702.369	2.012.052	1.867.557
Cataluña	6.759.009	10.077.851	9.270.724	7.340.473
Valencia	3.500.214	5.382.263	5.084.857	3.871.187
Extremadura	834.931	1.115.523	1.007.975	883.836
Galicia	2.380.547	2.896.244	2.459.761	2.118.001
Madrid	6.837.642	12.688.902	9.305.578	6.511.731
Murcia	984.701	1.650.059	1.183.365	873.760
Navarra	636.536	866.699	755.010	715.601
País Vasco	2.236.843	3.238.724	3.022.614	2.244.722
La Rioja	310.870	416.848	339.589	242.776
Ceuta	55.888	100.952	100.259	68.819
Melilla	45.053	57.646	73.238	60.019
Total	40.511.885	61.116.470	50.474.390	41.771.971

Fuente: Instituto Nacional de Estadística

La información de la Tabla 2 y la Tabla 3 representan el total de viajes realizados por cada Comunidad Autónoma, sin distinguir si dichos viajes se han realizado dentro de la misma C. A. o fuera. A continuación se muestran el número de viajes que se han realizado a otras CC. AA., que, como se puede observar, dicha cifra disminuye considerablemente pues, en este sentido, no se tienen en cuenta los desplazamientos entre las provincias de una misma Comunidad Autónoma (Tabla 4).

Tabla 4: Total viajes por Comunidad Autónoma a otra Autonomía (2019).

Posición	Comunidad Autónoma emisora	Viajes	% sobre total
5	Andalucía a otra Comunidad Autónoma	5.440.461	6,5%
8	Aragón a otra Comunidad Autónoma	3.120.498	3,8%
11	Asturias a otra Comunidad Autónoma	2.055.153	2,5%
14	Baleares a otra Comunidad Autónoma	1.400.337	1,7%
15	Canarias a otra Comunidad Autónoma	1.332.791	1,6%
16	Cantabria a otra Comunidad Autónoma	1.216.447	1,5%
7	Castilla y León a otra Comunidad Autónoma	5.039.465	6,1%
6	Castilla-La Mancha a otra Comunidad Autónoma	5.231.884	6,3%
2	Cataluña a otra Comunidad Autónoma	7.591.304	9,1%
4	Comunidad Valenciana a otra Comunidad Autónoma	5.976.019	7,2%
12	Extremadura a otra Comunidad Autónoma	1.933.575	2,3%
10	Galicia a otra Comunidad Autónoma	2.170.488	2,6%
1	Madrid a otra Comunidad Autónoma	27.626.345	33,3%
9	Murcia a otra Comunidad Autónoma	2.657.189	3,2%
13	Navarra a otra Comunidad Autónoma	1.671.093	2,0%
3	País Vasco a otra Comunidad Autónoma	7.396.290	8,9%
17	La Rioja a otra Comunidad Autónoma	861.097	1,0%
18	Ceuta a otra Comunidad Autónoma	210.828	0,3%
19	Melilla a otra Comunidad Autónoma	144.199	0,2%
	Total	83.075.463	

Fuente: Instituto Nacional de Estadística

Considerando este nuevo criterio, la clasificación varía. Comunidad de Madrid es la C. A. que emite un mayor número de desplazamientos (33,3)%, seguida de Cataluña (9,1)% y País Vasco (8,9)%.

4.3. Motivación de la Movilidad en España

La relevancia de las características personales y regionales son factores determinantes que afectan a los motivos por los cuales se lleva a cabo la movilidad por el país. En efecto, las razones por las que se realizan movimientos pueden clasificarse, según el INE, en:

- Negocios y otros motivos profesionales
- Congresos, ferias y convenciones
- Desplazamiento al centro de trabajo
- Otros motivos profesionales
- Motivos personales
- Ocio, recreo y vacaciones
- Turismo de sol y playa
- Turismo de naturaleza
- Turismo cultural
- Turismo gastronómico
- Turismo deportivo
- Turismo termal y de bienestar
- Otro tipo de turismo de ocio
- Visitas a familiares y amigos
- Otros motivos
- Turismo de compras
- Desplazamiento al centro de estudios habitual
- Otros motivos de educación y formación
- Salud
- Motivos religiosos o peregrinaciones
- Otros

El presente trabajo se centra en la movilidad en España, considerando como causa todos los motivos expuestos anteriormente. Además, no se ha considerado las Ciudades Autónomas de Ceuta y Melilla, por ser su información incompleta.

De este modo, la matriz de adyacencia resultante consiste en una matriz cuadrada no simétrica, pues se tiene una red dirigida, con tantas filas y columnas como número de vértices tenga la red, esto es, 17. Como ya se comentó en el Capítulo 2, cada elemento $a_{ij} = 1$ cuando haya una arista que una el vértice i con el vértice j , es decir, la Comunidad Autónoma i con la Comunidad Autónoma j . Una vez identificado el valor de los elementos de la matriz, se han considerado nulos los trayectos realizados dentro de una misma C. A, evitando de este modo los loops ($a_{ii} = 1$), de acuerdo a los datos del Anexo I (Tablas 21, 22, 23, 24, 25 y 26).

4.4. Análisis de la movilidad mediante Ucinet

4.4.1 Datos utilizados

Se va a representar gráficamente y analizar las redes resultantes haciendo uso de los datos correspondientes a los viajes realizados (Anexo I), considerando todos los motivos comentados en el apartado anterior. Por otro lado, debido a la información incompleta referente a los movimientos entre Comunidades Autónomas, no se tienen registrados algunos desplazamientos que se realizan de unas CC. AA. a otras.

En consecuencia, el contenido que a continuación se desarrolla incluye la cantidad de viajes realizados por los habitantes de las distintas Comunidades Autónomas, a excepción de Ceuta y Melilla, a otras CC. AA. suponiendo el total de alojamientos, total de motivos, total de transportes, total de viajes y total de noches (Figura 7).

Para la representación gráfica de la red y el cálculo de los estadísticos se ha utilizado la aplicación informática “Ucinet”. Además, Ucinet tiene incorporada la herramienta “NetDraw” de modo que se pueda visualizar las redes. Cabe destacar que Ucinet permite exportar directamente los datos de dicha red, sin necesidad de realizar ningún cálculo adicional.

4.4.2 Visualización de la Red 2015

En la Figura 10 se muestra la red de desplazamiento del año 2015 haciendo uso de los datos correspondientes a la Tabla 21 del Anexo I.

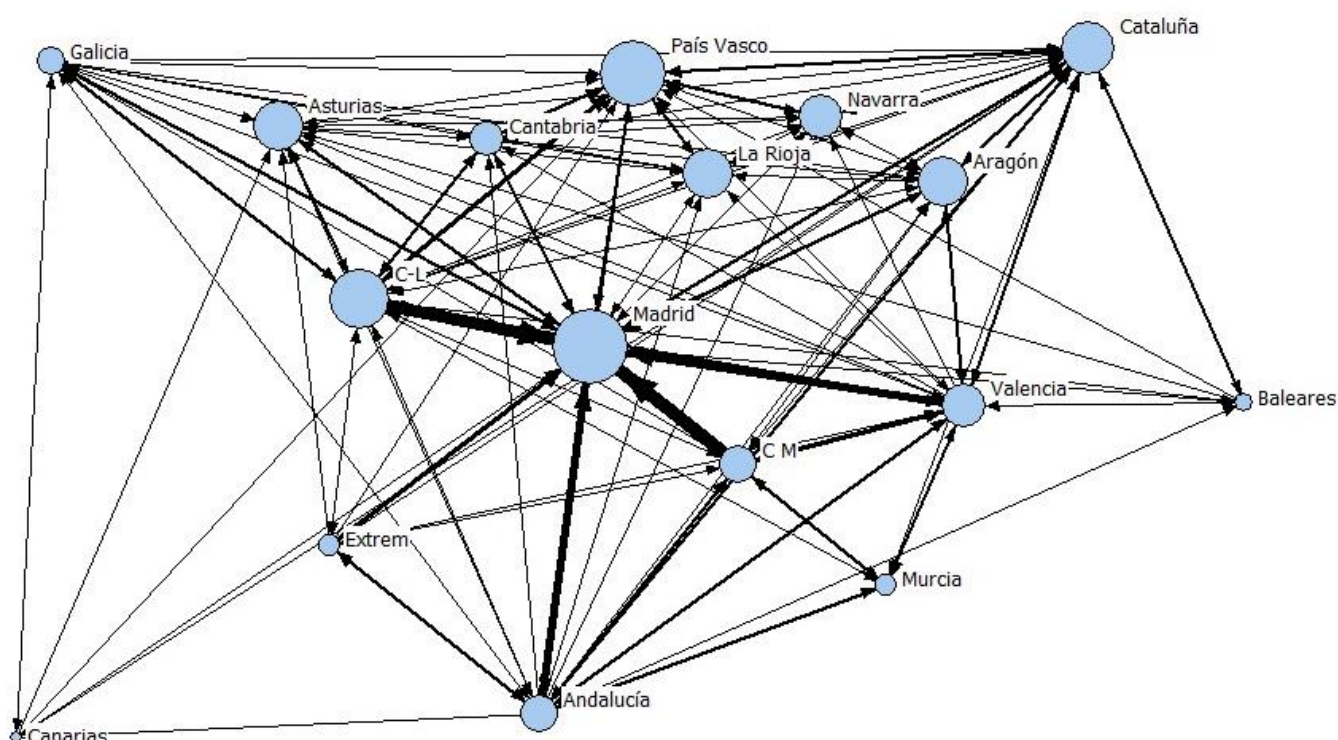


Figura 10: Visualización de la Red (2015)

Nota: El tamaño de los nodos es proporcional al número de conexiones (grado de entrada) que cada C. A. tiene con otras. El grosor de los arcos dirigidos entre nodos es en función del peso de cada C. A., es decir, representa el volumen de los desplazamientos que una C. A. recibe (denominado en inglés, *in-strength*) (a mayor grosor, mayor el número de desplazamientos que se han producido dentro de ese año considerado de la muestra).

En la Figura 10 se puede observar como la Comunidad Autónoma de Madrid es la Autonomía que recibe la mayor parte de los desplazamientos (26.431.373) siendo su procedencia de las CC. AA. de Castilla y León (supone un 21,7%), Castilla-La Mancha (24,5%) y Andalucía (12,3%), principalmente. No se observa una gran diferencia en los arcos dirigidos del resto de Comunidades Autónomas. País Vasco es la segunda C. A. con mayor entrada de desplazamientos (6.794.606) desde Castilla y León (22,2% de los mismos) y Cantabria (18%), primordialmente.

Por otra parte, Madrid, seguida del País Vasco, son las CC. AA. receptoras de desplazamientos de un mayor número de Comunidades Autónomas (16 y 14, respectivamente).

4.4.3 Visualización de la Red 2016

En la Figura 11 se obtiene la red resultante haciendo uso de los valores en el año 2016.

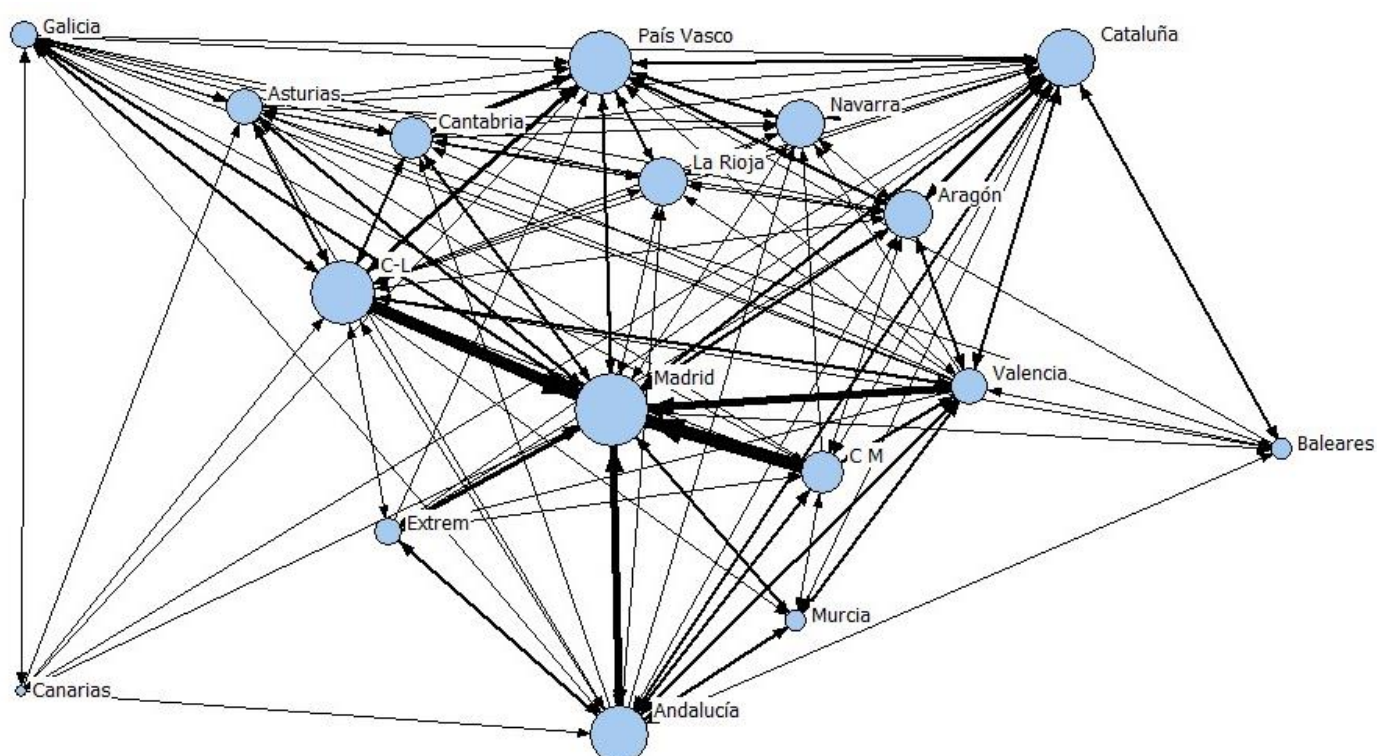


Figura 11: Visualización de la Red (2016)

En la Figura 11 se observa como la Comunidad Autónoma de Madrid es la que recibe nuevamente la mayor parte de los movimientos (27.143.294), principalmente desde Castilla y León (23,5%), Castilla-La Mancha (24,9%), Comunidad Valenciana (12,4%) y Andalucía (12,7%). No se observa una gran diferencia en los arcos dirigidos del resto de Comunidades Autónomas. No obstante, Cataluña resulta ser la segunda C. A. que sigue a la Comunidad de Madrid, con 7.366.710 desplazamientos.

Por otra parte, Madrid, seguida de Cataluña y Castilla y León, son las CC. AA. receptoras de desplazamientos de un mayor número de Comunidades Autónomas (16, 14, 14, respectivamente).

4.4.4 Visualización de la Red 2017

La red resultante para el año 2017 se observa en la Figura 12.

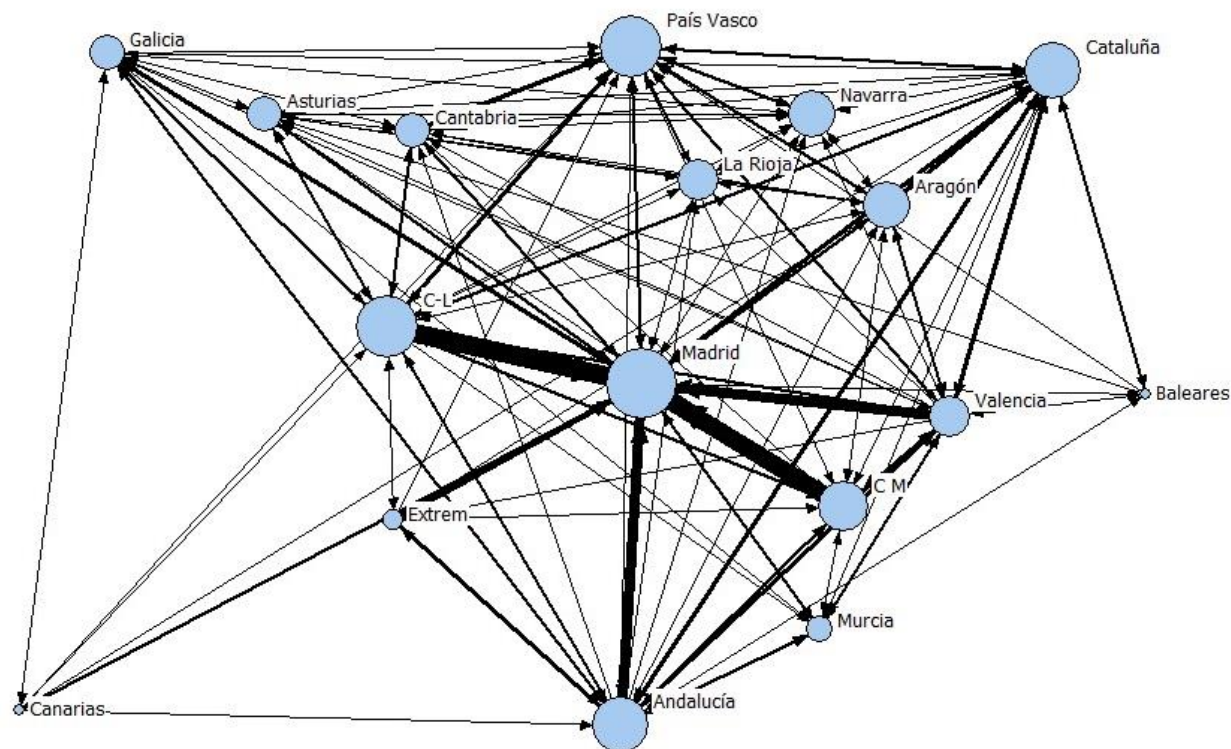


Figura 12: Visualización de la Red (2017)

En la Figura 12 se observa como la C. A. de Madrid es la que recibe la mayor parte de los desplazamientos (29.216.584), desde Castilla y León (supone el 24,1%), Castilla-La Mancha (22,8%), Comunidad Valenciana (12,8%) y Andalucía (12,4%). No se observa una gran diferencia en los arcos dirigidos del resto de Comunidades Autónomas. País Vasco es la segunda C. A. receptora de un mayor número de desplazamientos, 7.292.915.

Por otra parte, Madrid, seguida de País Vasco y Castilla y León, son las CC. AA. receptoras de desplazamientos de un mayor número de Comunidades Autónomas (16, 14, 14, respectivamente).

4.4.5 Visualización de la Red 2018

En la Figura 13 se expone el caso para el año 2018.

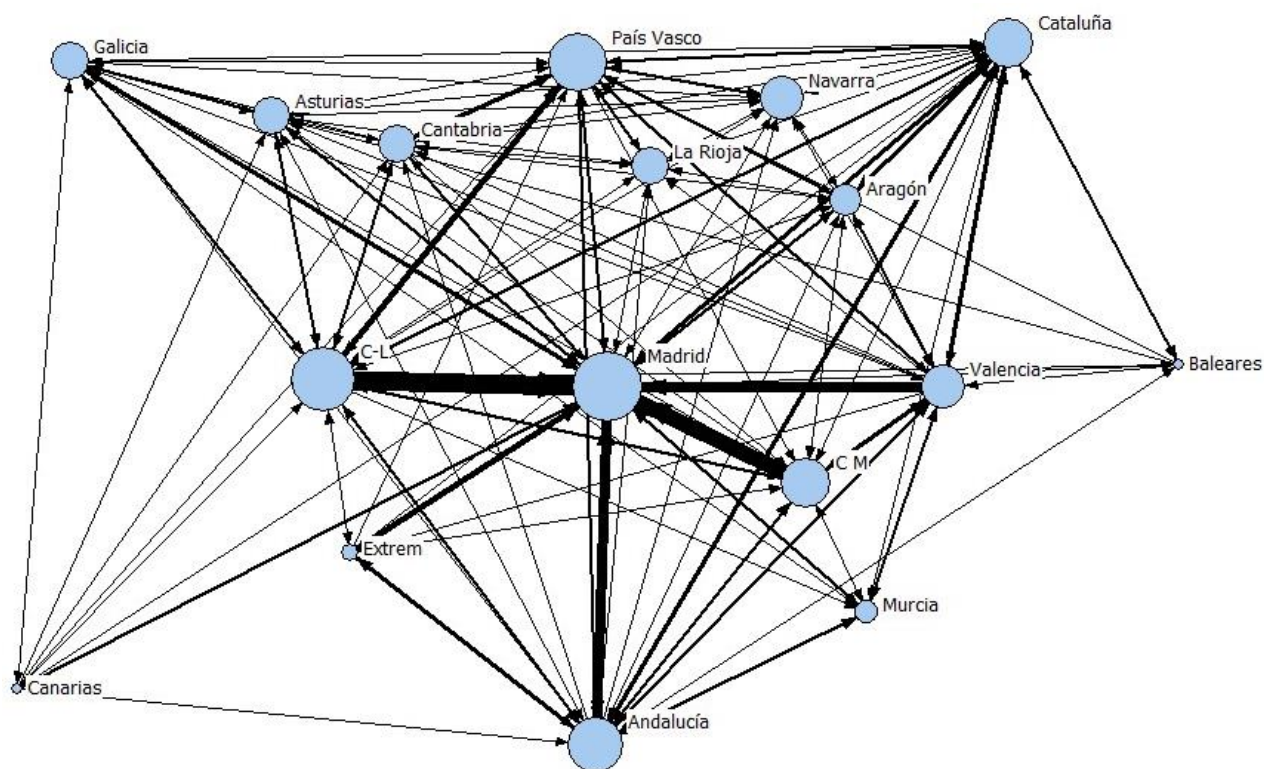


Figura 13: Visualización de la Red (2018)

En la Figura 13 se observa como la Comunidad de Madrid es la C. A. que recibe la mayoría de los desplazamientos (27.820.135), cifra que se ha visto decrementada respecto al año anterior, recibiendo dichos desplazamientos principalmente desde Castilla y León (representa el 24,8%) y Castilla-La Mancha (21%). Cataluña es la segunda C. A. receptora de un mayor número de desplazamientos, 7.358.475.

Por otra parte, Madrid, seguida de Castilla y León, son las CC. AA. receptoras de desplazamientos de un mayor número de Autonomías (16 y 15, respectivamente).

4.4.6 Visualización de la Red 2019

La red que se va a exponer a continuación corresponde a los valores últimos disponibles proporcionados por el Instituto Nacional de Estadística que representa los viajes entre Comunidades Autónomas considerando el total de motivos por los que se viaja, el total de alojamientos posibles, el total de transportes usados para la realización de dichos viajes y el total de noches que se han hospedado los viajeros.

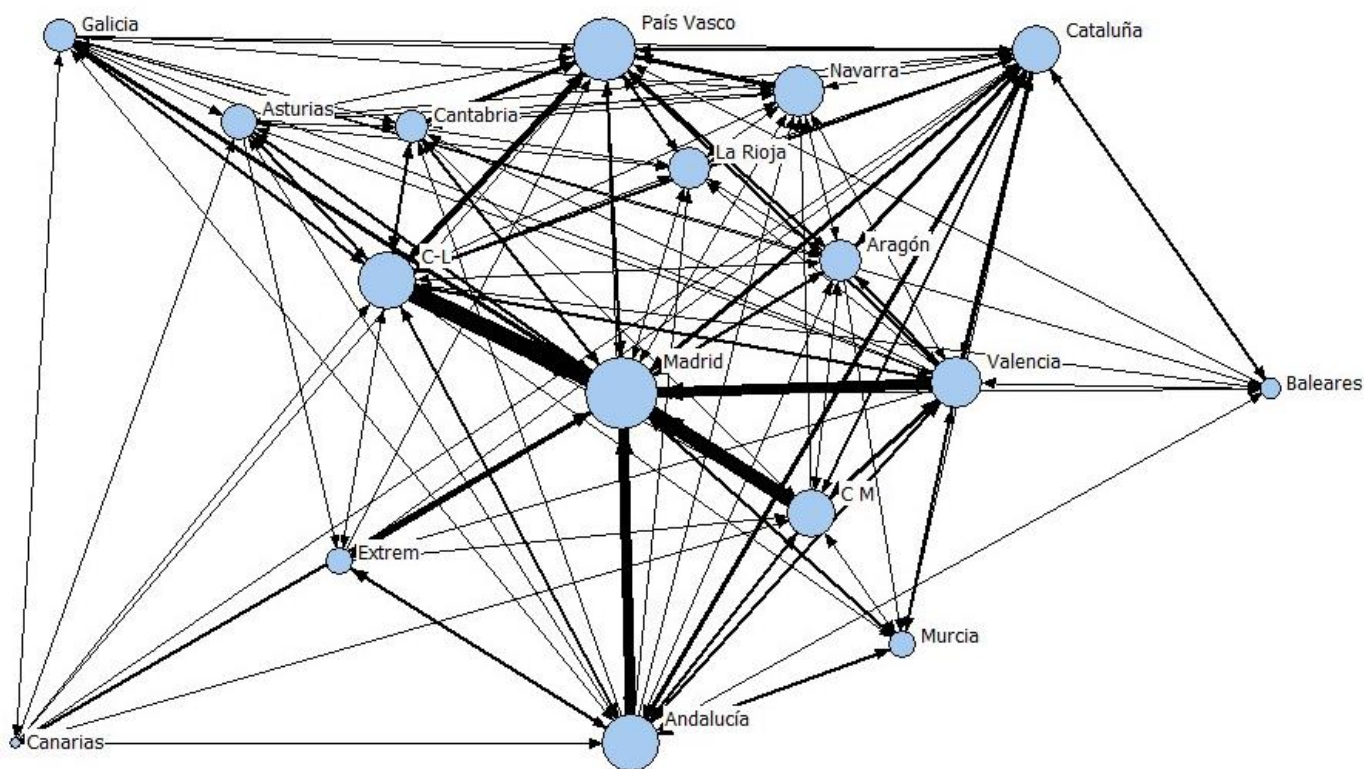


Figura 14: Visualización de la Red (2019)

En la Figura 14 se observa como la C. A. de Madrid es la que recibe la mayor parte de los desplazamientos (27.619.435), desde Castilla y León (23,9%), Castilla-La Mancha (19,4%), Comunidad Valenciana (13,2%) y Andalucía (13,3%), principalmente. No se observa una gran diferencia en los arcos dirigidos del resto de Comunidades Autónomas. País Vasco es la segunda C. A. receptora de un mayor número de desplazamientos, 7.323.373.

Por otra parte, Madrid, seguida de País Vasco, son las CC. AA. receptoras de desplazamientos de un mayor número de Autonomías (16 y 14, respectivamente).

Realizando una comparativa de las redes estudiadas, se puede observar los cambios que se han ido produciendo según ha avanzado el tiempo.

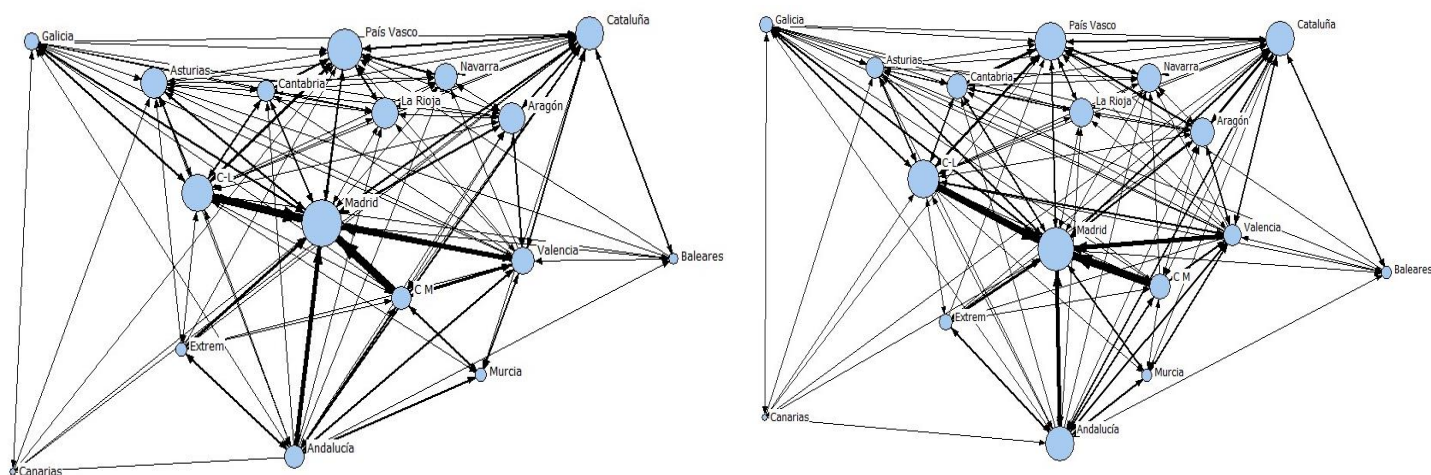


Figura 15: Visualización Red 2015 y 2016 (comparativa)

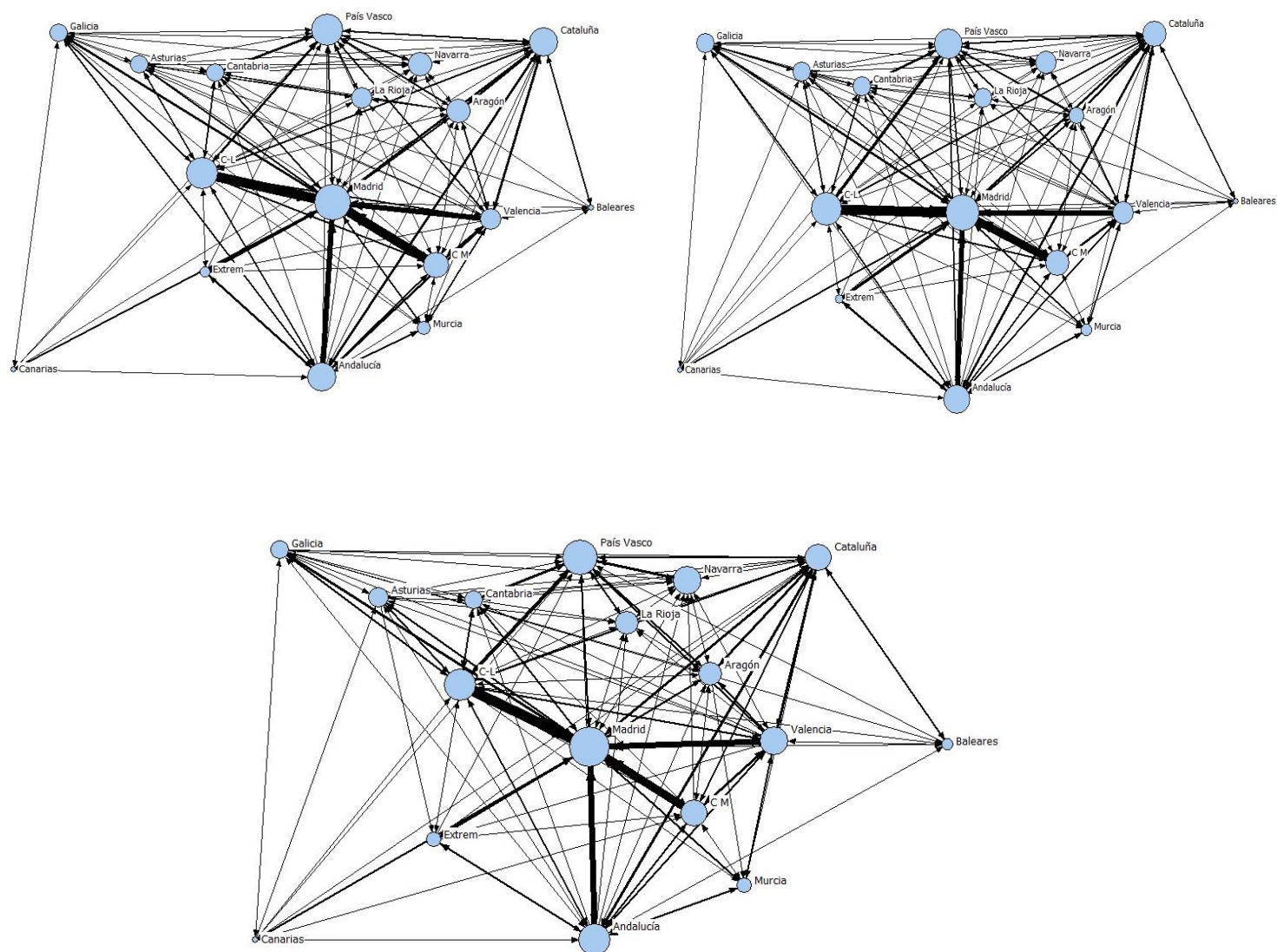


Figura 16: Visualización Red 2017, 2018 y 2019 (comparativa)

4.4.7 Visualización de la Red Promedio (2015-2019)

En la Figura 17 se representa gráficamente la red de los desplazamientos promedio durante el periodo 2015-2019, que puede resultar ser más completa que las mostradas anteriormente, ya que se suavizan las variaciones anuales de los flujos.

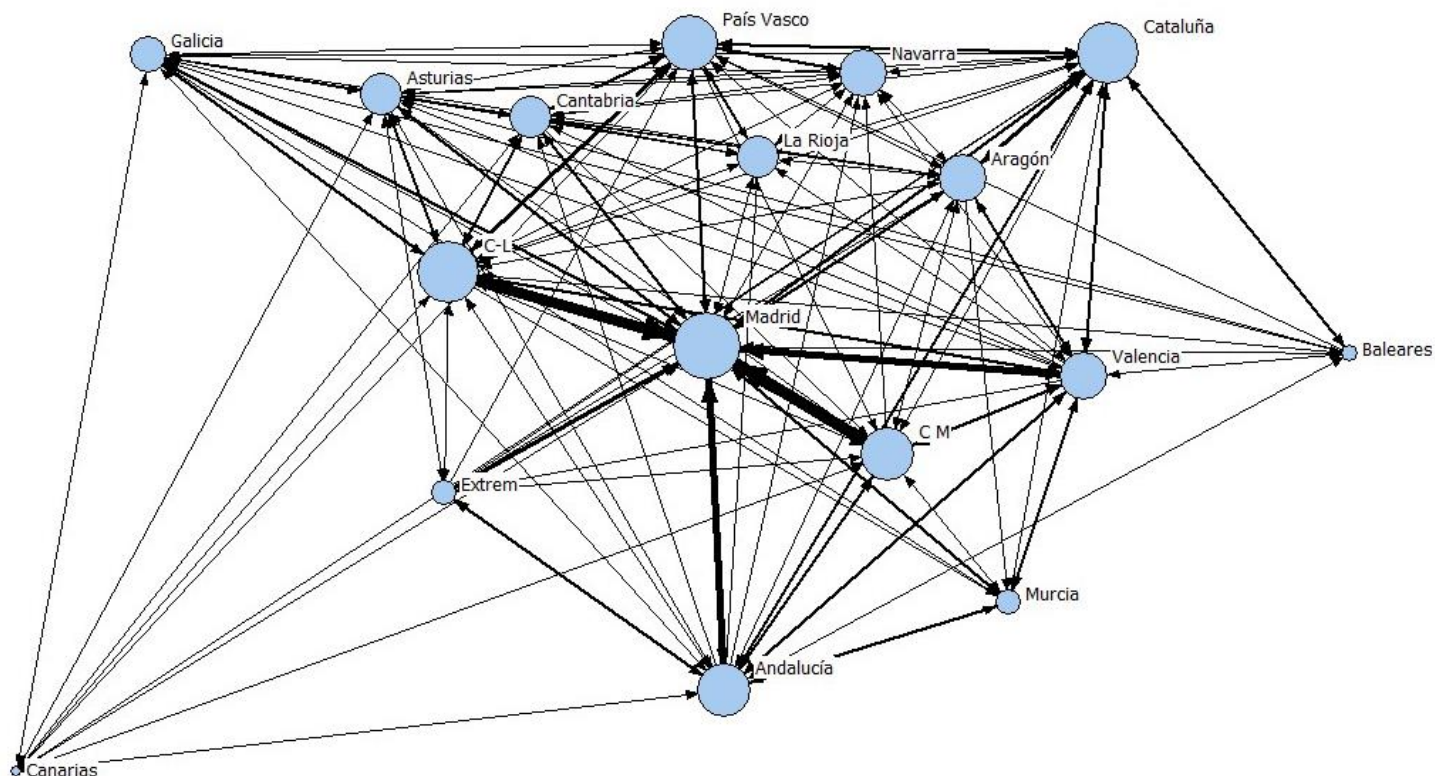


Figura 17: Visualización de la Red (Valores Promedio)

En la Figura 17 se observa como la C. A. de Madrid es la que recibe la mayor parte de los desplazamientos (27.646.164), principalmente desde Castilla y León (23,6%), Castilla-La Mancha (22,5%), Comunidad Valenciana (12,8%) y Andalucía (12,7%). No se observa una gran diferencia en los arcos dirigidos del resto de Comunidades Autónomas. País Vasco es la segunda C. A. receptora de un mayor número de desplazamientos, 7.165.538.

Por otra parte, Madrid, seguida de Cataluña y Castilla y León, son las CC. AA. receptoras de desplazamientos de un mayor número de Comunidades Autónomas (16, 15 y 15, respectivamente).

4.5. Comentarios adicionales

Una vez identificado todas las redes correspondientes a los periodos que comprenden de 2015 a 2019, junto con la red relativa a los valores promedio de éstas, se puede comprobar como, desde el año 2015 en adelante, las Comunidades Autónomas de Castilla-La Mancha y Castilla y León tienen un flujo abundante hacia la capital de España. Esto puede explicarse debido a la gran cantidad de primeras y segundas residencias que hay en estas CC. AA. pertenecientes a los habitantes de Madrid.

Además, se observa como el flujo receptor y emisor del nodo correspondiente a la Comunidad Autónoma de Madrid va aumentando según avanzan los años, es decir, hay una mayor concentración de los desplazamientos hacia la C. A. de Madrid.

Respecto a las CC. AA. de Canarias y Baleares, la movilidad que reciben desde otras Comunidades Autónomas es principalmente por motivos de ocio y turismo. Este hecho se respalda con la información obtenida del Ministerio de Transportes, Movilidad y Agenda Urbana el cual señala como las principales rutas en España, para el año 2012, las siguientes (Figura 18).



Figura 18: Principales rutas por avión (2012)

Hay que señalar además del papel impulsor de movimiento la figura que ha tenido Renfe Alta Velocidad Española (AVE) en España. La red de Alta Velocidad Española, así como Alvia o Avant, pertenecientes a la misma operadora (Renfe), han hecho posible viajar a gran velocidad por diferentes trayectos a lo largo del país. La ruta actual de los trayectos que realiza se muestra en la Figura 19.



Figura 19: Trayectos de AVE en España

Se puede observar una clara similitud entre los servicios ferroviarios de alta velocidad y la evolución mostrada haciendo uso de las redes obtenidas mediante la aplicación informática Ucinet. En este sentido, Renfe ha permitido una conexión entre Castilla-La Mancha, Castilla y León, Andalucía, Comunidad Valenciana, Cataluña y Madrid.

4.6. Estadística de la red de desplazamiento (2015-2019)

Una vez identificada las diversas redes que comprende desde el año 2015 hasta el año 2019, junto a la red realizada con los valores promedio de éstas, se procederá a comentar los estadísticos que se obtienen haciendo uso de la aplicación informática Ucinet, a través de la herramienta “Múltiples medidas de toda la Red” (en

inglés y de acuerdo a cómo aparece en Ucinet: “*Multiple whole network measures*”) que caracterizan la topología de las redes.

Como puede comprobarse, en todas las tablas aparecerán los mismos estadísticos que se definieron en la Capítulo 3: “Principales estadísticos de la Red”. En cuanto a sus valores, éstos irán variando según el alteración de los datos que se pueden observar en las tablas del Anexo I.

Tabla 5: Resultados Red 2015

2015	
id	Estadísticos 2015
Nº nodos	17
Nº enlaces	162
Grado Medio	9,529
Í. Centralización (grado)	0,317
Í. Centralización (grado) salida	0,43
Í. Centralización (grado) entrada	0,43
Densidad	0,596
Conectividad	1
Fragmentación	0
Distancia Media	1,404
Diámetro	2
Compacidad	0,789
Amplitud	0,202

Tabla 6: Resultados Red 2016

2016	
id	Estadísticos 2016
Nº nodos	17
Nº enlaces	171
Grado Medio	10,058
Í. Centralización (grado)	0,292
Í. Centralización (grado) salida	0,395
Í. Centralización (grado) entrada	0,395
Densidad	0,629
Conectividad	1
Fragmentación	0
Distancia Media	1,371
Diámetro	2
Compacidad	0,814
Amplitud	0,186

Tabla 7: Resultados Red 2017

2017	
id	Estadísticos 2017
Nº nodos	17
Nº enlaces	172
Grado Medio	10,117
Í. Centralización (grado)	0,3
Í. Centralización (grado) salida	0,391
Í. Centralización (grado) entrada	0,391
Densidad	0,632
Conectividad	1,000
Fragmentación	0
Distancia Media	1,368
Diámetro	2
Compacidad	0,816
Amplitud	0,184

Tabla 8: Resultados Red 2018

2018	
id	Estadísticos 2018
Nº nodos	17
Nº enlaces	176
Grado Medio	10,352
Í. Centralización (grado)	0,283
Í. Centralización (grado) salida	0,375
Í. Centralización (grado) entrada	0,375
Densidad	0,647
Conectividad	1
Fragmentación	0
Distancia Media	1,352
Diámetro	2
Compacidad	0,824
Amplitud	0,176

Tabla 9: Resultados Red 2019

2019	
id	Estadísticos 2019
Nº nodos	17
Nº enlaces	171
Grado Medio	10,058
Í. Centralización (grado)	0,283
Í. Centralización (grado) salida	0,395
Í. Centralización (grado) entrada	0,395
Densidad	0,629
Conectividad	1
Fragmentación	0
Distancia Media	1,371
Diámetro	2
Compacidad	0,814
Amplitud	0,186

Tabla 10: Resultados Red Valores Promedio

Promedio	
id	Estadísticos (promedio)
Nº nodos	17
Nº enlaces	192
Grado Medio	11,294
Í. Centralización (grado)	0,208
Í. Centralización (grado) salida	0,313
Í. Centralización (grado) entrada	0,313
Densidad	0,706
Conectividad	1
Fragmentación	0
Distancia Media	1,294
Diámetro	2
Compacidad	0,853
Amplitud	0,147

4.6.1 Resultados de los Estadísticos de la Red 2015

Número de nodos: 17. Son todas las CC. AA. que se han considerado. Este es un dato que no se va a ver modificado en el resto de tablas resultantes al utilizar, en todos los casos y para todos los años, las mismas diecisiete Comunidades Autónomas.

Número de enlaces: 162. Todas las conexiones entre Comunidades Autónomas que tiene la red.

Grado medio: 9,529. Una Comunidad Autónoma tiene relación con una media de nueve CC. AA. distintas. En realidad esto no sucede así, puesto que todas las CC. AA. están relacionadas, produciéndose viajes entre todas ellas. Ésto resulta por la existencia de datos faltantes.

Índice de centralización de grado: 0,316. Al ser un valor lejano a 1, se trata de una red bien conectada y poco centralizada.

Índice de centralización de grado de salida: 0,43. Considerando únicamente los arcos de salida, la red sigue siendo conectada y poco centralizada

Índice de centralización de grado de entrada: 0,43. Sucede lo mismo respecto la índice de centralización de grado de salida. Para todas las redes, el índice de centralización de grado de salida y el índice de centralización de grado de entrada poseerán el mismo valor pues el grado máximo de entrada es el mismo que el grado máximo de salida.

Densidad: 0,595. Se trata de una red moderadamente conectada.

Conectividad: 1. Representa una red completamente conectada entre sí. Esto sucede puesto que todas las Comunidades Autónomas están conectadas, es decir, se dan viajes entre todas ellas.

Distancia media: 1,404. Representa la distancia media desde un nodo al resto dentro de la red.

Diámetro: 2. Supone la máxima distancia que separa los dos nodos más alejados de la red. Este es un dato que no se va a ver modificado en el resto de tablas resultantes.

4.6.2 Resultados de los Estadísticos de la Red 2016

Número de enlaces: 171. Se han registrado un mayor número de viajes, ya sea por una mayor exactitud en la

muestra realizada o porque realmente se dieron más viajes en 2016.

Grado medio: 10,058. Una Comunidad Autónoma tiene relación con una media de diez CC. AA. distintas. De nuevo, este estadístico se ha visto incrementado respecto al año anterior.

Índice de centralización de grado: 0,291. Se considera una red bien conectada y poco centralizada.

Índice de centralización de grado de salida e Índice de centralización de grado de entrada: 0,395. La red es regular, considerando únicamente los grados de entrada y salida.

Densidad: 0,628. Se trata de una red moderadamente conectada, levemente mejor que el periodo anterior.

Distancia media: 1,371. La distancia media de un nodo al resto dentro de la red se ha visto decremada respecto al periodo anterior.

4.6.3 Resultados de los Estadísticos de la Red 2017

Se procede a comentar los valores más significativos que se ven modificados de un año a otro.

Número de enlaces: 172. Van incrementando el número de viajes a medida que van pasando los años.

Grado medio: 10,117. De nuevo, una Comunidad Autónoma tiene relación con una media de diez CC. AA. distintas.

Índice de centralización de grado: 0,300. Se trata de un valor muy similar al registrado en la red del año anterior.

Índice de centralización de grado de salida e Índice de centralización de grado de entrada: 0,391. Su valor no se ha visto modificado ligeramente respecto al año 2016. Sigue siendo una red regular donde todos los nodos tienen el mismo grado de entrada y de salida.

Densidad: 0,632. Refleja una red moderadamente conectada, mejor a la del año 2015 y 2016.

Distancia media: 1,367, representando que la distancia media entre dos nodos es menor que en periodos anteriores.

4.6.4 Resultados de los Estadísticos de la Red 2018

Número de enlaces: 176. De nuevo, se ha visto incrementado el valor del número de arcos de la red.

Grado medio: 10,352. Cada Comunidad Autónoma sigue teniendo una relación con una media de diez CC. AA. distintas.

Índice de centralización de grado: 0,283. Sigue teniendo un valor muy similar al registrado en la red del año anterior.

Índice de centralización de grado de salida e Índice de centralización de grado de entrada: 0,375. Considerando los nodos de entrada y salida por separado, la red es regular, pues su valor es cercano a 0.

Densidad: 0,647. Refleja una red moderadamente conectada, mejor a la del año 2015, 2016 y 2017.

Distancia media: 1,352. Con un valor muy similar registrado en el año anterior.

4.6.5 Resultados de los Estadísticos de la Red 2019

Número de enlaces: 171. Con un valor muy similar al del año anterior.

Grado medio: 10,058. Este valor ha disminuido significativamente, manteniéndose el hecho de que una Comunidad Autónoma tiene relación con una media de diez CC. AA. distintas.

Índice de centralización de grado: 0,283. Siendo, de nuevo, un valor muy similar al obtenido el año anterior.

Índice de centralización de grado de salida e Índice de centralización de grado de entrada: 0,395. Mismo valor que en el año 2016.

Densidad: 0,628. Supone una red moderadamente conectada, aún siendo su valor inferior al de periodos anteriores.

Distancia media: 1,371. Con valor muy similar al del año anterior.

4.6.6 Resultados de los Estadísticos de la Red Promedio (2015-2019)

Número de enlaces: 192. Todos los viajes entre Comunidades Autónomas que se dan.

Grado medio: 11,294. Una Comunidad Autónoma tiene relación con una media de once CC. AA. distintas. En realidad esto no sucede así, puesto que todas las CC. AA. están relacionadas, produciéndose viajes entre todas ellas. Ésto resulta por la existencia de datos faltantes.

Índice de centralización de grado: 0,208. Al ser un valor lejano a 1, se trata de una red bien conectada y poco centralizada.

Índice de centralización de grado de salida e Índice de centralización de grado de entrada: 0,313. Haciendo uso de la tabla de valores promedio, tras considerar los arcos de entrada y salida de forma individual, se obtiene una red regular donde todos los nodos tienen el mismo grado de entrada y salida, respectivamente.

Densidad: 0,706. En valores promedio, implica una red moderadamente conectada.

Conectividad: 1. Representa una red completamente conectada entre sí. Esto sucede puesto que todas las Comunidades Autónomas están conectadas, es decir, se dan viajes entre todas ellas.

Diámetro: 2. Se trata de un valor reducido al tener un valor en la densidad elevado.

Distancia media: 1,294. La distancia media desde un nodo al resto dentro de la red supera un arco, en valores promedio, debido a la existencia de datos faltantes.

4.7. Resultados de la distribución Ley de Potencias

La relevancia de una Comunidad Autónoma en PageRank depende de los desplazamientos que ésta recibe. En la Tabla 11 se muestra el ranking de cada C. A. en función del valor que tiene su PageRank. Así, se observa como la Comunidad de Madrid, seguida de Cataluña y País Vasco son las Comunidades Autónomas más relevantes. Por el contrario, las CC. AA. menos relevantes son Baleares, Canarias y La Rioja.

Tabla 11: PageRank

Relevancia	CC. AA	PageRank
1	Madrid	0.235
2	Cataluña	0.091
3	País Vasco	0.089
4	Valencia	0.078
5	Castilla y León	0.076
6	Andalucía	0.075
7	Castilla-La Mancha	0.066
8	Aragón	0.043
9	Navarra	0.034
10	Galicia	0.032
11	Murcia	0.030
12	Asturias	0.029
13	Extremadura	0.027
14	Cantabria	0.025
15	Baleares	0.022
16	Canarias	0.021
17	La Rioja	0.019

En la Figura 20 se representa gráficamente la red en función del valor de PageRank por Comunidad Autónoma. De este modo, el tamaño de los nodos es proporcional al valor que tiene su PageRank.

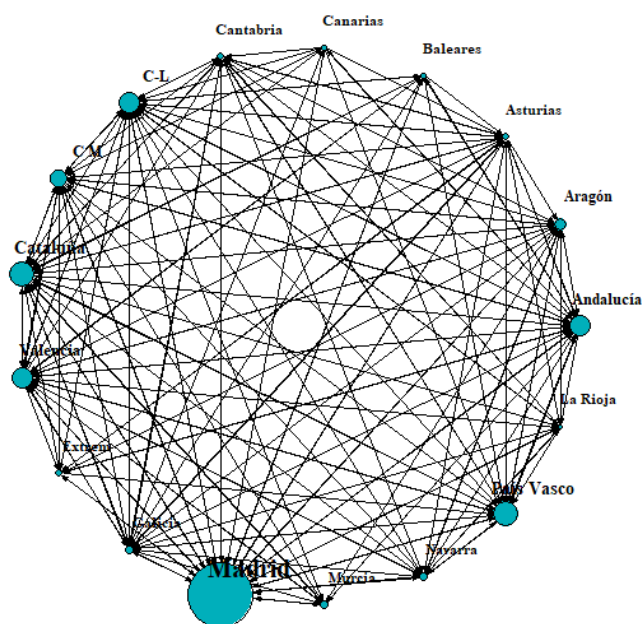


Figura 20: Red según PageRank

Gráficamente, tras agregar un vector con los valores de PageRank de cada Comunidad Autónoma e identificar el máximo valor, en la Figura 21 se muestra una nueva red en función de la C. A. más influyente, siendo, en este caso, la Comunidad de Madrid.

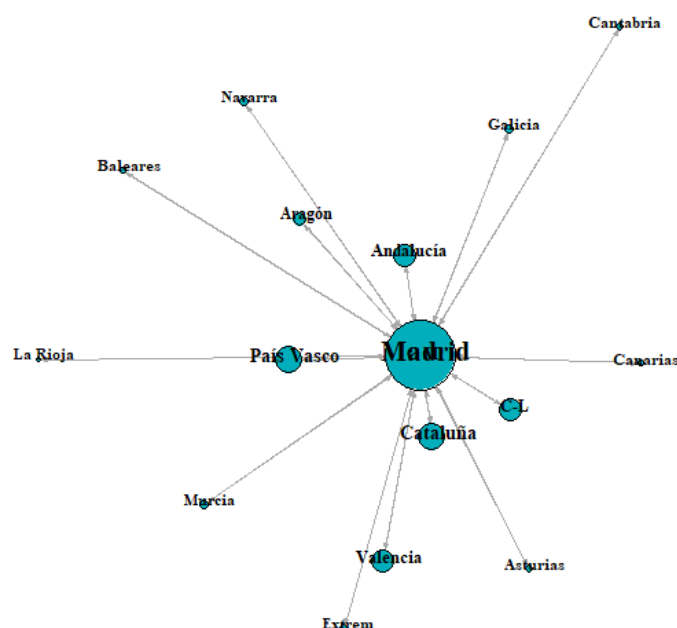


Figura 21: Red según PageRank de la Autonomía más influyente

A modo de conclusión, se puede afirmar que la Comunidad Autónoma más influyente en España en referencia a la movilidad de personas es la Comunidad de Madrid.

4.8. Resultados del Contraste de Bondad de Ajuste

Se va a analizar si la distribución Ley de Potencias (en inglés, *Power-Law*) se ajusta a los datos correspondientes a la movilidad entre CC. AA. (promedio) (Tabla 26), a partir del contraste de bondad de ajuste basado en el estadístico de Kolmogorov–Smirnov (KS statistic). En la Tabla 12 se muestran los resultados, pudiéndose afirmar, con un nivel de significación del 10%, que la distribución de ley de Potencias se ajusta al grado de la red promedio (Estadístico KS = 0.126; p-valor = 0.997).

Tabla 12: Ajuste de la ley de Potencias según grado

	Incertidumbre de los parámetros				
	α	x_{\min}	probabilidad logarítmica	estadístico KS	p-valor
Grado total	5.943	23	-26.008	0.126	0.997

Nota: α representa el exponente de la distribución de Potencias; x_{\min} es el mínimo valor con el que la distribución de Potencias se ajustó al modelo; *probabilidad logarítmica*: referente a los parámetros que se han ajustado a la distribución; *estadístico KS* representa un estadístico de prueba de un test KS que compara la distribución Power-Law con el vector de entrada (grado de entrada de la red, en este caso) (un menor valor supone un ajuste mejor); finalmente, un p-valor inferior a 0.05 indica que la prueba rechaza la hipótesis referente a que los datos podrían haberse extraído de la distribución ajustada a la ley de Potencias.

La Figura 22 muestra la relación que existe entre el grado de la red y la función de distribución acumulativa (en inglés, *cumulative distribution function, CDF*), donde se observa gráficamente los resultados tras realizar el análisis de la red según su ajuste a una distribución Power-Law.

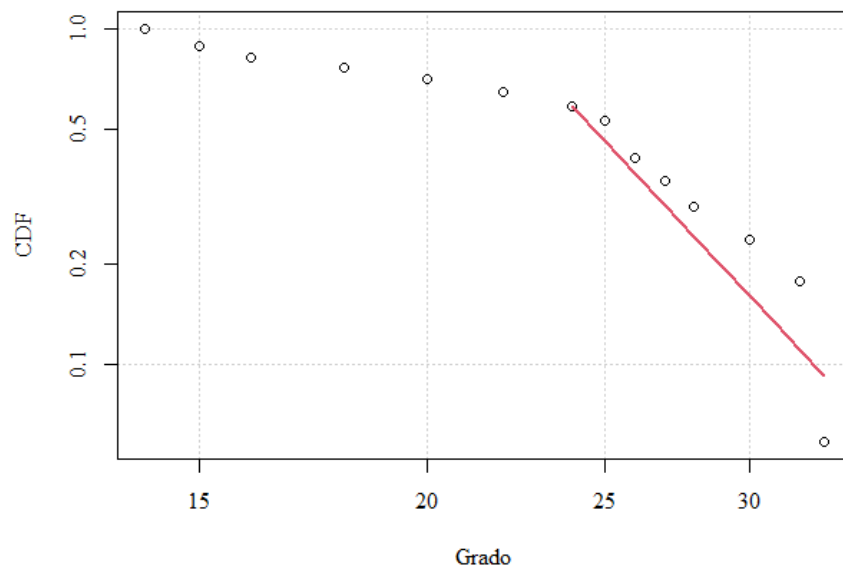


Figura 22: Resultados Power-Law

5 MODELOS ERGM

En el presente capítulo se va a exponer los modelos Exponenciales de Grafos Aleatorios (Exponential Random Graph Models, ERGM). Los modelos ERGM se caracterizan por ser modelos estadísticos que permiten analizar la probabilidad que tendrán los pares de nodos de una red comparando la red observada con los modelos ERGM.

5.1. Introducción a los Modelos ERGM

Los modelos ERGM (Snijders et al., 2006) se utilizan para la especificación de la distribución probabilística de las distintas redes que se desee estudiar. Estos modelos explican la probabilidad de una red observada como una función de variables, tanto endógenas como exógenas. En el análisis de redes, los atributos afectan a la estructura de la red, y la estructura afecta a los atributos, dando lugar a una relación mutua entre ambos. Es por ello la necesidad de controlar los efectos exógenos y endógenos puesto que la variación de éstas afecta a la estructura de la red en su conjunto.

Hay que señalar además que los modelos ERGM parten de unos supuestos teóricos:

- Las redes emergen localmente.
- Los patrones de la red estudiada se pueden entender como evidencia de procesos en marcha.
- Y las redes están estructuradas estocásticamente, es decir, que evolucionan aleatoriamente en función del tiempo.

Las variables endógenas, que pueden estar representadas a través de la reciprocidad, conectividad, popularidad o triangulación, entre otros, describen la presencia o ausencia de vínculos entre los nodos de la red. En este sentido, las variables endógenas generan un modelo de la estructura de la red a estudiar, mediante la creación de la red a partir de diversas configuraciones (o patrones locales de vínculos).

Respecto a las variables exógenas, pueden ser atributos no controlables e inherentes por los nodos o las diversas relaciones exógenas en la red, y son variables que no tienen relación con la estructura de la red, pero que influyen en la forma que tendrá la misma.

Los modelos ERGM permiten producir un conjunto de resultados en forma de coeficientes y niveles de significancia que reflejan el cambio en la probabilidad de un vínculo asociado con el cambio de una variable endógena o exógena. Por lo tanto, los modelos ERGM permiten obtener estimaciones de máxima semejanza para los parámetros del modelo especificado dado un conjunto de datos, así como la definición de modelos para una red que incluya elementos tales como mutualidad, relaciones de triangulación, etc., la evaluación de modelos de convergencia, coeficientes individuales y realizar comparaciones entre distintos modelos o simular nuevas redes en base a las distribuciones de probabilidad del modelo previamente estudiado con ERGM.

5.2. Descripción de los Modelos ERGM

Sea Y una matriz de adyacencia donde $y_{ij} = 1$ si los elementos están relacionados mediante un arco. Sea X la matriz de covariables que incorpora las medidas estructurales de la red así como los atributos de los nodos y

arcos, los modelos ERGM se pueden expresar como:

$$P_{\theta,\gamma}(Y = y|X) = \frac{e^{\theta^T g(y,X)}}{k(\theta,\gamma)}$$

donde:

- Y : representa la realización de una red, siendo una variable similar a una variable aleatoria.
- y : aplicación a la red observada.
- θ : especifica el vector de coeficientes.
- γ : simboliza el espacio de grafos (o redes) posibles.
- $g(y, X)$: vector del modelo estadístico de la red.
- $k(\theta, \gamma)$: representa un escalar de normalización, es decir, una cantidad respecto a todas las redes posibles, restringiendo a las redes con el mismo conjunto de nodos y .

Además, la ecuación de los modelos ERGM se puede describir de una forma alternativa:

$$\text{logit}(Y_{ij} = 1|y_{ij}^c) = \theta^T \delta(y_{ij})$$

donde:

- Y_{ij} : observación de la variable aleatoria genérica.
- y_{ij} : representa la relación que tienen los nodos i y j en la red y . De nuevo, si $y_{ij} = 1$ los elementos están relacionados mediante un arco.
- y_{ij}^c : simboliza el estado de todos los pares de vértices (nodos) en la red y que no sean (i, j) .
- $\delta(y_{ij})$: medida de cómo varían los estadísticos de la red en función de si el vértice ij está, o no, conectado a los demás.

5.3. Estimación y análisis del modelo ERGM

A través de la estimación de los modelos ERGM al análisis del movimiento de residentes en España, se podrán explorar e identificar las relaciones entre las CC. AA. y algunas de las características que determinan su funcionamiento, es decir, la determinación de los flujos emisores y receptores.

5.3.1 Variables endógenas

5.3.1.1 Arcos

Los arcos (denominado en inglés, *edges*) se refiere al número de vínculos entre nodos existente en la red.

5.3.1.2 Reciprocidad

Las conexiones recíprocas (denominado en inglés, *mutual*) hace referencia a la tendencia que tienen los nodos de la red a formar relaciones recíprocas (si el nodo i se relaciona con el nodo j , muy posiblemente j decida relacionarse con i).

5.3.1.3 Triángulos

Los triángulos (denominado en inglés, *triangles*) representa la relación existente entre los nodos i, j, k , de modo que un triángulo se define, en redes no dirigidas, como el conjunto $\{(i, j), (j, k), (k, i)\}$. En redes dirigidas, un triángulo se define como cualquier conjunto de tres arcos entre (i, j) y (j, k) y (k, i) o (i, k) .

5.3.2 Variables exógenas

5.3.2.1 Presencia de costa en la Comunidad Autónoma

Un motivo que puede influir en el aumento de los desplazamientos puede estar relacionado con los destinos de playa o costa. Por esta razón, se ha introducido la variable exógena “Costa” según tenga (o no) la C. A. costa en su región. La variable “Costa” es de tipo binaria (1 = “Presencia costa”; 0 = “Ausencia costa”).

Las CC. AA. que sí presentan costa son: Andalucía, Principado de Asturias, Islas Baleares, Canarias, Cantabria, Cataluña, Comunidad Valenciana, Galicia, Región de Murcia y País Vasco.

Por el contrario, aquellas Comundidades Autónomas sin costa son: Aragón, Catilla-La Mancha, Castilla y León, Extremadura, La Rioja, Comunidad de Madrid y Cominidad Foral de Navarra.

5.3.2.2 Presencia de estación de esquí en la Comunidad Autónoma

Otro motivo que puede contribuir a un aumento de los desplazamientos se encuentra en los destinos de montaña, con la finalidad de esquiar. Se introduce la variable exógena “Esquí”, variable binaria (1 = “Presencia estación esquí”; 0 = “Ausencia estación de esquí”), identificando las CC. AA. donde se hayan las siguientes estaciones de esquí:

- Cordillera Cantábrica: Valgrande-Pajares (Asturias), Fuentes de Invierno (Asturias), San Isidro (León), Leitariegos (León), El Morredero (León), Alto Campoo (Cantabria), Lunada (Burgos) y Manzanaeda (Orense).
- Pirineos: Astún (Huesca), Candanchú (Huesca), Formigal (Huesca), Panticosa (Huesca), Cerler (Huesca), Baqueira Beret (Lérida), Boí Taüll Resort (Lérida), Espot Esquí (Lérida), Port Ainé (Lérida), Port del Comte (Lérida), La Molina (Gerona), La Masella (Gerona), Tavascán (Lérida), Vallter 2000 (Gerona) y Vall de Núria (Gerona).
- Sistema Central: Valdesquí (Comunidad de Madrid), Puerto de Navacerrada (Comunidad de Madrid), La Pinilla (Segovia) y La Covatilla (Salamanca).
- Sistema Ibérico: Valdezcaray (La Rioja), Valle del Sol (Burgos), Javalambre (Teruel) y Valdelinares (Teruel).
- Sierra Nevada.

Es decir, las Comunidades Autónomas a las que pertenecen estas provincias y que por tanto tendrán asignado el atributo de estación de esquí son: Andalucía, Aragón, Principado de Asturias, Cantabria, Castilla y León, Cataluña, Comunidad de Madrid y La Rioja.

5.3.2.3 Fiestas más conocidas y otros acontecimientos

Se ha incorporado la variable exógena “Festividades”, atributo asignado a las CC. AA. que poseen las fiestas más conocidas en España y otros acontecimientos mostrados en el siguiente listado:

- Carnaval de Cádiz, carnaval de Tenerife y carnaval de Gran Canaria.
- Las fallas de Valencia.
- La Semana Santa de Andalucía y Castilla y León.
- La Feria de Abril de Sevilla.

- Las Fiestas de San Isidro en Madrid.
- El Camino de Santiago.
- Y La Tomatina en Buñol, Valencia.

La variable “Festividades” es de tipo binaria (1 = “Presencia festividades”; 0 = “Ausencia festividades”).

5.3.2.4 Presencia del AVE

Otro motivo que puede explicar un posible aumento de los desplazamientos recae en la variable explicativa “AVE”. De este modo, la presencia de rutas de AVE entre CC. AA. puede suponer un aumento en la movilidad entre estos territorios.

Cabe destacar que en este apartado se tendrán en cuenta las Comunidades Autónomas que conecten con otras mediante la red ferroviaria de Renfe AVE. De este modo, como pudo comprobarse en la Figura 19 del capítulo anterior y aplicando el criterio comentado previamente, queda excluida la Comunidad Autónoma de Galicia pues, aún teniendo presencia de AVE en su territorio, su trayecto no conecta con otras CC. AA. restringiendo, de este modo, la movilidad únicamente por el interior de Galicia.

En consecuencia, las CC. AA. que sí quedan conectadas son: Andalucía, Castilla-La Mancha, Castilla y León, Comunidad Valenciana, Comunidad de Madrid, Cataluña y Aragón.

La variable “AVE” es de tipo binaria (1 = “Presencia AVE”; 0 = “Ausencia AVE”).

5.3.2.5 Distinción de Regiones

Finalmente, se ha introducido la variable explicativa “Region”. Su finalidad es distinguir tres regiones en el territorio español, mediante las cuales pueda ser más factible la movilidad entre CC. AA. que pertenezcan a una misma región por su proximidad. Éstas son:

- Norte: Aragón, Asturias, Cantabria, Cataluña, Galicia, Navarra, País Vasco y La Rioja.
- Sur: Andalucía, Canarias y Murcia.
- Centro: Baleares, Castilla León, Castilla-La Mancha, Valencia, Extremadura y Madrid.

La variable “Region” es de tipo categórica y representa las tres regiones del territorio español (1 0 = “Norte”; 0 1 = “Sur”; 0 0 = “Centro”, siendo esta última la variable de referencia de las otras dos).

5.4. Simulación con RStudio

Se va a estudiar la red desde la perspectiva de las relaciones entre las Comunidades Autónomas, es decir, cómo los nodos interactúan entre sí, haciendo uso de la aplicación informática “RStudio”, herramienta que permite la computación estadística basada en un lenguaje de programación R.

Antes de comenzar a mostrar los resultados obtenidos, es preciso apuntar la notación utilizada en los resultados obtenidos en la estimación del modelo:

- Valor estimado. Un valor negativo corresponde a la no aleatoriedad en la formación del enlace.
- MCMC: Cadenas de Markov de Montecarlo (en inglés, *Markov Chain Monte Carlo*). Métodos de simulación que permite generar muestras y estimar cantidades de interés.
- (**): p-valor $\leq 0,01$; (*): $0,01 < \text{p-valor} \leq 0,05$; (*): $0,05 < \text{p-valor} \leq 0,1$. La presencia de (*) representa que el factor es significativo, influyendo en la generación de arcos.
- AIC: Criterio de Información de Akaike (en inglés, *Akaike Information Criterion*): mide la calidad

relativa de un modelo estadístico. Un menor valor del mismo proporcionará una mejor solución.

- BIC: Criterio de Información Bayesiano (en inglés, *Bayesian Information Criterion*): criterio para seleccionar modelos entre un conjunto de los mismos. Un menor valor del mismo proporcionará una mejor solución.

En primer lugar, se procede a calcular y simular el modelo en su forma más simple, considerando únicamente la existencia de arcos, que equivale al modelo de Erdős-Renyi. Tras simular el código en RStudio, de la Tabla 13 se obtiene que el número de arcos totales existentes en la red son 192. Haciendo uso de la ecuación logit expresada previamente se puede interpretar el coeficiente obtenido en la Tabla 13:

$$\begin{aligned} \text{logit}(Y_{ij} = 1|y_{ij}^c) &= \theta^T \delta(y_{ij}) \\ &= 0,8755 \times \text{cambio en el } n^{\circ} \text{ de arcos} \\ &= 0,8755 \times 1 \end{aligned}$$

donde *cambio en el n° de arcos* supone la unidad pues, tras añadir un nuevo arco a la red, el número total de arcos se ve incrementado en 1. De este modo, la probabilidad de ocurrencia de un enlace es la inversa de logit, es decir, un 0,70:

$$\begin{aligned} \frac{\exp(0,8755)}{(1 + \exp(0,8755))} \\ = 0,70 \end{aligned}$$

Tabla 13: Resultados del modelo (ERGM) número de arcos

Parámetros	Términos ERGM	Valor estimado	Error	MCMC	p-valor	Pr(> p)
Efectos estructurales de la red	edges	0.8755	0.1331	0	6.579	<1e-04 ***
Diagnóstico del modelo	AIC	331.6				
	BIC	335.2				

A continuación, se procede a añadir el término “mutual” (reciprocidad) (Tabla 14), que representa la existencia de las relaciones $i \rightarrow j$ y $j \leftarrow i$. De los resultados se obtiene que las relaciones de reciprocidad ascienden a 81, con efecto positivo y significante, indicando la existencia de más relaciones recíprocas en este modelo que las esperadas de un modelo aleatorio con 192 arcos. Además, realizando el mismo procedimiento que anteriormente, se obtiene que la probabilidad de crear un enlace recíproco entre dos nodos es 0,816.

$$\begin{aligned} \frac{\exp(1,4893)}{(1 + \exp(1,4893))} \\ = 0,816 \end{aligned}$$

Tabla 14: Resultados del modelo (ERGM)

Parámetros	Términos ERGM	Valor estimado	Error	MCMC	p-valor	Pr(> p)
Efectos estructurales de la red	mutual	1.4893	0.1742	0	8.549	<1e-04 ***
Diagnóstico del modelo	AIC	306.3				
	BIC	309.9				

Incorporando al modelo inicial el término de “triangle”, se puede medir la probabilidad de la red de tener nodos adyacentes interconectados como un conjunto de tres arcos, también denominado triangulación. Se

obtiene que el número asciende a 2.223. En este caso, la probabilidad de formación de triángulos en la red es 0,507.

$$\frac{\exp(0,032)}{(1 + \exp(0,032))} = 0,507$$

Tabla 15: Resultados del modelo (ERGM)

Parámetros	Términos ERGM	Valor estimado	Error	MCMC	p-valor	Pr(> p)
Efectos estructurales de la red	triangle	0.031944	0.002805	0	11.05	<1e-04 ***
Diagnóstico del modelo	AIC	313.4				
	BIC	317				

Por otra parte, se va a hacer uso de las variables exógenas previamente comentadas. Al trabajar con una red dirigida, es necesario realizar un estudio de la influencia que tiene cada atributo en los arcos de entrada y salida de los nodos, es decir, en la simulación se ha considerado la influencia de estos atributos sobre los nodos según sean factores significativos sobre los arcos emisores, sobre los arcos receptores y de forma conjunta.

En la Tabla 16 se muestra la estimación del modelo `ergm_1`, que representa la influencia de las covariables en la generación de arcos. Un primer aspecto que se puede abordar es cómo la covariable “Costa” no es un factor significativo y por tanto no influye en la generación de los arcos de la red. Ocurre lo mismo con la covariable “Esqui”. En este sentido, la presencia de costa y de estaciones de esquí en las Comunidades Autónomas no son razón suficiente por la cual aumenta el número de movimientos de esas CC. AA. Sin embargo, la presencia del AVE sí influye en un aumento de los movimientos, teniendo un efecto significativo. Las covariables “Festividades” y “Region.Norte” influyen de forma menos significativa, de acuerdo al número de (*). Finalmente, la interpretación de la variable exógena “Region.Sur” indica que la probabilidad de que una C. A. del sur forme un enlace frente a que una C. A. del centro forme un enlace es menor, de acuerdo al signo negativo del término ERGM.

Tabla 16: Resultados del modelo (ERGM) con efectos exógenos

Parámetros	Términos ERGM	Valor estimado	Error	MCMC	p-valor	Pr(> p)
Efectos estructurales de la red	edges	-2.6959	0.5872	0	-4.591	<1e-04 ***
	mutual	0.7959	0.5448	0	1.461	0.1440
Efectos de los atributos de los nodos	nodefactor.AVE.1	1.8162	0.4201	0	4.323	<1e-04 ***
	nodefactor.Costa.1	0.5252	0.2867	0	1.832	0.0670 .
	nodefactor.Esqui.1	0.3804	0.2986	0	1.274	0.2027
	nodefactor.Festividades.1	0.9683	0.3326	0	2.912	0.0036 **
	nodefactor.Region.Norte.1	0.7200	0.3435	0	2.096	0.0361 *
	nodefactor.Region.Sur.1	-0.8838	0.4097	0	-2.157	0.0310 *
Diagnóstico del	AIC	235.3				

modelo	BIC	264.1
--------	-----	-------

A continuación, se procede a analizar los resultados obtenidos de la simulación considerando la influencia de las variables exógenas en la generación de arcos que inciden sobre los nodos (*ergm_2*). En la Tabla 17 se observa como la única covariable significativa es “AVE”. De este modo, la presencia de líneas de AVE en una C. A. influye en un aumento de los movimientos que recibe ésta. Sin embargo, no se puede afirmar lo mismo del resto de covariables pues, al no ser factores significativos, no afectan a la formación de enlaces que inciden sobre las CC. AA.

Tabla 17: Resultados del modelo (ERGM) con efectos exógenos (entrada)

Parámetros	Términos ERGM	Valor estimado	Error	MCMC	p-valor	Pr(> p)
Efectos estructurales de la red	edges	-1.37977	0.42143	0	-3.274	0.001060 **
	mutual	2.27299	0.41814	0	5.436	< 1e-04 ***
Efectos de los atributos de los nodos	nodeifactor.AVE.1	1.54494	0.45811	0	3.372	0.000745 ***
	nodeifactor.Costa.1	-0.18964	0.35662	0	-0.532	0.594885
	nodeifactor.Esqui.1	0.21447	0.37074	0	0.578	0.562934
	nodeifactor.Festividades.1	0.09186	0.37477	0	0.245	0.806380
	nodeifactor.Region.Norte.1	0.77192	0.43506	0	1.774	0.076015 .
	nodeifactor.Region.Sur.1	-0.31559	0.47543	0	-0.664	0.506819
Diagnóstico del modelo	AIC	281.6				
	BIC	310.4				

Finalmente, se procede a analizar la situación contraria a la descrita anteriormente: cómo influyen las variables explicativas en un aumento de los arcos que emiten los nodos (*ergm_3*). Atendiendo a los resultados obtenidos, las variables exógenas que influyen en mayor medida en la generación de los arcos que emiten las Comunidades Autónomas son la presencia de AVE y la existencia de festividades. En menor medida, la existencia de costa también influye en el aumento del movimiento que emiten estas CC. AA. Las covariables “Esqui” y “Region.Norte” no contribuyen a un aumento de los enlaces que emiten los nodos, no siendo factores significativos. La interpretación de la covariable “Region.Sur” indica que la probabilidad de que una C. A. del centro forme un enlace frente a que una C. A. del sur forme un enlace es mayor, de acuerdo al signo negativo del término ERGM. Sin embargo, no tiene gran influencia en la red, atendiendo al rango donde se encuentra su p-valor.

Tabla 18: Resultados del modelo (ERGM) con efectos exógenos (salida)

Parámetros	Términos ERGM	Valor estimado	Error	MCMC	p-valor	Pr(> p)
Efectos	edges	-2.6164	0.4975	0	-5.259	< 1e-04 ***

estructurales de la red	mutual	2.4862	0.4663	0	5.331	< 1e-04 ***
	nodeofactor.AVE.1	1.9192	0.4891	0	3.924	< 1e-04 ***
Efectos de los atributos de los nodos	nodeofactor.Costa.1	1.1936	0.4004	0	2.981	0.00287 **
	nodeofactor.Esqui.1	0.3723	0.4268	0	0.872	0.38298
	nodeofactor.Festividades.1	2.0192	0.5124	0	3.941	< 1e-04 ***
	nodeofactor.Region.Norte.1	0.5038	0.4705	0	1.071	0.28433
	nodeofactor.Region.Sur.1	-1.4300	0.6020	0	-2.375	0.01753 *
Diagnóstico del modelo	AIC	235.4				
	BIC	264.2				

Finalmente, atendiendo a los valores de los estadísticos AIC y BIC, el modelo `ergm_1` es el que proporciona una mayor calidad del ajuste, entre todos los modelos estudiados.

A continuación, se va a hacer uso de la función de Contraste de Bondad de Ajuste (en inglés, *Goodness-of-Fit*, GOF) perteneciente a ERGM de forma que se pueda comparar la red simulada con la red observada. Dicha comparación es posible puesto que la función GOF simula diversas redes haciendo uso de los parámetros del modelo inicial utilizado, obteniendo, de este modo, medidas de estas redes. En este caso, los p-valores que se obtienen son la diferencia entre los estadísticos pertenecientes a la red observada y los referentes a la red simulada. Un p-valor reducido supone un posible problema ante el correcto ajuste de los estadísticos del modelo. Se ha empleado para esta última simulación el código de `ergm_1`.

En la Tabla 19 se observa los p-valores individuales de cada covariable. Atendiendo a estos resultados, las variables exógenas “AVE” y “Costa” son las que mejor se ajustan a los modelos simulados.

Tabla 19: Resultados GOF (p-valor)

Términos ERGM	p-valores
edges	0.9136283
mutual	0.5606255
nodeofactor.AVE.1	0.9661010
nodeofactor.Costa.1	0.9885409
nodeofactor.Esqui.1	0.5313997
nodeofactor.Festividades.1	0.6989611
nodeofactor.Region.Norte.1	0.3604792
nodeofactor.Region.Sur.1	0.5128918

En las Figuras 23 y 24 se representa gráficamente el procedimiento GOF. Los diagramas de caja representan los valores de la red simulada, siendo la línea en negrita el valor de la red original. Idealmente, éstos dos deberían superarse.

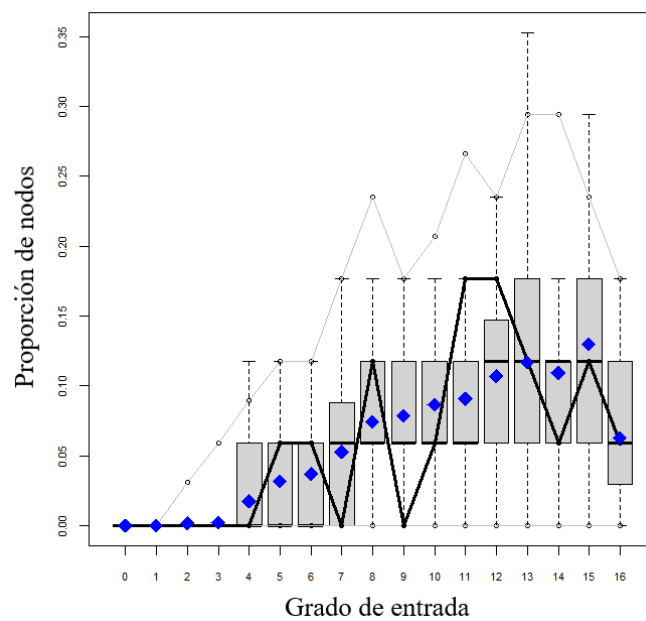


Figura 23: Resultados GOF arcos incidentes

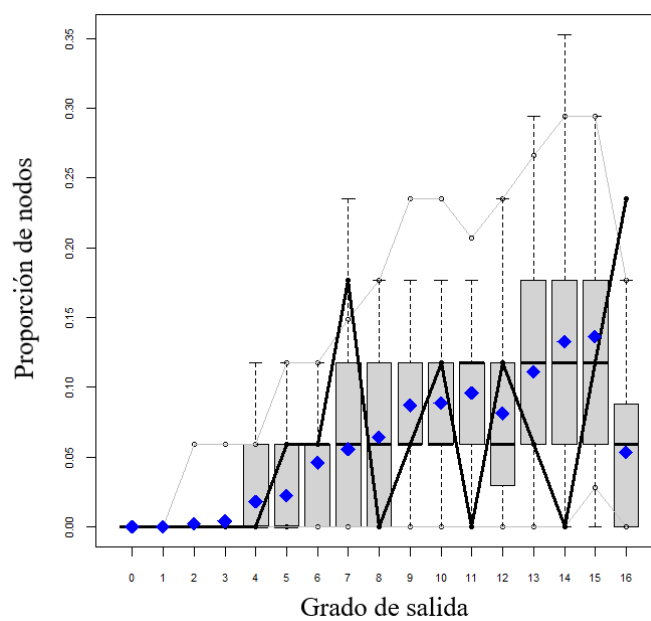


Figura 24: Resultados GOF arcos emisores

Atendiendo al grado de la red, se percibe cómo el ajuste del modelo original no resulta ser consistente de acuerdo al procedimiento GOF.

6 EFECTOS DEL VIRUS SARS-CoV-2 SOBRE LA MOVILIDAD

En este capítulo se analiza la vulnerabilidad de la red a través de la situación producida por el SARS-CoV-2, tanto de forma global como nacional.

En la Figura 25 se muestra gráficamente los casos de SARS-CoV-2 globalmente, atendiendo a los rangos de colores (a mayor intensidad, peor la situación).

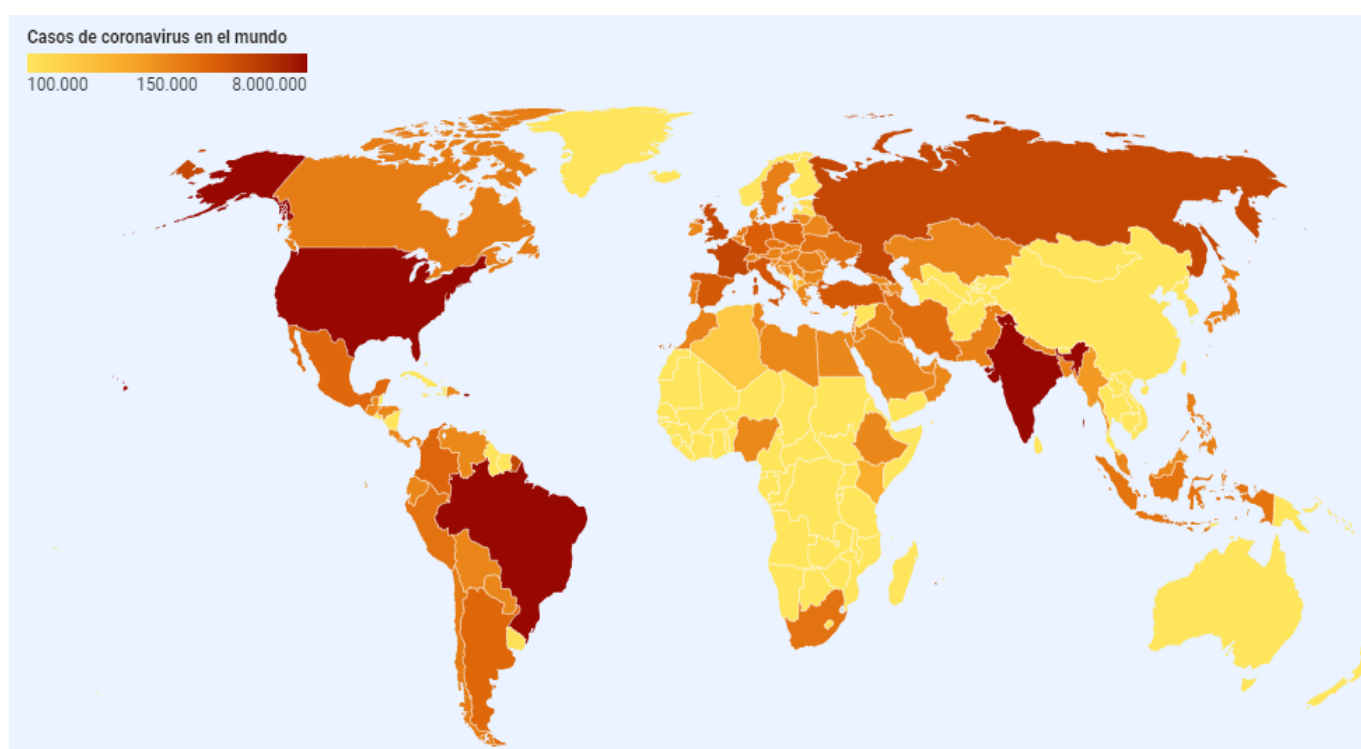


Figura 25: Casos de SARS-CoV-2 en el mundo

Uno de los impactos más evidentes de la pandemia producida por el SARS-CoV-2 es la restricción en el número de desplazamientos a nivel nacional e internacional. Según la Organización Mundial del Turismo (OMT), en 2020 se produjo una caída de las llegadas internacionales en un 74%. Los destinos de todo el mundo recibieron mil millones de llegadas internacionales menos respecto a 2019. Para tener una referencia de la catástrofe producida por la SARS-CoV-2 en España, donde el sector terciario es el más importante en la economía, tras la crisis financiera del año 2009, alcanzando una tasa de variación anual del PIB de -4,36% en el segundo trimestre del año, en 2020 se obtuvo, en el mismo trimestre, una variación anual del PIB de -21,60%. Económicamente, las pérdidas se multiplican por más de diez veces a las registradas durante la crisis económica mundial de 2009.

El SARS-CoV-2 ha afectado directamente a la movilidad, puesto que las medidas implantadas por la mayoría de los países han supuesto, principalmente, el cierre temporal de sus fronteras, además de pruebas obligatorias de tests negativos y cuarentena en el caso de querer viajar a otros lugares. Todo ello ha supuesto el decrecimiento de los viajes internacionales y nacionales. Para tener una idea global de lo que ha supuesto la pandemia respecto a la movilidad, los descensos en las llegadas de personas en las distintas regiones del mundo fueron las siguientes durante el año 2020: Asia y el Pacífico (84%), África y Oriente Medio (75%),

Europa (70%) y las Américas (69%).

De acuerdo a todo lo anterior, desde el inicio de la pandemia se puede observar la gran similitud que ha tenido el aumento de los contagios y fallecimientos con la disminución de los viajes que se han producido tanto de forma global como internacional. Se puede observar en la Figura 26 el número de casos de SARS-CoV-2 que ha habido desde el inicio de la pandemia, siendo su comienzo en marzo del año 2020. En la Figura 26 se puede observar que tras el mes de abril del año 2020 los casos por SARS-CoV-2 comenzaron a incrementar de forma exponencial afectando a todos los países del mundo.

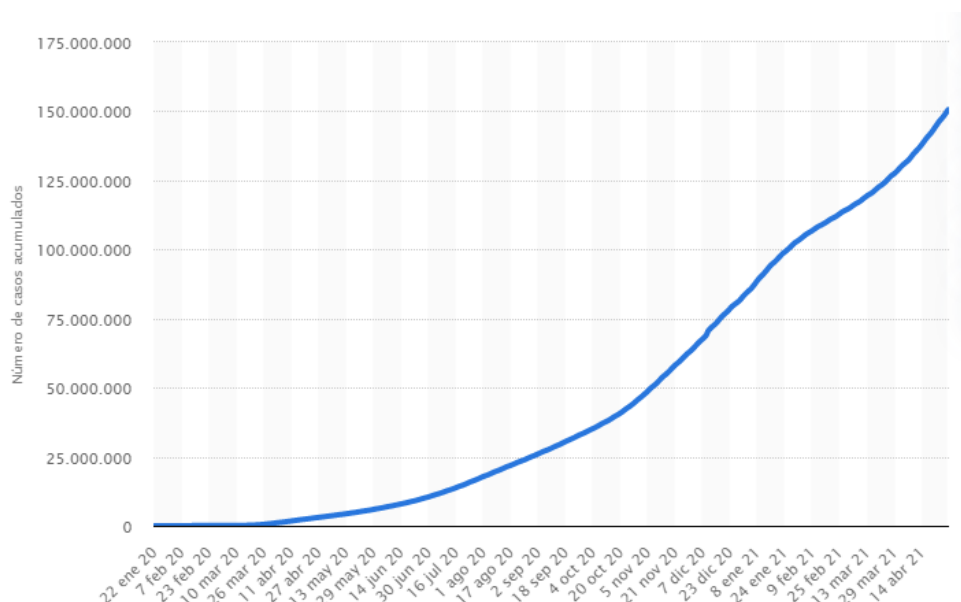


Figura 26: Casos de SARS-CoV-2 en el mundo (acumulado)

Fuente: Statista.com

Como se señalaba al principio, la movilidad internacional que recibió España en el año 2020 disminuyó considerablemente, pues a principios del año 2020 aún seguían produciéndose desplazamientos en la frontera, sin considerar Ceuta y Melilla, para los que no se tienen datos, de unos 4.069.159 desplazamientos. Sin embargo, el SARS-CoV-2 se detectó en países europeos a mediados de marzo, alcanzando los 1.970.060 desplazamientos internacionales. En cuanto a las cifras registradas correspondientes al mayor derrumbe, ésta queda marcada en el mes de junio. En la Tabla 20 se puede observar el desplome que tuvieron los desplazamientos a lo largo del año 2020 en España respecto a los desplazamientos internacionales que recibieron las CC. AA. españolas. Por un lado, en el mes de febrero se produjo el mayor número de desplazamientos, debido, principalmente, a que todavía no se había detectado el virus. Por otro lado, en el mes de junio se registró el menor número de viajes internacionales a causa de la reciente apertura de la movilidad nacional e internacional. Además, en los meses de abril y mayo no se registraron movimientos internacionales debido al cierre de fronteras.

Tabla 20: Movilidad internacional en España (2020-2021)

	Andalucía	Aragón	Asturias	Baleares	Canarias	Cantabria	CL	C-M	Cataluña	Valencia	Extrem	Galicia	Madrid	Murcia	Navarra	País Vasco	La Rioja	Total
2021M02	28661	4426	Sin Datos	13149	74280	Sin Datos	9843	Sin Datos	54020	28066	Sin Datos	12870	31580	2206	Sin Datos	11742	Sin Datos	270843
2021M01	49716	6407	2823	15166	86569	2119	22203	4382	85265	49727	11674	19321	45559	6775	Sin Datos	18105	Sin Datos	425811
2020M12	60577	Sin Datos	Sin Datos	22083	207064	4008	22721	22469	97305	71865	Sin Datos	28072	47885	8194	Sin Datos	22015	Sin Datos	614258
2020M11	49227	4517	Sin Datos	20102	160662	1743	11734	4363	58794	53946	Sin Datos	24116	32799	6655	Sin Datos	16638	Sin Datos	445296
2020M10	143378	11221	5904	84139	136896	11821	36343	3801	230580	176971	Sin Datos	55000	39575	16843	Sin Datos	44551	Sin Datos	997023
2020M09	179214	9406	9216	119279	92586	10494	34073	7557	276064	185256	13417	56096	48123	17762	Sin Datos	62406	Sin Datos	1120949
2020M08	364111	23339	23917	453794	240666	27731	71257	8750	459271	430092	12517	99155	60816	40117	Sin Datos	81181	Sin Datos	2396714
2020M07	359725	19397	12003	588431	215600	29246	57580	7619	436742	383013	Sin Datos	83676	76239	40628	Sin Datos	92316	Sin Datos	2402215
2020M06	14183	Sin Datos	Sin Datos	28382	2201	Sin Datos	Sin Datos	Sin Datos	76634	31487	Sin Datos	Sin Datos	20522	1906	Sin Datos	10097	Sin Datos	185412
2020M05	Sin Datos	Sin Datos	Sin Datos	Sin Datos	Sin Datos	Sin Datos	Sin Datos	Sin Datos	Sin Datos	Sin Datos	Sin Datos	Sin Datos	Sin Datos	Sin Datos	Sin Datos	Sin Datos	Sin Datos	Sin Datos
2020M04	Sin Datos	Sin Datos	Sin Datos	Sin Datos	Sin Datos	Sin Datos	Sin Datos	Sin Datos	Sin Datos	Sin Datos	Sin Datos	Sin Datos	Sin Datos	Sin Datos	Sin Datos	Sin Datos	Sin Datos	Sin Datos
2020M03	302928	20633	Sin Datos	116018	459947	Sin Datos	39370	Sin Datos	406683	253718	Sin Datos	35058	251590	27848	Sin Datos	56267	Sin Datos	1970060
2020M02	626653	39860	Sin Datos	185554	1170238	Sin Datos	83648	Sin Datos	966356	528712	Sin Datos	78075	518594	68750	Sin Datos	72839	Sin Datos	4339279
2020M01	605451	35300	Sin Datos	103772	1101164	6894	72290	22625	870157	433023	Sin Datos	83310	611184	52061	Sin Datos	71928	Sin Datos	4069159

Fuente: Instituto Nacional de Estadística

Una vez identificada la situación de los desplazamientos internacionales a España, se procederá a introducir la situación nacional respecto a la movilidad entre CC. AA. En este sentido, entre las medidas que se implantaron desde el mes de marzo de 2020 están el cierre de fronteras, el cierre perimetral de CC. AA., de provincias y de municipios, además del confinamiento de sus habitantes. No obstante, desde el inicio de la pandemia, tanto el Gobierno de España como las distintas Comunidades Autónomas fueron implantando distintas medidas individualmente para intentar frenar el avance del SARS-CoV-2, teniendo éstas más o menos éxito considerando los resultados obtenidos respecto a los contagios y fallecimientos.

A continuación, se procede a mostrar el avance que tuvo el SARS-CoV-2 en España desde la detección de los primeros contagios (Figura 27).

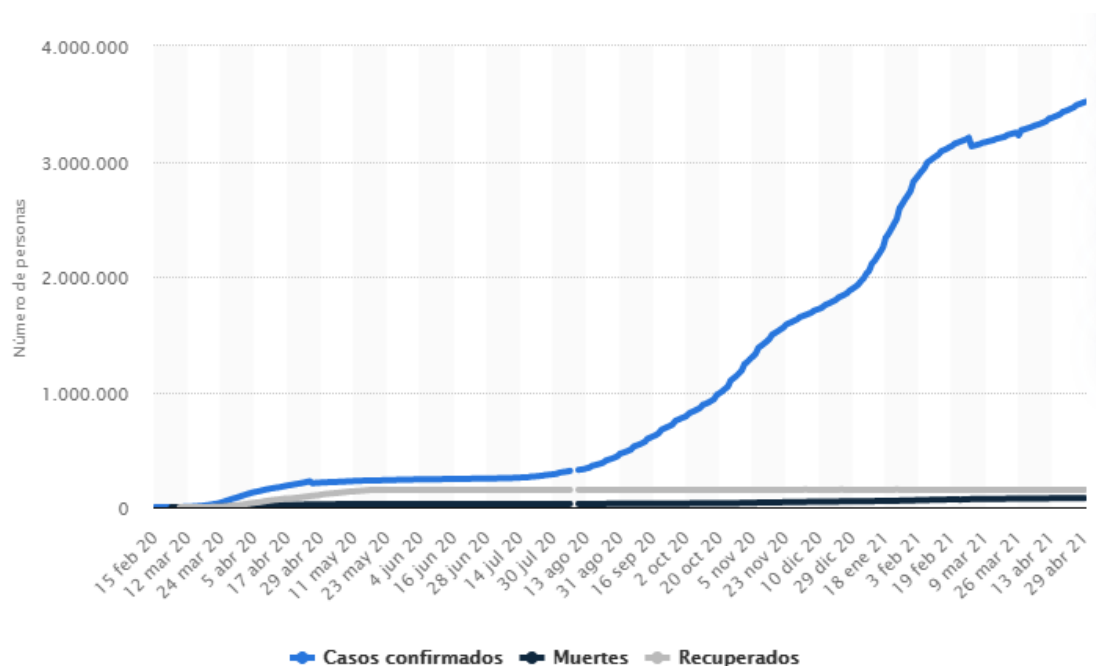


Figura 27: Casos de SARS-CoV-2 en España (acumulado)

Fuente: Statista.com

No obstante, se puede observar el desarrollo del SARS-CoV-2 en España en función de los casos notificados en cada fecha mostrada por cada 100.000 habitantes (Figura 28).



Figura 28: Casos de SARS-CoV-2 en España (desarrollo)

Fuente: RTVE

De acuerdo a todo lo anterior, se puede realizar una comparativa (ver Figura 29) entre el aumento del número de contagios por el virus SARS-CoV-2 y el descenso de los movimientos por CC. AA. pues, a medida que iban incrementando los casos, las restricciones implementadas fueron más estrictas.

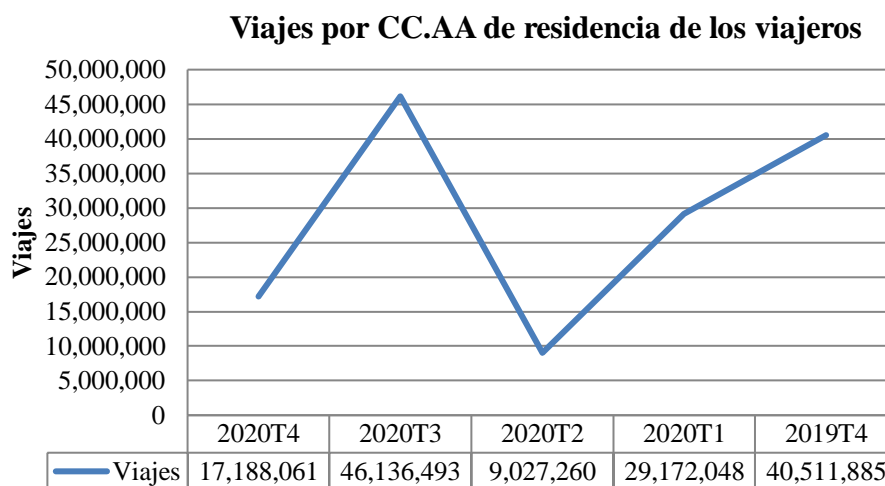


Figura 29: Viajes por CC. AA. de residencia de los viajeros

Fuente: Instituto Nacional de Estadística

Realizando un desglose de la Figura 29 en función de las Comunidades Autónomas de residencia de los viajeros, se obtiene la Tabla 20.

Tabla 21: Movilidad nacional en España (2019 - 2020)

Comunidad Autónoma	Viajes				
	2020T4	2020T3	2020T2	2020T1	2019T4
Total	17.188.061	46.136.493	9.027.260	29.172.048	40.511.885
Andalucía	2.471.841	7.400.428	1.859.361	4.952.181	6.514.810
Aragón	580.896	1.471.804	250.991	972.053	1.420.515
Asturias	344.815	948.533	268.865	628.254	752.366
Baleares	341.545	656.040	191.105	700.353	945.883
Canarias	1.043.871	1.659.064	407.207	1.021.806	1.447.914
Cantabria	145.912	455.897	138.088	351.724	515.067
Castilla-León	1.391.331	2.395.669	485.573	1.810.376	2.492.339
Castilla la Mancha	974.438	1.691.687	427.811	1.272.583	1.840.757
Cataluña	2.438.568	7.496.517	1.311.910	5.114.868	6.759.009
Valencia	2.430.667	4.220.883	949.687	2.812.720	3.500.214
Extremadura	536.866	899.894	123.160	587.418	834.931
Galicia	866.644	1.978.028	380.710	1.598.931	2.380.547
Madrid	2.173.601	9.468.850	1.147.133	4.679.206	6.837.642
Murcia	261.162	1.261.240	340.335	560.728	984.701
Navarra	299.531	702.533	196.033	494.575	636.536
País Vasco	746.428	3.039.143	469.264	1.355.988	2.236.843
La Rioja	99.200	324.149	70.384	174.523	310.870
Ceuta	21.941	38.122	Sin Dato	40.903	55.888
Melilla	18.803	28.011	Sin Dato	42.858	45.053

Fuente: Instituto Nacional de Estadística

Como puede comprobarse observando las Figuras 27, 28 y 29 y la Tabla 21, y teniendo en consideración que la primera vez que se declaró el confinamiento por la pandemia del SARS-CoV-2 fue el 14 de marzo de 2020, la movilidad aminoró de una forma drástica tras la detección de los primeros casos en España, además de la implementación del Estado de Alarma, teniendo España para ese momento 4.200 casos positivos y alzándose la cifra de fallecidos a las 120 personas. En este sentido, el confinamiento domiciliario duró hasta el día 9 de abril de 2020 (perteneciente al segundo trimestre del año), periodo en el cual la cifra de desplazamientos se encuentra en su valor más reducido observando la Figura 29. Posteriormente, tras la reducción de restricciones

durante el tercer trimestre del año, periodo al que pertenecen los meses de verano y coincidiendo con las vacaciones en los colegios y universidades, la cifra de desplazamientos por el territorio español aumentó considerablemente, de acuerdo a lo observado en la Figura 29, sumando los 46.136.493 trayectos. Sin embargo, tras la finalización del tercer trimestre de 2020, la cifra de contagios por SARS-CoV-2 no ha dejado de incrementar puesto que a fecha 1 de julio de 2020, al comienzo de las vacaciones, se registraba en España 149 casos diarios, mientras que a finales de agosto se registraron 3.829 casos, unas 25 veces más. De acuerdo a todo lo anterior, se puede afirmar que esta permisividad en la movilidad durante las vacaciones de 2020 no contribuyó a reducir el número de contagios y fallecimientos por SARS-CoV-2 en España, más bien todo lo contrario. Avanzando en el tiempo, tras permitir la movilidad durante las vacaciones de Navidad, aún estando en una franja con un riesgo alto (Figura 28), a la finalización de ésta se produjo la tercera ola de la pandemia. Durante ese periodo, se volvieron a endurecer las medidas de movilidad, suponiendo un nuevo cierre perimetral de los municipios en la mayoría de las provincias españolas. Mediante la implementación de estas medidas, se consiguió disminuir el número de contagios y fallecimientos. Sin embargo, el riesgo sigue siendo alto a día de hoy.

En consecuencia a todo lo comentado anteriormente, se ha establecido una situación de desconocimiento respecto a cómo será el avance del SARS-CoV-2. A un corto plazo, las expectativas de mejora son muy variadas, puesto que no se conoce con seguridad cómo avanzará el virus. Las perspectivas de recuperación para 2021 establecen tres escenarios de lo que podría suceder: el primer panorama prevé, por el 45% de los encuestados por la UNWTO (Organización Mundial del Turismo), una mejora; el 25% espera que la situación se mantenga similar a lo ya vivido, mientras que el 30% espera un empeoramiento a lo largo de 2021.

Sin embargo, la vacunación podría suponer una ayuda para volver a la “normalidad” y conseguir frenar el avance del SARS-CoV-2. Tras el comienzo de la campaña de vacunación en diciembre de 2020 en Estados Unidos, Reino Unido, Rusia y China, la vacuna ha logrado llegar al resto de países con el fin de inmunizar a la población y reducir el número de contagios, así como el número de fallecidos. Mundialmente se ha alcanzado un total de 2.650 millones de dosis administradas; no obstante, la campaña de vacunación no se está realizando de forma homogénea por los siete continentes. En el mes de junio de 2021 el porcentaje de población con al menos una dosis administrada era el siguiente: América del Norte (41,01%), Europa (38,11%), Latinoamérica y Caribe (25,75%), Asia (21,98%), Oceanía (15,9%) y África (2,44%). Como se puede observar, hay una clara diferencia en relación al reparto de vacunas.

En España, la vacunación comenzó el 27 de diciembre de 2020 inmunizando a los ancianos y sanitarios. A fecha 21 de junio de 2021, los datos de vacunación en España alcanzaron las 36.357.745 dosis administradas, suponiendo un total de 14.431.683 personas con pauta completa (un 30,4%). El ritmo de vacunación en España se puede observar en la Figura 30.

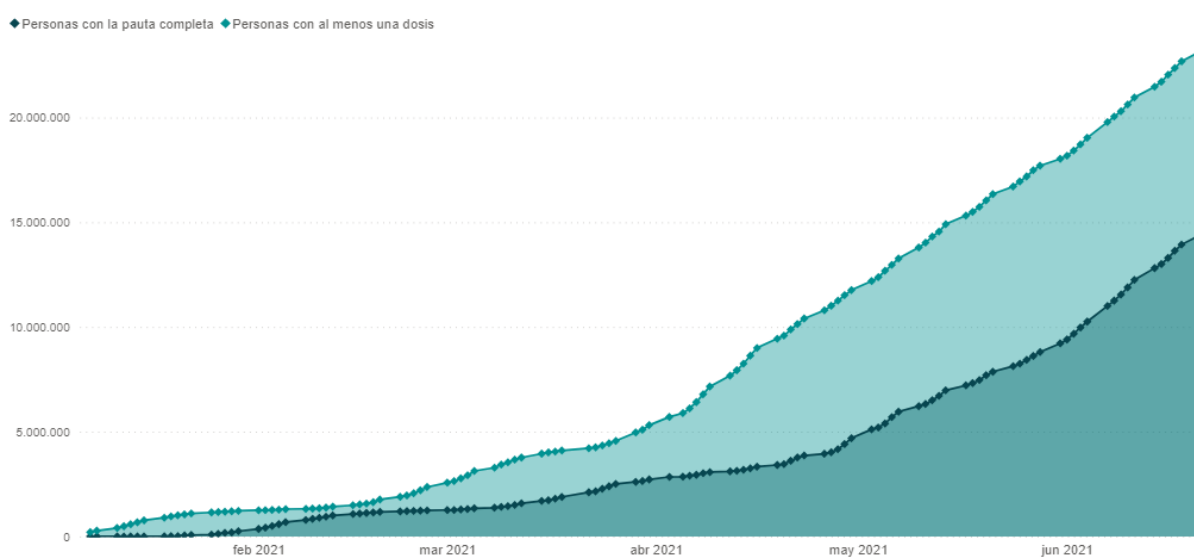


Figura 30: Evolución de personas vacunadas en España

Fuente: Ministerio de Sanidad

El virus SARS-CoV-2 ha creado un panorama de incertidumbre, puesto que en una encuesta realizada en octubre de 2020 por la UNWTO, el 79% de los encuestados opinaban que habría una recuperación en 2021,

hecho que no ocurrirá de momento, puesto que la situación hoy en día es similar a meses anteriores, considerando las nuevas cepas y mutaciones del virus que se detectan, además de la incertidumbre respecto a la protección a largo plazo de las vacunas. Muchos expertos consideran que no habrá una mejora para alcanzar un nivel de movilidad como el que había antes de la pandemia, considerando este periodo en el cuarto trimestre de 2019, hasta el año 2023 o incluso posteriormente.

7 CONCLUSIONES

Las redes complejas resultan de gran interés en diferentes ámbitos de la vida real, y gracias a Internet se puede acceder a una gran cantidad de datos, que tras su análisis concreto, proporcionan información y resultados con conclusiones coherentes para una adecuada toma de decisiones.

El primer objetivo del presente trabajo ha sido describir la metodología utilizada, Redes Complejas, con el fin de comprender su tipología y sus estadísticos característicos.

El segundo objetivo ha sido presentar las redes resultantes haciendo uso del flujo de movimientos que se produce en el territorio español entre las diecinueve Comunidades Autónomas a excepción de Ceuta y Melilla, de las que no se dispone información. Conviene indicar que los resultados obtenidos reproducen una aproximación de la realidad, sin llegar a una exactitud completa pues ciertas Comunidades Autónomas no tienen registrados los flujos que reciben o emiten a otras CC. AA. De acuerdo a todo lo anterior, considerando la información registrada por el Instituto Nacional de Estadística, se ha representado la movilidad entre las CC. AA. para los periodos entre 2015 y 2019, además de una red adicional con los valores promedio de estos periodos.

Del análisis realizado se destaca el hecho de que las distintas rutas del AVE han permitido conectar las zonas de Andalucía, Castilla y León, Castilla-La Mancha, Valencia, Madrid y Cataluña y fortalecer la centralidad de Madrid en los flujos de población.

Haciendo uso de los estadísticos de la red, se ha comprobado como las redes estudiadas durante el periodo 2015-2019 estaban bien conectadas pero poco centralizadas, siguiendo un modelo de red regular. Además, se ha obtenido mediante el análisis PageRank que la Comunidad Autónoma más relevante de la red es Madrid. Por otro lado, se ha verificado que la distribución de Ley de Potencias se ajusta al grado de la red promedio.

Posteriormente, con el propósito de explicar la relevancia de la red estudiada, se ha introducido la metodología de los modelos ERGM, para analizar las posibles variables que resultan determinantes para la formación de flujo en las redes de movilidad en España. En consecuencia, se ha desarrollado un listado con atributos que tienen, o no, las Comunidades Autónomas: ya sea la presencia de costa, de estaciones de esquí, de fiestas o acontecimientos destacados, la presencia o ausencia del AVE y la región donde se encuentra ubicada cada C. A. Tras las simulaciones realizadas, las variables explicativas “Costa” y “Esquí” no son influyentes en el aumento de desplazamientos que se producen entre las Comunidades Autónomas, a diferencia de las covariables “AVE” y “Festividades”. Adicionalmente, la variable explicativa “Region” influye levemente en el aumento de la movilidad entre las CC. AA.

Finalmente, se ha analizado la situación actual causada por el SARS-CoV-2, pues la pandemia, que ha alcanzado más de 3.764.651 contagios y 80.689 fallecidos, ha provocado un desplome en los movimientos en todos los continentes. Por este motivo, tanto el Gobierno de España como el de las CC. AA. establecieron restricciones en la movilidad, limitando los desplazamientos dentro de los municipios, provincia o Comunidad Autónoma, según fuese la fase del Estado de Alarma. Con la llegada de la campaña de vacunación, la situación ha mejorado. Sin embargo, sigue existiendo una situación de incertidumbre respecto a la evolución del SARS-CoV-2.

En síntesis, mediante la metodología aplicada de Redes Complejas se ha mostrado gráficamente la movilidad de residentes en España, pudiendo visualizar la red resultante, y se ha realizado un análisis de las variables que resultan determinantes en los desplazamientos entre Comunidades Autónomas.

REFERENCIAS BIBLIOGRÁFICAS

- Almende, B. & Thieurmel, B. (2019). *visNetwork: Network Visualization using “vis.js” Library* (R package version 2.0.9). <https://cran.r-project.org/package=visNetwork>
- Altmark, S. (2016). *Estudio del gasto en turistas de cruceros en Uruguay para la temporada 2010-2011 mediante el análisis de redes*. 1–34.
- Álvarez-Vaz, R. & Altmark, S. (2019). Study of Expenditure on Cruise Tourists in Uruguay for the 2010-2011 Season Using Network Analysis. *Cuadernos Del CIMBAGE*, 21, 27–64.
- Analytictech. *Cohesion Dyadic and Whole Network*.
- Barabási, A. & Albert, R. (1999). Emergence of Scaling in Random Networks. *Science*, 286(5439), 509 LP – 512. <https://doi.org/10.1126/science.286.5439.509>
- Bellman, R. (1958). On a routing problem. *Quarterly of Applied Mathematics: Vol. 16 (1)* (pp. 87–90). <https://doi.org/10.1090/qam/102435>
- Bender-deMoll, S. (2019). *ndtv: Network Dynamic Temporal Visualizations* (R package version 0.13.0). <https://cran.r-project.org/package=ndtv>
- Bojanowski, M. (2015). *intergraph: Coercion Routines for Network Data Objects*. <http://mbojan.github.io/intergraph>
- Borgatti, S. P., Everett, M. G. & Freeman, L. C. (2002). *UCINET 6*. Harvard, MA: Analytic Technologies.
- Cabrer, B., Serrano, G. & Simarro, R. (2009). Flujos migratorios y movilidad del capital humano. *Investigaciones Regionales*, 16, 5–41. <https://doi.org/28912353001>
- Caparrini, F. S. *Introducción a las redes complejas*. Último acceso fecha 4 Julio, 2021, de <http://www.cs.us.es/~fsancho/?e=80>
- Casleton, E., Nordman, D. & Kaiser, M. (2017). A local structure model for network analysis. *Statistics and Its Interface*, 10(2), 355–367. <https://doi.org/10.4310/SII.2017.v10.n2.a15>
- Cordón García, O. (2013). Redes Aleatorias. En *Redes y Sistemas Complejos* (p. 52). <https://sci2s.ugr.es/sites/default/files/files/Teaching/GraduatesCourses/RedesSistemasCompejos/Tema05-1-ModelosdeRedesAleatorias-13-14.pdf>
- Cordón García, O. (2013). Redes Sociales. En *Redes y Sistemas Complejos* (p. 92). <https://sci2s.ugr.es/sites/default/files/files/Teaching/GraduatesCourses/RedesSistemasCompejos/Tema03-RedesSociales-13-14.pdf>
- Cordón García, O. (2013). Algoritmos de Poda y Visualización de Redes. En *Redes y Sistemas Complejos* (p. 97).
- Cranmer, S. J. & Desmarais, B. A. (2011). Inferential network analysis with exponential random graph models. *Political Analysis*, 19(1), 66–86. <https://doi.org/10.1093/pan/mpq037>
- Cross Goma, A. (2018). Aplicación de análisis de redes complejas y caracterización de la red de metro de Madrid. En *Escuela Técnica Superior de Ingeniería*. Universidad de Sevilla.
- Csardi, G. & Nepusz, T. (2006). *The igraph software package for complex network research* (Complex Systems; p. 1695). InterJournal. <https://igraph.org>

- Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik 1959 1:1*, 1(1), 269–271. <https://doi.org/10.1007/BF01386390>
- El Diario. *Así avanza la vacunación: mapas y gráficos sobre su evolución en España y el mundo*. Último acceso fecha 4 Julio, 2021, de https://www.eldiario.es/sociedad/vacuna-covid-mapas-graficos-proceso-vacunacion-espana-mundo-julio-2_1_6782953.html
- El Mundo. (n.d.). *Nuevo récord del turismo español en 2019: 83,7 millones de llegadas y un 2,8% más de gasto / Economía*. Último acceso fecha 4 Julio, 2021, de <https://www.elmundo.es/economia/2020/02/03/5e37d7a9fc6c83a8068b45d2.html>
- Erdős, P. & Rényi, A. (1960). On the Evolution of Random Graphs. *Publication of the Mathematical Institute of the Hungarian Academy of Science*, 17–61. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.153.5943>
- Feria Toribio, J. M. (2006). La delimitación y organización espacial de las áreas metropolitanas españolas: una perspectiva desde la movilidad residencia-trabajo. *Ciudad Y Territorio Estudios Territoriales (CyTET)*, 42(164), 189–210.
- Fernández López, S., Pérez Esparrells, C. & Vaquero García, A. (2009). Movilidad internacional de la Universidad Española: Análisis regional e institucional del programa Sócrates-Erasmus. *Revista de Estudios Regionales*, 02, 143–172.
- Gandrud, C., Allaire, J. & Russel, K. (2017). *networkD3: D3 JavaScript Network Graphs from R* (R package version 0.4). <https://cran.r-project.org/package=networkD3>
- Gillespie, C. (2009). *The powerLaw package : Examples*. 1–10.
- Gillespie, C. (2015). Fitting Heavy Tailed Distributions: The {powerLaw} Package. *Journal of Statistical Software*, 64, 1–16. <http://www.jstatsoft.org/v64/i02/>
- González-Leonardo, M. (2020). Migraciones internas, inmigración exterior y emigración de españoles hacia el extranjero: un balance por nivel educativo. ¿Es España un país de baja movilidad? *Documents d'Anàlisi Geogràfica*, 66(3), 591. <https://doi.org/10.5565/rev/dag.596>
- Gutiérrez-Moya, E., Lozano, S. & Adenso-Díaz, B. (2020). Analysing the structure of the global wheat trade network: An ERGM approach. *Agronomy*, 10(12). <https://doi.org/10.3390/agronomy10121967>
- Hosterltur. *El turismo es el sector que más riqueza aporta a la economía española*. Último acceso fecha 4 Julio, 2021, de https://www.hosterltur.com/130893_el-turismo-el-sector-que-mas-riqueza-aporta-a-la-economia-espanola.html
- Hunter, D. R., Goodreau, S. M. & Handcock, M. S. (2008). Goodness of fit of social network models. *Journal of the American Statistical Association*, 103(481), 248–258. <https://doi.org/10.1198/016214507000000446>
- Hunter, D. R., Handcock, M. S., Butts, C. T., Goodreau, S. M. & Morris, M. (2008). ergm: A package to fit, simulate and diagnose exponential-family models for networks. *Journal of Statistical Software*, 24(3), 1–29. <https://doi.org/10.18637/jss.v024.i03>
- Hunter, D., Handcock, M., Butts, C., Goodreau, S. & Morris, M. (2018). *ergm: Fit, Simulate and Diagnose Exponential-Family Models for Networks* (R package version 3.9.4). The Statnet Projec. <http://www.statnet.org>
- Instituto Nacional de Estadística. *INEbase: Cuenta satélite del turismo en España*. Último acceso fecha 4 Julio, 2021, de https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=estadistica_C&cid=1254736169169&menu=ultiDatos&idp=1254735576863

- Instituto Nacional de Estadística. *INEbase: Estadísticas de migraciones*. Último acceso fecha 4 Julio, 2021, de https://www.ine.es/dynsgs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736177000&menu=ultiDatos&idp=1254735573002
- Kamada, T. & Kawai, S. (1989). An algorithm for drawing general undirected graphs. *Information Processing Letters*, 31(1), 7–15. [https://doi.org/10.1016/0020-0190\(89\)90102-6](https://doi.org/10.1016/0020-0190(89)90102-6)
- Latapy, M. & Magnien, C. (2008). Complex network measurements: Estimating the relevance of observed properties. *Proceedings - IEEE INFOCOM*, 2333–2341. <https://doi.org/10.1109/INFOCOM.2007.227>
- Lozano, S. (2021). Conceptos básicos de redes complejas. En *Métodos Cuantitativos Avanzados de Gestión*.
- Lozano, S. (2021). Introducción al análisis de redes complejas. En *Métodos Cuantitativos Avanzados de Gestión*.
- Lozano, S. (2021). Caracterización de redes complejas. En *Métodos Cuantitativos Avanzados de Gestión*.
- Lozano, S. (2021). Modelos de Redes Modelos de Redes. En *Métodos Cuantitativos Avanzados de Gestión*.
- Lozano, S. & Gutiérrez, E. (2018). A complex network analysis of global tourism flows. *International Journal of Tourism Research*, 20(5), 588–604. <https://doi.org/10.1002/jtr.2208>
- Morris, M., Handcock, M. & Hunter, D. (2008). Specification of Exponential-Family Random Graph Models: Terms and Computational Aspects. *Journal of Statistical Software*, 24(4), 1548. <https://doi.org/10.18637/jss.v024.a04>
- Opsahl, T. (2009). *Structure and Evolution of Weighted Networks*. University of London (Queen Mary College), London, UK. <http://toreopsahl.com/publications/thesis/>
- OXFAM. *Contaminación de la atmósfera: causas y soluciones | Ingredientes que Suman*. Último acceso fecha 4 Julio, 2021, de https://blog.oxfamintermon.org/contaminacion-de-la-atmosfera-causas-y-soluciones/?gclid=CjwKCAjwuIWHBhBDEiwACXQYsenros4z_hO-HwHQD7ME9bbAddG215fv--WBZhA20QYUoOY8lzfBoCXAEQAyD_BwE
- Palacios, D. & Villalobos, C. (2016). Redes académicas al interior de las escuelas chilenas: Un estudio exploratorio utilizando Modelos Exponenciales de Grafos Aleatorios (ERGM). *Revista Hispana Para El Análisis de Redes Sociales*, 27(2), 33–44. <https://doi.org/10.5565/rev/redes>
- Pedersen, T. (2021). *ggraph: An Implementation of Grammar of Graphics for Graphs and Networks* (R package version 2.0.5). <https://cran.r-project.org/package=ggraph>
- R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. <https://www.r-project.org/>
- RDocumentation. *fit_power_law function*. Último acceso fecha 8, 2021, de https://www.rdocumentation.org/packages/igraph/versions/1.2.6/topics/fit_power_law
- RTVE *Curva de casos y muertes por Coronavirus en España*. Último acceso fecha 4 Julio, 2021, de <https://www.rtve.es/noticias/20210702/curva-contagios-muertes-coronavirus-espana-dia-dia/2010514.shtml>
- RTVE. (2021). *La vacunación en el mundo, última hora*. <https://www.rtve.es/noticias/20210704/vacuna-coronavirus-mundo/2073422.shtml>
- Santana, E. *Machine Learning con R*. Último acceso fecha 4 Julio, 2021, de <http://apuntes-r.blogspot.com/>
- Snijders, T., Pattison, P., Robins, G., & Handcock, M. (2006). New specifications for Exponential Random Graph Models. *Sociological Methodology*, 36(1), 99–153. <https://doi.org/10.1111/J.1467-9531.2006.00176.X>

- Stanford Human Evolutionary Ecology and Health. *Introduction to ERGMs | Social Network Analysis for Anthropologists*. Último acceso fecha 7 Julio, 2021, de <https://eehh-stanford.github.io/SNA-workshop/ergm-intro.html>
- Statista. *Viajes internos y al extranjero de los españoles 2008-2020*. Último acceso fecha 4 Julio, 2021, de <https://es.statista.com/estadisticas/495296/viajes-internos-y-al-exterior-de-los-residentes-en-espana/>
- Tutte, W. T. (1963). How to Draw a Graph. *Proceedings of the London Mathematical Society*, s3-13(1), 743–767. <https://doi.org/10.1112/PLMS/S3-13.1.743>
- Vicente Cestero, E. & Mateos Caballero, A. (2018). *Data science y redes complejas : métodos y aplicaciones* (E. C. de E. R. A. SA (ed.); reimpressa).
- Watts, D. J. & Strogatz, S. H. (1998). Collective dynamics of ‘small-world’ networks. *Nature*, 393, 440–442. <https://doi.org/10.1038/30918>
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>
- Wickham, H., François, R., Henry, L. & Müller, K. (2021). *dplyr: A Grammar of Data Manipulation* (R package version 1.0.5). <https://cran.r-project.org/package=dplyr>
- World Tourism Organization. *2020: el peor año de la historia del turismo, con mil millones menos de llegadas internacionales*. Último acceso fecha 4 Julio, 2021, de <https://www.unwto.org/es/news/2020-el-peor-ano-de-la-historia-del-turismo-con-mil-millones-menos-de-llegadas-internacionales>

Anexo II: Código en R

```
#####Librerias#####
library(igraph)
library(statnet)
library(intergraph)
library(ergm)
library (network)
library(sna)
library(ggraph)
library(visNetwork)
library(threejs)
library(networkD3)
library(ndtv)
library(ggplot2)
library(readr)
library(dplyr)
library(tnet)
library(powerLaw)

#####Datos#####
dat=read.table(file="C:/Users/ANA/Desktop/TFG/Excel/Viajes.csv",header=TRUE, sep = ";"
, row.names=1, check.names=FALSE, nrow=18)
m=as.matrix(dat)

#Atributos
att<-read.table(file="C:/Users/ANA/Desktop/TFG/Excel/Atributos.csv",header=T, sep =
";", stringsAsFactors=FALSE, check.names=FALSE, nrow=18)

#Red
net <- network(m, vertex.attr = att,vertex.attrnames = colnames(att),directed = T, hyper
= FALSE, loops = FALSE, multiple = FALSE, bipartite = FALSE)
which.matrix.type(net)
which.matrix.type(as.matrix.network(net,matrix.type="adjacency"))

#Formato Red
```



```
ggraph(net) + geom_edge_link() + geom_node_point() + geom_node_text(aes(label=name),
  repel = TRUE)+ geom_edge_link(width=1, colour="black") + geom_node_point(size=4,
  colour="#00AFBB") + theme_graph()
```

```
#####Page Rank#####
```

```
# grafica objeto de grafo como ejemplo
```

```
plot(net, layout= layout.fruchterman.reingold, vertex.size=8, vertex.label.dist=0.4,
  vertex.color="#00AFBB", edge.arrow.size=0.5, edge.color="black",vertex.label.dist=0.9,
  vertex.label.color="black",vertex.label.dist=6,
  vertex.label.font=c(2),vertex.label.dist=0.9,
  vertex.label.color="black",vertex.label.dist=6, vertex.label.font=c(2))
```

```
# calcula page rank
```

```
pr.tmp <- data.frame(pr.value= page.rank(net)$vector)
```

```
pr <- data.frame(pr.desc=row.names(pr.tmp), pr.val = pr.tmp[,1])
```

```
head(pr[order(pr$pr.val,decreasing = T),],17)
```

```
# grafica interacciones usando valor page.rank para tamaños en grafico
```

```
V(net)$label.cex = 0.6 + pr$pr.val*3 # tamaño de letra segun page.rank
```

```
a <- layout_on_sphere(net)
```

```
plot(net,layout=a, vertex.size=pr$pr.val*100, vertex.label.dist=2,
  vertex.label.color="black", vertex.label.dist=2, vertex.label.font=c(2),
  vertex.color="#00AFBB", edge.arrow.size=0.1, edge.color="black")
```

```
b <- layout_in_circle(net)
```

```
plot(net,layout=b, vertex.size=pr$pr.val*100, vertex.label.dist=2,
  vertex.label.color="black",vertex.label.dist=2, vertex.label.font=c(2),
  vertex.color="#00AFBB", edge.arrow.size=0.1, edge.color="black")
```

```
plot(net,layout=layout.fruchterman.reingold, vertex.size=pr$pr.val*100,
  vertex.label.dist=2, vertex.label.color="black",vertex.label.dist=2,
  vertex.label.font=c(2),vertex.color="#00AFBB", edge.arrow.size=0.1, edge.color="black")
```

```
# GRAFICANDO PERSONA MAS IMPORTANTE
```

```
# Se agrega vector al objeto grafo, con valores de PageRank (pr$pr.val y grafo vienen de
  script anterior)
```

```
V(net)$page.rank.vector <- pr$pr.val
```

```
# Se identifica el valor máximo de page rank
```

```
max.pr <- max(V(net)$page.rank.vector)
```

```
# Se crea objeto g2 con grafo de persona más influyente
```

```
g2 <- subgraph.edges(net, E(net)[inc(V(net)[page.rank.vector>=max.pr])])
```

```
# gráfico del objeto g2
```

```

V(g2)$label.cex = 0.6 + V(g2)$page.rank.vector * 3 # tamaño de letra usando PageRank
plot(g2,          vertex.size=V(g2)$page.rank.vector*100,          vertex.label.dist=0.4,
vertex.color="#00AFBB",          edge.arrow.size=0.1,          vertex.label.dist=0.9,
vertex.label.color="black", vertex.label.dist=6, vertex.label.font=c(2))

#####Power Law#####
#Degree
deg.dist <- degree_distribution(net, cumulative=T, mode="in")
deg.dist
d<-degree(net)
d
m_bl=displ$new(d+1)###conpl$ si es continua la variable
est = estimate_xmin(m_bl)
estim<-m_bl$setXmin(est)
estim
plot(m_bl, bg="black",panel.first=grid(),family="serif",xlab= "Grado",ylab= "CDF")
lines(m_bl, col=2, lwd=2)
fit<- power.law.fit(d, "R.mle")
fit

#####ERGM#####
#edges
ergm0<-ergm(net~edges)
summary(net ~ edges)
summary(ergm0)
all.equal(network.density(net), plogis(coef(ergm0)[[1]]))
plogis(coef(ergm3)[[ 'edges' ]])

#mutual
ergm1<-ergm(net~mutual)
summary(ergm1)
summary(net~mutual)

#triangle
ergm2<-ergm(net~triangles)
summary(ergm2)
summary(net~triangles)
plogis(coef(ergm2)[[ 'triangles' ]])
mcmc.diagnostics(ergm2)
gof_1=gof(ergm2)
gof_1

```

```

plot(gof_1)

#edges+mutual+triangle
##no funciona bien
ergm3<-ergm(net~edges+mutual+triangles)
summary(ergm3)
summary(net~edges+mutual+triangles)

#####Simulaciones ERGM#####
#AVE#
ave <- net %v% "AVE"
summary(ave)
a1 <- ergm(net~ edges + nodefactor("AVE"))
summary(a1)
a2 <- ergm(net~ edges + nodeifactor("AVE"))
summary(a2)
a3 <- ergm(net~ edges + nodeofactor("AVE"))
summary(a3)

#Costa#
costa <- net %v% "Costa"
summary(costa)
b1 <- ergm(net~ edges + nodefactor("Costa"))
summary(b1)
b2 <- ergm(net~ edges + nodeifactor("Costa"))
summary(b2)
b3 <- ergm(net~ edges + nodeofactor("Costa"))
summary(b3)

#Esquí#
esqui <- net %v% "Esqui"
summary(esqui)
c1 <- ergm(net~ edges + nodefactor("Esqui"))
summary(c1)
c2 <- ergm(net~ edges + nodeifactor("Esqui"))
summary(c2)
c3 <- ergm(net~ edges + nodeofactor("Esqui"))
summary(c3)

#Festividades#
festividades <- net %v% "Festividades"

```

```

summary(esqui)
d1 <- ergm(net~ edges + nodefactor("Festividades"))
summary(d1)
d2 <- ergm(net~ edges + nodeifactor("Festividades"))
summary(d2)
d3 <- ergm(net~ edges + nodeofactor("Festividades"))
summary(d3)

#Region#
region <- net %v% "Region"
summary(region)
e1 <- ergm(net~ edges + nodefactor("Region"))
summary(e1)
e2 <- ergm(net~ edges + nodeifactor("Region"))
summary(e2)
e3 <- ergm(net~ edges + nodeofactor("Region"))
summary(e3)

#Desplazamientos
ergm_1 <- ergm(net ~ edges + mutual + nodefactor("AVE")+ nodefactor("Costa")+
nodefactor("Esqui")+ nodefactor("Festividades") + nodefactor("Region.Norte")+
nodefactor("Region.Sur"))
summary(ergm_1)

#Desplazamientos nodos salida
ergm_2 <- ergm(net ~ edges + mutual + nodeifactor("AVE")+ nodeifactor("Costa")+
nodeifactor("Esqui")+ nodeifactor("Festividades") + nodeifactor("Region.Norte")+
nodeifactor("Region.Sur"))
summary(ergm_2)

#Desplazamientos nodos entrada
ergm_3 <- ergm(net ~ edges + mutual + nodeofactor("AVE")+ nodeofactor("Costa")+
nodeofactor("Esqui")+ nodeofactor("Festividades") + nodeofactor("Region.Norte")+
nodeofactor("Region.Sur"))
summary(ergm_3)

#####GOF#####
gof1=gof(ergm_1)
par(mfrow=c(1,2))
plot(gof1)
gof1
mcmc.diagnostics(ergm_1)

```

